

Analysis of field investigative interviews of children conducted by specially trained police investigators

Trond Myklebust



**NORWEGIAN POLICE
UNIVERSITY COLLEGE**

Dissertation submitted for the degree of PhD

Department of Psychology
Faculty of Social Sciences
University of Oslo
2009

© **Trond Myklebust, 2009**

*Series of dissertations submitted to the
Faculty of Social Sciences, University of Oslo
No. 197*

ISSN 1504-3991

All rights reserved. No part of this publication may be reproduced or transmitted, in any form or by any means, without permission.

Cover: Inger Sandved Anfinsen.
Printed in Norway: AiT e-dit AS, Oslo, 2009.

Produced in co-operation with Unipub AS.
The thesis is produced by Unipub AS merely in connection with the thesis defence. Kindly direct all inquiries regarding the thesis to the copyright holder or the unit which grants the doctorate.

*Unipub AS is owned by
The University Foundation for Student Life (SiO)*

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	2
ABSTRACT	4
LIST OF PAPERS	6
INTRODUCTION	7
Interviewers` utterances	8
Childrens` age and gender	9
Interviewers` level of competence	9
Structure of the interview	10
FIIC in Norway	10
THE PRESENT DISSERTATION STUDIES	11
SUMMARY OF PAPERS	13
Summary of paper 1	13
Summary of paper 2	13
Aim of the study	13
Hypothesis	14
Method	14
Results	15
Summary of paper 3	15
Aim of the study	15
Hypothesis	16
Method	16
Results	17
Summary of paper 4	17
Aim of the study	17
Method	17
Results	18
GENERAL DISCUSSION	19
FUTURE PERSPECTIVES	25
CONCLUSIONS	25
REFERENCES	27
Paper I	
Paper II	
Paper III	
Paper IV	

ACKNOWLEDGEMENTS

First and foremost I would like to thank my supervisor and mentor Roald A. Bjørklund. Thank you for introducing me to the world of investigative interviewing and forensic psychology, always being available, with never ending effort and optimism trying to teach me how to write concise, and for always carefully considering my drafts and presentations. I have very much appreciated your wise comments and sound advice.

The specially trained interviewers in the studies should know that I greatly appreciate the trust and generosity they have shown by sending me their field investigative interviews of children for analysis.

I wish to thank the Director of Public Prosecution for their interest in the project, and I also want to thank my respective contact persons in each of the Police Departments.

My every-day workplace during the PhD period was the Norwegian Police University College (NPUC). Thanks for funding and supporting me through this period. And to all my great colleagues and friends at the Research Department: A sincere thanks to each and one of you for contributing to a motivating, creative and joyful environment.

Sincere appreciation is extended to my colleagues at the NPUC library. Your assistance in providing me with the articles and literature, has been the best thinkable. I also want to thank the librarians at the National library, Interlibrary Loans - Borrowing office (Innlånstjenesten), for your assistance and effective cooperation with the NPUC library.

It is joyful to do a PhD when you have a colleague in England, such as Gavin Oxburgh, genuine interested in your work, agreeing that the best reward of all is the lovely shuffle of filled pages in your hand. Thank you Gavin for your feedbacks on drafts, our discussions, your support, and for making me laugh. I look forward to continue our cooperation and research.

I want to thank several specialists, within their respective areas, for generously sharing their knowledge, discussions, and their sound advice and feedback. This includes of course all my colleagues at NPUC, The National Criminal Investigation Service (Kripas), my co-author on the first article: Laurence Alison (University of Liverpool), Ray Bull (University of Leicester), Julie Cherryman (University of Portsmouth), Becky Milne (University of

Portsmouth), Martine Powell (Deakin University), Kristin Kvigne & John Ståle Stamnes (Interpol).

My sincere thanks go to friends and family who have always supported me.

The doctoral process has been a lot about me and little about the people around me.....thank you Siren for your patience.

Oslo, June 2009.

Trond Myklebust

ABSTRACT

The research of field investigative interviews of children (FIIC) are mainly studies of individual factors by the children and interviewers, largely driven by a concern for non-contamination of the children's memory of the alleged offence in the interaction between the child and the interviewer. During the course of the present research, 100 videotaped FIIC, conducted by special trained interviewers, have been analysed and include some of the most prominent variables that are considered vital in the literature related to FIIC. The main objective of the thesis was to identify how specially trained interviewers conduct FIIC and study the factors facilitating the length of the interviewee's responses. Four central areas of interest were studied.

Focusing on the structure of the interviews, the first study sought to examine the standard of FIIC set against a structured interview model developed in England & Wales. The English model attempts to accord with psychological principles that lead to effective interviewing and so, if appropriately followed, such interviews should enhance the elicitation of more accurate material. Set against these principles of best practice, the analysis of the Norwegian FIIC indicated that a number of inappropriate and ineffective strategies appeared to be used in the police interviews. For example, in the literature, there is an agreement that open questions, as opposed to closed questions, are more likely to elicit longer accounts. The second study assessed the effect of interview training on police officers' use of open and closed questions in FIIC. In all interviews the mean number were 20 open and 217 closed questions, corresponding to an open-closed question ratio (OCR) of 1:10. Contrary to our hypothesis, analyses of variance (ANOVA) showed no main effect of competence. In the distribution of questions throughout the interview there were a descending number of open questions while the distribution of closed questions showed an inverted U-distribution with most frequent use of closed questions in the middle part of the interviews.

Focusing on the children's responses, the third study reviewed and analysed some of the most prominent variables considered to facilitate the interviewees' responses in the literature of FIIC. Of all the variables, the categories of the interviewers' utterances had most impact on the children's responses with the open questions eliciting the longest answers. The variable to follow was the children's age, with the oldest children yielding longer responses than the younger children to the open questions. Contrary to our hypothesis, the interviewers'

competence, childrens' gender, nor time had the expected impact on the length of the childrens' responses.

Finally, we wanted to analyse if some of the often sited variables affecting FIIC also affects the outcome of the case as judged by the prosecutors or the courts. One hundred FIIC were divided into one of the three different legal outcome possibilities in child sexual abuse cases: (i) insufficient evidence to proceed (IEP); (ii) convictions; or (iii) acquittals by the court. The results indicate that the courts decisions are affected by the length of the children's responses in their testimonies. Amongst the female interviewees older than 10 years, there were no cases of acquittals and the convicted cases were overrepresented. The childrens' response to open questions was found to be the main difference between the three FIIC outcomes. The childrens' verbal competence effect clearly demonstrates the importance for interviewers, the prosecution services, and the courts to conduct content analysis of the interviews, and improve their procedures in evaluation of FIIC.

Conclusion: In the analysis of the most prominent variables in FIIC the conclusion of the present thesis could be summarised around three main findings. Firstly, the interviewer's use of open questions was demonstrated to be the benchmark in FIIC, eliciting the longest answers compared to any of the other variables. Secondly, interviewing skills are not static. Even if the interviewers do have the knowledge and the research-based recommendations are endorsed, they are not implemented in the interviewers' way of conducting FIIC in practice. Thirdly, there is a verbal competence effect in court. The results indicate that the courts decisions are affected by the length of the children's responses to the interviewer's open questions.

LIST OF PAPERS

- 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.
- Myklebust, T., & Bjørklund, R. A. (2006). The effect of long-term training on police officers' use of open and closed questions in field investigative interviews of children (FIIC). *Journal of Investigative Psychology and Offender Profiling*, 3, 165-181.
- Myklebust, T., & Bjørklund, R. A. (in press). Factors affecting the length of responses in field investigative interviews of children (FIIC) in child sexual abuse cases. *Psychiatry, Psychology and Law*.
- Myklebust, T., & Bjørklund, R. A. (2009). The Child Verbal Competence Effect in Court: A Comparative Study of Field Investigative Interviews of Children (FIIC) in Child Sexual Abuse Cases. *Journal of Investigative Psychology and Offender Profiling*, 6 (2), 117-128.

"...It is a capital mistake to theorise before one has data. Insensibly one begins to twist facts to suit theories, instead of theories to suit facts..." (p.119)"

Doyle A.C. "A Scandal in Bohemia".

In (1989). *The original illustrated "Strand" Sherlock Holmes.*

INTRODUCTION

The aim of any investigative interview is to elicit the most accurate and detailed account of the alleged offence in a manner that does not place undue stress on the interviewee.

In most child sexual abuse (CSA) cases field investigative interviews of children (FIIC) plays a crucial role in these investigations and are considered to be the evidence in chief, primarily because other evidence is typically unavailable. Historically, interest in the testimonies of children by legal professions and social scientists has reflected specific judicial events, changes in the organisation of the judicial system, and the social conditions of the area (Ceci & Bruck, 1993; 1995). From the early research by Stern (1903/04) and Varendonck (1911) a considerable amount of studies have been conducted. A large body of research has been conducted under experimental conditions (Ceci, Loftus, Leichtman, & Bruck, 1994; Faller, 1996; Finnila, Mahlberg, Santtila, Sandnabba, & Niemi, 2003; Goodman, Quas, Batterman-Faunce, Riddelsberger, & Kuhn, 1994; Landström, Granhag, & Hartwig, 2007; Melinder, Scullin, Gunnerød, & Nyborg, 2005) while other researchers raise questions favouring field studies as the most convenient research method (Cederborg, Orbach, Sternberg, & Lamb, 2000; Cherryman, 2000; Davies, Westcott, & Horan, 2000; Fisher, Geiselman, & Raymond, 1987; Garven, Wood, Malpass, & Shaw, 1998; Hershkowitz, Horowitz, & Lamb, 2005; Korkman, Santtila, Blomqvist, & Sandnabba, 2008; Lamb, et al., 1996; Powell, 2002; Sternberg, et al., 1996; Thoresen, Lønnum, Melinder, Stridbeck, & Magnussen, 2006). The studies have generated a remarkable consensus about children's capacities and deficiencies. Although children can clearly remember incidents they have experienced, a variety of factors in FIICs could influence the children's willingness and ability to express the information, the interviewers' ability to elicit-, and the quality of the information provided in the FIIC.

Research conducted on FIIC are mainly studies of individual factors by the children and interviewers, largely driven by a concern for non-contamination of the childrens` memory of the alleged offence in the interaction between the child and the interviewer. In CSA cases the implicit assertion in FIIC is to gain long, detailed and accurate responses from the interviewed children. These accounts, unhampered by the questions or any personal influence from the interviewer, are also called free narratives (Dale, Loftus, & Rathbun, 1978; Fisher, 1995; Hershkowitz, Horowitz, Lamb, Orbach, & Sternberg, 2004; Hershkowitz, Lamb, Sternberg, & Esplin, 1997; Lamb, et al., 1996; Lamb, Sternberg, & Esplin, 1994).

Interviewers` utterances

From the literature, one of the most prominent variables related to FIIC are the utterances of the interviewer. Developed from laboratory research, there is an agreement that regardless of the age of the children, the cognitive capacity of the interviewee, or the length of the delay between events and the interview, open questions are more likely to elicit longer and accurate accounts (Dent, 1986, 1991; Dent & Stephenson, 1979; G. S. Goodman & Aman, 1990; G. S. Goodman, Bottoms, Schwartz-Kenney, & Rudy, 1991; Hutcheson, Baxter, Telfer, & Warden, 1995; Lamb & Fauchier, 2001; Oats & Shrimpton, 1991; Orbach & Lamb, 2001; Ornstein, Gordon, & Larus, 1992). In comparison, those children who answer closed questions tend to use single or fewer words, often with a limited number of response alternatives (Poole & Lamb, 1998; Richardson, Dohrenwend, & Klein, 1965). Although free narrative reports are longer and more detailed, they usually do not provide all the information the interviewer requires. To focus upon already known or previously revealed information from the offender, the interviewer tends to ask closed questions. Despite the agreement about the advantage of open questions, there still remain discrepancies over how to best describe types of questions and their respective characteristics. From a linguistic perspective the system of categorising questions could be done with a lexical perspective focusing on the *phrasing* of the question (e.g. *`selection`*, or *`yes/no`*) or with focus on the *function* of a question (e.g. *`invitation`*, *`leading`*, or *`suggestive`*). Oxburgh, Myklebust and Grant (submitted) found the divergent description in the literature to be mainly within the questions categorised by their function. To prevent this discrepancy the classification system developed by Richardson et al. (1965), focusing on the lexical phrasing of the questions, has been used in the present thesis.

Childrens` age and gender

Several studies indicate that the children's age and gender also affect the responses in FIIC. According to Foster-Cohen (1990) and Harley (2008), linguistic development of children is multi-dimensional and non-uniform with different aspects of communicative ability progressing at different rates for different children. Furthermore, as children grow older the length, informativeness, and complexity of their recall memory increases (Fivush, 1997, 1998; Poole & Lamb, 1998; Saywitz & Camparo, 1998; Schneider & Pressley, 1997) with the vocabularies of young children often being more limited and less descriptive than those of older children and adults (Brown, 1973; Morison, Moir, & Kwansa, 2000; Walker, 1999). Young children in their normal day-to-day interaction with adults rely on degrees of scaffolding to provide narratives and it can be argued that at some stages of their development the youngest children may be unable to produce a narrative without appropriate scaffolding. The danger of adult scaffolding in interview situations is contamination of memory. Some studies also indicate a gender effect of the interviewed children. Hershkowitz, Horowitz and Lamb (2005; Lamb, Hershkowitz, Orbach, & Esplin, 2008) examined all CSA investigations conducted in Israel between 1998 and 2002 and reported the interviewed girls to provide more details than the boys.

Interviewers` level of competence

Focusing on the importance of the use of open questions to elicit long narratives, the interviewers` level of competence was introduced as an important variable. Fisher, Geiselman and Raymond (1987) observed that the interviewers` level of competence affected the responses in interviews of adults. The authors recommended formal, scientifically based training of the police officers at the institutional level. Fisher and Geiselman (1992) denoted their interview method as the cognitive interview technique (CIT) and observed that interviewers trained in this model obtained longer responses in interviews of adults compared with interviewers without such training. They suggested that training programmes will progress most efficiently if interviews are divided into intensive short sessions, rather than one large session, keeping the presentation component at a short duration with the efficacy of the training programme profiting by extended feedback to the individual interviewers (Fisher & Geiselman, 1992; Fisher et al., 1987). Contrary, Lamb and his colleagues argues that long-time improvement in the quality of investigative interviews are observed only when the training is distributed over time with follow-up supervision and feedback to the interviewers in the period of the studies (Lamb, Sternberg,

Orbach, Esplin, & Mitchell, 2002; Lamb, Sternberg, Orbach, Hershkowitz, et al., 2002; Orbach, et al., 2000; Sternberg, Lamb, Davies, & Westcott, 2001).

Structure of the interview

Prior to the 1990s, the investigative procedure was simply not geared towards childrens' cognitive and linguistic abilities or to their vulnerability regarding suggestibility (Memon, Vrij, & Bull, 1998). The fact that the investigative procedure might intimidate children and make them less likely to be able to recount their alleged experience was completely overlooked. The structure of the interviews (i.e. how the interviews should be planned and conducted), has been discussed by several authors, with some arguing that the interviewer should follow highly structured sequences (APSAC, 2002; Fisher & Geiselman, 1992; Hindman, 1987; Home Office, 1992; 2002; 2007; Orbach, et al., 2000; Queensland Family Services, 1992; Steller & Boychuk, 1992; The National Crime Faculty, 1998; Yuille, Hunter, Joffe, & Zaparnuik, 1993). In these structured, best-practice interview protocols, there are four recommendations that are common: (i) the elicitation of free-narrative accounts from the child witness; (ii) a good rapport between the child and the interviewer; (iii) the provision of clear groundrules; and (iv) an open-minded approach considering alternative hypothesis.

FIIC in Norway

Between 1992 and 1994, the Bjugn case (Riksadvokaten, 1994) revealed a lack of sufficient knowledge among the interviewers on how best to conduct FIIC. The allegations in this case, and the criticism of the interviewers' competence and style of questioning, was similar to other highly publicised CSA cases in the past decades, such as the McMartin pre-school and Kelly Michaels case in the United States (State v. Buckley, 1990; Garven et al., 1998; State v. Michaels, 1988; State v. Michaels, 1993; State v. Michaels, 1994; Ceci & Bruck, 1995), the Orkney inquiries in the United Kingdom (Clyde, 1992) and the Roum case in Denmark (Nielsen, 2001). In Norway, the criticism of the Bjugn case led to a new regulation and guideline stating that FIIC should be conducted only by qualified interviewers with special training in interviewing children (Justisdepartementet, 1998). Based on an assumption that specialised training of police officers would elicit more information from the interviewed children compared to non-specific interview training, substantial resources and effort has been

made in Norway to increase the competence of police officers conducting FIIC. At the Norwegian Police University College (NPUC), the interview training is based on scientifically and research-based theory (PHS, 2002; 2003; Gamst & Langballe, 2004). The FIIC interviewers are the most qualified interviewers in Norway and have dedicated most of their professional careers to the area of investigative interviews of children.

THE PRESENT DISSERTATION STUDIES

The present studies have analysed videotaped FIIC conducted by specially trained interviewers. For ethical, practical and methodological reasons, most of the research on child witnesses has been conducted in experimental laboratories. Accuracy is much more difficult to establish in the field than in a laboratory analogue context because investigative interviewers rarely know what really happened in the alleged case. Instead, among others, three different approaches and strategies have been used to compare the laboratory results with the field studies. The first has been by comparing the responses provided by a suspect who has confessed to the alleged offence, with the responses provided by children in the accompanying FIIC to see whether there is convergence. Second, there are some cases with a video or audio recording of the actual abuse, which allow researchers to contrast that with the child's accounts. Third, by looking at the individual cases and FIIC for contradiction in the child's account. When a child contradicts him- or herself, we know s-/he said something that is inaccurate because both accounts cannot be true. We do not know which is accurate and which is inaccurate, but the number of contradictions can be counted as a proxy measure of accuracy.

Laboratory settings might not provide the police officers with the same sense of urgency and motivation often experienced in the field (Fisher, 1995). Experimental interviewers are rarely given extensive background information of the event, whereas in real investigations ideas and hypotheses about the incident, or antecedent perceptions of the interviewee, may influence the way the interviews are conducted (Mortimer, 1994; Oxburgh, Williamson, & Ost, 2006). Furthermore, the interviewers' awareness of being monitored may affect their regular way of conducting the interview. In total, this favours field studies when analysing how investigative interviews are conducted and the ecological validity of previous studies using experimental conditions is, therefore, questioned (Geiselman, Fisher, Cohen, Holland, & Surtes, 1986),

together with the use of personnel without any formal investigative or interviewer training (Fisher, Geiselman, Raymond, Jurkevich, & Warhaftig, 1987).

The main objective of the present dissertation is to analyse FIIC conducted by specially trained police investigators and to identify the factors facilitating the interviewee's responses in CSA cases.

SUMMARY OF PAPERS

Summary of paper 1

Focusing on the structure in the interviews, this was the first study of investigative interviews in Norway and sought to examine the state of FIIC set against a structured interview model developed in England & Wales. The structured PEACE model (mnemonic for Preparation and planning, Engage and explain, Account, Closure, & Evaluate) are based on psychological principles that lead to effective interviewing and so, if appropriately followed, should enhance the elicitation of accurate material (Milne & Bull, 1999; Clarke & Milne, 2001). A sample of eleven interviews was divided into 5-minute sequences. Questions in each of the sequences were then classified according to Richardson, Dohrenwend and Klein's (1965) definition of open and closed questions. The amount of talking time of each conversant was also measured. Finally, other significant features were extracted for qualitative examinations. The analysis revealed significant discrepancies from the PEACE model with a variety of inappropriate techniques, known to promote ineffective interviews, being employed. The results conclude that the interviewers spent as much time talking as the interviewees. There was a significant majority of closed questions throughout the interviews. At the outset the interviewer asked many identification questions and then sought to confirm these through yes-no questions towards the end. There was hardly any variation in the interview patterns regardless of the children age groups. An obvious limitation of the results was the small sample size. However, there was nothing unusual or distinct about the interviews and a similar trend was found in the later studies.

Summary of paper 2

Aim of the study

Focusing on the interviewers' questions, the purpose of the study was to assess the effect of long-term training on police officers' use of open and closed questions in investigative interviews of children. Fisher, et al., (1987) introduced the open-closed question ratio (OCR) in order to characterize the proportion of open questions in relation to the number of closed questions in an interview. They argued that a high portion of open questions, described as a high OCR, would enhance the quality of the information received in a police interview and introduced the interviewers level of competence as an important variable generating a higher OCR. The training of Norwegian interviewers emphasise the use of structure and sequences

in the interview analogues to the training of police officers in Europe and America (Walker & Hunt, 1998). Structured interviews might be characterized by police officers using open questions in the beginning of the interview followed by closed questions to follow up information already given by the interviewee as free narrative. Contrary, other authors studying child-adult communication argue that a dialogical model accounting for a more active child (i.e. one who introduces terms of references and the topics to be discussed during the interview), provides a more realistic description of the structure of an ordinary field interview. Dialogue based interviews do not display a similar structure, reflecting that the child and police officer mutually engage with each other during questions and answers (Clark & Brennan, 1991; Manning, 1988; Watzlawick, Bavelas, & Jackson, 1967). The studies of methods used in police interviews do not have common definitions of which part of the interviews should be included. All the questions in the complete interview might be essential for the interviewer's collection and validation of the information given by the interviewee. Consequently, the total lengths of the interviews were analysed.

Hypothesis

By analysing 100 FIIC undertaken by the most qualified interviewers in Norway we expected that extensive training would generate more frequent use of open questions in FIIC, resulting in a higher OCR, compared to less trained police officers. Secondly, we hypothesis that the most trained police officers will use more open questions during the first phases of the interviews compared to the less trained police officers.

Method

At the Norwegian Police University College, education in the investigative interviewing of children could be divided into two groups based upon the educational level. Police officers at *Level 2* have substantial more theoretical and practical training compared with officers at *Level 1*. Fifty police officers from each group participated in the study with a video recorded FIIC. The transcribed interviews were electronically digitised and the interviewers' questions were classified according to the content dictionary first developed by Richardson, Dohrenwend and Klein (1965). The numbers of open and closed questions were calculated in each interview. Differences between the groups (Level 1/ Level 2) and the kind of questions

(open/closed) were analysed in a mixed between – within subject ANOVA design. To study how the pattern of questioning changed over time, the videos were used to divide the interviews into three equally long tertiaries. The numbers of open and closed questions were calculated for each of the three tertiaries, generating six cell means for each interview. Gender distribution among officers and children were analysed by chi-square non-parametric test and the other differences between the two groups were analysed by student's t-tests.

Results

This was the first study to assess the effect of long-term training on police officers in Norway. In all interviews, the mean numbers were 20 open and 217 closed questions, corresponding to an OCR of 1:10. Contrary to our prediction there was a non-significant difference in the use of open and closed questions between the two groups of interviewers. The most competent interviewers used 22 open questions and the less competent used 19 open questions and the difference was not statistically significant. ANOVA showed a significant effect of question but no main effect of competence. The number of open questions in the last tertiary was about half of the number as in the first tertiary. The distribution of closed questions showed an inverted U-distribution, with most frequent use of closed questions in the middle part of the interviews.

Summary of paper 3

Aim of the study

Empirical studies of FIIC indicate a strong correlation between the number of words in the childrens' responses and the information contained in the utterances (Lamb, et al., 1996; Sternberg, et al., 1996). In this study, 100 FIIC were analysed to reveal central variables facilitating the interviewees' responses in CSA cases. The literature indicates that several variables affect the children's responses in FIIC. However, the different studies only include a limited number of the above mentioned variables. The purpose of the study was to include all the variables. Focusing on the interviewer, the interviewed child and the context of the interview, the following variables were included; the childrens' age and gender, the interviewers' utterances, the interviewers level of competence, and duration of the interview.

Hypothesis

We assumed that all these variables would affect the length of the children's responses. According to the literature we hypothesised that the competence of the interviewer and the category of the questions would have the most impact on the length of the responses, while the children's age and gender, and the portions of the interview would have less influence on the responses in FIIC.

Method

The study included one hundred FIIC conducted by separate specially trained police interviewers. Each interviewer met the following characteristics: (i) the interviewed child had to be between 6 and 16 years of age (school-age); (ii) no deviation in the child's linguistic development or skills; (iii) information about the interviewers' gender, age, training level (low/high), and the number of practiced FIIC; (iv) information about the age and gender of the interviewed child; (v) the interview had to be fully video recorded; (vi) the interview should be completely transcribed; and (vii) the FIIC should be the only interview of the child in the case.

By median split method, the children were divided into one of two age groups; Young (6-10 yrs) and Old (11-16 yrs). The competence of the interviewers consists of two aspects, one is the theoretical competence and the other is the practical competence. The combination of low/high theory (T-/T+) and low/high practice (P-/P+) gave four competence groups. The transcripts of the interviews were analysed with each category of interviewers' utterances measured against each child witnesses' respective answer. Each of the interviewers' utterances was classified according to the categories by Richardson et al. (1965). The mean number of words in each of the six categories of interviewer utterances, and in the corresponding answer from the child, was calculated for each of the tertiary in all interviews and used as cell means for subsequent statistical analysis. The numbers of words in the children's answers were used as the dependent variable.

Results

We observed the mean numbers of words in the children's answers dependent upon the six different categories of interviewer utterances. ANOVA revealed the responses in the oldest group of children to be significantly longer than in the youngest group, with post hoc showing a significant difference between the two age groups only to the open questions. A temporal effect was demonstrated for the six categories, affecting only the number of words in answers to open questions, where the number of words in the first tertiary was significantly longer than in the second and third tertiary. Analysis of the interviewers' level of competence revealed a difference in the number of words in the children's answers to the open questions, with children interviewed by interviewers with high theoretical and low practical competence producing significant longer answers than interviewers with high theoretical and high practical competence. Of all the variables, the categories of the interviewer utterances had most impact on the length of the children's responses. Derived from the children's responses to the open questions, the variable to follow was the children's age, with the oldest children yielding longer responses than the younger children. Contrary to our hypothesis, the competence of the interviewers did not have the expected impact on the length of the responses. Neither did the children's gender, nor time as a temporal effect in the FIIC.

Summary of paper 4

Aim of the study

A review of several studies indicate that factors related to the category of questions, the competence of the interviewer, and the age and gender of the interviewed child, all affect the FIIC. The purpose of this study was to analyse if the same factors also affect the outcome of the FIIC, either as evaluated by the prosecutor as insufficient evidence to proceed or by the court judged as acquittals or convictions.

Method

The study included 100 FIIC conducted by separate police officers were included in the study. Each interviewer met the same individual characteristics as in paper 3. In addition the case should be completed either as 'IEP' or proceeded by the court. The outcome of the proceeding in the court should be known. Each of the interviewers' utterances was classified in

according to the categories first developed by Richardson et al. (1965). The mean number of words in the children's answers to the six categories of interviewer utterances was calculated in all interviews and used as cell means for subsequent statistical analysis. By median split half method the children were divided into the 'young' children group (below 11 yrs) and the 'old' children group (above 10 yrs). Including gender as an independent variable gave four groups of children. For the interviewers the combination of low/high theory and low/high practice gave four competence groups. The lengths of the FIIC were divided into short and long interviews.

Results

We found that none of the old girls were represented in the acquitted outcome and that they were overrepresented in the convicted outcome. The fact that the court always judges the cases with old girls as convicted outcomes has never been observed before. ANOVA revealed a significant main effect of the three FIIC outcomes (IEP/acquitted/convicted). The number of words in the children's answers in the 'convicted' category was significantly higher than the two other FIIC outcomes. The number of words in the children's answers to open questions was significantly different from all the other categories with the convicted outcome being 1.7 and 1.5 times longer than the acquitted and IEP outcome, respectively. ANOVA did not show any effect of the interviewers' competence, neither any effect of the four groups of children on the open responses, nor for the length (long/short) of the interviews.

GENERAL DISCUSSION

The FIIC is a unique conversational context due to the roles assumed by the participants, the style and content, the participants themselves and the motive for the conversation. In contrast to everyday conversations children in FIIC have to talk to unfamiliar adults about potentially sensitive topics in a very formal setting. Children's ability to be informative experts about their own experiences, like their reliability and suggestibility, is influenced by a number of variables, most significantly the ways in which the interviewers steer the conversation (Lamb & Brown, 2006).

As demonstrated in the present studies, the quality measurement in all FIIC is the form of questions used by the interviewer, with open questions eliciting the longest answers compared to any of the other variables. Study one demonstrated that the interviewers did not follow best practice guidelines or instructions, eliciting free narratives from the interviewed children. Instead a variety of inappropriate techniques, known to promote suggestibility and ineffective communication were employed. As demonstrated in the qualitatively analysis, the interviews starts with the interviewer presenting seven or more information units such as; date and time for the interview, role of the police and the interviewer, the interview room and technical descriptions, people monitoring the interview, the importance of the child telling the truth, talking loud and clear etc. After this the interviewer follow up with a series of questions where the interviewer already know or should know the answers, such as; the child's address, members of the family, name of school, which class they attend and name of favourite teacher and subject. This is often followed by questions regarding the child's leisure activities. From the very start of the interviews there are very little opportunities given to the interviewed child's free narrative, indicating a lack of the interviewers awareness of the basic communicative principles in FIIC. An explanation for why the interviewers are giving this introduction, with an overwhelming amount of information, could be that section 12 of the Norwegian regulations (Justisdepartementet, 1998) contains more or less a checklist of *what* the interviewer should inform and admonish the child to do.

"Section 12 How to conduct the interview

Prior to the interview, the witness shall as a general rule receive information about who is going to attend the interview and where they will be located. As a general rule, the witness shall also be told that the interview will be videotaped or recorded on tape.

The person who conducts the interview must admonish the witness to tell the whole truth, cf. section 128 of the Criminal Proceedings Act. The admonition must be adapted to the age and mental development of the witness.

The interview must be adapted to the age and mental development of the witness, and to the circumstances in general..." (Justisdepartementet, 1998, p.10) [my translation]

This in contrast to the guideline in England & Wales (Home Office, 2007) where the interviewer get research based support in *how* the interview should be conducted. The Norwegian regulation (section 12) states that the interview must be adapted to the age and mental development of the interviewed child. Despite the importance, there are no procedures for the interviewers or court in how to evaluate the interviewed child's verbal competence, conceptual understanding and maturity. I question on what background the judges, in charge of the interviews, in most of their written reports from the monitored FIIC states "the child seemed developmentally mature to his/her age".

In all the interviews the children were told the importance of telling the truth. In England & Wales this is done by presenting the child for a scenario, adapted to their age, about a boy or a girl at their own age that does something wrong while playing, and tries to get away with it by denying the fact when confronted by their parents. The interviewers ask open questions to get the children to analyse the situation and why the child in the story acted as he or she did. From the children's answers the interviewer's introduces the importance of being honest and telling the truth in the present interview situation. In this way the interviewers get the children to reflect, explain, and give long statements at an early stage of the interview. Contrary, the Norwegian interviewers informed the children the importance of telling the truth. The majority of the interviewers asked a closed yes-no question, such as: "Do you know what it is to tell the truth". When asked, all the children answered "yes" to this question. Additionally, most interviewers asked concrete closed questions about the truth, such as:

Interviewer: "If I said that your trouser is green, would that be a truth or a lie?"
Child: "A lie."
Interviewer: "Yes, because your trouser is red?"
Child: "Yes."
Interviewer: "Now, I know that you know the difference of truth and lie."
Child: "Yes."

Once again this emphasises the importance of having an interviewer who has a great range of questions at her or his disposal and is properly prepared to present the interviewee with varied

open questions. To achieve this, the interviewer has to collect information about the child outside the interview situation, so the interviewer becomes able to assess the child's perception of reality and the reliability of the information provided by the child. Our findings in the first three papers support Egan (2002), who argues that those who ask closed questions find themselves using more and more questions; in other words, one closed question begets another.

As demonstrated in study three the average length of children's responses increases as children grow older. This is the challenge with interviewing young children. Their responses is short, which means that the ball is falling back on the interviewer's court very fast, and the interviewer have to come up with another question very quickly. If the interviewer have not worked hard and prepared very well s/he will probably fall back on asking a closed instead of an open question. As demonstrated in the present studies the problem is not so much the children's inability to respond to open questions, but it is actually the interviewer's inability to come up with a good open question in response to a short answer from the child. And because it is short it is often not a particularly clear narrative, and the interviewer becomes unsure, led a little bit a drift and not confident in what to do. Uncertainty for an investigative interviewer is the one thing that almost guarantees him or her to ask closed questions (Egan, 2002; Shepherd, 2007). So reducing that uncertainty is perhaps one of the most important goals of the training and ongoing practice, in providing people with a set of alternatives that they can use in that situation.

In the literature the training of FIIC interviewers has been studied within two traditions. One tradition has been the short-term training (Fisher & Geiselman, 1992) and the other has been long-term training tradition (Lamb, Sternberg, Orbach, Esplin, et al., 2002; Lamb, Sternberg, Orbach, Hershkowitz, et al., 2002). We hypothesized that the extensive long-term training, including both several theoretical courses and a large number of FIIC, would improve the competence of the police officers interviewing skills. Further, subsequently increase the use of open questions and longer responses from the interviewed children for the most trained and experienced police officers. Contrary, we found no difference in the use of open questions between the interviewers at different competence levels. For the children's responses, the interviewers with high theoretical knowledge and low practical competence revealed nearly double the length as the interviewers with both high theoretical and practical knowledge. Several explanations to this unexpected result have been considered. It could be that

interviewers with low practical competence are more prepared in their interviews and have based their plan on their theoretical knowledge, while the interviewers with high level of practice are more distanced to their theoretical competence, leaning more on their practical experience from their previous interviews. Some authors have argued the importance of systematic supervision of interviewers in the field (Lamb, Sternberg, Orbach, Hershkowitz, et al., 2002), and that what happens at the local police station after the formal theoretical training is as important as the initial training itself. We have not included any evaluation of the training programmes in investigative interviewing at the Norwegian Police University College. However, both the law and the training programmes for police officers conducting FIIC focus on the use of scientifically based interviewing techniques and strongly emphasises the interviewers to achieve the child's free narrative by the use of open questions (Act of 22 May 1981 No. 25; Myklebust, 2005). The problem is this gap between knowledge and practice. Interviewers do not implement best practice even when they know them and think they are complying. Our results questions whether the training programmes could be improved. We advocate that supervision may generate a group atmosphere of co-operation between several police officers, which concentrates on performing optimal interviews. Several studies (e.g. Cannon-Bowers, Salas & Converse, 1993; Hardin & Higgins, 1996) have indicated that team-based cooperation generates better communication strategies compared to individual activity. At an organisational level, the question of effective and optimised learning and training has attracted substantial attention in general work and organisational psychology (Argyris & Schön, 1974, 1978; Deutsch, Coleman, & Marcus, 2006; Garvin, 2000). Procedures, routines and methods used successfully by other professional organisations (Ericsson, Charness, Feltovich, & Hoffman, 2006; McGregor, 1960) should in a higher degree be evaluated and implemented within the police. Such organisations, with systematic training and experience in evaluation of their assignments and employees, are in several sectors including aviation; with evaluations of training, authorisations and re-authorisations of pilots (RNoAF, 2004), and within the health service, oil sector and shipping. All with specific requirements in training and authorisations.

Project-management and use of teams are already used with success within some parts of the police organisation. It is the model used by the National Criminal Investigation Service (Kripos) both in cases run by them and in their assistance to the 27 local police districts in Norway. At the case level the "Lommemannsaken" (internationally presented and known as 'the Pocket-Man Case') (Skjønsfjell, 2008) could serve as an example. The name was

developed from the suspect's modus operandi (1975-2006). He cut holes in his trouser pocket and asked young boys to help him get something (keys, coins) out of his pocket. Instead, the unsuspecting boys would end up touching the offender's penis. While many of the boys managed to run away some of the boys were forced to perform oral sex and other sexual activities on him. Operating in different police districts all over Norway resulted in the evidence in most cases being found insufficient to proceed to court. It was first after several years some investigators saw the modus and link, resulting in a group of investigators being designated to run what became a national investigation, reopening cold cases back to 1975. According to the investigative team many of the FIIC in the case were of poor tactical quality and many of the children had to be re-interviewed. The information from the interviews conducted by a team of interviewers, dedicated to the case, was considered to be quantitatively- and qualitatively better, than most of the interviews conducted by the single interviewers in each of the involved police districts (Skjønsvoll, 2008). The positive effect of teams and project management is also known and used by the police in other fields, such as narcotics, economic crimes, and crime prevention. For most interviewers the preparation for FIIC is a lonely process. Much energy and preparation time is used in organise for all the monitoring participants being informed and meeting at the same time for the interview. The judge, prosecutor, defendant's lawyer, the child's lawyer, and (if necessary) the technician running the recording equipment, is the rest of the team in most FIIC. Their main focus and representation is around the interviewed child. To assist the interviewer s/he should have a dedicated monitoring co-interviewer focusing on the interviewer's performance. During the pause in the end of the interview, while the interviewer is consulting the judge and the rest of the legal representatives in the monitoring room (challenging the lines of inquiry or presented statements), the trained co-interviewer could assist the interviewer by coming up with alternative questions. In this way it will be easier for the interviewer to open up, and implement more open questions in the FIIC.

Taken together, the recommendation for the police service is that the guided interview training for police officers should be as similar as the real FIIC. It should be conducted by interviewers working in the investigative team, in the same environment and under the same conditions as their real investigative interviews. In the real FIIC the interviewers should be assisted by peer consultants, monitoring the FIIC together with the legal representatives.

The PEACE model are based on psychological principles that lead to effective interviewing and so, if appropriately followed, should enhance the elicitation of accurate material. PEACE is an acronym identifying the steps of the model which provide an interview structure; Planning and preparation, Engage and explain, Account, Closure, & Evaluation (for details see; Milne & Bull, 1999; Clarke & Milne, 2001). Following a structured interview model such as the PEACE model of interviewing, the officers have to realise the importance of the P (planning) and E (evaluation). In the evaluation processes of their conducted interviews, individual and ongoing group supervisions have to play a major role in the implementation of the interviewers' knowledge into their FIIC practice.

The results demonstrated that the children were verbally competent throughout the whole interview. Children in the oldest age group (11-16 yrs) gave longer answers than the younger children (6-10 yrs). The open questions generated the longest responses in both groups of children. For the interviewers, this means that the same questions will have the same effect on the length of the answers between the two age groups. This result is of operational importance to the interviewers in their planning and preparation of the FIIC. The verbal competence of the children is not only important for the interviewers, it is also having an effect on the process of the cases in court.

In study four, the findings indicate that the courts' decisions are affected by the length of the children's responses in their testimonies. Among girls older than 10 years, there were no cases of acquittals and an over-representation of convictions. In the categories of open questions, the responses were about twice as long in the convicted- compared to the acquitted outcome.

This finding, denoted *the children's verbal competence effect*, has never been reported before. It raises interesting questions regarding what characterises the content of the interviews leading to conviction compared to acquittals. Further, which interpretation and evaluation methods are used by the members of the courts in their decision-making processes when presented for the videotaped FIIC. For the investigators the result operationally demonstrates the importance of open questions and free narratives in FIIC, indicating; "the longer the answer, the higher the likelihood for a convicted outcome of the trial". The verbal competence effect clearly demonstrates the importance for interviewers, the prosecution services, and the courts to conduct content analysis of the interviews, and improve their procedures in evaluation of FIIC. It is my clear opinion that in the years to come, more focus on this is needed both by the academics and the practitioners in the field of investigative interviewing.

FUTURE PERSPECTIVES

The reported results in this thesis opens up for at least two interesting approaches. In the work with the project a large database of FIIC has been developed. One approach could be to continue the content analysis of the FIIC with other variables than the ones in the present studies. Based on the findings in study four, analysis of obtained investigation relevant information would be of interest. Another direction could be an experimental approach to find more effective ways to conduct FIIC. More concrete, the experimental studies could be designed in the development of effective training methods for the interviewers. Such approach could be to compare different organisational models. This could be done as pre-post studies at the same police department or by comparing police departments that have implemented different models in their way of conducting FIIC.

CONCLUSIONS

Whether or not an FIIC results in free narrative from the interviewed child says a great deal about the conversation, and in the cause of that it is the interviewer's responsibility to make sure that the communication with the interviewed child goes well. The interviewer is the one who ultimately determines whether the FIIC is going to be an effective information gathering exercise. In the analysis of the most prominent variables in FIIC the conclusion of the present thesis could be summarised around three main findings.

Firstly, the interviewers' use of open questions was demonstrated to be the quality measurement in FIIC, eliciting the longest answers compared to any of the other variables. The OCR in study two and three was 1:10. Obviously police officers who conduct FIIC need to use more open questions in their interviews. And the only way to do that is for the interviewers to plan, practice, evaluate and continue to evaluate their interview practices.

Secondly, interviewing skills are not static. Even if the interviewers do have the knowledge and the research-based recommendations are endorsed, they are not implemented in the interviewers' way of conducting FIIC in practice. The studies demonstrate that the interviewers, however skilful, get rusty and the interviewers with most practical competence are less effectively getting the shortest responses from the interviewed children. In the suggested improvement of the interviewers' FIIC, two models have been presented. The first,

at the individual interviewers' level, suggested improvement in the training programmes by including supervision of the interviewers with the interviews conducted similar to the real FIIC, at the local police stations. The second suggested model was at the organisational level. By organising teams of investigators working together with dedicated peer consulting co-interviewers, monitoring and focusing on the interviewer's performance, coming up with alternative questions during the FIIC.

Thirdly, there is a verbal competence effect in court. The results indicate that the courts decisions are affected by the length of the children's responses in their testimonies. The children's responses to open questions was found to be 1.9 and 2.3 times longer in the convicted cases compared to acquittals or cases with insufficient evidence to proceed to court. This emphasises the importance of open questions in the interviews, and for both the interviewers and members of court to implement content analysis of the interviews and improve their evaluation procedures of FIIC.

Contrary to the popular myth, children do respond when they are questioned. The key point I would like to emphasise is:

Whether or not children give a long response depends, not only on their limitation, but also on the interviewers' limitations and the interviewers' ability to, in a sense, take advantage of the children's strength and get them to function at the best of their abilities. Whereas, what I often saw in the analysis for this thesis were children functioning below their best effort. Interviewers, teaching institutions, academia and research environments, legislative authorities, and the court of law, are all responsible for improving the quality of procedures in connection with FIIC. The results in the present thesis show that there is a need for improvement.

REFERENCES

- Act of 22 May 1981 No. 25 *Criminal Procedure Act Concerning Legal Procedures in Criminal Cases [Straffeprosessloven - Lov av 22.mai 1981 nr.25 om rettergangsmåten i straffesaker]*.
- APSAC (2002). *Guidelines on investigative interviewing in cases of alleged child abuse*. Elmhurst, IL. : APSAC National Operations Manager.
- Argyris, C., & Schön, D. A. (1974). *Theory in Practice: Increasing Professional Effectiveness*. San Fransisco: Jossey-Bass.
- Argyris, C., & Schön, D. A. (1978). *Organizational Learning. A Theory of Action Perspective*. Reading, Massachusetts: Addison Wesley.
- Brown, R. (1973). *A first language*. Cambridge, MA: Harvard University Press.
- Cannon-Bowers, J. A., Salas, E., & Converse, S. (1993). Shared Mental Models in Expert Team Decision Making. In N. J. Castellan (Ed.), *Individual and Group Decision Making* (pp. 221-245). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Ceci, S. J., & Bruck, M. (1993). Suggestibility of the child witness: a historical review and synthesis. *Psychological Bulletin*, 113(3), 403-439.
- Ceci, S. J., & Bruck, M. (1995). *Jeopardy in the Courtroom. A scientific analysis of children`s testimony*. Washington: American Psychological Association.
- Ceci, S. J., Loftus, E., Leichtman, M. D., & Bruck, M. (1994). The role of source misattributions in the creation of false beliefs among preschoolers. *International Journal of Clinical and Experimental Hypnosis*, 42(4), 304-320.
- Cederborg, A. C., Orbach, Y., Sternberg, K. J., & Lamb, M. E. (2000). Investigative interviews of child witnesses in Sweden. *Child Abuse & Neglect*, 24(10), 1355-1361.
- Cherryman, J. (2000). *Police investigative interviewing: Skill analysis and concordance of evaluations*. University of Portsmouth, Portsmouth.
- Clark, H. H., & Brennan, S. E. (1991). Grounding in Communication. In L. B. Resnick, J. M. Levine & S. D. Teasley (Eds.), *Perspectives on Socially Shared Cognition*. Washington, D.C.: American Psychological Association.

- Clarke, C., & Milne, R. (2001). *National Evaluation of the PEACE Investigative Interviewing Course. Police Research Award Scheme. Report NO:PRAS/149*. London: National Crime Faculty.
- Clyde, J. J. (1992). *The Report of the Inquiry into the Removal of Children from Orkney in February 1991*. Edinburgh: HMSO.
- Dale, P. S., Loftus, E. F., & Rathbun, L. (1978). The influence of the form of the question on the eyewitness testimony of preschool children. *Journal of Psycholinguistic Research*, 7(4), 269-277.
- Davies, G. M., Westcott, H. L., & Horan, N. (2000). The impact of questioning style on the content of investigative interviews with suspected child sexual abuse victims. *Psychology, Crime & Law*, 6, 81-97.
- Dent, H. R. (1986). An experimental study of the effectiveness of different techniques of questioning mentally handicapped child witnesses. *British Journal of Clinical Psychology*, 25, 13-17.
- Dent, H. R. (1991). Experimental studies of interviewing child witnesses. In J. Doris (Ed.), *The suggestibility of children's recollections* (pp. 138-146). Washington, DC: American Psychological Association.
- Dent, H. R., & Stephenson, G. M. (1979). An experimental study of the effectiveness of different techniques of questioning child witnesses. *British Journal of Social and Clinical Psychology*, 18, 41-51.
- Deutsch, M., Coleman, P. T., & Marcus, E. C. (2006). *The Handbook of Conflict Resolution. Theory and Practice* (2 ed.). San Francisco: Jossey-Bass.
- Doyle, A. C. (1989). *The original illustrated "Strand" Sherlock Holmes*. Ware: Wordsworth.
- Egan, G. (2002). *The Skilled Helper: A Problem-Management & Opportunity-Development Approach to Helping* (7 ed.). California: Brooks/Cole.
- Ericsson, K. A., Charness, N., Feltovich, P., & Hoffman, R. (2006). *The Cambridge Handbook of Expertise and Expert performance*. New York: Cambridge University Press.
- Faller, K. C. (1996). Interviewing Children Who May Have Been Abused: A Historical Perspectiv and Overview of Controversies. *Child Maltreatment*, 1(2), 83-95.

- Finnila, K., Mahlberg, N., Santtila, P., Sandnabba, N. 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. (1995). Interviewing victims and witnesses of crime. *Psychology, Public Policy and Law*, 1(4), 732-764.
- Fisher, R. P., & Geiselman, R. E. (1992). *Memory-enhancing techniques for investigative interviewing. The cognitive interview*. Springfield, IL: Charles C. Thomas Publisher.
- Fisher, R. P., Geiselman, R. E., & Raymond, D. S. (1987). Critical analysis of police interview techniques. *Journal of Police Science and Administration*, 15(3), 177-185.
- Fisher, R. P., Geiselman, R. E., Raymond, D. S., Jurkevich, L. M., & Warhaftig, M. L. (1987). Enhancing enhanced eyewitness memory: Refining the cognitive interview. *Journal of Police Science and Administration*, 15(4), 291-297.
- Fivush, R. (1997). Event memory in early childhood. Studies in developmental psychology. In N. Cowan (Ed.), *The development of memory in childhood*. Hove: Psychology Press. Taylor & Francis.
- Fivush, R. (1998). Children's recollections of traumatic and nontraumatic events. Trauma, memory, and suggestibility in children. *Development and Psychopathology*, 10, 699-716.
- Foster-Cohen, S. (1990). *The communicative competence of young children: a modular approach*. London: Longman.
- Gamst, K. T., & Langballe, Å. (2004). *Barn som vitner - En empirisk og teoretisk studie av kommunikasjon mellom avhører og barn i dommeravhør. Utvikling av en avhørmetodisk tilnærming [Children as witnesses - An empirical and theoretical study of the communication between the interviewer and the child in investigative interviews of children. Development of a methodological approach]* Oslo: Det utdanningsvitenskapelige fakultet.
- 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(3), 347-359.

- Garvin, A. (2000). *Learning in action: A guide to putting the learning organization to work*. Boston: Harvard Business School Press.
- Geiselman, R. E., Fisher, R. P., Cohen, G., Holland, H. L., & Surtes, L. (1986). Eyewitness Responses to Leading and Misleading Questions Under the Cognitive Interview. *Journal of Police Science and Administration*, 14(1), 31-39.
- Goodman, G., Quas, J., Batterman-Faunce, J., Riddelsberger, 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., & Aman, C. (1990). Children's use of anatomically detailed dolls to recount an event. *Child Development*, 61, 1859-1871.
- Goodman, G. S., Bottoms, B. L., Schwartz-Kenney, B. M., & Rudy, L. (1991). Children's testimony about a stressful event: Improving children's reports. *Journal of Narrative and Life History*, 1(1), 69-99.
- Hardin, C. D., & Higgins, E. T. (1996). Shared reality: How social verification makes the subjective objective. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition: Foundations of social behaviour* (pp. 28-84). New York: Guilford Press.
- Harley, T. (2008). *The Psychology of Language* (3 ed.). London: Psychology Press.
- Hershkowitz, I., Horowitz, D., & Lamb, M. E. (2005). Trends in children's disclosure of abuse in Israel: A national study. *Child Abuse & Neglect*, 29, 1203-1214.
- Hershkowitz, I., Horowitz, D., Lamb, M. E., Orbach, Y., & Sternberg, K. J. (2004). Interviewing youthful suspects in alleged sex crimes: a descriptive analysis. *Child Abuse & Neglect*, 28, 423-438.
- Hershkowitz, I., Lamb, M. E., Sternberg, K. J., & Esplin, P. W. (1997). The relationships among interviewer utterance type, CBCA scores, and the richness of children's responses. *Legal and Criminological Psychology*, 2, 169-176.
- Hindman, J. (1987). *Step by step. Sixteen steps toward legally sound sexual abuse investigation*. Ontario: Alexandria Associates.
- Home Office (1992). *Memorandum of Good Practice on Video Recorded Interviews with Child Witnesses for Criminal Proceedings*. London: HMSO.

- Home Office (2002). *Achieving Best Evidence in Criminal Proceedings: Guidance for Vulnerable or Intimidated Witnesses, Including Children*. London: HMSO
- Home Office (2007). *Achieving Best Evidence in Criminal Proceedings: Guidance on Interviewing Victims and Witnesses, and Using Special Measures*. London: HMSO.
- 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.
- Justisdepartementet (1998). *Rundskriv G-70/98 om forskrift 2 oktober 1998 nr 925 om dommeravhør og observasjon og om endringer i straffeprosessloven, straffeloven og påtaleinstruksen [Regulations concerning out of court judicial examination and observation]*. Oslo: Det Kongelige Justis- og politidepartement.
- Korkman, J., Santtila, P., Blomqvist, T., & Sandnabba, N. K. (2008). Failing to keep it simple: Language use in child sexual abuse interviews with 3-8- year- old children *Psychology, Crime & Law, 14*(1), 41-60.
- Lamb, M. E., & Brown, D. A. (2006). Conversational apprentices: Helping children become competent informants about their own experiences. *British Journal of Developmental Psychology, Crime & Law, 24*, 215-234.
- Lamb, M. E., & Fauchier, A. (2001). The effects of question type on self-contradiction by children in the course of forensic interviews. *Applied Cognitive Psychology, 15*, 483-491.
- Lamb, M. E., Hershkowitz, I., Orbach, Y., & Esplin, P. W. (2008). *Tell Me What Happened. Structured Investigative Interviews of Child Victims and Witnesses*. Chichester: John Wiley & Sons, Ltd.
- Lamb, M. E., Hershkowitz, I., Sternberg, K. J., Esplin, P. W., Hovav, M., Manor, T., et al. (1996). Effects of investigative utterance types on Israeli children's responses. *International Journal of Behavioral Development, 19*(3), 627-637.
- Lamb, M. E., Sternberg, K. J., & Esplin, P. W. (1994). Factors influencing the reliability and validity of statements made by young victims of sexual maltreatment. *Journal of Applied Developmental Psychology, 15*, 255-280.
- Lamb, M. E., Sternberg, K. J., Orbach, Y., Esplin, P. W., & Mitchell, S. (2002). Is ongoing feedback necessary to maintain the quality of investigative interviews with allegedly abused children? *Applied Developmental Science, 6*(1), 35-41.

- Lamb, M. E., Sternberg, K. J., Orbach, Y., Hershkowitz, I., Horowitz, D., & Esplin, P. W. (2002). The effects of intensive training and ongoing supervision on the quality of investigative interviews with alleged sex abuse victims. *Applied Developmental Science, 6*(3), 114-125.
- Landström, S., Granhag, P. A., & Hartwig, M. (2007). Children`s live and videotaped testimonies: How presentation mode effects observers` perception, assessment and memory. *Legal and Criminological Psychology, 12*, 333-347.
- Manning, P. K. (1988). *Symbolic Communication. Signifying Calls and the Police Response*. London: The MIT Press.
- McGregor, D. (1960). *The human side of enterprise*. New York: McGraw-Hill.
- Melinder, A., Scullin, M. H., Gunnerød, V., & Nyborg, E. (2005). Generalizability of a two-factor measure of young children's suggestibility in Norway and USA. *Psychology, Crime & Law, 11*(2), 123-145.
- Memon, A., Vrij, A., & Bull, R. (1998). *Psychology and Law - Truthfulness, Accuracy and Credibility*. Maidenhead, Berkshire, England: McGraw-Hill Publishing Company.
- Milne, R., & Bull, R. (1999). *Investigative interviewing: Psychology and practice*. Chichester: Wiley.
- Morison, S., Moir, J., & Kwansa, T. (2000). Interviewing children for the purpose of research in primary care. *Primary Health Care Research and Development, 1*, 113-130.
- Mortimer, A. (1994). *Police investigative interviewing: Pre-interview thinking and written interview behaviour*. Paper presented at the 4th European Conference on Psychology and Law.
- Myklebust, T. (2005). Dommeravhør av barn i Norge - intensjon og fakta [Investigative interviews of children in Norway - intention and facts]. In M. Egge & J. Strype (Eds.), *Politirollen gjennom 100 år. Tradisjon og endring [The role of the police in 100 years. Tradition and changes]* (pp. 142-172): PHS Forskning.
- Nielsen, B. G. (2001). *Straffesystemet i børneperspektiv. Seksuelle overgrep mot born i familien. [Criminal law administration from a child`s persepective. Child sexual abuse cases within the family]*. København: Jurist- og Økonomforbundets Forlag.

- Oats, K., & Shrimpton, S. (1991). Children's memories for stressful and non-stressful events. *Medical Science and Law*, 31(1), 4-10.
- Orbach, Y., Hershkowitz, I., Lamb, M. E., Sternberg, K. J., Esplin, P. W., & Horowitz, D. (2000). Assessing the value of structured protocols for forensic interviews of alleged child abuse victims. *Child Abuse & Neglect*, 24(6), 733-752.
- Orbach, Y., & Lamb, M. E. (2001). The relationship between within-interview contradictions and eliciting interviewer utterances. *Child Abuse & Neglect*, 25, 323-333.
- 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.
- Oxburgh, G., Myklebust, T., & Grant, T. (Submitted). The Question of Question Types in Police Interviews: A review of the literature from a psychological and linguistic perspective. *International Journal of Speech Language and the Law*.
- Oxburgh, G., Williamson, T., & Ost, J. (2006). Police officers' Use of Emotional Language During Child Sexual Abuse Investigations. *Journal of Investigative Psychology and Offender Profiling*, 3, 35-45.
- PHS (2002). *Kursplan i utdanning i etterforskning av seksuelle overgrepssaker trinn III [Course description for education in investigation of sexual abuse cases and investigative interviews of children and vulnerable witnesses - Step III]*. Oslo: Politi­høgskolen [Norwegian Police University College].
- PHS (2003). *Kursplan i utdanning i etterforskning av seksuelle overgrepssaker trinn II [Course description for education in investigation of sexual abuse cases and investigative interviews of children and vulnerable witnesses - Step II]*. Oslo: Politi­høgskolen [Norwegian Police University College].
- Poole, D. A., & Lamb, M. E. (1998). *Investigative interviews of children: A guide for helping professionals*. Washington, DC: American Psychological Association.
- Powell, M. B. (2002). Specialist Training in Investigative and Evidential Interviewing: Is it Having Any Effect on the Behaviour of Professionals in the Field? *Psychiatry, Psychology and Law*, 9(1), 44-55.
- Queensland Family Services (1992). *Interviewing Children and Recording Evidence. Education, Health and Police Services*. Queensland: Family Services.

- Richardson, S. A., Dohrenwend, B. S., & Klein, D. (1965). *Interviewing: Its forms and functions*. London: Basic Books, Inc., Publishers.
- Riksadvokaten (1994). *Riksadvokatens gjennomgang av den såkalte Bjugn-saken. [The Director of Public Prosecutions. An evaluation of the Bjugn case]*. Oslo: Riksadvokaten.
- RNoAF, C. o. S. (2004). *Military Flight Regulations*. Oslo: RNoAF.
- Saywitz, K., & Camparo, L. (1998). Interviewing child witnesses: A developmental perspective. *Child Abuse & Neglect*, 22, 825-843.
- Schneider, W., & Pressley, M. (1997). *Memory development between two and twenty* (2 ed.). Mahwah, NJ: Erlbaum.
- Shepherd, E. (2007). *Investigative interviewing. The Conversation Management Approach*. Oxford: Oxford University Press. .
- Skjønsvjell (2008, 25-26 November). *The Pocketman*. Paper presented at the Homicide and Serial Sex Crimes Conference, Interpol General Secretariat, Lyon.
- State v. Buckey, Superior Court, Los Angeles County, California, #A750900 (1990).
- State v. Michaels, Superior Court, Essex County, New Jersey (1988).
- State v. Michaels, 264 N.J. Super 579, 625 A.D. 2d 489 (N.J. Super Ad, 1993).
- State v. Michaels. 136 N.J. 299, 642 A.2d 1372 (N.J., 1994).
- Steller, M., & Boychuk, T. D. (1992). Children as witnesses in sexual abuse cases: Investigative Interview an Assessment Techniques. In H. Dent & R. Flin (Eds.), *Children as witnesses* (pp. 47-73). Chichester: Wiley.
- Stern, W. (1903/1904). *Beiträge zur Psychologie der Aussage*. Leipzig: Verlag von Johann Ambrosius Barth.
- Sternberg, K. J., Lamb, M. E., Davies, G. M., & Westcott, H. L. (2001). The Memorandum of Good Practice: theory versus application. *Child Abuse & Neglect*, 25, 669-681.
- Sternberg, K. J., Lamb, M. E., Hershkowitz, I., Esplin, P. W., Redlich, A. D., & Sunshine, N. (1996). The relation between investigative utterance types and the informativeness of child witnesses. *Journal of Applied Developmental Psychology*, 17, 439-451.

- The National Crime Faculty (1998). *A practical guide to investigative interviewing*.
Bramshill: Training and Development Unit.
- 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(6), 629-640.
- Varendonck, J. (1911). Les témoignages d'enfants dans un procès retentissant. *Archives de Psychologie*, 11, 129-171.
- Walker, A. G. (1999). *Handbook on questioning children. A linguistic perspective* (2 ed.).
Washington, DC: American Bar Association: Center on Children and the Law.
- Walker, N. E., & Hunt, J. S. (1998). Interviewing child victim-witnesses: How you ask is what you get. In C. P. Thompson, D. J. Herman, J. D. Read, D. Bruce, D. G. Payne & M. P. Toglia (Eds.), *Eyewitness memory: Theoretical and applied perspectives* (pp. 55-87). Mahwah, NJ: Lawrence Erlbaum Associates.
- Watzlawick, P., Bavelas, J. B., & Jackson, D. D. (1967). *Pragmatics of Human Communication. A study of interactional patterns, pathologies, and paradoxes*. New York: W. W. Norton & Company.
- Yuille, J. C., Hunter, R., Joffe, R., & Zaparniuk, J. (1993). Interviewing children in sexual abuse cases. In G. S. Goodman & B. L. Bottoms (Eds.), *Child victims, child witnesses: Understanding and improving testimony* (pp. 95-115). New York: Guilford.

THE CURRENT STATE OF POLICE INTERVIEWS WITH CHILDREN IN NORWAY: HOW DISCREPANT ARE THEY FROM MODELS BASED ON CURRENT ISSUES IN MEMORY AND COMMUNICATION?

TROND MYKLEBUST^{a,*} and LAURENCE ALISON^b

^a*Asker & Baerum Police, Box 415, E-3.11, N-1302 Sandvika, Norway*

^b*Centre for Investigative Psychology, Department of Psychology, Eleanor
Rathbone Building, University of Liverpool, Liverpool, L69 7ZA,
United Kingdom*

This study sought to examine the state of police interviews with children in Norway set against the various models that have been developed in the United Kingdom. UK models attempt to accord with psychological principles that lead to effective interviewing and so, if appropriately followed, should enhance the elicitation of accurate material. Set against these principles of best practice, an analysis of 11 police interviews indicated that a number of inappropriate and ineffective strategies appear to be used in police interviews in Norway. These included a preponderance of closed questions, a large number of instructions, paraphrasing and not giving the child enough time to talk. Results suggest that the Norwegian system needs to be made more aware of the psychological vulnerabilities that can lead to suggestibility in children.

Key words: police interviewing, children as witnesses, Norwegian police.

INTRODUCTION

Since the Cleveland inquiry in United Kingdom (1987), the Bjugn case in Norway (1995) and reports from Rochdale and Orkney relevant professionals and the public have been aware that investigative interviews need to be conducted very carefully with young children. Poor interviewing can lead to suggestibility and the elicitation of erroneous information. An individual can be influenced by subtle suggestions, expectations, stereotypes, and leading questions that can either unconsciously alter memories

* Corresponding author.

(Ceci, 1996) or exert social pressure on the individual to alter aspects of the information.

Prior to the 1990s, the investigative procedure simply was not geared towards children's cognitive and linguistic abilities or to their vulnerability regarding suggestibility (Memon, Vrij and Bull, 1998). The fact that the investigative procedure might intimidate children and make them less likely to be able to recount their alleged experience was completely overlooked. In the UK in the last decade several methods have been suggested by different professionals working with children, to improve the results of investigative interviews (Home Office, London, 1997; Queensland, 1992; Hindman, 1987; Steller and Boychuk 1992, in Bray 1993). All have a number of recommendations in common. The research upon which this is based suggests several strategies for improving children's performance during investigative interviews (Ceci and Bruck, 1996).

STRATEGIES FOR IMPROVING CHILDREN'S PERFORMANCE IN THE INTERVIEW

Open Versus Closed Questions

Open questions give, on average, two and a half times as many details and words than closed questions. Interviewees also give longer and more qualitative detailed responses (Sternberg, Lamb, Hershkowitz, Yudilevitch, Orbach, Esplin and Hovav, 1997; Geiselman, 1993). While open questions result in broad accounts of the sequence of events, closed questions are often used by interviewers to narrow down the focus of the inquiry. Closed questions are of three main types: (i) identification: A question requiring the identification of person, place, group time, etc., (ii) selection: A closed-alternative question, where the subject has to select one from the two or more possible responses suggested by the interviewer, (iii) Yes-no type: A question that can be answered satisfactorily with a "Yes" or "No".

Children's response to specific questions are much less accurate than the responses to open-ended questions. Dent and Stephenson (1979) reported that 9% of the information provided by 10 to 11 year-olds in response to free recall questions was inaccurate, compared with 19% in response to specific questions.

Regardless of the experimental procedures, the ages studied, the cognitive capacity of the subjects, or the length of the delay between events and the interview, open ended questions are more likely to elicit accurate accounts (Dent, 1991; Hershkowitz, 1997, Hershkowitz, Lamb, Sternberg and Esplin, 1997; Lamb, Sternberg and Esplin, 1994).

Children choose what information they will report in response to open-ended questions, while specific questions focus on details that were never encoded or are no longer remembered. This raises the risk of error because both children and adults show poorer memories as critical details become more peripheral (Goodman *et al.*, 1987; Steward, Bussey, Goodman and Saywitz, 1993). Secondly, children generally try to answer specific questions that we know they cannot possibly answer accurately (Poole and White, 1993; Pratt, 1990), and adults and children both often try to answer strange questions, such as in Winer *et al.*'s experiment (Winer, Rasnake and Smith, 1987); "Is a cup sadder than an orange?"

Yes/No Questions

Some researchers believe that yes/no questions are more problematic than other types of specific questions. Poole and Lindsay (1995) claim that conversational convention dictates that children should try to answer questions and be co-operative. As a consequence of this, some children will frequently say, "yes" to yes-no questions. In Poole and Lindsay's experiment children erroneously responded, "yes" to 62% of the questions about demonstrations they had never experienced (Poole *et al.*, 1995).

Identification Type

Most identification questions start with the letters "wh". Some of them are more straightforward and easier to answer than others: "Who do you want to play with?" is, for example, easier to answer than, "Why do you want to play with David?" An analysis of transcript data by Aldridge, Timmins and Wood (1997) suggests that many more "what"-questions are asked than any other forms of questions. The next most frequent types are "who"-, "where"-, and "how"-questions, followed by far fewer "why"-, and, "when"-questions. The number of "why"-questions is not insignificant. In their data, there were only a few transcripts that did not contain at least one "why"-question. Aldridge *et al.* found that children performed better on questions of the forms "what, where, and who", than they did with "how, when and why".

Selection Type

In daily conversations, adults frequently interrupt children or tolerate only short lapses before jumping in with specific questions. For most children adults decide what to discuss and how much time that is spent on each conversation. The interview will therefore violate many of the conversational rules that children are used to. For example:

- Children believe that adults know everything and thus it is children that ask questions and not adults (Dent and Flin, 1992).
- Very young children often assume that because one adult (the perpetrator or one of the parents has been told) knows what took place, other adults must already know (Toglia, Ceci and Ross, 1992).
- Many children are warned not to speak to strangers and for most interviewees, the interviewer is perceived as a stranger (Bull, 1992).
- Most children are brought up in a culture where they generally are instructed not to talk about certain topics, sex included, in public. The whole conversation or interview might therefore be of an intimate nature (Bull, 1992).

The child, under the false impression that the adult knows the answer to the questions, might therefore interpret questions of selection as a test.

Walker, Lunning, and Eilts (1996) asked kindergarteners, second graders, and fifth graders to watch a videotape and answer 36 forced-choice questions such as, "Did you see a little girl or a little boy on the video?" The placement of the correct answer was counterbalanced – it appeared equally often in the first or second position in the question, or not at all. As predicted, the kindergarteners were significantly less accurate than the older children, but children in all three age groups showed a response set favouring selection of the second option. Kindergarteners answered correctly only 17.5% of the time when neither option was correct, whereas second and fifth graders more often provided a correct "neither" response (37.5% and 48.3% of the time, respectively). Thus even older children answered correctly less than half of the time when the correct option was not embedded in the question.

Repeating Questions

Repeating questions might signal to the child that their first answer was incorrect. If asked more than once the child might change his or her answer to please the interviewer. Poole and White (1991; 1993) concluded that repeating open-ended questions within an interview is relatively harmless, but that repeating closed or specific questions is particularly dangerous. These results are consistent with a finding by Memon and Vartoukian (1996), that children's accuracy increased slightly when interviewers repeated open-ended questions within an interview but decreased slightly when interviewers repeated specific questions.

Repeating the Interviewee's Answer – Paraphrasing

During language development children often obtain feedback from adults. This process is called "paraphrasing". Paraphrasing occurs when adults

repeat what a child has said but change the original phrase so that it makes grammatical sentence. According to Harley (1995) this is done repetitively every day.

According to Donaldson (1978) children are used to being corrected when they are wrong, and believe that adults have the correct answer when they change the child's sentence or correct them. In interview situations many interviewers rephrase the interviewee's sentences in order to see if they have understood the interviewee. Children might, however, interpret this as the adult providing the correct response even if the interviewee meant something different in the statement.

The Influence of Questioning on Children's Understanding of the Difference between Lies and the Truth

Children's understanding of the truth seems to vary, depending on how adults ask them questions (Huffman, Warren and Frazier, 1997). The ability to distinguish between lies and the truth varies as a function of two factors; the form of the question asked and the topic selected for discussion. Research on children's understanding of these concepts (Walker, 1994) has shown that interviewers can improve children's perceived competency by asking them concrete questions about the truth, rather than asking them to define the truth. Abstract questions, that are not grounded in a content domain, such as: "What does it mean to tell the truth?" may confuse even school aged children (Walker, 1994).

The Importance of Saying "I don't know"

Some studies have shown that children are more accurate when interviewers advise them that it is appropriate to say "I don't know". This also stresses the importance of warning the interviewee and informing him or her that they are not expected to know the answer to every question (Poole *et al.*, 1995).

Stereotypes

Leichtman and Ceci (1995) demonstrated how children are influenced by stereotypes. For example if they have been told a particular individual is clumsy and are then asked to consider who might have accidentally damaged something then they will apportion blame to the individual they were told was clumsy, regardless of whether they have any evidence. Suggestive interviewing exacerbates this effect and Ceci and Bruck (1993) take such findings as an indication that, not only can children form powerful social stereotypes, but also that this information interacts with suggestive questioning to produce false accounts.

Interviewer Bias

The topic of “interviewer bias” or “experimenter bias” has been researched throughout the century. Rice (1929) provided one of the earliest studies. He studied 12 experienced interviewers from various social service agencies in New York City. Rice was struck by how some of the interviewers’ beliefs influenced the contents of the reports they obtained from the interviewee.

Interviewer biases are reflected in the atmosphere of the interview. Sometimes interviewers provide much encouragement during the interview in order to put the children at ease and to provide a highly supportive environment. Such encouraging statements, however, can quickly lose their impartial tone, when a biased interviewer selectively reinforces only particular responses from the interviewee. This occurs by positively acknowledging statements from the interviewee, and by head nodding, smiling, and positive statements such as “Yes, That’s right, etc.”.

Another process by which this may occur is through Ceci *et al*’s (1996) claim that, interviewers may not ask children open-ended questions, but resort to a barrage of very specific questions, many of which are repeated or leading. Repeated interviewing can facilitate this effect. With some authorities estimating that, on average, child witnesses may be questioned up to 12 times during the course of an investigation (Whitcomb, 1992), there is plenty of room for long sequences of continually suggestive questioning.

Interruption during the Interview

According to Fisher, Geiselman, and Raymond (1987), frequent interruptions cause two main problems for the victim or witness. Firstly, interruptions may prevent maximum recall since focus is switched from the memory of the crime scene to the interviewer’s question. Secondly, after several interruptions the victim or witness might begin to expect further interruption from the interviewer. Less effort may then be made in recalling the information. According to Bjørklund (1996), interviewers who ask questions that change between the different sensory or temporal modalities, are more engaged in confirming their own assumptions than listening to the victim or witnesses’ observations.

HYPOTHESIS

In Norway there is no particular model or “Guide to Interviewing”. The legal framework, and especially the Prosecution Instructions and Criminal Procedure, are the only formal guideline for investigators. Based on this we

anticipated that the Norwegian interviews would deviate from the model of best practice and would, therefore, show a preponderance of many of the aforementioned problems associated with interviewing children.

METHOD

Background

This study is based on Police interviews of children suspected as victims of sexual abuse. Because of the sensitivity of the data, the Director of Public Prosecution in Norway has given this project strict restriction of anonymity of identity of the people mentioned in the interviews, the interviewee and the interviewer, and police departments involved. With a total population of 4.4 million (Statistics Norway, 1999) and 7,567 police officers (Hammer, 1999), Norway is a small country compared with UK. Analysing interviews and discussion of the interview methods used by police investigators are therefore based on co-operation and confidence from the police investigators involved, that their identity remain anonymous.

Therefore, few demographic details can be given regarding either interviewers or interviewees suffice as to say that the analysed interviews represent victims from all over the country, interviewed by several interviewers, and from different police departments.

Samples

Eleven interviews were collected from different Police Departments in Norway. The interviews were obtained randomly. This was done by contacting the investigators' superiors at different Police Departments, asking them for a sample of interviews with children in the age categories: 5 years, 7–9 years, and > 11 years.

Eleven interviews were randomly picked out, to be used in the analysis. This, of course is a very small sample upon which to make general observations. However, from the first author's experience, observations, and discussions with other police investigators, there was nothing unusual about these interviews. The eleven interviews used in the analysis were therefore deemed representative investigative interviews with children in Norway.

Each case consisted of:

1. A videocassette of the interview.
2. An audiocassette of the interview.
3. The interviewer's personal notes from the preparation, used as a support tool during the interview.

4. A transcription of the interview based on the audiocassette.
5. A document informing about the status of the case, and the result if the case had been to court.

The documents regarding the result of each of case were put in an envelope and were not opened until after the analysis.

Based on Piaget's Stages of Cognitive Development the interviews were divided into three age categories representing the Pre operational, Concrete operational, and Formal operational stages. Based on the interviews the age categories are: 5 years, 7–9 years and > 11 years.

Procedure

Content dictionary

A content dictionary was developed around Richardson, Dohrenwend and Klein's (1965) definition of open and closed questions. Interviews were divided into 5-minute sequences. Questions in each of the sequences were then classified and counted for each section. The amount of talking time of each conversant was also measured. Finally, other significant features were extracted for a qualitative examination of the text.

RESULTS

Inter-rater Reliability

An independent rater randomly selected and analysed four of the eleven interviews. The result of the analysis of the number of questions, gave an inter-rater reliability of .86. The result of the time-analysis for the interviews, gave an inter-rater reliability of .79. There were a few differences between the raters for the open questions and closed – identification – questions, where some of the open questions were rated as identification questions, and some of the identification questions were rated as open questions. However, the reliability coefficient for this was .91, indicating a good level of agreement

Collective Results

Because of the small number of cases examined here meaningful statistical comparisons could not be carried out. Rather, general trends can be observed at a descriptive level between and within individual interview schedules.

The collective mean duration of time the interviewee and interviewer speaks is nearly the same, at 29% and 31% respectively. The rest of the interview contains pauses (40%).

The results of the analysis of the mean values for questions and instructions are presented in figure 1:

The mean number of closed questions is substantially higher than the number of open questions and instructions for the interviews. There are almost no open questions. The mean number of instructions peaks in the beginning and at the end of the interviews.

Dividing the closed questions into identification, yes-no, and selection type gave the results presented in figure 2:

The number of identification-type questions are higher than the yes-no questions in the beginning of the interview. However, identification drops in the last 30% of the interview. Here the two graphs cross each other, and while the numbers of yes-no questions are nearly the same as in the rest of the interview, the number of identification questions decrease.

Four Periods

There are three significant stages of change throughout the interviews resulting in four distinct periods. There is a clear change of topic and questioning style at around 20%, 30%, and 80% of the interviews. This is reflected both in the overall graph (see Figure 1) as well as in each age

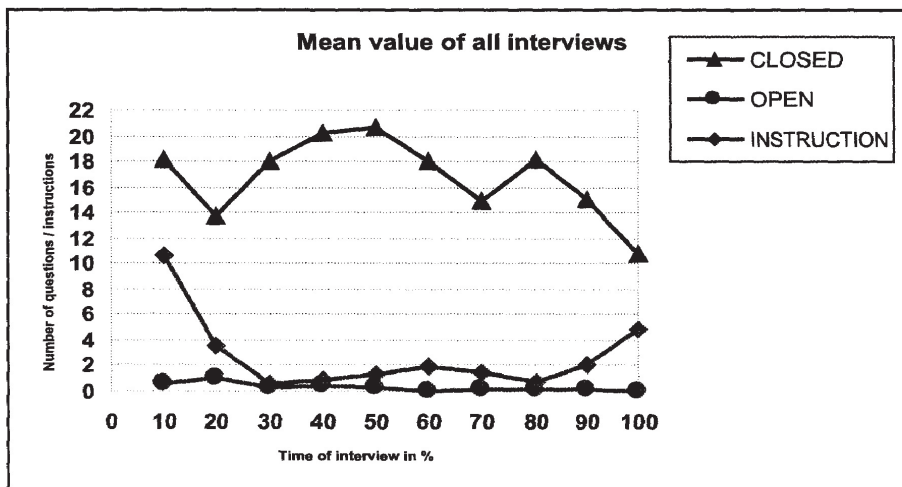


Figure 1 Shows the number of closed questions, open questions, and instructions divided into the percentages of the interviews.

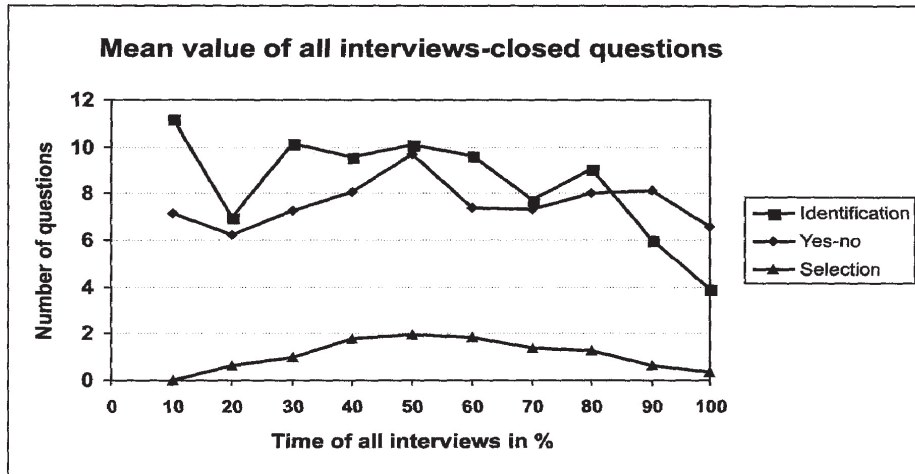


Figure 2 Shows the mean number of closed questions divided into the three categories identification, yes-no, and selection. These are divided into the percentages of the interviews.

band (see figs 3a–3f). These four periods relating to the three changes are considered below:

Period 1: 1st 20%

Characteristic of this period is a high proportion of instructions and an above average number of closed questions. The mean number of instructions is at its highest ($x=10.6$, where the mean number for the total interview is 2.81). The mean number of closed questions for this period is 18.2, 1.43 above the total average.

All interviews started with the interviewer giving the following information:

- Date and time of the interview.
- Explanations about the role of the police in the investigation.
- Informing the interviewee that he or she is used to interviewing children.
- When he or she talks to children it is usually done in a specially designed room, like “this” room, where the interviewee is now.
- Presentation of the interview-room, with information about the camera, microphones, one-way mirror.
- Presentation of the people behind the mirror, who is observing the interview.
- The importance of telling the truth and of talking loud enough so that the tape will pick the sound up.

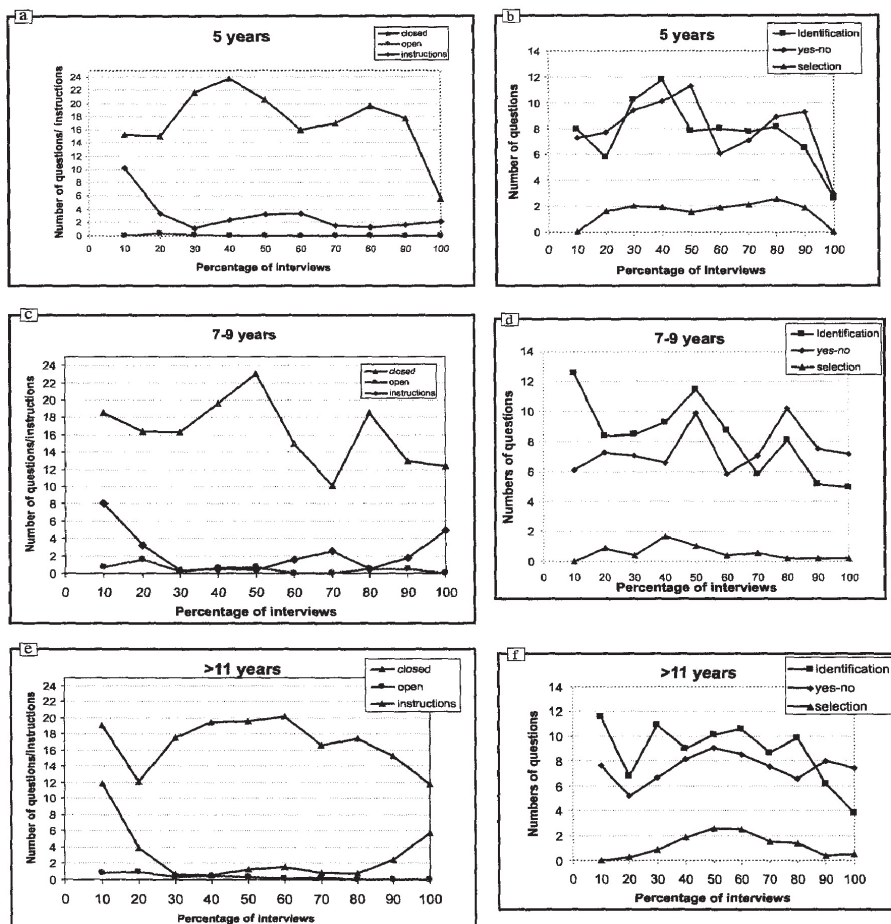


Figure 3a-f Shows the number of questions/instructions for the three different age categories.

The number of closed questions is also high at this stage, particularly in relation to identifications. These included:

- Age.
- Address.
- Who they live with.
- Number and names of brothers and sisters.
- If they have a pet, and if they have – the name of the pet.
- Which school and class they attend.

Whilst such questions may, in small doses, be appropriate in order to clarify and calibrate certain evidential issues, the preponderance of such questions throughout the rest of the interview strongly suggests a strategy that is operating more as a process of interviewer bias than as an objective search for the facts of the case. Even if one were to remove these instructional closed questions from the analysis the number of closed questions remains high throughout. Indeed, the overall peak of closed questioning is at the halfway stage where the interview pair are discussing crucial details directly relevant to the alleged offence.

Questions about the incident and the suspicion of sexual abuse, are never asked in this sequence. The number of yes-no questions are higher in the first period for the 5-year-olds than for the other groups. Where the 7–9 and > 11 groups are asked identification questions such as “What is your address?”, and “Who do you live with there?”. The group of 5 year-olds get the question phrased in a confirming fashion:

Interviewer: ‘I know that you live in XX street number XX?’
 Child: ‘Yes’.
 Interviewer: ‘Who do you live with there?’
 Child: ‘My brother and mother. . . .’
 Interviewer: ‘And your father?’
 Child: ‘Yes.’

In all the interviews children were told about the importance of telling the truth. All the children were asked a closed yes-no question such as; “Do you know what it is to tell the truth?”. All the interviewee’s answered yes to this question. Additionally, all were asked concrete questions about the truth:

Interviewer: ‘If I said that you came here by plane, would that be truth or a lie?’
 Child: ‘A lie.’
 Interviewer: ‘Because you got here by?’
 Child: ‘Car, my mother drove me. . . .’

Questions like this were used for all interviewers regardless of age. None of the interviewer’s asked abstract questions, such as “could you tell me what lie is?”

Period 2: 20–30%

In this period the number of instructions decreases from a mean of 3.6 to 0.6. The number of closed questions increases from the end of period one, with a mean of 13.8 to 17.9. Relating to the other three periods, the number of open questions is highest in period 2, even if the total number of open-questions does not exceed 2.

In the group of 5 year-olds there was an increasing number of closed questions (from an average of 15.1 to 21.7). The score for the > 11 group increased from 12.1 to 17.6. The number of closed questions for the 7–9 group remains constant during this period (16.4 to 16.3).

Period 3: 30%–80%

The number of closed questions steadily increases from 17.9, to peak at 20.7. This occurs at approximately the half-way mark. From there on the number gradually decreases. There is no significant difference between the different age groups for the numbers of closed questions (means = 19.7 (5 years), 17.1 (7–9 years) and 18.4 (> 11 years) respectively). The number of identification questions is higher than the average of yes-no questions for the whole period. There were no major differences in the average number of identification and yes-no questions between the age groups (means of identification = 8.9 (5 years), 8.7 (7–9 years), 9.8 (> 11 years); means of yes-no = 8.8 (5 years), 7.7 (7–9 years), 7.7 (> 11 years).

During this period the interviewer asks detailed questions about the incident to clarify:

- What happened?
- Who was involved?
- How did it happen?
- Where did it happen?
- When did it happen?

This echoes Aldridge *et al.*'s (1997) work where they found that most frequent questions asked were; "What, who, where, and how". However, in all cases examined here the specific focus of such questioning was centred around the sexual act and not any details associated with clarifying what led up to the allegation or what occurred afterwards.

Questions are asked about who else the child has discussed this with and answers generally concern another family member, or a friend (as was the case in one of the over 11 year- old group). However, none of the interviewees were asked any detailed questions about circumstances surrounding these conversations, such as for how long, how many times, or how the person they told reacted to the questions.

In one example, a child offers the following statement:

'... then he climbed into my bed and asked me if I wanted to fuck.'

The last word is usually not in a 5 year-olds vocabulary. It may of course relate to a word that the suspect used. If the interviewee does have a number of elder brothers or friends, there is, however, the possibility that

the word has been learnt through telling others about the event. This may also be the case in other interviews where various parts of anatomy are referred to in adult terms or where legal-terms are used that are unlikely to be within the child's vocabulary. In most cases the interviewer made no attempt to follow this up.

Qualitative analysis of the identification questions and yes-no questions, showed that the interviewers were consistently paraphrasing the majority of the interviewee's answers;

Interviewee: '... and then we ran away from him and back to the beach ...'
 Interviewer: 'So you ... and then ran from the forest back to the beach?'
 Interviewee: 'Yes'

The use of paraphrasing in the interviews seems to influence the communication process, where the interviewer several times rephrases the interviewee's initial answer, and then employs a yes-no question. For example:

Interviewee: 'We were talking about the temperature in the water ...'
 Interviewer: 'You were talking about the temperature in the water ...?'
 Interviewee: 'Yes.'

The mean number for yes-no questions, in this period is 75.7. Through period three, the interviewers' closed questions appear to indicate a set of prior beliefs regarding the event rather than an objective assessment of what may have occurred. This happens through all age-categories, however the interviews of youngest group demonstrate this most clearly. This group demonstrates a point at which identification and yes-no curves cross each other four times during the period. The crossing of these curves seem to be turning points where the interviewer has made up his or her mind about the incident, and uses yes-no questions to support his or her conclusions. For the 7-9 and > 11 groups, these changes are characteristic of period four. For the youngest group, it appears that the interviewers divided the questions into smaller chunks in period 3, asking detailed questions about a special part of the incident, and then asking yes-no questions to support his or her conclusion. This process occurs four times. The hypothesis that the interviewer has already come to a conclusion is perhaps also supported by the fact that the interviewer had a break in this phase, around 60%-80% of the interview. During the break they consulted the observers, to see if they had any questions they would like to ask the interviewee. This might also explain the drop in questions around this period, and the increase of instructions just before and after the short break.

The number of selection questions is higher in period three than in any other period. These questions were mostly used in the interviews to specify number of incidents, actions, and times.

Interviewer: 'Did this happen before, or after the trip to Stockholm?'

Interviewee: 'Before.'

Interviewer: 'Are we talking about 10 or are we talking about 100 times?'

The mean value for all interviews is 1.5 selection-questions for period 3.

Period 4: 80–100%

Across all interviews the number of closed questions decreases from 18.3 to 8. Identification questions decrease from 9.4 in the third period to 5 in this period, whilst the average number of yes-no questions is nearly the same for both periods (7.95 and 7.3 respectively). Characteristic for this period is an overall decrease in closed questions but, in particular, in the identification type.

After the break, in period three or four, it seems that the interviewers are confirming their conclusions by asking more yes-no questions;

Interviewer: 'So you are saying that he touched your penis and held his hand over your mouth?'

Interviewee: 'Yes.'

Interviewer: 'Then he asked you if you wanted to fuck?'

Interviewee: 'Yes.'

When the interviewer appeared to be satisfied with the interviewee's answer, the interviewer frequently asked if the interviewee had any questions they would like to ask, or any comments. For the child this may have been a signal that the interview would soon be over.

The Significance of the Four Periods

General trends across the interviews broadly correspond to four phases. Period one involves a large proportion of instructions and a fairly high number of closed questions. Whilst period two does incorporate a relatively higher number of open questions, they are still very infrequent. Period three appears to be the point at which the interviewers are confirming statements, since a very high number of closed questions are asked and there is a great deal of paraphrasing. Finally, period four results in a gradual decrease in closed questioning and an increase in instructions. At the broadest level then the interviews in this sample represent two instructional phases occurring at the beginning and towards the end of the interview with a very limited period of open questions followed by a barrage of closed confirming questions. These phases may be referred to as (i) opening instructions, (ii) limited opportunity for free narrative (iii) interviewer confirming phase (iv) closing instructions. It is important to note that, throughout, the general strategy involves continual sequences of closed questions. These sequences

may, therefore reflect interviewer bias in that very little opportunity is given over to free narrative and other than the two instructional phases, the main corpus of the questioning involves consistent closed questioning, paraphrasing and potentially confirming prior beliefs. Therefore, it appears that the interviewers may have been satisfied after phase three that they had the full picture of events and then merely had to bring closure through a final sequence of instructional elements.

Summary of Results

- Overall, the interviewer spent as much time talking as the interviewee
- There was a significant majority of closed questions throughout the interviews.
- There were almost no open-questions throughout
- Instructions peaked at the beginning, and before and after the breaks in the interview.
- There was hardly any variation in these patterns regardless of age groups.

DISCUSSION

Although there is little in the way of explicit training packages in Norway, the law does state that in relation to conducting interviews (Prosecution Instruction §8–6):

... The witness shall explain himself orally and shall be encouraged to tell everything he knows about the case, *in his own way, in his own time and using his own words. Only after this may specific questions be asked* ... Questions which by their content or their form might lead towards a reply in a particular direction must not be asked, unless they are used to test the reliability of information which the witness has previously given or other special grounds make it permissible.

The results of this study suggest that even at this basic level and, perhaps because of a lack of awareness of the ways in which particular strategies can lead to suggestibility, these interviews fall short of that requirement.

A variety of inappropriate techniques, known to promote suggestibility and ineffective interviewing, were employed.

Evaluation of results from this analysis concludes that the interviewers:

- Spent as much, if not more, time talking during the interview.
- Asked few open-questions.
- Asked a majority of closed questions.

- Gave many instructions in the beginning, before and after the breaks in the interview.
- Paraphrased answers
- At the outset asked many identification questions and then sought to confirm these through yes-no questions towards the end.
- Ignored potentially different needs of the children as a function of the child's age.

Even at the outset of the interviews, because of the attention paid to instructions, all of the children had to absorb a large amount of information about the “dos and don'ts” of the interview. This overload of information may have influenced encoding and retrieval ability throughout the rest of the interview. The rationale for such a pattern may, however, lie in the legal requirements:

‘§12: Before the examination begins the witness should normally be told who is present at the interview and where they are. The witness should also, as a rule, be informed when a video and/or sound recording is being made. The person carrying out the interview should admonish the witness to tell the truth according to the rules laid out in the Criminal procedure Act. The admonishing should be given in a manner, which is appropriate to the age of the witness, their development stage, and any other relevant circumstances . . .’

Regulations concerning out of court judicial examination and observation,
2nd October 1998 no. 925

However, it also notes that the interviewer should take the children's developmental stage into consideration. It is perhaps significant then that all the children, regardless of age, were given the same number of instructions phrased in exactly the same way.

According to the MGP (Memorandum of Good Practice, Home Office, 1997) or PEACE model (PEACE standing for Plan and prepare, Engage and explain, Account stage, Closure of interview, Evaluation of interview), the main aim of the first phase of the interview is to build up a rapport between the interviewer and the interviewee:

‘ . . . this phase serves a number of important additional functions. If used correctly, it should supplement the interviewer's knowledge about the child's social, emotional and cognitive development, and particularly about his or her communication skills an degree of understanding . . .’

Memorandum of Good Practice (Home Office, 1997) pp. 15–16.

Instead, information concerning how the interviewee should speak and behave during the interview, as well as the question-answer format, might give rise to poor communication. Instead of building rapport, this may have lead to anxiety about the process and concern over what to say and what

not to say. The question-answer format might also promote the wrong impression of the rules of communication for rest of the interview. The interviewee might be prepared only to answer closed questions, to avoid what he or she may see as "unnecessary detail" and thereby hand over total control to the interviewer. This might also explain why the children seem to be restricted on the few occasions when they were asked open-ended questions.

Additionally, the interviews contained significant periods of paraphrasing. This method of recapping, provided by the interviewer in "adult language", could result in errors. It is known that this process can lead to acquiescence on the part of the child and, subsequently, to potentially erroneous information (Harley, 1995). This, coupled with identification questions at the outset followed by confirming "yes-no" questions towards the end may also be indicative of the "interviewer bias" phenomenon. Such a pattern may reflect the interviewer's preconceptions about what occurred. What is then covered in the interview is simply a search for confirmation of a particular theory rather than objective search for the truth.

CONCLUSION

This is the first study of a representative sample of investigative interviews with children in Norway. The analysis revealed significant discrepancies from the PEACE model, a model based on relevant psychological theories of memory. Inappropriate and ineffective structures appear to be used in Norwegian police interviews, with a preponderance of closed question and instructions.

An obvious limitation of these results is the small sample size. However, there was nothing unusual or distinctive about these interviews. They were selected randomly and were carried out by different interview teams in different constabularies and in different parts of Norway. The chances of all 11 interviews containing such problematic features might, therefore, be expected to be very low. Thus, our results are likely to reflect a robust effect.

Moreover, given the relative lack of attention to the psychological underpinnings of children's testimony in Norway, one would perhaps hardly expect interviews to be carried out any worse or any better than the interviews that used to be carried out in the UK prior to the introduction of such knowledge.

According to McGough (1996) if an initial interview is improperly conducted, the child's account can never be re-examined effectively. Given that the investigative interview is often the single most important component of any trial involving a child witness these are very real concerns. In considering these issues it may be appropriate to try and implement the rules of best practice already established in the UK.

References

- Aldridge, M., Timmins, K. and Wood, J. (1997). Children's understanding of legal terminology: Judges get money at pet shows, don't they? *Child Abuse Review*, **6**, 141–146.
- Aldridge, M. and Wood J. (1996). Children's understanding and use of body part terminology: A guide for child witness interviewers. In Wood, J. (Ed). *Research Papers in Linguistic*, **8**, 1–14.
- Bjorklund, D. F. and Harnishfeger K. K. (1985). The evolution of inhibition mechanisms and their role in human cognition and behaviour. In F. N. Dempster and C. J. Brainerd (Eds.), *Interference and inhibition in cognition* (pp. 141–173). New York: Academic Press.
- Bjørklund, R. A. (1996). *Politipsykologi*, Oslo: Vett og Viten.
- Blakar, R. M. (1982). *Communication: A social perspective on clinical issues*. Oslo: Universitetsforlaget.
- Blakar, R. M. (1972). Språket som maktmiddel eller avsenderen som aktiv skaper. *Tidsskrift for samfunnsforskning*, **13**, 199–222.
- Blakar, R. M. (1975). *Språk er makt*. Oslo: Pax forlag.
- Bray, M. (1993). *Barnets stemme. Om seksuelle overgrep mot barn*. Oslo: Pax Forlag.
- Bruck, M. and Ceci, S. J. (1996). Issues in the scientific validation of interviews with young children. Comment on Interviewing young children about body touch and handling. *Monograph of the Society for Research in Child Development*, **61**, 3–4 Serial No. 248.
- Bull, R. (1992). Obtaining evidence expertly: The reliability of interviews with child witnesses. *Expert Evidence*, **1**, 5–12.
- Ceci, S. J. (1994). Cognitive and social factors in children's testimony. In Sales, B. and Vandebos, G. (Eds.), *Psychology in litigation and legislation: Master lectures* (pp. 14–54). Washington, DC: American Psychological Association.
- Ceci, S. J. and Bruck, M. (1996). *Jeopardy in the Courtroom. A Scientific analysis of children's testimony*. Washington, DC: American Psychological Association.
- Ceci, S. J., Crotteau-Huffman, M., Smith, E. and Loftus, E.W. (1994). Repeatedly thinking about non-events. *Consciousness and Cognition*, **3**, 388–407.
- Ceci, S. J. and Bruck, M. (1993). Suggestibility of the child witness: a historical review and synthesis. *Psychological Bulletin*, **113**, 403–439.
- Ceci, S. J., Leichtman, M. D. and Putnick, M. E. (1992). *Cognitive and social factors in early deception*. Hillsdale, NJ: Erlbaum.
- Criminal Procedure Act – Act of 22 May 1981 No 25. With Subsequent amendments, most recently by Act of 17 July 1998, No 56. Royal Ministry of Justice and the Police.
- Dent, H.R. (1991). Experimental studies of interviewing child witnesses. In J. Doris (Ed.). *The suggestibility of children's recollections*. Washington, DC: American Psychological Association.
- Dent, H. R. and Flin, R. (1992). *Children as Witnesses*. Chichester: Wiley.
- Dent, H. R. and Stephenson, G. M. (1979). An experimental study of the effectiveness of different techniques of questioning child witnesses. *British Journal of Social and Clinical Psychology*, **18**, 41–51.
- Donaldson, M. (1978). *Children's mind*. London: Fontana Press.
- Fisher, R. P., Geiselman, R. E. and Raymond, D. S. (1987). Critical analysis of police interview techniques. *Journal of Police Science and Administration*, **15**, 177–185.
- Geiselman, R. E., Saywitz, K. J. and Bornstein, G. K. (1993). Effects of cognitive questioning techniques on children's recall performance. In Goodman, G. S. and Bottoms, B. L. (Eds.) *Child victims, child witnesses: Understanding and improving testimony*. New York: Guilford Press.
- Goodman, G. S., Aman, C. and Hirschman, J. (1987). Child sexual and physical abuse: Children's testimony. In Ceci, S. J., Toglia, M. P. and Ross, D. F. (Eds.), *Children's eyewitness memory*. New York: Springer-Verlag.
- Hammer, P. personal communication, August 16 1999, The Norwegian Home Office, p: + 47 22 24 90 90.
- Harley, T. H. (1995). *The psychology of language, from data to theory*. East Sussex: Psychological Press.

- Hershkowitz, I. (1997). The dynamics of interviews yielding plausible and implausible allegations of child sexual abuse. In Poole, D. A. and Lamb, M. E. (1998). *Investigative Interviews of Children. A guide for helping professionals*. Washington: American Psychological Association.
- Hershkowitz, I., Lamb, M. E., Sternberg, K. J. and Esplin, P. W. (1997). The relationships among interviewer utterance type, CBCA scores and the richness of children's responses. *Legal and Criminological Psychology*, **2**, 169–176.
- Hindman, J. (1987). *Step by step. Sixteen steps toward legally sound sexual abuse investigation*. Ontario: Alexandria Associates
- Home Office (1997). *Memorandum of Good Practice on video recorded interviews with child witnesses for criminal proceedings*. London: Her Majesty's Stationery Office.
- Huffman, M. L., Warren, A. R. and Frazier, S. (1997). *The effect of a true/lie discussion on children's accuracy and competence*. Paper presented at the 23rd annual conference of the Association for Moral Education, November 1997, Atlanta, USA.
- Lamb, M. E., Sternberg, K. J. and Esplin, P. W. (1995). Making children into competent witnesses: Reactions to the amicus brief in re Michaels. *Psychology, Public Policy and the Law*, **1**, 438–449.
- Lamb, M. E., Sternberg, K. J. and Esplin, P. W. (1994). Factors influencing the reliability and validity of statements made by young victims of sexual maltreatment *Journal of Applied Developmental Psychology*, **15**, 255–280.
- Leichtman, M. D. and Ceci, S. J. (1995). The effects of stereotypes and suggestions on preschoolers' reports. *Developmental Psychology*, **31**, 568–578.
- McGough, L.S. (1996). Commentary: Achieving real reform – The case for American interviewing protocols. *Monographs of the Society for Research in Child Development*, **61**, 188–203.
- Memon, A., Vrij, A. and Bull, R. (1998). *Psychology and Law: Truthfulness, Accuracy and Credibility*. London: McGraw-Hill Publishing Company.
- Memon, A. and Vartoukian, R. (1996). The effects of repeated questioning on young children's eyewitness testimony. *British Journal of Psychology*, **87**, 403–415.
- Memon, A., Cronin, O., Eaves, R. and Bull, R. (1996). An empirical test of the mnemonic components of the Cognitive Interview. In Davies, G., Lloyd Bostock, S., McMurrin, M. and Wilson, C. (Eds.), *Psychology, law and criminal justice: International developments in research and practice*. New York: Walter de Gruyter.
- The Penal Code – Alminnelig Borgerlig Straffelov (Straffeloven) av 22. mai 1902 nr 10, med endringer, sist ved lov av 11. juli 1999 nr 39. Oslo: Cappelen Akademiske Forlag.
- Poole, D. A. and Lamb, M. E. (1998). *Investigative Interviews of Children. A guide for helping professionals*. Washington: American Psychological Association.
- Poole, D. A. and Lindsay, D. S. (1995). Interviewing pre-schoolers. Effects of non-suggestive techniques, parental coaching and leading questions on reports of non experienced events. *Journal of Experimental Child Psychology*, **60**, 129–154.
- Poole, D. A. and 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.
- Poole, D. A. and White, L. T. (1991). Effects of question repetition on the eyewitness testimony of children and adults. *Developmental Psychology*, **27**, 975–986.
- Pratt, C. (1990). On asking children and adults, bizarre questions. *First Language*, **10**, 167–175.
- Prosecution Instruction – Forskrift av 2. oktober 1998 nr 926, om ordningen av påtalemyndigheten, (påtaleinstruksen), Oslo: Cappelen Akademisk Forlag, Lovdata.
- Queensland Family Services (1992). Interviewing Children and Recording Evidence. *Education, Health and Police Services*. Queensland: Family services.
- Regulations concerning out of court judicial examination and observation – Rundskriv G-70/98 om forskrift 2 oktober 1998 nr 925 om dommeravhør og observasjon og om endringer i straffeprosessloven, straffeloven og påtaleinstruksen. Det Kongelige Justis- og politidepartement.
- Richardson, G. C. (1993). *The child witness: A linguistic analysis of child sexual abuse testimony*. Unpublished doctoral dissertation.

- Richardson, S. A., Dohrenwend, B. S. and Klein, D. (1965). *Interviewing: Its Forms and Functions*. London: Basic Books.
- Rice, S. A. (1929). Interviewer bias as a contagion. *American Journal of Sociology*, **35**, 421–423.
- Statistics Norway (1999). Folketallet i fylkene, 1. juli 1999 (On line). Available: <http://www.ssb.no>.
- Sternberg, K. J., Lamb, M. E., Hershkowitz, I., Yudilevitch, L., Orbach, Y., Esplin, P. W. and Hovav, M. (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. and Saywitz, K. J. (1993). Implications of developmental research for interviewing children. *Child Abuse and Neglect*, **17**, 25–37.
- Toglia, M., Ross, D. and Ceci, S. (1992). The suggestibility of children's memory. In Howe, M., Brainerd, C. and Reyna, V. (Eds.), *The Development of Long-term Retention*. New York: Springer Verlag.
- Walker, A. G. (1994). *Handbook on questioning children: A linguistic perspective*. Washington, DC: American Bar Association Center on Children and the Law.
- Walker, N. E., Lunning, S. M. and Eilts, J. L. (1996). *Do children respond accurately to forced choice questions? Yes or no*. Paper presented at Recollections of Trauma: Scientific Research and Clinical Practice, NATO Advanced Study Institute, Port de Bourgenay, France
- Whitcomb, D. (1992). *When the child is a victim*. Washington, DC: National Institute of Justice.
- Winer, G. A., Rasnake, L. K. and Smith, D. A. (1987). Language versus logic: Responses to misleading classificatory questions. *Journal of Psycholinguistic Research*, **16**, 311–327.

The Effect of Long-Term Training on Police Officers' Use of Open and Closed Questions in Field Investigative Interviews of Children (FIIC)

TROND MYKLEBUST^{1,*} and ROALD A. BJØRKLUND²

¹Norwegian Police University College, Box 5027, Majorstua, N-0301 Oslo, Norway

²Department of Psychology, University of Oslo, Box 1094, Blindern, N-0317 Oslo, Norway

Abstract

This study intends to assess the effect of long-term training on police officers' use of open and closed questions in investigative interviews of children. One hundred field investigative interviews of children were divided into two groups based on the interviewers' level of competence (training and experience). The police officers' questions were classified into groups of either open or closed. In all interviews, the mean numbers were 20 open and 217 closed questions, corresponding to an open–closed question ratio of 1 : 10. The most competent interviewers used 22 open questions and the less competent used 19 open questions, but the difference was not statistically significant. Analyses of variance (ANOVA) showed a significant effect of question but no main effect of competence. ANOVA revealed a descending number of open questions during the interviews, whilst the distribution of closed questions showed an inverted U distribution with most frequent use of closed questions in the middle part of the interviews. Possible strategies of enhancing the use of open questions are discussed. Copyright © 2006 John Wiley & Sons, Ltd.

Key words: investigative interviews; interview training; police; questions

INTRODUCTION

The purpose of the present study is to assess the effect of long-term theoretical and practical training on the police officers' use of open and closed questions in investigative interviews of children. The effect of training has previously been investigated in several papers (Aldridge & Cameron, 1999; Fisher, Geiselman, & Amador, 1989; Lamb, Sternberg, Orbach, Hershkowitz, Horowitz, & Esplin, 2000) but so far the conclusion is not clear.

Stern, (1903/04) was the first to distinguish between the two kinds of interview, denoted as *Bericht* and *Verhör*. Stern argued that the quality of the interviewee's statements was

*Correspondence to: Trond Myklebust, Norwegian Police University College, P.O. Box 5027, Majorstua, N-0301 Oslo, Norway.
E-mail: trond.myklebust@pfs.no

dependent on how the interviewer phrased his or her questions. He categorised Bericht by questions asked so the interviewee was allowed to give his or her account, in his or her own way, unhampered by questions or by any personal influence from the interviewer. These accounts are also called free narratives and are elicited by the use of open questions (e.g. 'Tell me what happened'). The free narrative, both from children and adults, has been recognised as informative and accurate information about the interviewees' lives (Dale, Loftus, & Rathbun, 1978; Dent & Stephenson, 1979; Fisher, 1995; Hershkowitz, Horowitz, Lamb, Orbach, & Sternberg, 2004; Hershkowitz, Lamb, Sternberg, & Esplin, 1997; Lamb, Hershkowitz, Sternberg, Esplin, Hovav, Manor, & Yudilevitch, 1996; Lamb, Sternberg, & Esplin, 1994). In the interviews denoted as Verhör, the interviewees are required to answer a set of pre-arranged questions, also called closed questions, with a varying degree of suggestibility (e.g. 'What kind of jacket was he wearing?' or 'Did he wear a green jacket?').

The difference between Bericht and Verhör allows the classification of questions into two categories; open or closed, respectively. The open questions encourage the interviewees' free narrative and facilitate multiple-word responses (Poole & Lamb, 1998; Richardson, Dohrenwend, & Klein, 1965). There is an agreement in the literature that regardless of age, cognitive capacity of the interviewee, or length of the delay between events and the interview, open questions are more likely to elicit longer and accurate accounts (Dale, Loftus, & Rathbun, 1978; Dent, 1986, 1991; Dent & Stephenson, 1979; Goodman & Aman, 1990; Goodman, Bottoms, Schwartz-Kenney, & Rudy, 1991; Hershkowitz *et al.*, 1997; Hutcheson, Baxter, Telfer, & Warden, 1995; Lamb, 1994; Lamb & Fauchier, 2001; Lamb, Sternberg, & Esplin, 1994; Oats & Shrimpton, 1991; Orbach & Lamb, 2001; Ornstein, Gordon, & Larus, 1992). In comparison, those who answer closed questions tend to use single or fewer words, often with a limited number of response alternatives (Poole & Lamb, 1998; Richardson, Dohrenwend, & Klein, 1965). Although free narrative reports are longer and more detailed, they usually do not provide all the information that the investigative interviewer requires. To focus upon already known or previously revealed information from the offender, the interviewer tends to ask closed questions. The competence of child witnesses is often doubted on the grounds that children are susceptible to suggestions, compliant, and often misled by closed questions (Garven, Wood, Malpass, & Shaw, 1998; Loftus, 1979; Schooler & Loftus, 1993). Suggestibility refers to the degree which the encoding, storing, and reporting of events can be influenced by a range of internal and external factors (Ceci & Bruck, 1995). Children choose what information they will report in response to open questions, whilst closed questions raise the risk of error in the question's focus on details that were never encoded or are no longer remembered by the child (Goodman, Aman, & Hirshman, 1987; Goodman & Aman, 1990). Regardless of the resolution of the various controversies concerning children's suggestibility, most researchers agree that the manner in which children are questioned can have profound implications on what they remember and this, in turn, increases the importance of careful interviewing by use of open questions (Brainerd & Ornstein, 1991; Lamb, Sternberg, & Esplin, 1994; Saywitz, 1988). Sternberg, Lamb, Hershkowitz, Esplin, Redlich, & Sunshine (1996) reported that open questions yielded responses four times longer and three times richer than responses to closed questions. Despite many methodological differences across studies, consistent results have continually emerged. The younger the child, the less information and fewer details are supplied in their free reports of witnessed events (Aldridge, 1992; Pipe, Lamb, Orbach, & Esplin, 2004). For young children, some studies (Davies, Westscott, & Horan, 2000; Hutcheson *et al.*, 1995) have

found that closed questions, as compared with open questions, might provide longer and more detailed answers. However, only 2% of the questions in Davies, Westscott, and Horan (2000) were classified as open and there is a methodological limitation to the study with only a small sample size (36 interviews) from only one UK police force. The study has also limited information concerning the interviewers' training and experience. Nevertheless, the findings confirm that both quantity and quality of answers are significantly influenced by the types of questions asked, together with the aspects of the interviewer's style.

There is no universally accepted way to categorise different kinds of closed questions. Richardson, Dohrenwend, and Klein (1965) divided the closed questions into three sub-categories: (i) identification; (ii) selection; and (iii) yes–no responses. Lamb *et al.* (1996) increased the sub-categories of closed questions into seven. However, in the statistical analysis, they reduced the number of categories to three: (i) directive; (ii) leading; and (iii) suggestive. In general, the interviewer may reasonably expect the open and closed questions to influence the length and accuracy of the interviewee's responses. Fisher, Geiselman, and Raymond (1987) introduced the open–closed question ratio (OCR) in order to characterise the proportion of open questions in relation to the number of closed questions in an interview. The authors reported an excessive use of closed questions with 3 open and 26 closed questions in the typical investigative interview, giving an OCR of almost 1:9. They argued that a high portion of open questions, described as a high OCR, would enhance the quality of the information received in a police interview. Fisher, Geiselman, and Raymond (1987) introduced the interviewer's level of competence as an important variable in generating a higher OCR. They reported that police officers do not use scientifically based knowledge when they interview people. They suggested introducing formal, scientifically based training to the police officers at the institutional level, with more focus on the psychology of memory and results from other scientifically based studies. The authors recommended that the institutionalised interview training should be given at two levels; at basic police training upon entry to the service and at specialist training for more experienced investigators. They stated that their interview method, which they called the cognitive interview technique, could be learned within a few hours and that the method lengthens the effective field interview time by only a few minutes (Fisher & Geiselman, 1992). Fisher, Geiselman, and Amador (1989) and George and Clifford (1992) monitored experienced police officers' interviews with victims and witnesses of crime, before and after training in the cognitive interview technique. Prior to the training, the interviewers asked primarily closed questions. The training was then conducted over four multiple short sessions as per the Fisher, Geiselman, and Amador (1989) study and two as per the George and Clifford (1992) study. The training consisted of background lectures on good and poor interview techniques, tape-recorded samples of evaluated police interviews, and role-play exercises with individual feedback. The interviews conducted by trained interviewers increased the OCR from 1:20 to 1:2 (George & Clifford, 1992). The result supported the hypothesis that short-term training increased the OCR. Contrary to Fisher, Geiselman, and Raymond (1987), Lamb and his colleagues have argued that long-time improvements in the quality of the interviews are observed only when training is distributed over time (Lamb *et al.*, 2000; Lamb, Sternberg, Orbach, Esplin, & Mitchell, 2002). Following the development of an intensive 1-week training course, Aldridge and Cameron (1999) compared the investigative interviewing skills of police officers and social workers who had undertaken the training with a control group who did not receive training. This was carried out over a 9-month period. They did not find any significant effect of the 1-week intensive training course on the OCR. The authors explained the ineffectiveness of the short

course as a consequence of the discrepancy between the complexity of the task of interviewing children and the time available on the course. The structure of the interviews (i.e. how the interviews should be planned and conducted) has been discussed by several authors, with some arguing that the interviewer should follow highly structured sequences (American Professional Society on the Abuse of Children, 1990; Fisher & Geiselman, 1992; Hindman, 1987; Home Office, 1992, 2002; Orbach, Hershkowitz, Lamb, Sternberg, Esplin, & Horowitz, 2000; Queensland Family Services, 1992; Steller & Boychuck, 1992; The National Crime Faculty, 1998; Yuille, Hunter, Joffe, & Zaparniuk, 1993). Other authors studying child–adult communication argue that a dialogical model accounting for a more active child (i.e. one who introduces terms of references and the topics to be discussed during the interview) provides a more realistic description of the structure of an ordinary field interview (Clark & Brennan, 1991; Manning, 1988; Walker, 1999; Watzlawick, Bavelas, & Jackson, 1967). The training of Norwegian interviewers emphasise the use of structure and sequences in the interview analogues to the training of police officers in Europe and America (Walker & Hunt, 1998). Structured interviews might be characterised by police officers using open questions in the beginning of the interview followed by closed questions to follow up information already given by the interviewee as free narrative. Dialogue-based interviews on the other hand do not display a similar structure, reflecting that the child and police officer mutually engage with each other during question and answer. Studies reporting on the effect of structure in interviews are sparse. Myklebust and Alison (2000) used a four-sequenced division and reported that the number of open and closed questions changed differently during the interviews. They reported a declining number of open questions and an inverted U distribution of closed questions. Wright and Alison (2004) reported that the number of open-ended questions declined and the number of closed questions ascended during field interviews.

After the Bjugn case in Norway, substantial effort has been made in this country to increase the skills of police officers who interview children. The Bjugn case involved seven adults, in the small community of Bjugn, being arrested in 1992 on suspicion of sexual abuse and rape of children. During the subsequent investigation, the police conducted more than 550 interviews of 220 witnesses and conducted 61 judicial hearings of 40 children. For six of the seven suspects implicated, charges were dropped. After a two-and-a-half-month long trial, the defendant was acquitted in 1994 (Kringstad, 1997; Riksadvokaten, 1994; Wiig & Brøgger, 2003). The allegations in this case and the criticism of the interviewers' competence, style of questioning, and the time from the alleged abuse to when the interview takes place, are similar to other highly publicised sexual abuse cases in the past decades, such as the McMartin pre-school and Kelly Michaels case in the US (Ceci and Bruck, 1995; Garven *et al.*, 1998; *State v Buckey*, 1990; *State v Michaels*, 1988, 1993, 1994), the Cleveland, Rochdale, and Orkney inquiries in the UK (Clyde, 1992) and the Roum case in Denmark (Nielsen, 2001).

One of the consequences following the Bjugn case was to increase the quality of the interview training for police officers conducting interviews of children. At the Norwegian Police University College, the interview training is based on scientifically and research-based theory. The basic police training is a 3-year, full-time educational programme, with theoretical and practical training for subsequent service and further specialisation. The specialisation of investigative interviewing of children is offered at two subsequent levels, denoted as *Level 1* and *Level 2*, respectively. The interviewers at Level 1 receive theoretical and practical training, including formal instructions on legal regulations, principles of cognition and memory enhancement of vulnerable witnesses, principles of structured

interviewing, children's linguistic abilities, suggestibility, tape-recorded samples of good/bad interviewing techniques, and role-playing exercise with group feedback by experts (PHS, 2003). Level 2 includes a more thorough formal instruction on some of the topics from Level 1. In addition, they get expert feedback on their own and other participants' conducted interviews (PHS, 2002). Both the law and the training programmes for police officers conducting investigative interviews strongly emphasises the use of open questions (Act of 22 May 1981, No. 25; Myklebust, 2005). The interviewers who have attended both levels are the most qualified interviewers in Norway and have dedicated most of their professional careers to the area of investigative interviews of children. These interviewers have received a thorough knowledge of the field, with emphasis on the use of open questions. The interviewers in Level 1 are also specially trained. However, they have not attended to similar number of theoretical lectures and they have conducted a smaller number of *field investigative interviews of children* (FIIC). In this study, we have used police officers from these two levels of education in order to assess the effect of long-term institutionalised training on the use of open and closed questions in FIIC. To our knowledge, the total length of the basic and specialised training programme in the present study is the longest institutionalised training programme that has been studied in order to assess any training effect on OCR in FIIC.

We have analysed the interviewers' questions using videotaped FIIC conducted by experienced interviewers. For ethical, practical, and methodological reasons, most of the research on child witnesses has been conducted in experimental laboratories. Accuracy is much more difficult to establish in the field than in laboratory analogue context because investigative interviewers seldom know what really happened in the alleged case. Furthermore, laboratory settings might not provide the police officers with the same sense of urgency and motivation often experienced in the field (Fisher, 1995). Experimental interviewers are rarely given extensive background information of the event, whereas in real investigations, ideas and hypotheses about the incident, or antecedent perceptions of the interviewee, may influence the way the interviews are conducted (Mortimer, 1994; Oxburgh, Williamson, & Ost, 2006). Furthermore, the interviewer's awareness of being monitored may affect their regular way of conducting the interview. In total, this favours field studies when analysing how investigative interviews are conducted and we question the ecological validity of previous studies using experimental conditions (Geiselman, Fisher, Cohen, Holland, & Surtes, 1986), or the use of personnel without any formal investigative or interviewer training (Fisher, Geiselman, Raymond, Jurkevich, & Warhaftig, 1987). The studies of methods used in police interviews do not have common definitions of which part of the interviews should be included. By only analysing the interviewers' questions regarding the alleged offence, the analysis excludes the introductory and closing sequence of the interview (Aldridge & Cameron, 1999; Lamb, Hershkowitz, Sternberg, Esplin, Hovav, Manor, & Yudilevitch, 1996). All the questions in the complete interview might be essential for the interviewer's collection and validation of the information given by the interviewee. Consequently, the total length of the interviews should be analysed. The interviewed children in our study are all school children and the literature emphasises the use of open questions also for this age group (Dent & Stephenson, 1979; Lamb, Hershkowitz, Sternberg, Esplin, Hovav, Manor, & Yudilevitch, 1996).

Fisher, Geiselman, and Raymond (1987) argue that interviews of children should be conducted by trained interviewers who use open questions. In this study, we will assess interviews undertaken by the most qualified police officers in Norway. First, we hypothesise that extensively trained police officers will generate more frequent use of open ques-

tions in real field interviews of children, as compared with less-trained police officers. Consequently, the null hypothesis is that there will be no difference between the two groups of police officers as regards the use of open and closed questions independently of training level. Training of police officers who conduct interviews focus on the use of open questions in the first part of the interviews. Second, we hypothesise that the more trained police officers will use more open questions during the first phases of the interviews compared with the less-trained police officers.

METHOD

Sample

At the Norwegian Police University College, education in the investigative interviewing of children could be divided into two groups based on the educational level. Police officers at *Level 2* have more substantial theoretical and practical training compared with officers at *Level 1*. Police officers from both groups were requested to participate in the study.

Inclusion criteria for the officers were:

- (1) they should select one of their interviews and report their gender, age, hours of theoretical training, number of practiced interviews, and the age/gender of the interviewed child;
- (2) the interview should be fully video-recorded, including all sequences from the child entering the interview room until leaving the room;
- (3) the interview should be completely transcribed into a written document and the transcriptions should be approved by the judge in charge of the interview according to the national regulations; and
- (4) the video-recorded interview should be the only interview of the child in the case, so the child's total amount of information in the case should be indicated on the video.

We selected the first 50 interviewers we received from the two groups of police officers, denoted as Level 1 and Level 2, respectively.

Procedure

Each of the 100 transcriptions were validated with the actual performance on the authentic videotapes. First, the transcriptions were electronically digitalised in order to allow computer-based analyses of all the words in the interviews. Second, the transcriptions were compared word by word with the videotapes by each of the authors separately. Any disagreements between the authors were subsequently discussed before any modifications were done to the electronic version of the transcriptions. A total number of 23,756 questions were controlled and corrections were made in 48 cases (0.2%). The interviews were then anonymised in each of the two groups of interviewers and the interviewers' questions in the transcribed interviews were classified by one of the authors. This was done according to the criterion of *open* and *closed* questions in the content dictionary first developed by Richardson, Dohrenwend, and Klein (1965) and further elaborated by Poole and Lamb (1998). An *open question* requires more than a few words for an adequate response (e.g. 'Tell me about that') in the Norwegian language. A *closed question* can be answered adequately in a few words. This group of questions includes a limited number of response alternatives. The closed question is either:

- (1) *identification* (e.g. person, place, group, time, number, etc.): who, where, when, how many, or which questions (e.g. 'What was his name?');
- (2) *selection*: fixed-alternative question; or
- (3) *yes-no* type questions answered by 'yes' or 'no', possibly supplemented by a redundant phrase, such as 'I think so' or 'I doubt it'.

Open questions deal with broader topics than closed questions and allow children more flexibility to choose which aspects of an event they will describe—e.g. 'Tell me about the person you met' is more open than 'What did he look like?'.

We validated the classification of questions by randomly choosing 10 of the interviews in Level 1, together with 10 of the interviews in Level 2—this gave a total of 3830 questions. The results from the 20 validated interviews gave an inter-rater reliability coefficient of 0.96, indicating a high level of agreement between the authors. The validation revealed disagreement in 154 questions classified as 16 open/closed and 138 open/identification questions, respectively.

The numbers of open and closed questions were calculated in each interview. Differences between the groups (Level 1/Level 2) and the kind of questions (open/closed) were analysed in a mixed between- and within-subject analysis of variance (ANOVA) design. Effect size is indicated by partial Eta squared. To study how the pattern of questioning changed over time, the videos were used to divide the interviews into three equally long tertiaries. Consequently, even if the duration of the interviews varied between 14 and 104 minutes ($M = 51$ minutes; $SD = 22.25$), they were normalised into three tertiaries. The numbers of open and closed questions were calculated for each of the three tertiaries, generating six cell means for each interview. Gender distribution amongst officers and children were analysed by Chi-square non-parametric test and the other differences between the two groups were analysed by Student's *t*-tests. Five per cent was used as the alpha level in all the statistical tests. The study was approved by the Norwegian Director of the Public Prosecution and conducted according to the standards of the Norwegian Regional Committee of Research Ethics.

RESULTS

Table 1 shows the demographic information from the 100 investigative interviews divided into two groups of interviewers, (Level 1 and Level 2). The mean duration of the interviews were 50 minutes in the Level 1 ($SD = 21.55$) and 53 minutes ($SD = 23.00$) in the Level 2 group ($t = 0.81$, $p > 0.05$). The mean age of the interviewers was 34 years ($SD = 5.85$) in the Level 1 group and 37.9 years ($SD = 6.62$) for Level 2 ($t = 3.07$, $p < 0.01$). The officers in Level 2 had 32 hours more formal training than the officers in Level 1 group, giving a significant difference ($p < 0.001$) between the two groups, with 139 hours of training for Level 1 ($SD = 0$ (similar length of education amongst all officers in the group)) and 171 hours for Level 2 ($SD = 0$), respectively. The mean number of conducted interviews was 11.8 ($SD = 7.00$) for Level 1 and 26.2 ($SD = 16.10$) for Level 2. The difference in the number of conducted interviews was significant ($t = 5.78$, $p < 0.001$). The distribution of gender amongst the interviewers and children were almost identical in the two groups with 43 female interviewers and 35 interviewed girls in the Level 1 group, and 44 female interviewers and 36 interviewed girls in Level 2, respectively (both p 's > 0.05 , Chi-square test). The mean age of the interviewed children was 10 years 2 months

Table 1. Demographic information and significant differences between the two groups of police officers (Level 1 and Level 2)

	Interviewers' level of training		Level of significance
	Level 1	Level 2	
Number of interviews	50	50	ns*†
Duration (minutes)	50 (21.6)	53 (23.0)	ns*†
Age of interviewer (years)	34.0 (5.9)	37.9 (6.6)	**†
Formal training (hours)	139 (0)	171 (0)	***†
Conducted interviews	11.8 (7.0)	26.2 (16.1)	***†
Gender of interviewer	43♀ 7♂	44♀ 6♂	ns*‡
Gender of children	35♀ 15♂	36♀ 14♂	ns*‡
Age of children	10 years 2 months (2.0)	10 years 3 months (2.2)	ns*†

SDs are in parentheses.

*ns, non-significant ($p > 0.05$); ** $p < 0.01$; *** $p < 0.001$.

†Student's *t*-test.

‡Chi-square test.

Table 2. The distribution of the total number of questions divided into the two groups of police officers (Level 1 and Level 2)

Type of questions	Interviewers' level of training		Total
	Level 1	Level 2	
Open	1083	960	2043
Closed			
Identification	3730	3677	7407
Selection	530	458	988
Yes/No	7051	6267	13318
Total	12394	11362	23756

($SD = 2.03$) and 10 years 3 months ($SD = 2.22$) in the Level 1 and Level 2 group, respectively. The age of the children was not significantly different in the two groups ($t = 0.24$, $p > 0.05$).

Table 2 shows the total number of the interviewers' questions distributed into the categories of questions for the Level 1 and Level 2 groups. The total number of questions was 23,756. A total of 2043 questions were classified as *open* questions and 21,713 as *closed* questions. Of the closed questions, 7407 questions were categorised as identification type, 988 questions as selection type, and 13,318 questions as yes/no type. Table 3 shows the mean number of open and closed questions distributed for the two groups. A 2(Level 1–Level 2) \times 2(open–closed questions) mixed between- and within-subject ANOVA revealed a significant main effect of question type. The mean number of closed questions was 217.1 ($SD = 111.3$) and the mean number of open questions was 20.4 ($SD = 15.2$), $F(1,98) = 330.4$, $p < 0.001$, partial Eta squared = 0.771. ANOVA did not show any difference between the number of questions used in each group, Level 1, $M = 227.2$ ($SD = 125.2$), Level 2, $M = 247.9$ ($SD = 107.7$), $F(1,98) = 0.78$, $p > 0.05$, partial Eta squared = 0.008. The interaction between the two groups of police officers and the two kinds of question was not significant, $F(1,98) = 0.53$, $p > 0.05$, partial Eta squared = 0.005.

Table 3. The distribution of the mean number of open and closed questions divided into the two groups of police officers (Level 2 and Level 1)

Type of questions	Interviewers' level of training		Level of significance
	Level 2	Level 1	
Open	19.2 (15.08)	21.7 (15.36)	ns
Closed	208.0 (121.86)	226.2 (100.14)	ns

SDs are in parentheses.
ns, non-significant.

Table 4. The distribution of the mean number of open and closed questions after the interviews are divided into three tertiaries

Type of questions	Tertiaries		
	First	Second	Third
Open	8.8 (5.7)	6.7 (6.0)	4.9 (4.9)
Closed	66.9 (40.7)	78.3 (41.5)	71.9 (38.1)

SDs are in parentheses.

Table 4 shows the means of open and closed questions after the interviews are divided into three tertiaries. The total number of open questions were highest in the first tertiary, ($M = 8.8$; $SD = 5.7$) and decreased in the second ($M = 6.7$; $SD = 6.0$) and third tertiary ($M = 4.9$; $SD = 4.9$). For all the interviews, the number of closed questions was highest in the second tertiary ($M = 78.3$; $SD = 41.5$), and lowest in the first ($M = 66.9$; $SD = 40.7$) and third tertiary ($M = 71.9$; $SD = 38.0$). The interaction between the type of question and the tertiaries was significant, $F(1,2) = 14.4$, $p < 0.001$. ANOVA revealed a significant interaction between the type of questions and the tertiaries, $F(1,2) = 17.3$, $p < 0.001$, partial Eta squared = 0.230. Post-hoc analysis showed that the number of open questions was different in all three tertiaries (all t 's > 5.0 , all p 's < 0.001). Similar analysis of the closed questions showed that the number of questions in the second tertiary was significantly different from the first ($t = 4.5$, $p < 0.001$) and third tertiary ($t = 2.4$, $p < 0.05$), but the number of closed questions was not different in the first and third tertiary ($t = 1.9$, $p > 0.05$). We did not observe any significant time dependent difference in the use of open and closed questions between the two level groups.

DISCUSSION

We have analysed 100 transcripts of FIIC in their entirety in order to assess the effect of long-term training in the interviewers' use of open and closed questions. Three main findings have been observed. The first finding was that a significant smaller number of open questions were found compared to the number of closed questions. The second finding was a non-significant difference in the use of open and closed questions between the two groups of interviewers. The third observation was an effect of the number of questions distributed during the interview. The number of open questions in the last tertiary was

about half of the number in the first tertiary. The distribution of closed questions was similar to an inverted U with the highest number of questions in the second tertiary.

Fisher, Geiselman, and Raymond (1987) formulated the significant difference between the number of open and closed questions as an OCR. In our study, we have found an overall OCR of 1:10, which is about 10 times as many closed questions compared with the total number of open questions. In their pioneering study, Fisher, Geiselman and Raymond (1987) examined 11 tape-recorded interviews and reported an OCR of 1:9. Later, Lamb *et al.* (1996) studied 22 interviews and classified the questions into 10 categories. Based upon their results, we have calculated an OCR of 1:38. Fisher, Geiselman and Raymond (1987) analysed the interviewers' utterances with regard to the total length of the interviews. A different method was used by Aldridge and Cameron (1999), who investigated the types of questions used by the interviewers during the free-narrative and questioning stages of the interview. Their definition and set of criteria was based upon the step-wise approach dividing the interview into four stages, denoted by rapport, free narrative, open-ended questions, and specific questions. Lamb and his colleagues used another method by analysing a *substantive portion* of each interview. The substantive portion was operationally defined by Lamb *et al.* (2000) as the portion of the interview during which the incidents under investigation were discussed. Their analysis of the interview protocol starts with the first scripted substantive question and continues until the interviewee is shifting focus to a neutral non-substantive topic at the end of the interview. Lamb *et al.*'s (1996; 2000) limitation of the examination to the substantive portion of the interviews does not include any further description on how to compare the substantive portion of the interview with the total length of the entire interview. We have not found any studies reporting that the result from the substantive portion of an interview diverges from the entire length, but it is still unclear how to compare the different studies. Craig, Scheibe, Raskin, Kircher, and Dodd (1999) analysed 48 interviews and used six categories when they classified all the interviewers' utterances. Based upon their findings, we have calculated the OCR to be close to 1:4. Davies, Westcott, and Horan (2000) reported 2% of open questions, corresponding to an OCR of 1:50. They examined 36 interviews; however, it is unclear if they included the entire lengths of the interviews. All the previously mentioned studies clearly report an essential low number of open questions compared with the number of closed questions. Since Stern (1903/04) and Pear and Wyatt (1914) advocated that children are reliable only when they respond spontaneously, the use of closed questions has been considered as susceptible to suggestion, whilst open questions facilitates the children's free narrative responses. In the present study, the interviewers asked open questions in all the interviews even though the amount has to be considered quite low. Aldridge and Cameron (1999) reported that only 11 of the sample of 27 interviews in their study contained requests for free narratives. Similarly, Davies, Wilson, Mitchell, and Milsom (1995) found that in 28% of their 40 evaluated FIIC, there were no free narratives. In summary, FIIC does not facilitate children's free narrative by open questions. The predominant use of closed questions is widespread. It occurs internationally and across different groups of interviewers (Aldridge & Cameron, 1999; Cederborg, Orbach, Sternberg, & Lamb, 2000; Craig *et al.*, 1999; Davies, Westcott, & Horan, 2000; Fisher *et al.*, 1987; Lamb, Sternberg, & Esplin, 1995; Lamb *et al.*, 1996; Moston, Stephenson, & Williamson, 1993; Santtila, Korkman, & Sandnabba, 2004; Sternberg, Lamb, Davies, & Westcott, 2001).

The main theory that has been advocated in order to explain the relatively low number of open questions has addressed the competence of the police officers that have conducted

the interviews. The competence theory has been studied in several papers. Fisher, Geiselman, Raymond, Jurkevich, and Warhaftig (1987) were the first to examine the effect of training on the competence of the interviewers. They conducted an experiment that included 16 undergraduate students as mock victims and three undergraduate college students as mock interviewers. The authors reported a significant effect of training on the number of correct answers from the interviewed students, but did not report the OCRs before and after the training. Since the study only included adult students in an experimental condition, it remains unclear if the findings are valid compared to real-life field interviews of children. There have been numerous studies of FIIC, starting with Sternberg, Lamb, Hershkowitz, Yudilevitch, Orbach, Esplin, and Hovav (1997); however, Lamb and his colleagues argued for extensive interview training. They analysed 21 transcripts of experienced interviewers before and after completing one of four training conditions. They concluded that OCR improved from 1:15 in the pre-trained officers group to an OCR of 1:8 for the group of officers completing the training conditions. They found that didactic workshops and instructions in the utilisation of structured interview procedures had little effect on the number of open questions. By contrast, intensive training in the use of a structured interview protocol followed by continuing supervision and feedback resulted in the reported improvement in both the OCR and the number of details in the children's statements.

The importance of supervision was further demonstrated in Lamb *et al.* (2002) when they analysed 74 interviews in an experimental condition, with and without supervision. Thirty-seven of the interviews were conducted under supervision and the other 37 after the supervision had been terminated. The OCRs were reported to be equivalent to 1:2 and 1:4 in the conditions with and without supervised interviews, respectively. The authors concluded that the positive effect of supervision on the number of open questions decreased after supervision and hypothesised that the effect of training obtained under supervision ceased in the post-supervision condition. Lamb *et al.* (2002) did not discuss any theory related to the effect of supervision, hence it is not clear if the feedback served to improve the interviewers' use of open questions or if other aspects of the more global term 'supervision' could be alternative candidates to elicit a free, narrative account amongst the interviewers.

To our knowledge, all the authors who have studied the OCR have explained the relatively low number of open questions due to inexperienced interviewers and they have argued that better training may increase the use of open questions. The training theory has been studied within two traditions. One tradition has been the short-term training (Fisher & Geiselman, 1992) and the other has been long-term training tradition (Lamb *et al.*, 2000, 2002). In the present study, we hypothesised that the extensive long-term training, including both several theoretical courses and a large number of FIIC, may improve the competence of the police officers' interviewing skills and subsequently increase the use of open questions amongst the more trained and experienced police officers. However, we did not observe any support for the long-term training effect hypothesis on the OCR. The police officers in the Level 2 group received significantly more theoretical and practical training compared with the other group. The police officers in the Level 2 group were older and may have had more of a routine compared with the other officers, however, interestingly, they did not use more open questions. The non-significant effect of long-term training in the present study may invalidate any simple theory stating that training results in a change in OCR in such a way that the relative number of open questions increases.

Analysis of the total number of questions distributed into tertiaries showed an inverted U distribution with most questions posed in the second tertiary. The distribution of the

open questions descended during the three subsequent tertiaries, with 8.8, 6.7, and 4.9 questions, respectively. According to the theory for structural models, the OCR should change during the interview in such a way that the proportion of open questions declines, with the proportion of closed questions ascending. Our results support the declining slope of open questions, but not the increasing slope of closed questions during the interview. For the closed questions, we found an inverted U distribution and we divided the normalised interviews into three equal-length time tertiaries. One might argue that another division of the interviews might better reflect a highly structured model; however, several *ad hoc* divisions of the results obtained in the present study did not reveal any distribution of closed questions diverting from the inverted U distribution.

Our results could be vulnerable to several factors. *First*, it is unclear whether the quality of the training that the officers in the Level 2 group had received was in accordance with the requirements of optimal training in modern interviewing techniques. We have not included any evaluation of the training programmes in police interviewing at the Norwegian Police University College; however, since the Bjugn case in 1994, the focus on scientifically based interviewing techniques at the Police University College has increased substantially. Both the law and the training programmes for police officers conducting investigative interviews strongly emphasises the use of open questions (Act of 22 May 1981, No. 25; Myklebust, 2005). Compared with other countries, we have no information to consider the Norwegian training programmes in interviewing to be sub-standard or flawed. We believe that the significant larger number of child interviews conducted amongst the officers in the advanced group has been followed up by more discussions and evaluations appropriate to stimulate more attention on the child and on the co-operation with the child. *Second*, even if the quality of the interview training of the police officers has been good enough, the situational conditions at the local area of the interview could counteract with the interviews to be conducted according to the recommended standards. One of the authors of the present study has himself conducted a large number of police interviews and it is not easy to argue that situational conditions may explain the non-significant effect of long-term training in the use of open questions between the two groups of officers. The situation as such, with shortage of planning time, absence of information about the child, and so forth, may influence the quality of the interview, but conversely, may not affect the use of open question as such. *Third*, the police officers have themselves chosen the interviews that have been included in the study and perhaps the younger officers have used more recourses to select an outstanding FIIC. *Finally*, one might object that the present study does not include officers who have received sufficient long-term training. The significant age difference between the two groups indicates that the police officers have been occupied with conducting child interviews for different periods. In addition, the requirements to be included in the advanced groups reflect a sustained motivation and adherence to working with children.

The present findings do not support the long-term training theory, which suggests that training will generate more frequent use of open questions. We consider that the results reported by Lamb *et al.* (2002) are quite interesting since they observed a greater OCR in the case of supervision compared with post-supervision. Lamb *et al.* (2002) argued that supervision may facilitate the use of open questions; however, we advocate that supervision may generate a group atmosphere of co-operation between several police officers, which concentrates on performing optimal interviews. Several studies (e.g. Cannon-Bowers, Salas, & Converse, 1993; Hardin & Higgins, 1996) have indicated that team-

based co-operation generates better communication strategies compared to individual activity.

Ericsson, Krampe, and Tesch-Romer (1993) formulated a theory of deliberate practice that includes focus on co-operation between group members in a team. They suggested that deliberate practice was an effortful activity motivated by the goal of improving performance. This requires effort and attention from the learner and involves activities selected by a supervisor or instructor to further learning. Irrespective of the specific skill or current level of performance, which is to be improved, instructions from an authority figure are rated more important than other structured activities (Ericsson, Krampe, & Tesch-Romer, 1993; Sternberg, Forsythe, Hedlund, Horvath, Wagner, Williams, Snook, & Grigorenko, 2000). Hodges and Starkes (1996) and Helsen, Starkes, and Hodges (1998) studied a group of athletes, with their results supporting the deliberate practice theory. They argued that in addition to the instructor, the most critical part of producing skilled performers was finding individuals who were highly motivated and likely to persist over the long duration required to produce a field expert. We assume that the long-term training and use of supervisors in groups might explain why Lamb and his colleagues reported an increased OCR (Lamb *et al.*, 2000; Lamb, Sternberg, Orbach, Esplin, & Mitchell, 2002).

Supervision and team-based interviews demand appropriate skills and knowledge, both by the interviewers and supervisors. We suggest that the FIIC training programmes at the Norwegian Police University College have more focus on efficient team-based communication skills. The present findings indicate that individual, practical, and formal training does not increase the use of open questions in interviews. The effect of team-based interviewing could be investigated by comparing interviews performed by one officer with interviews performed by several officers.

Team-based interviews might have some shortcomings compared with interviews conducted by an individual officer. A set of disadvantages may be attributed the circumstances that two officers may confuse a child more than one officer if they constantly change to be in charge or stimulate the child to give different narrative recall of the alleged case. Consequently, team-based interviews require that the police officers are extensively trained in co-operation skills.

At an organisational level, the question of effective and optimised training has attracted substantial attention in general work and organisational psychology. Argyris and Schön (1974, 1978) developed a theory called *action-science* to explore the discrepancy between what police officers say they do and what they actually do during an interview. According to this theory, it is predicted that people act in certain ways under certain conditions when guided by certain values in order to achieve desired consequences. The theory proposes that interviewers should first try to change their tactics if things go wrong in the interview; a phenomenon they term single-loop learning. If the change in tactics does not work, the interviewer has to reconsider and reanalyse the assumptions, values, beliefs, and norms that influenced the negative action in the interview. The latter is termed *double-loop learning* and shows that people often believe that they are acting in accordance with a set of beliefs (espoused theory). However, because of individual differences, they are in fact often acting in ways that contradict their espoused theories (theories-in-use) (Argyris & Schön, 1974, 1978; Deutsch & Coleman, 2000). To achieve double-loop learning, we argue that the guided interview training should be similar to the real FIIC, conducted by interviewers working in the team, in the same environment and under the same conditions as real investigative interviews. For each of the interviewers, we also argue that what

happens at the local police station after the formal training is as important as the initial training itself. At the institutional level, a programme and framework for ongoing and regular support, evaluation, and feedback is required to achieve double-loop learning and long-time effect of the interview training.

In conclusion, the present study reports a low number of open questions in field interviews of children and we failed to find any effect of extensive long-term training, including both several training courses and a large number of field interviews, as a function of extensively trained police officers' use of open questions compared to less-trained officers. Analysis of interviews in order to find any time effects during the interviews does not support a structured model. We argue that a team-based approach to interviews of children would increase the number of open questions, which, as a consequence, would yield more accurate information from the interviewee.

REFERENCES

- Act of 22 May 1981, No. 25. Criminal Procedure Act Concerning Legal Procedures in Criminal Cases [Straffeprosessloven—Lov av 22. mai 1981 nr. 25 om rettergangsmåten i straffesaker.]
- Aldridge, J. (1992). The further training of professionals dealing with child witnesses. In H. Dent, & R. Flin (Eds.), *Children as witnesses* (pp. 231–244). Chichester, UK: Wiley.
- Aldridge, J., & Cameron, S. (1999). Interviewing child witnesses: Questioning techniques and the role of training. *Applied Developmental Science*, 3, 136–147.
- American Professional Society on the Abuse of Children (APSAC). (1990). *Guidelines for psychosocial evaluations of suspected sexual abuse in young children*. Chicago, IL: APSAC.
- Argyris, C., & Schön, D. A. (1974). *Theories in practice: Increasing professional effectiveness*. San Francisco, CA: Jossey-Bass.
- Argyris, C., & Schön, D. A. (1978). *Organizational learning*. Reading, MA: Addison-Wesley.
- Brainerd, C., & Ornstein, P. A. (1991). Children's memory for witnessed events: The developmental backdrop. In J. Doris (Ed.), *The suggestibility of children's recollections* (pp. 10–20). Washington, DC: American Psychological Association.
- Cannon-Bowers, J. A., Salas, E., & Converse, S. (1993). Shared mental models in expert team decision making. In N. J. Castellan (Ed.), *Individual and group decision making: Current directions* (pp. 221–246). Hillsdale, NJ: Erlbaum.
- Ceci, S. J., & Bruck, M. (1995). *Jeopardy in the Courtroom. A scientific analysis of children's testimony*. Washington, DC: American Psychological Association.
- 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.
- Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127–149). Washington, DC: American Psychological Association.
- Clyde, J. J. (1992). *The report of the inquiry into the removal of children from Orkney in February 1991*. Edinburgh: HMSO.
- Craig, R. A., Scheibe, R., Raskin, D. C., Kircher, J. C., & Dodd, D. H. (1999). Interviewer questions and content analysis of children's statements of sexual abuse. *Applied Developmental Science*, 3, 77–85.
- Dale, P. S., Loftus, E. F., & Rathbun, L. (1978). The influence of the form of the question of the eyewitness testimony of preschool children. *Journal of Psycholinguistic Research*, 7, 269–277.
- Davies, G., Westscott, H. L., & Horan, N. (2000). The impact of questioning style on the content of investigative interviews with suspected child sexual abuse victims. *Psychology, Crime and Law*, 6, 81–97.
- Davies, G., Wilson, C., Mitchell, R., & Milsom, J. (1995). *Videotaping children's evidence: An evaluation*. London, UK: Home Office.
- Dent, H. R. (1986). An experimental study of the effectiveness of different techniques of questioning mentally handicapped child witnesses. *British Journal of Clinical Psychology*, 25, 13–17.

- Dent, H. R. (1991) Experimental studies of interview child witnesses. In J. Doris (Ed.), *The suggestibility of children's recollections* (pp. 138–146). Washington, DC: American Psychological Association.
- Dent, H. R., & Stephenson, G. M. (1979). An experimental study of the effectiveness of different techniques of questioning child witnesses. *British Journal of Social and Clinical Psychology*, *18*, 41–51.
- Deutsch, M., & Coleman, P. T. (2000). *The handbook of conflict resolution: Theory and practice*. San Francisco, CA: Jossey-Bay.
- Ericsson, K. A., Krampe, R. T., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*, 363–406.
- Fisher, R. P. (1995). Interviewing victims and witnesses of crime. *Psychology, Public Policy and Law*, *1*(4), 732–764.
- Fisher, R. P., & Geiselman, R. E. (1992). *Memory-enhancing techniques for investigative interviewing. The cognitive interview*. Springfield, IL: Charles C. Thomas Publisher.
- 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*, *5*, 722–727.
- Fisher, R. P., Geiselman, R. E., & Raymond, D. S. (1987). Critical analysis of police interview techniques. *Journal of Police Science and Administration*, *15*(3), 177–185.
- Fisher, R. P., Geiselman, R. E., Raymond, D. S., Jurkevich, L. M., & Warhaftig, M. L. (1987). Enhancing enhanced eyewitness memory: Refining the cognitive interview. *Journal of Police Science and Administration*, *15*(4), 291–297.
- 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*(3), 347–359.
- Geiselman, R. E., Fisher, R. P., Cohen, G., Holland, H., & Surtes, L. (1986). Eyewitness responses to leading and misleading questions under the cognitive interview. *Journal of Police Science and Administration*, *14*(1), 31–39.
- George, R., & Clifford, B. (1992). Making the most of witnesses. *Policing*, *8*, 185–198.
- Goodman, G. S., & Aman, C. (1990). Children's use of anatomically detailed dolls to recount an event. *Child Development*, *61*, 1859–1871.
- Goodman, G. S., Aman, C., & Hirshman, J. (1987). Child sexual and physical abuse: children's testimony. In S. J. Ceci, M. P. Toglia, & D. P. Ross (Eds.), *Children's eyewitness memory* (pp. 1–23). New York: Springer-Verlag.
- Goodman, G. S., Bottoms, B. L., Schwartz-Kenney, B. M., & Rudy, L. (1991). Children's testimony for a stressful event: Improving children's reports. *Journal of Narrative and Life History*, *1*, 69–99.
- Hardin, C., & Higgins, E. T. (1996). Shared reality: How social verification makes the subjective objective. In R. M. Sorrentino, & E. T. Higgins (Eds.), *Handbook of motivation and cognition: Foundations of social behaviour* (3rd ed., pp. 28–42). New York: Guilford Press.
- Helsen, W. F., Starkes, J. L., & Hodges, N. J. (1998). Team sports and the theory of deliberate practice. *Journal of Sport and Exercise Psychology*, *20*, 12–34.
- Hershkowitz, I., Lamb, M. E., Sternberg, K. J., & Esplin, P. W. (1997). The relationships among interviewer utterance type, CBCA scores, and the richness of children's responses. *Legal and Criminological Psychology*, *2*, 169–176.
- Hershkowitz, I., Horowitz, D., Lamb, M. E., Orbach, Y., & Sternberg, K. J. (2004). Interviewing youthful suspects in alleged sex crimes: A descriptive analysis. *Child Abuse and Neglect*, *28*, 423–438.
- Hindman, J. (1987). *Step by step. Sixteen steps toward legally sound sexual abuse investigation*. Ontario: Alexandria Associates.
- Hodges, N. J., & Starkes, J. L. (1996). Wrestling with the nature of expertise: A sport specific test of Ericsson, Krampe and Tesch-Romer's (1993) theory of deliberate practice. *International Journal of Sport Psychology*, *27*, 400–424.
- Home Office. (2002). *Achieving best evidence in criminal proceedings: Guidance for vulnerable or intimidated witnesses, including children*. London: Home Office Communication Directorate.
- Home Office in conjunction with Department of Health. (1992). *Memorandum of good practice on video recorded interviews with child witnesses for criminal proceedings*. London: HMSO.

- Hutcheson, G., 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.
- Kringstad, H. (1997). *Bjugnformelen* [The Bjugnformula]. Oslo: Tiden Norsk Forlag.
- Lamb, M. E. (1994). The investigation of child sexual abuse: An interdisciplinary consensus statement. *Child Abuse and Neglect, 18*, 1021–1028.
- Lamb, M. E., & Fauchier, A. (2001). The effects of question type on self-contradiction by children in the course of forensic interviews. *Applied Cognitive Development, 15*, 483–491.
- Lamb, M. E., Hershkowitz, I., Sternberg, K. J., Esplin, P. W., Hovav, M., Manor, T., & Yudilevitch, L. (1996). Effects of investigative utterance types on Israeli children's responses. *International Journal of Behavioural Development, 19*(3), 627–637.
- Lamb, M. E., Sternberg, K. J., & Esplin, P. W. (1994). Factors influencing the reliability and validity of statements made by young victims of sexual maltreatment. *Journal of Applied Developmental Psychology, 15*, 255–280.
- Lamb, M. E., Sternberg, K. J., & Esplin, P. W. (1995). Making children into competent witnesses: Reactions to the amicus brief in re Michaels. *Psychology, Public Policy, and Law, 1*, 494–520.
- Lamb, M. E., Sternberg, K. J., Orbach, Y., Esplin, P. W., & Mitchell, S. (2002). Is ongoing feedback necessary to maintain the quality of investigative interviews with allegedly abused children? *Applied Developmental Science, 6*(1), 35–41.
- Lamb, M. E., Sternberg, K. J., Orbach, Y., Hershkowitz, I., Horowitz, D., & Esplin, P. W. (2000). The effects of intensive training and ongoing supervision on the quality of investigative interviews with alleged sex abuse victims. *Applied Developmental Science, 6*(3), 114–125.
- Loftus, E. F. (1979). *Eyewitness testimony*. Cambridge, MA: Harvard University Press.
- Manning, P. K. (1988). *Symbolic communication. Signifying calls and the police response*. London: The MIT Press.
- Mortimer, A. (1994). *Police investigative interviewing: Pre-interview thinking and written interview behaviour*. Paper presented at the 4th European Conference on Psychology and Law, Barcelona.
- Moston, S., Stephenson, G. M., & Williamson, T. (1993). The incidence, antecedents and consequences of the use of the right to silence during police questioning. *Criminal Behaviour and Mental Health, 3*, 30–47.
- Myklebust, T. (2005). Dommeravhør av barn i Norge—intensjon og fakta [Investigative interviews of children in Norway—intention and facts]. In M. Egge, & J. Strype (Eds.), *Polititrollen gjennom 100 år. Tradisjon og endring* [The role of the police in 100 years. Tradition and changes] (pp. 142–172.). Oslo: PHS Forskning 2005: 2.
- Myklebust, T., & Alison, L. (2000). The current state of police interviews with children in Norway: How discrepant are they from models based on current issues in memory and communication. *Psychology, Crime and Law, 6*, 331–351.
- Nielsen Beth, G. (2001). *Straffesystemet i børneperspektiv. Seksuelle overgrep mot barn i familien* [Criminal law administration from a child's perspective. Child sexual abuse cases within the family]. København: Jurist- og Økonomforbundets Forlag.
- Oats, K., & Shrimpton, S. (1991). Children's memories for stressful and non-stressful events. *Medical Science and Law, 31*, 4–10.
- Orbach, Y., Hershkowitz, I., Lamb, M. E., Sternberg, K. J., Esplin, P. W., & Horowitz, D. (2000). Assessing the value of structured protocols for forensic interviews of alleged child abuse victims. *Child Abuse and Neglect, 24*(6), 733–752.
- 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., Gordon, B. N., & Larus, D. M. (1992). Children's memory for a personally experienced event: Implications for testimony. *Applied Cognitive Psychology, 6*, 49–60.
- Oxburgh, G. E., Williamson, T., & Ost, J. (2006). Police officers' use of emotional language during child sexual abuse investigations. *Journal of Investigative Psychology and Offender Profiling, 3*, 35–45.
- PHS. (2002). *Kursplan i utdanning i etterforskning av seksuelle overgrepssaker trinn III* [Course description for education in investigation of sexual abuse cases and investigative interviews of children and vulnerable witnesses—Step III]. Oslo: Politihøgskolen.
- PHS. (2003). *Kursplan i utdanning i etterforskning av seksuelle overgrepssaker trinn II* [Course description for education in investigation of sexual abuse cases and investigative interviews of children and vulnerable witnesses—Step II]. Oslo: Politihøgskolen.

- Pear, T. H., & Wyatt, S. (1914). The testimony of normal and mentally defective children. *British Journal of Psychology*, 3, 388–419.
- Pipe, M. E., Lamb, M. E., Orbach, Y., & Esplin, P. W. (2004). Recent research on children's testimony about experienced and witnessed events. *Developmental Review*, 24, 440–468.
- Poole, D. A., & Lamb, M. E. (1998). *Investigative interviews of children: A guide for helping professionals*. Washington, DC: American Psychological Association.
- Queensland Family Services. (1992). *Interviewing children and recording evidence*. Education, Health and Police Services. Queensland: Family Services.
- Richardson, S. A., Dohrenwend, B. S., & Klein, D. (1965). *Interviewing: Its forms and functions*. London: Basic Books, Inc., Publishers.
- Riksadvokaten. (1994). *Riksadvokatens gjennomgang av den såkalte Bjugn-saken* [The Director of Public Prosecutions. An evaluation of the Bjugn case.] Oslo: Riksadvokaten.
- Santtila, P., Korkman, J., & Sandnabba, K. (2004). Effects of interview phase, repeated interviewing, presence of a support person, and anatomically detailed dolls on child sexual abuse interviews. *Psychology, Crime and Law*, 10, 21–35.
- Saywitz, K. J. (1988). The credibility of child witnesses. *Family Advocate*, 10, 38.
- Schooler, J. W., & Loftus, E. F. (1993). Multiple mechanisms mediate individual differences in eyewitness accuracy and suggestibility. In J. M. Puckett, & H. W. Reese (Eds.), *Mechanisms of everyday cognition* (pp. 177–203). Hillsdale, NJ: Lawrence Erlbaum Associates.
- State v Buckley*, Superior Court, Los Angeles County, California, #A750900 (1990).
- State v Michaels*, Superior Court, Essex County, New Jersey (1988).
- State v Michaels*, 264 N.J. Super 579, 625 A.D. 2d 489 (N.J. Super Ad, 1993).
- State v Michaels*. 136 N.J. 299, 642 A. 2d 1372 (N.J., 1994).
- Steller, M., & Boychuck, R. (1992). Children as witnesses in sexual abuse cases. In H. Dent, & R. Flin (Eds.), *Children as witnesses* (pp. 47–73). Chichester, UK: Wiley.
- Stern, W. (1903/04). *Beiträge zur Psychologie der Aussage*. Leipzig: Verlag von Johann Ambrosius Barth.
- Sternberg, R. J., Forsythe, G. B., Hedlund, J., Horvath, J. A., Wagner, R. K., Williams, W. M., Snook, S. A., & Grigorenko, E. L. (2000). *Practical intelligence in everyday life*. Cambridge, UK: Cambridge University Press.
- 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 types and the informativeness of child witnesses. *Journal of Applied Developmental Psychology*, 17, 439–451.
- Sternberg, K. J., Lamb, M. E., Hershkowitz, I., Yudilevitch, L., Orbach, Y., Esplin, P. W., & Hovav, M. (1997). Effects of introductory style on children's abilities to describe experiences of sexual abuse. *Child Abuse and Neglect*, 21, 1133–1146.
- The National Crime Faculty. (1998). *A practical guide to investigative interviewing*. Bramshill: Training and Development Unit.
- Walker, A. G. (1999). *Handbook on questioning children. A linguistic perspective*. Washington: ABA Center on Children and the Law.
- Walker, N. E., & Hunt, J. S. (1998). Interviewing child victim-witnesses: How you ask is what you get. In C. P. Thompson, D. J. Herrman, J. D. Read, D. Bruce, D. G. Payne, & M. P. Togli (Eds.), *Eyewitness memory: Theoretical and applied perspectives* (pp. 55–87). Mahwah, NJ: Lawrence Erlbaum Associates.
- Watzlawick, P., Bavelas, J. B., & Jackson, D. D. (1967). *Pragmatics of human communication. A study of interactional patterns, pathologies, and paradoxes*. New York: W. W. Norton & Company.
- Wiig, C., & Brøgger, Sr, J. (2003). *Hekseprosess. Bjugn-saken i et juridisk og kulturhistorisk perspektiv* [Witch trial. The Bjugn case in a legal and cultural historical perspective.] Trondheim: Communicatio.
- Wright, A. M., & Alison, L. (2004). Questioning sequences in Canadian police interviews: Constructing and confirming the course of events? *Psychology, Crime and Law*, 10, 137–154.
- Yuille, J. C., Hunter, R., Joffe, R., & Zaparniuk, J. (1993). Interviewing children in sexual abuse cases. In G. S. Goodman, & B. L. Bottoms (Eds.), *Child victims, child witnesses: Understanding and improving testimony* (pp. 95–115). New York: Guilford Press.

Factors affecting the length of responses in field investigative interviews of children (FIIC) in child sexual abuse cases

TROND MYKLEBUST^{1,*} and ROALD A. BJØRKLUND²

¹Norwegian Police University College, Box 5027, Majorstua, N-0301 Oslo, Norway

²Department of Psychology, University of Oslo, Box 1094, Blindern, N-0317 Oslo, Norway

Abstract

In this study some of the most prominent variables considered to facilitate the interviewees' responses in the literature of field investigative interviews of children (FIIC) are all reviewed and included in the analysis of one hundred FIIC. Analysis of variance (ANOVA) showed that of all the variables the categories of the interviewer's utterances had most impact on the children's responses with the open questions eliciting the longest answers. The variable to follow was the children's age, with the oldest children yielding longer responses than the younger children to the open questions. Contrary to our hypothesis, the interviewer's competence, children's gender, nor time had the expected impact on the length of the children's responses. Possible explanations for the results are discussed.

Keywords: Investigative interviews, FIIC, CSA, questions, answers, interview training, police.

INTRODUCTION

Empirical studies of field investigative interviews of children (FIIC) indicate a strong correlation between the number of words in the children's responses and the information contained in the utterances (Lamb, Hershkowitz, Sternberg, Esplin, et al. 1996) ; Sternberg, Lamb, Hershkowitz, Esplin, et al. 1996). With no other evidence typically available in child sexual abuse (CSA) cases, the implicit assertion in FIIC is to gain long, detailed and accurate responses from the interviewed children. Myklebust and Bjørklund (2009) demonstrated that the courts appear to be affected by

*Corresponding author: Trond Myklebust, Norwegian Police University College, P.O Box 5027, Majorstua, N-0301 Oslo, Norway. E-mail: trond.myklebust@phs.no.

children using many words in their testimonies. In the present study, 100 FIIC are analysed to reveal central variables facilitating the interviewees' responses in CSA cases. Prior to the mid 1990s there had been no systematic field studies of the ways in which investigative interviews of children were conducted. The researches of FIIC are mainly studies of individual factors by the children and interviewers, largely driven by a concern for non-contamination of the children's memory of the alleged offence in the interaction between the child and the interviewer. We will review some of the most prominent variables that have been considered in the literature related to FIIC, and try to include them in the present study.

The interviewers' utterances have almost always been an important variable. Lamb, Hershkowitz, Sternberg, Esplin, et al. (1996) was the first to examine the relation between type of investigative utterances and children's responses in FIIC. The authors analysed 22 FIIC conducted by 12 Israeli investigators and observed five types of investigative utterances; (i) *open* or *invitational*; (ii) *facilitative* (non-suggestive encouragements to continue talking); (iii) *directive* (in which the child's attention was refocused on something s/he earlier mentioned); (iv) *leading* (attention on thing or event s/he had not previously mentioned), and (v) *suggestive* (those which hinted an expected response), respectively. As predicted by the authors, the open questions (invitations) yielded responses significantly longer than any of the other three types of utterances that focused the child's attention: *direct*, *leading* or *suggestive*. Sternberg et al. (1996) used the same five categories of utterances and studied 45 interviews conducted by six male detectives in the United States. The authors compared children as victims of single versus multiple incidents and found the responses to open questions significantly longer than responses to closed questions. The term *closed* question refers to specific questions that have a limited number of response alternatives. The advantage of open questions was also the main finding in the Lamb, Hershkowitz, Sternberg, Boat, et al. (1996) examination of 24 interviews when the authors evaluated the effect of interviews conducted with or without anatomically detailed dolls. Korkman, Santtila, Westeråker, Sandnabba (2008) analysed their interviews using slightly different coding procedures introducing two suggestive categories (*specific* and *unspecific*). Richardson, Dohrenwend, and Klein (1965) were some of the first to identify characteristics of the question types to consider the effect at the level of the individual questions. They introduced one of the first classification systems tending to divide the closed questions into three sub categories, denoted '*identification*', '*selection*' or '*yes/no*'. Aldridge and Cameron (1999) defined five types of questions as: (i) *free reports*; (ii) *open*; (iii) *specific*; (iv) *leading*; and (v) *non-leading*. In a similar vein, Davies, Westcott, and Horan (2000) analysed interviews according to four question characteristics: (i) *open ended*; (ii) *closed*; (iii) *specific, yet not leading*; and, (iv) *leading*. Cederborg, Orbach, Sternberg, and Lamb (2000) described four question types: (i) *invitation*; (ii) *directive*; (iii) *option-posing*; and (iv) *suggestive*. Despite the agreement about the advantage of open questions as regards to the length and details in children's responses (Dent, 1991; Hershkowitz, Lamb, Sternberg, & Esplin, 1997; Lamb, 1994; Lamb, Hershkowitz, Sternberg, Esplin, et al., 1996; Poole & Lamb, 1998; Sternberg et al., 1996), there still remain discrepancies over how to best describe types of questions and their respective characteristics (see Oxburgh, Myklebust & Grant (submitted), for an overview). From a linguistic perspective, the system of categorising questions could be done with a lexical perspective focusing on the phrasing of the question (e.g. *selection*, or *yes/no*) or with focus on the function of a question (e.g. *invitation*, *leading*, or *suggestive*). Derived from this, the closed questions in the evaluation studies by Lamb et al., and Korkman et al. (op.cit.), could be categorised by their function (*directive*, *leading* and *suggestive*), while Richardson et al., (1965) could be

categorised by their lexical phrasing (*identification, selection or yes/no*). Oxburgh, Myklebust and Grant (submitted) found the characterisations of closed categories to differ within the literature. The authors found i.e. the 'wh...' lexis of 'who', 'what', 'where', 'which', and 'when', categorised as *closed-specific* by Loftus (1982), as *open* by CPTU (1992), and as both *open* and *closed* by Shepherd (2007). The authors found the divergent description in the literature to be mainly within the questions categorised by their function. To prevent this discrepancy we have used the classification system developed by Richardson et al. (1965), focusing on the lexical phrasing of the questions, in the present analysis of the Norwegian children's responses to the interviewer's utterance types.

Several studies indicate that the children's age and gender also affect the responses in FIIC. According to Foster-Cohen (1990) and Harley (2009) linguistic development of children is multi-dimensional and non-uniform with different aspects of communicative ability progressing at different rates for different children, with children moving from more direct, literal understandings and use of language, to a less direct, less immediate usage. As presented in table 1, studies of children from Israel (Lamb, Hershkowitz, Sternberg, Esplin et al., 1996), United States (Sternberg et al., 1996) and Finland (Korkman et al., 2008) have all observed an effect of the children's age on the number of words in the children's responses in FIIC.

Table 1 A comparison of FIIC studies from Israel, United States, and Finland.

	Number of interviews	Age of children	Number of words in response to				
			Invitation	Facilitators	Directive	Leading	Suggestive
Israel. Lamb, M. E., Hershkowitz, I., Sternberg, K. J., Esplin, P. W., Hovav, M., Manor, T., & Yudilevitch, L (1996) ¹ .	22	5-11 yrs	15.8	9.3	5.6	4.9	5.1
US. Sternberg, K. J., Lamb, M. E., Hershkowitz, I., Esplin, P. W., Redlich, A. D., & Sunshine, N. (1996)	45	4-12 yrs	30.0 ^{a)}	12.4 ^{a)}	5.6 ^{a)}	5.4 ^{a)}	5.6 ^{a)}
			17.1 ^{b)}	15.6 ^{b)}	6.5 ^{b)}	5.4 ^{b)}	4.3 ^{b)}
US. Lamb, M. E., Hershkowitz, I., Sternberg, K.J., Boat, B., & Everson, M. D. (1996)	24	4-12 yrs	30.8 ^{c)}	12.9 ^{c)}	5.9 ^{c)}	3.7 ^{c)}	5.5 ^{c)}
			7.7 ^{d)}	5.6 ^{d)}	3.9 ^{d)}	3.5 ^{d)}	4.8 ^{d)}
Finland. Korkman, J., Santtila, P., & Sandnabba, N.K. (2006) ² .	27	Mean 6 yrs	6.4	–	6.0 ^{e)} 4.7 ^{f)}	–	2.9
Finland. Korkman, J., Santtila, P., Westeråker, M., & Sandnabba, N.K. (2008) ³ .	43	3-8 yrs	4.8	6.0	4.1	3.0	3.7 ^{g)}

²⁾ Categories are denoted; Invitations, Directive, Clarifications, and Specific- suggestive utterances. Their Clarification category is equivalent to the Directive category in Lamb et al ¹.

³⁾ Categories are denoted; Invitations, Facilitators, Directive, Option-posing and Specific suggestive and Unspecific suggestive.

^{a)} Interviews of children as victims of single incidents. ^{b)} Interviews of children as victims of multiple incidents.

^{c)} Interviews of children without anatomical dolls. ^{d)} Interviews of children with anatomical dolls used by the interviewer.

^{e)} Directive category. ^{f)} Clarification category; equivalent to the Directive category in Lamb et al ¹.

^{g)} The two suggestive categories (specific and unspecific) are merged into one category in the present table.

The age of the children in table 1 were 5 to 11 years in the Israeli-, 4-12 years in the US-, and 3-8 years in the Finnish sample. The Finnish studies by Korkman et al. revealed shorter answers than the other FIIC studies comparing the categories of questions and the length, by the number of words, in the answers. In their first study the *invitation* elicited the longest answer with mean 6.4 words, followed by *directive utterances* with *specific suggestive* giving the shortest responses with mean 2.9 words. The children in their second study responded with the shortest responses, 3.0 words, to *option posing* questions (closed questions), while *facilitators* and *invitations* (open questions) rendered the longest responses by the child with 6.0 and 4.8 words, respectively.

The low number of words in the respective responses could be explained by the age of the children, supporting the theories that children develop at vastly different rates and as children grow older the length, informativeness, and complexity of their recall memory increases (Fivush, 1997, 1998; Poole & Lamb, 1998; Saywitz & Camparo, 1998; Schneider & Pressley, 1997) with the vocabularies of young children often being more limited and less descriptive than those of older children and adults (Brown, 1973; Morison, Moir, & Kwansa, 2000; Walker, 1999). The results presented in table 1 support these theories, with the oldest Israeli and US children providing longer answers than the younger children in the Finnish studies. A study of children from England and Wales (Davies et al., 2000), showed that open questions were most effective with older children (12 – 14 years) but not with younger children (4 – 11 years). The younger children provided more information in response to specific yet not leading or closed questions. One reason for this might be the natural and often unconsciously provided support (or scaffolding) to a child's narrative production by an adult (see Reynolds & Evans, 2009). Young children in their normal day to day interaction with adults rely on degrees of scaffolding to provide narratives and it can be argued that at some stages of their development, as in the Davies, et al. study (op.cit), the youngest children may be unable to produce a narrative without appropriate scaffolding. The danger of adult scaffolding in interview situations is contamination of memory. With no effect of open questions to the youngest children, the latter result is of operational importance to the interviewers in their planning and preparations of questions for their interviews, with the most effective categories of questions being different for young and old children, respectively.

In their comparison of legal outcomes in 100 CSA cases, Myklebust and Bjørklund (2009) revealed an effect of the interviewed children's gender and age. In all cases that progressed to court, the FIIC was considered to be the evidence-in-chief. In the cases with FIIC of girls older than 10 years, all led to convictions of the suspect in court. This gender effect is in line with the studies by Hershkowitz, Horowitz and Lamb (2005); Lamb, Hershkowitz, Orbach, and Esplin (2008) examining all child abuse investigation conducted in Israel between 1998 and 2002. The authors reported the interviewed girls to provide more details than the boys. The present study will examine both the effect of age and gender on the facilitation of the interviewees' responses in CSA cases.

Fisher, Geiselman and Raymond (1987) observed that the interviewers' level of competence affected the responses in interviews of adult interviewees. The authors recommended formal, scientifically based training of the police officers at the institutional level. Fisher and Geiselman (1992) denoted their interview method as the cognitive interview technique (CIT) and observed that interviewers trained in this model obtained longer responses in interviews of adults compared with interviewers

without such training. They suggested that training programs will progress most efficiently if it is divided into intensive short sessions, rather than one big session, keeping the presentation component at a short duration with the efficacy of the training program profiting by extended feedback to the individual interviewers (Fisher & Geiselman, 1992; Fisher et al., 1987). In one of several studies of the effect of the CIT the training was conducted in four 60-minute group sessions, including lectures describing various components of the procedure and demonstrations of good and poor interviewing techniques (Fisher, Geiselman & Amador, 1989). Contrary, Lamb and his colleagues argue that long-time improvement in the quality of investigative interviews are observed only when the training is distributed over time (Lamb, Sternberg, Orbach, Esplin, & Mitchell, 2002; Lamb, Sternberg, Orbach, Hershkowitz, Horowitz, & Esplin, 2000). In their studies, the length of training varied between 3 to 5 full day sessions of initial training programs, with follow up supervision and feedback to the interviewers in the period of the studies (Lamb et al., 2002; Orbach, Hershkowitz, Lamb, Sternberg, Esplin, & Horowitz, 2000; Sternberg, Lamb, Davies, & Westcott, 2001). The interviewers' competence could be described according to their theoretical and practical level, based on the length of their formal, scientifically based training at the institutional level and the number of conducted FIIC (Fisher & Geiselman, 1992; Lamb, et al., 2002; Lamb, et al., 2000). The basic implicit assumption is that police officers who have received long and more specialised training generate longer answers from the interviewed children compared to officers with shorter interview training.

The portion of the interviews has also been focused on by researchers. The majority of FIIC studies are analysing the *substantive portion* of each interview, operationally defined as the portion of the interview during which the incidents under investigation were discussed (Lamb et al., 2008; Sternberg et al., 1996), thereby excluding any introductory sequence at the beginning of the interview, attempts to establish rapport with the child, and any attempts at the end of the interview to discuss neutral topics (Lamb, Hershkowitz, Sternberg, Boat, et al., 1996; Lamb, Hershkowitz, Sternberg, Esplin et al., 1996; Lamb, Sternberg, & Esplin, 1994). The *substantive portion* was also analysed by Korkman, Santtila, and Sandnabba (2006), while the first thirty 'on-topic' utterance pairs were coded in the second study (Korkman et al., 2008) keeping the number of question-answer pairs constant for short and longer interviews. We have not found any studies reporting that the result from the substantive portion of an interview diverges from the entire length, but it is still unclear how to compare the different studies. Myklebust and Bjørklund (2006) observed that FIIC lasted for an average of 51 minutes (SD 22.3, range 14-104 minutes). We hypothesise that the duration of a FIIC generates fatigue among the children that makes shorter utterances in last phase of the interview compared with the first phase. Subsequently, we will also examine any time-on-task effect on the children's responses.

The literature indicates that several variables affect the children's responses in FIIC. However, the different studies only include a limited number of the above mentioned variables. The purpose of the present study is to include all the variables. We assume that all the variables, i.e. the categories of questions, the children's age and gender, the competence of the interviewers, and the portions of the interview, will affect the length of the children's responses. According to the literature we hypothesise that the competence of the interviewer and the category of the questions will have the most impact on the length of the responses, while the children's age and gender, and the portions of the interview will have less influence on the responses in FIIC.

METHOD

Sample

The study included 100 FIIC conducted by separate special trained police interviewers, all statutorily mandated to conduct investigative interviews of children in Norway. Each interviewer met the following characteristics:

- (1) The child interviewed had to be 6-16 years of age (school-age).
- (2) No deviation in the child's linguistic development or skills.
- (3) Information about the interviewers' gender, age, training level (low/high), and the number of practiced FIIC.
- (4) Information about the age and gender of the interviewed child.
- (5) The interview had to be fully video recorded, from the child entering the interview room until leaving the room.
- (6) The interview should be completely transcribed and approved by the judge according to the national Norwegian regulations.
- (7) The FIIC should be the only interview of the child in the case.

The theoretical competence of the Norwegian police officers conducting the interviews could be divided in two groups, 'high' and 'low', respectively. The interviewers with high theoretical competence (T+) have substantially more in-depth and broader theoretical training sessions at the Norwegian Police University College (NPUC), compared to the interviewers with low theoretical competence (T-) (PHS, 2002;2003). The interviewers' practical experience, quantified by the mean number of conducted interviews, were significantly different between the two groups of children for interviewers at 'T-', with 15 (SD = 7.7) and 9 (SD = 5.2) conducted interviews for the Young and Old groups of children respectively ($t = 2.90, p < 0.01$). For the 'T+' interviewers, the number of conducted interviews were 29 (SD =16.5) and 23 (SD = 15.5) for the respective two groups ($t = 1.13, p > 0.05$).

By median split method, the children were divided into one of two age groups, denoted 'Young' and 'Old', respectively. In the Young group, the age span of the children was 6-10 years (N=52, mean = 8.6 yrs, SD = 1.3) and for the Old group, the age span of the children was 11-16 years (N=48, mean = 12.2 yrs, SD = 0.9). Table 2 shows the demographic information and significant differences between the two groups of children.

Table 2. Demographic information and significant differences between the two groups of children (Old / Young).

	Children's age group		Level of significance
	Young	Old	
Number of interviews	52	48	
Gender of children	33 ♀ 19 ♂	38 ♀ 10 ♂	ns* ^{b)}
Mean duration (minutes)	48 (17.8)	55 (25.9)	ns* ^{a)}
<u>Interviewers with 'low theoretical competence' (T-):</u>			
Number of interviewers	27	23	
Mean number of conducted interviews	15 (7.7)	9 (5.2)	** ^{a)}
<u>Interviewers with 'high theoretical competence' (T+):</u>			
Number of interviewers	25	25	
Mean number of conducted interviews	29 (16.5)	23 (15.5)	ns* ^{a)}

SDs are in parantheses.

ns*, non-significant ($p > 0.05$); ** $p < 0.01$

^{a)} Students t-test,

^{b)} Chi-square test,

The number of children in the two age groups was not statistically different, neither was the distribution of the children's gender within the groups (all p 's > 0.05 , Chi-square test). The mean duration of the interviews were 48 minutes in the Young group (SD = 17.8) and 55 minutes (SD = 25.9) in the Old group ($t = 1.69$, $p > 0.05$). Of the interviewers at the lowest theoretical competence level (T-) 27 had conducted interviews that were categorised as being from the Young group of children, while 23 were in the Old group. For the number of interviewers at the highest theoretical competence level (T+) the numbers were 25 in both groups of children.

Procedure

All the transcripts of the interviews were electronically digitised and validated word by word to ensure their accuracy and completeness with the video. All the validations were conducted by the authors and corrections were made in 332 (1.1%) of the interviewers utterances and in 843 (2.7%) of the children's answers. In the process of digitising, all the transcriptions were anonymised accordingly. The transcripts were then analysed with each category of interviewer's utterances measured against each child witnesses' respective answer. We identified a total number of 31,412 interviewer utterances with the children's respective answers. Each of the interviewers' utterances were classified in one of the following six categories, first developed by Richardson et al. (1965) and later modified by Myklebust and Bjørklund (2006):

1. **Instruction:** Information given by the interviewer to the child during the interview. The information is not directly related to the case that is investigated but about the interview itself, e.g., date and time the interview is conducted, equipment used during the interview, people present in the monitoring-room, the importance of answering the questions loud and truthfully, and information about the context the interview is held.
2. **Open question:** A question that requires more than a few words for an adequate response, e.g., "Tell me about that...", "Describe for me...", "Explain...". The child's focus is not delimited more than in a general way.

3. Identification: (person, place, group, time, number etc.); who, where, when, how many, or which questions, e.g., “What was his name?”.
4. Selection: fixed-alternative question, e.g., “Was the door open or closed?”
5. Yes-no: question answered by “yes” or “no”, e.g., “Do you know his name?”, possibly supplemented by a redundant phrase such as, “I think so”, or “I doubt it”.
6. Facilitators: Utterances like “okay”, restatements of the child’s previous utterance, and non-suggestive words of encouragement that are designed to encourage the child to keep talking.

Closed question was defined as category 3, 4 and 5. The closed questions include a limited number of response alternatives and could be answered adequately in a few words. *Open questions* deal with broader topics than closed questions and allow children more flexibility to choose which aspects of an event they will describe, e.g “Tell me about the person you met” [open question] and “What kind of jacket did he wear?” [closed, identification question].

Table 3. The distribution of the six different categories of interviewer utterances in 100 FIIC

Interviewer utterances	Number of interviewer utterances
Instruction	3682
Yes-no	13312
Selection	993
Identification	7405
Facilitators	3977
Open question	2043
TOTAL	31412

Table 3 shows the distribution of the six different categories of interviewer utterances.

We validated the classification of the interviewers’ utterances by randomly choosing 10 of the interviews in each of the two interviewers competence groups (T-/T+), providing a total of 6,438 utterances. The result from the 20 validated interviews gave an inter-rater reliability coefficient of 0.96, indicating a high level of agreement between the authors. The validation revealed disagreement in 238 questions classified as 147 open/identification questions, 69 open/facilitator and 22 instructions/yes-no, respectively. In the total material the number of words in the children’s answers to each of the investigators utterances was counted in all the interviews. The duration of the interviews varied between 14 and 104 minutes (M = 51 min, SD = 22.25). In order to study any time pattern development during the interviews, each of them was standardized into three parts, denoted as ‘tertiaries’ and the childrens’ number of words were counted in each of the tertiaryes for the six categories of utterances. Based upon the interviewers’ theoretical and practical training the interviewers were divided into four groups studying the effect of competence on the number of words in the children’s answer.

Data analyses

The mean number of words in each of the six categories of interviewer utterances, and in the corresponding answer from the child, was calculated for each of the tertiarys in all interviews and used as cell means for subsequent statistical analysis. The numbers of words in the children's answers were used as the dependent variable. To perform the statistical analysis we used SPSS v. 16.0. T-tests and analysis of variance (ANOVA) were used to conduct the analysis. Five percent was used as alpha level in all the statistical tests. The study was approved by the Director of the Public Prosecution and conducted according to the standards of the Norwegian regional committee of research ethics

RESULTS

The number of words in the children's answer was affected by several factors. Analysis of variance (ANOVA) revealed a main effect of the categories of utterances on the number of words in their answers during all interviews ($F(5,594) = 83.148$, $p < 0.01$, partial Eta squared = 0.412). The mean number of words for the respective categories of utterance was for instruction 2.9, yes-no 6.1, selection 8.5, identification 10.4, facilitators 7.2 and open questions 37.5 words, respectively. Post hoc analysis revealed that the number of words in the children's responses to open questions (mean = 37.5) was significantly different from the five other categories (all p 's < 0.001). In addition, the number of words to instructions (mean = 2.9) was statistically different from identification (mean = 10.4) (mean difference = 7.4, standard error = 2.0, $p < 0.01$).

Table 4. Mean number of words in children's answers related to six categories of interviewer utterances for the two age groups.

Interviewer utterances	Young children (N = 52)		Old children (N = 48)		Level of significance
	Mean	SD	Mean	SD	
Instruction	2.7	3.6	2.2	3.2	ns*
Yes-no	5.0	2.7	7.3	6.0	ns*
Selection	6.7	5.0	10.3	12.3	ns*
Identification	8.9	4.5	11.9	7.5	ns*
Facilitators	7.8	10.0	6.5	5.0	ns*
Open question	30.9	30.8	45.7	30.7	***
TOTAL	10.1	7.5	14.2	8.3	*

ns*, non-significant ($p > 0.05$); * $p < 0.05$; *** $p < 0.001$

ANOVA revealed a main effect of the age group ($F(1,98) = 10.529$, $p < 0.01$, partial Eta squared = .062). The younger children used a mean number of 10.1 words (SD = 7.5) in each answer while the older children used 14.2 words (SD = 8.3). In addition, the interaction between age group and categories of utterances was significant, $F(5,588) = 5.162$, $p < 0.001$, partial Eta squared = 0.042. Table 4 shows the number of words for the two groups of children distributed to the six categories of utterances. Post hoc analysis showed only a significant difference on the number of words between the two age groups on open questions (mean difference = 14.77, standard error = 2.79, $p < 0.001$).

ANOVA did not show any effect of the children's gender on the number of words in their answers (all p 's > 0.05).

Temporal effects (time) on number of words in the children's answers were examined by ANOVA combining tertiaries (3), utterances (6) and age group (2). The results gave a main effect of tertiaries ($F(2,297) = 3.32, p < 0.05, \text{partial Eta squared} = .022$). The mean number of words in the answers was 13.9, 12.8 and 10.5 for the first, second and third tertiary, respectively. Post hoc analysis showed that the first and third tertiary was significantly different (mean difference = 3.4, standard error = 1.5, $p < 0.05$). The temporal effect affected the number of words differently in the six categories of utterances ($F(10,1679) = 6.43, p < 0.001, \text{partial Eta squared} = .037$).

Table 5 shows how the temporal factor affected the mean number of words in the children's answers in response to the six categories of utterances. Post hoc analysis showed that time only affected the children's number of words on open questions. The number of words was 49.4, 34.1 and 28.9 in the first, second and third tertiary, respectively. The number of words in the first tertiary was significantly different from the two other tertiaries (both p 's < 0.001), while the difference between the second and the third was not significant. There was no effect of age group on the temporal effect on open question answers.

Table 5. Mean numbers of words (SD in parentheses) in the children's answers related to the different categories of utterances divided into tertiaries

Categories	Tertiaries			Total
	1 st	2 nd	3 rd	
Instruction	2.3 (3.7)	4.8 (8.7)	2.2 (6.3)	3.0 (6.4)
Yes-no	5.4 (4.4)	6.9 (6.5)	6.0 (5.4)	6.1 (5.5)
Selection	7.5 (10.5)	9.6 (13.4)	7.9 (8.5)	8.4 (11.0)
Identification	7.6 (5.6)	11.9 (9.8)	11.6 (8.0)	10.4 (8.2)
Facilitators	9.0 (13.4)	6.9 (7.0)	5.3 (4.6)	7.1 (9.2)
Open question	49.4 (58.2)	34.1 (28.7)	28.9 (25.2)	37.5 (41.0)
Total	13.9 (30.4)	12.8 (17.9)	10.5 (15.0)	12.4 (22.2)

The competence of the interviewers consists of two aspects, one is the theoretical competence and the other is the practical competence. By means of median split half method the investigators were divided into two groups. Those who have conducted less than 15 interviews were denoted 'practical low' (P-) and interviewers with more or equal to 15 interviews as 'practical high' (P+). We used the information about the interviewers theoretical competence to divide them into 'theoretical low' (T-) or 'theoretical high' (T+). The combination of low/high theory (T-/T+) and low/high practice (P-/P+) gave four competence groups with 37, 13, 13, and 37 interviewers in the (T-/P-), (T-/P+), (T+/P-), and (T+/P+) category, respectively. Table 6 shows the combination of theoretical and practical training among the one hundred interviewers according to our classification.

Table 6. Distribution of 100 hundred police interviewers into four groups (Group 1 to Group 4) of interviewers based upon their theoretical competence (T+/T-) and practical competence ('P-' less than 15 interviews / 'P+' more or equal to 15 interviews) and the mean number of words in the children's answers.

Competence level	Number of interviewers	Mean number of words in the children's total answers (SD in parentheses)	Mean number of words in the children's answers to the open questions (SD in parentheses)
Group 1 (T-/P-)	37	13.9 (20.5)	41.9 (34.1)
Group 2 (T-/P+)	13	11.6 (21.7)	42.9 (40.8)
Group 3 (T+/P-)	13	13.1 (22.6)	49.2 (36.9)
Group 4 (T+/P+)	37	10.1 (12.4)	28.5 (19.9)
TOTAL	100	12.1 (18.5)	38.0 (31.5)

Analysis of the interviewers' competence showed a main effect on the number of words in the children's answers ($F(3,96) = 2.83, p < 0.05$, partial Eta squared = .015). The children used a mean number of 13.9 words in 'Group 1'(T-/P-), 11.6 words in 'Group 2'(T-/P+), 13.1 words in 'Group 3'(T+/P-), and 10.1 words in 'Group 4'(T+/P+). Post hoc analysis revealed that only Group 1 (mean = 13.9 words) and Group 4 (mean = 10.1 words) was different from the other (mean difference 3.8, standard error =1.3, $p < 0.05$). The interaction between the competence groups and categories of utterances was also significant on the number of words in the children's answers ($F(15,576)=1.75, p < 0.05$, Eta squared = .044). Post test analysis showed that the children's mean number of words after open questions was 49.2 in Group 3 and 28.5 in Group 4 (mean difference 20.7, standard error = 4.5, $p < 0.01$).

DISCUSSION

In the present study variables presented in the literature, as factors affecting the length of the children's answers in FIIC, were all included in the analysis. One hundred FIIC were analysed in their entire length. First, we observed the mean numbers of words in the children's answers dependent upon the six different categories of interviewer utterances. Second, the responses in the oldest group of children was found to be significantly longer, counted by words, than in the youngest group, with post hoc showing a significant difference between the two age groups only to the open questions. A temporal effect was demonstrated for the six categories, affecting only the number of words in answers to open questions, where the number of words in the first tertiary was significantly longer than in the second and third tertiary. Analysis of the interviewers' level of competence revealed a difference in the number of words in the children's answers to the open questions, with children interviewed by interviewers with high theoretical and low practical competence producing significant longer answers than interviewers with high theoretical and high practical competence. Of all the

variables, the categories of the interviewer utterances had most impact on the length of the children's responses. Derived from the children's responses to the open questions, the variable to follow was the children's age, with the oldest children yielding longer responses than the younger children. Contrary to our hypothesis, the competence of the interviewers did not have the expected impact on the length of the responses. Neither did the children's gender, nor time as a temporal effect in the FIIC.

Of the interviewers' utterances, the open questions yielded responses approximately four and a half times longer than the closed (identification, selection, yes/no) questions. This result is in line with the Lamb et al. studies in table 1 (Lamb, Hershkowitz, Sternberg, Boat et al., 1996; Lamb, Hershkowitz, Sternberg, Esplin, et al., 1996; Sternberg et al., 1996;) with their responses to open questions providing approximately three and four times longer than the closed questions. Even if the mean number of words in our children's answers declined, the more literally leading the interviewers categories of utterances got, from 37.5 words for the "open-" to 6.1 words in the yes-no category, the only significant difference in number of words were the open questions being different to the other five categories.

The differences in number of words could be explained by several theories. One theory argues that there is a widespread consensus among researchers that the open and closed questions are probing two different memory systems. Adults and children asked to describe events from free recall, by interviewers asking open questions, are probing 'recall memory' while interviewers prompting the focus of the interviewee with closed questions, and especially option-posing questions such as selection and yes/no questions, shift from 'recall-' to 'recognition memory' (Dent, 1982; 1986; 1991; Dent & Stephenson, 1979; Lamb, et al., 1994; Oats & Shrimpton, 1991; Pipe, Lamb, Orbach, & Esplin, 2004). A second theory suggests (see reviews by Lamb & Brown, 2006) that in an interview, children are reliant on the interviewer to provide retrieval cues and to signal what information is important for the child to recall and report. From a linguistic perspective, the different types of questions asked will assist in the elaboration of the length of the answers. An open question requires the child to provide the sought after information without the interviewer pre-determining what the information consists of. In contrast, in the closed selection- or yes/no question, the interviewer pre-determines the information and the child is simply anticipated to respond to one of the suggested alternatives to the selection question, or simply "yes" or "no" to the 'yes/no' question. A third theory, derived from Sternberg, Lamb, Hershkowitz, Yudilevitch, Orbach, Esplin, and Hovav (1997), reveals the interviewer's style of questions will entrain the response style of the interviewed child, and by using a wide range of open questions throughout the interview, the child will provide longer answers. Sternberg et al. (1997) argues that rapport-building, with the use of open questions about the child's everyday life and a particular positive past event in the beginning of the interview, elicits two and a half times as many words in the response to the first utterances about the alleged incident under investigation. This compared to an introduction based on closed questions in the beginning of the interview, where the interviewer is building rapport with the child. With none of the closed questions being significant different, our results do not separate between the closed questions. Kahn and Cannell (1957) introduced an interview approach of beginning an interaction with an open question and following it with successively more closed questions. The approach was termed 'funnel sequence' and stated a significant difference in the length of answers to the different categories of closed questions. This theory is later advocated by professional investigative interview guidelines (APSAC, 2002;

Home Office, 2007; The National Crime Faculty, 1998). Our results do not support the funnel structure of the questions. Contrary, we state that the closed questions should only be asked after the interviewer have used a variety of open questions, giving the child several opportunities to describe the event in his or her own words.

Using the children's age as a variable the results revealed a main effect with the children in the oldest age group (11-16 yrs) giving longer answers than the younger children (6-10 yrs). The results in table 4 show that the mean number of words is higher in the responses to the questions for the old children, while the numbers of words in the instruction and facilitators categories is higher for the young group. Here, the open questions also generated the only significant difference. For the interviewers, this means that the same questions will have the same effect on the length of the answers between the two age groups. This result is of operational importance to the interviewers in their planning and preparation of the FIIC. However, for the instruction and facilitator categories, the children in the youngest group used more words. For the instruction category the difference is so small (0.5 words) that we consider it as at a level of chance. Although facilitators involve open encouragement to keep the interviewee talking, we found them to function much like the previous eliciting utterance. We raise questions about the designation of facilitators all the time as it may function as either of the other question categories. Until more research has been conducted on the use and effect of facilitators we recommend the interviewers to make use of the other question categories, emphasising the superiority of open question. Contrary to the present results, Davies et al. (2000) used time (the temporal length of the child's response) as the dependent variable and found longer answers to be provided in response to closed and specific, yet non-leading, questions for the 4-7 and 8-11 years olds compared to 12-14 year olds children. As Davies et al. reported they had several methodological limitations to their study, with a small sample from only one police force (Davies et al., 2000). Theoretically, children's level of linguistic development are often divided into three age groups; 'pre-schoolers (3-6 yrs)', 'school age (7-10 yrs)', and 'young adolescents (11-18 yrs)' (Morison et al., 2000; Poole & Lamb, 1998; Walker, 1999). The two age-groups in the present study fit into the two last categories, and also to the 'concrete operational period (7-11 yrs)' and 'formal operational period (11-15 yrs)' in Piaget's cognitive stage-theory of development (Miller, 1993; Piaget, 2007). We studied children between 6-16 years of age. Subsequently we do not know how younger children perform.

We divided the interviews into three equally long periods (tertiaries), and found a main effect of tertiary on the length of the responses, indicating that the children elicited different lengths of answers between the three periods. The length of answers in the beginning of the interview were significantly different from the last part, with a mean number of 13.9 and 10.5 words, respectively. Once again the open questions is the only category with a significant effect, with declining number of words in the tertiary with 49.4, 34.1 and 28.9 words (cfr. table 5), with the first and third tertiary being significantly different. This could be explained by the interview being a unique conversational context for children, requiring them to take on the role as expert informant about their knowledge and experience (Lamb & Brown, 2006). The cognitive demands on the child are high and the responses will be shorter as the interview prolongs, being more motivated and energetic in the beginning than in the end of the interview. Another explanation could be that the children felt the content being discussed in the last part being more difficult and demanding, than in the first part. However, this

would contradict most interview protocols recommending the interviewers to introduce the questions regarding the incidents under investigation rather early in the interview (Walker & Hunt, 1998). Even if the length of answers to the open questions is decreasing throughout the interview, the open questions reveal answers 8, 4 and 6 times longer than the closed questions for each of the tertiaries, respectively, with no age effects reported for the open questions. For the other categories the length in the answers are the same in all tertiaries.

We assumed that police officers with the highest theoretical and practical competence would conduct FIIC revealing the longest responses among the children. Contrary to our hypothesis, we found a significant difference between interviewers in Group 1 (T-/P-), with 13.9 words, and Group 4 (T+/P+) with 10.1 words in the children's responses. The children's mean number of words to the open questions were 49.2 words in Group 3 (T+/P-) and 28.5 in group 4 (T+/P+). The results, with the longest responses to the interviewers with low practical competence, we find to be interesting. Several explanations to this unexpected result have been considered. It could be that interviewers with low practical competence are more prepared in their interviews and have based their plan on their theoretical knowledge, while the interviewers with high level of practice are more distanced to their theoretical competence, leaning more on their practical experience from their previous interviews. Some authors have argued the importance of systematic supervision of interviewers (Lamb et al., 2002) in the field, and that what happens at the local police station after the formal theoretical training is as important as the initial training itself. We have not included any evaluation of the training programmes in investigative interviewing at the Norwegian Police University College. However, both the law and the training programmes for police officers conducting FIIC focus on the use of scientifically based interviewing techniques and strongly emphasises the interviewers to achieve the child's free narrative by the use of open questions (Act of 22 May 1981 No. 25; Myklebust, 2005). Our result questions whether the training programmes could be improved. Further discussions into this area are beyond the scope and remit for this research. The introduction of the interviewer's level of competence as an important variable was first suggested by Fisher et al. (1987). The authors introduced the open-closed question ratio (OCR) and reported an excessive use of closed questions with 3 open and 26 closed questions in the typical investigative interview, giving an OCR of almost 1:9. Davies et al. (2000) found in their study that only 2% of questions asked were classified as open. Korkman et al. (2006) found the same 2% as invitations, 22% as directives, 31% as option-posing and 31% suggestive. The interviews conducted by trained interviewers increased the OCR from 1:20 to 1:2 (George & Clifford, 1992). Lamb et al. found that ongoing group supervision and training in the use of a structured interview protocol produced a higher ratio of open questions than training providing only didactic instructions (Lamb et al., 2002). The need for more practice is one of the most commonly cited suggestions from both experts and trainee interviewers when asked how their training could be improved (Clarke & Milne, 2001; Guadagno, Powell & Wright, 2006; Powell, Fisher, & Hughes-Scholes, 2008a;2008b; Wright & Powell, 2006; 2007; Wright, Powell & Ridge, 2006; 2007; Aarons, Powell, & Browne, 2004). Our results, with the interviewers in Group 3 (T+/P-) and Group 4 (T+/P+), do not support these findings. It is possible that the theoretical training of the Norwegian officers is good, and that the shortage of time to plan may counteract the implementation of the interviewers' knowledge on how to conduct the interviews into practice in the FIIC. Powell and her colleagues examined the relationship between several background factors (e.g., experience in interviewing, knowledge of child development or law, job rank or status, interviewer gender) and interview

performance, and concluded that the recency of training is the only factor related to adherence to best-practice interview guidelines (Smith, Powell, & Lum, 2009). Therefore, adherence to open questions is the result of effective training (Powell, 2008; Powell, Fisher, & Wright, 2005; Smith et al., 2009). As demonstrated by Oxburgh, Myklebust et al. (submitted) several research papers, reports, books and guidance documents being produced by academic researchers, police training departments, and governmental institutions gives different explanations of what constitutes an open question. The divergent descriptions of open questions could confuse the police officers who conduct FIIC, and the academics and trainers bringing the discrepancies into the police training courses. To sum up, our results replicate other studies that indicate insignificant findings related to the competence of the interviewers. It is imperative that the interviewers are confident and competent at using open questions. The research into this area is scarce and we conclude that more research is needed in how to implement the interviewers' theoretical competence into practice.

In conclusion, the central aim of FIIC is to obtain as much relevant information as possible with as little suggestive- and specific prompting as possible from the interviewer. The literature indicates that several variables affect the children's responses in FIIC. The purpose of the present study was to include all these variables. The results demonstrated the benchmark in FIIC to be the interviewer's use of open questions, eliciting the longest answers compared to any of the other variables. The open-closed question ratio (OCR) in the present study was 1:10. Obviously police officers who conduct FIIC need to use more open questions in their interviews. Our results challenge better training programs for police officers who conduct FIIC. Both practitioners and academics have a fundamental role to play in implementing the knowledge of FIIC into practice.

ACKNOWLEDGEMENTS

The authors wish to thank the Director of Public Prosecution for their interest in the project, and the investigative interviewers and Police Departments that made their interviews available for study.

REFERENCES

- Act of 22 May 1981 No. 25. *Criminal Procedure Act Concerning Legal Procedures in Criminal Cases* [Straffeprosessloven - Lov av 22.mai 1981 nr.25 om rettergangsmåten i straffesaker].
- Aldridge, J., & Cameron, S. (1999). Interviewing child witnesses: Questioning techniques and the role of training. *Applied Developmental Science, 3*(2), 136-147.
- APSAC. (2002). *Psychosocial Evaluation of Suspected Sexual Abuse in Children* (2 ed.). Chicago: IL: American Professional Society on the Abuse of Children.
- Brown, R. (1973). *A first language*. Cambridge, MA: Harvard University Press.
- Cederborg, A. C., Orbach, Y., Sternberg, K. J., & Lamb, M. E. (2000). Investigative interviews of child witnesses in Sweden. *Child Abuse & Neglect, 24*(10), 1355-1361.
- Clarke, C., & Milne, R. (2001). *National Evaluation of the PEACE Investigative Interviewing Course. Police Research Award Scheme. Report NO:PRAS/149*. London: National Crime Faculty.
- CPTU. (1992). *Investigative interviewing: A guide to interviewing*. London: Central Planning & Training Unit.
- Davies, G. M., Westcott, H. L., & Horan, N. (2000). The impact of questioning style on the content of investigative interviews with suspected child sexual abuse victims. *Psychology, Crime & Law, 6*, 81-97.
- Dent, H. R. (1982). The effects of interviewing strategies on the results of interviews with child witnesses. In A. Trankell (Ed.), *Reconstructing the past: The Role of psychologists in criminal trials* (pp. 279-297). Stockholm: Norstedt.
- Dent, H. R. (1986). An experimental study of the effectiveness of different techniques of questioning mentally handicapped child witnesses. *British Journal of Clinical Psychology, 25*, 13-17.
- Dent, H. R. (1991). Experimental studies of interviewing child witnesses. In J. Doris (Ed.), *The suggestibility of children's recollections* (pp. 138-146). Washington, DC: American Psychological Association.
- Dent, H. R., & Stephenson, G. M. (1979). An experimental study of the effectiveness of different techniques of questioning child witnesses. *British Journal of Social and Clinical Psychology, 18*, 41-51.
- Fisher, R. P., & Geiselman, R. E. (1992). *Memory-enhancing techniques for investigative interviewing. The cognitive interview*. Springfield, IL: Charles C.Thomas Publisher.
- 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*(5), 722-727.
- Fisher, R. P., Geiselman, R. E., & Raymond, D. S. (1987). Critical analysis of police interview techniques. *Journal of Police Science and Administration, 15*(3), 177-185.
- Fivush, R. (1997). Event memory in early childhood. Studies in developmental psychology. In N. Cowan (Ed.), *The development of memory in childhood*. Hove: Psychology Press. Taylor & Francis.
- Fivush, R. (1998). Children's recollections of traumatic and nontraumatic events. Trauma, memory, and suggestibility in children. *Development and Psychopathology, 10*, 699-716.

- Foster-Cohen, S. (1990) *The communicative competence of young children: a modular approach*. London: Longman
- George, R., & Clifford, B. R. (1992). Making the most of witnesses. *Policing*, 8, 185-198.
- Guadagno, B. L., Powell, M. B., & Wright, R. (2006). Police officers' and legal professionals' perceptions regarding how children are, and should be, questioned about repeated abuse. *Psychiatry, Psychology and Law*, 13(2), 251-260.
- Harley, T. (2009) *The Psychology of Language* (3rd Ed.) Psychology Press.
- Hershkowitz, I., Horowitz, D., & Lamb, M. E. (2005). Trends in children's disclosure of abuse in Israel: A national study. *Child Abuse & Neglect*, 29, 1203-1214.
- Hershkowitz, I., Lamb, M. E., Sternberg, K. J., & Espelin, P. W. (1997). The relationships among interviewer utterance type, CBCA scores, and the richness of children's responses. *Legal and Criminological Psychology*, 2, 169-176.
- Home Office. (2007). *Achieving Best Evidence in Criminal Proceedings: Guidance on Interviewing Victims and Witnesses, and Using Special Measures*. London: Criminal Justice System.
- Kahn, R. L., & Cannell, C. F. (1957). *The dynamics of Interviewing. Theory, Technique and Cases*. New York: John Wiley & Sons, Inc.
- Korkman, J., Santtila, P., & Sandnabba, N. K. (2006). Dynamics of verbal interaction between interviewer and child in interviews with alleged victims of child sexual abuse. *Scandinavian Journal of Psychology*, 47, 109-119.
- Korkman, J., Santtila, P., Westeråker, M., & Sandnabba, N. K. (2008). Interviewing techniques and follow-up questions in child sexual abuse interviews. *European Journal of Developmental Psychology*, 5(1), 108-128.
- Lamb, M. E. (1994). The investigation of child sexual abuse: An interdisciplinary consensus statement. *Child Abuse & Neglect*, 18(12), 1021-1028.
- Lamb, M. E., & Brown, D. A. (2006). Conversational apprentices: Helping children become competent informants about their own experiences. *British Journal of Developmental Psychology, Crime & Law*, 24, 215-234.
- Lamb, M. E., Sternberg, K. J., & Esplin, P. W. (1994). Factors influencing the reliability and validity of statements made by young victims of sexual maltreatment. *Journal of Applied Developmental Psychology*, 15, 255-280.
- Lamb, M. E., Hershkowitz, I., Orbach, Y., & Esplin, P. W. (2008). *Tell Me What Happened. Structured Investigative Interviews of Child Victims and Witnesses*. Chichester: John Wiley & Sons, Ltd.
- Lamb, M. E., Hershkowitz, I., Sternberg, K. J., Boat, B. W., & Everson, M. D. (1996). Investigative interviews of alleged sexual abuse victims with and without anatomical dolls. *Child Abuse & Neglect*, 20(12), 1251-1259.
- Lamb, M. E., Sternberg, K. J., Orbach, Y., Esplin, P. W., & Mitchell, S. (2002). Is ongoing feedback necessary to maintain the quality of investigative interviews with allegedly abused children? *Applied Developmental Science*, 6(1), 35-41.
- Lamb, M. E., Sternberg, K. J., Orbach, Y., Hershkowitz, I., Horowitz, D., & Esplin, P. W. (2000). The effects of intensive training and ongoing supervision on the quality of investigative interviews with alleged sex abuse victims. *Applied Developmental Science*, 6(3), 114-125.

- Lamb, M. E., Hershkowitz, I., Sternberg, K. J., Esplin, P. W., Hovav, M., Manor, T., & Yudilevitch, L. (1996). Effects of investigative utterance types on Israeli children's responses. *International Journal of Behavioral Development, 19*(3), 627-637.
- Loftus, E. (1982) Interrogating eyewitnesses – good questions and bad. In R. Hogarth (Ed.), *Question framing and response consistency*, San Francisco: Josey-Bass.
- Miller, P. H. (1993). *Theories of Developmental Psychology* (3 ed.). New York: W.H Freeman and Company.
- Morison, S., Moir, J., & Kwansa, T. (2000). Interviewing children for the purpose of research in primary care. *Primary Health Care Research and Development, 1*, 113-130.
- Myklebust, T. (2005). Dommeravhør av barn i Norge - intensjon og fakta [Investigative interviews of children in Norway - intention and facts]. In M. Egge & J. Strype (Eds.), *Polititrollen gjennom 100 år. Tradisjon og endring [The role of the police in 100 years. Tradition and changes]* (pp. 142-172): PHS Forskning.
- Myklebust, T., & Bjørklund, R. A. (2006). The effect of long-term training on police officers' use of open and closed questions in field investigative interviews of children (FIIC). *Journal of Investigative Psychology and Offender Profiling, 3*, 165-181.
- Myklebust, T. & Bjørklund, R. A. (2009). The Child Verbal Competence effect in Court. A Comparative Study of Field Investigative Interviews of Children (FIIC) in Child Sexual Abuse Cases. *Journal of Investigative Psychology and Offender Profiling, 6* (2), 117-128
- Oats, K., & Shrimpton, S. (1991). Children's memories for stressful and non-stressful events. *Medical Science and Law, 31*(1), 4-10.
- Orbach, Y., Hershkowitz, I., Lamb, M. E., Sternberg, K. J., Esplin, P. W., & Horowitz, D. (2000). Assessing the value of structured protocols for forensic interviews of alleged child abuse victims. *Child Abuse & Neglect, 24*(6), 733-752.
- Oxburgh, G., Myklebust, T., & Grant, T. (Submitted). The question of question types in police interviews: A review of the literature from a psychological and linguistic perspective. *International Journal of Speech Language and the Law*.
- PHS. (2002). *Kursplan i utdanning i etterforskning av seksuelle overgrepssaker trinn III [Course description for education in investigation of sexual abuse cases and investigative interviews of children and vulnerable witnesses - Step III]*. Oslo: Politihøgskolen [Norwegian Police University College].
- PHS. (2003). *Kursplan i utdanning i etterforskning av seksuelle overgrepssaker trinn II [Course description for education in investigation of sexual abuse cases and investigative interviews of children and vulnerable witnesses - Step II]*. Oslo: Politihøgskolen [Norwegian Police University College].
- Piaget, J. (2007). *The Language and Thought of the Child, 3.ed.* New York: Routledge Classics.
- Pipe, M. E., Lamb, M. E., Orbach, Y., & Esplin, P. W. (2004). Recent research on children's testimony about experienced and witnessed events. *Developmental Review, 24*, 440-468.
- Poole, D. A., & Lamb, M. E. (1998). *Investigative interviews of children: A guide for helping professionals*. Washington, DC: American Psychological Association.
- Powell, M. B. (2008). Guide to effective training programs in the area of investigative interviewing of children. *Current Issues in Criminal Justice, 20*, 189-208
- Powell, M. B., Fisher, R. P., & Hughes-Scholes, C. H. (2008a). The effect of intra- versus post-interview feedback during simulated practice interviews about child abuse. *Child Abuse & Neglect, 32*, 213-227.

- Powell, M. B., Fisher, R. P., & Hughes-Scholes, C. H. (2008b). The effect of using trained versus untrained adult respondents in simulated practice interviews about child abuse. *Child Abuse and Neglect*, 32, 1007-1016.
- Powell, M. B., Fisher, R. P., & Wright, R. (2005). Investigative Interviewing. In N. Brewer & K. D. Williams (Eds.), *Psychology and Law: An Empirical Perspective* (pp. 11-42). NY: Guilford Publications.
- Reynolds, K.P. & Evans, M.A. (2009) Narrative performance and parental scaffolding of shy and nonshy children. *Applied Psycholinguistics* 30, 363--384.
- Richardson, S. A., Dohrenwend, B. S., & Klein, D. (1965). *Interviewing: Its forms and functions*. London: Basic Books, Inc., Publishers.
- Saywitz, K., & Camparo, L. (1998). Interviewing child witnesses: A developmental perspective. *Child Abuse & Neglect*, 22, 825-843. Shepherd, E. (2007). *Investigative interviewing. The Conversation Management Approach*. Oxford: Oxford University Press.
- Schneider, W., & Pressley, M. (1997). *Memory development between two and twenty* (2 ed.). Mahwah, NJ: Erlbaum. Smith, R. M., Powell, R. M., & Lum, J. (In press). The relationship between job status, interviewing experience, gender and police officers' adherence to open-ended questions. *Legal and Criminological Psychology*.
- Smith, R., Powell, M. B. & Lum, J. (2009). The relationship between job status, interviewing experience, gender and police officers' adherence to open-ended questions. *Legal and Criminological Psychology*, 14, 51-63.
- Sternberg, K. J., Lamb, M. E., Davies, G. M., & Westcott, H. L. (2001). The Memorandum of Good Practice: theory versus application. *Child Abuse & Neglect*, 25, 669-681.
- Sternberg, K. J., Lamb, M. E., Hershkowitz, I., Esplin, P. W., Redlich, A. D., & Sunshine, N. (1996). The relation between investigative utterance types and the informativeness of child witnesses. *Journal of Applied Developmental Psychology*, 17, 439-451.
- Sternberg, K. J., Lamb, M. E., Hershkowitz, I., Yudilevitch, L., Orbach, Y., Esplin, P. W., & Hovav, M. (1997). Effects of introductory style on children's abilities to describe experiences of sexual abuse. *Child Abuse & Neglect*, 21(11), 1133-1146.
- The National Crime Faculty. (1998). *A practical guide to investigative interviewing*. Bramshill: Training and Development Unit.
- Walker, A. G. (1999). *Handbook on questioning children. A linguistic perspective*. Washington, DC: American Bar Association: Center on Children and the Law.
- Walker, N. E., & Hunt, J. S. (1998). Interviewing child victim-witnesses: How you ask is what you get. In C. P. Thompson, D. J. Herman, J. D. Read, D. Bruce, D. G. Payne & M. P. Toglia (Eds.), *Eyewitness memory: Theoretical and applied perspectives* (pp. 55-87). Mahwah, NJ: Lawrence Erlbaum Associates.
- Wright, R., & Powell, M. B. (2006). Investigative interviewers' perceptions of their difficulty in adhering to open-ended questions with child witnesses. *International Journal of Police Science and Management*, 8(4), 316-325.
- Wright, R., & Powell, M. B. (2007). What makes a good investigative interviewer of children? A comparison of police officers' and experts' perceptions. *International Journal of Police Science and Management*, 30(1), 21-31.

- Wright, R., Powell, M. B., & Ridge, D. T. (2006). Child abuse investigation: An in-depth analysis of how police officers' perceive and cope with daily work challenges. *Policing: An International Journal of Police Strategies and Management*, 29(3), 498-512.
- Wright, R., Powell, M. B., & Ridge, D. (2007). What criteria do police officers use to measure the success of an interview with a child? *Psychology, Crime & Law*, 13(4), 394-404.
- Aarons, N. M., Powell, M. B., & Browne, J. (2004). Police perceptions of interviews involving children with intellectual disabilities: A qualitative inquiry. *Policing & Society*, 14(3), 269-278.

The Child Verbal Competence Effect in Court: A Comparative Study of Field Investigative Interviews of Children in Child Sexual Abuse Cases

TROND MYKLEBUST^{1,*} and ROALD A. BJØRKLUND²

¹Norwegian Police University College, Box 5027, Majorstua, N-0301 Oslo, Norway

²Department of Psychology, University of Oslo, Box 1094, Blindern, N-0317 Oslo, Norway

Abstract

The purpose of this study is to compare field investigative interviews of children (FIIC) with three different legal outcomes in child sexual abuse cases: (i) insufficient evidence to proceed (IEP); (ii) convictions; or (iii) acquittals by the court. One hundred FIIC were divided into one of the three outcome possibilities. Amongst the female interviewees older than 10 years, there were no cases of acquittals and the convicted cases were over-represented. The children's response to open questions was found to be the main difference between the three FIIC outcomes. The responses to these open questions were 1.9 and 2.3 times longer in the convicted cases compared to acquittals and IEP. Possible explanations for the result are discussed. Copyright © 2009 John Wiley & Sons, Ltd.

Key words: investigative interviews; police; questions; court; convictions

INTRODUCTION

In most child sexual abuse (CSA) cases, the field investigative interviews of children (FIIC) play a crucial role in the investigation. In the subsequent judicial process, the prosecutors consider the FIIC to be either sufficient or insufficient as evidence to proceed to court. For the FIIC considered sufficient to court the final verdict will end as either conviction or acquittal. The purpose of the present study is to compare the three outcome of the FIIC which could be either: (i) convictions; (ii) acquittals; or (iii) insufficient evidence to proceed (IEP).

In the cases that should be judged by the courts in the Nordic countries, the videotaped FIIC is presented to court as a substitute for the child victim being questioned and cross-examined by the members of the court during the trial (Fridriksdottir, 2008). In CSA cases, the videotaped FIIC is considered to be the evidence-in-chief, primarily because no other evidence is typically available. Based upon the videotapes, the court judges the cases as

*Correspondence to: Trond Myklebust, Norwegian Police University College, P.O. Box 5027, Majorstua, N-0301 Oslo, Norway.
E-mail: trond.myklebust@phs.no

convictions or acquittals. We have received the outcomes of 100 FIIC and of these, 31 are classified as IEP, 15 as acquittals and 54 as convictions. During the last five decades, a large amount of studies have focused on variables that characterise important aspects of FIIC. In the present study, we want to analyse if some of the often-sited variables affecting FIIC also affect the outcome as judged by the prosecutors or the courts.

It is well documented that the quality of the responses in FIIC depends upon how the interviewers phrase their questions. From the early research (Stern, 1903/1904), questions eliciting the interviewees to give their accounts in their own way without being negatively influenced by the interviewers, are demonstrated to yield long, detailed, and accurate responses. These accounts are also called 'free narratives' and are elicited by open questions (e.g. 'Tell me what happened'). Responses to open questions are demonstrated to be about four and a half times longer than the more leading closed questions (e.g. 'What was the colour of the car?' or 'Was he in the bedroom?') (Lamb, Hershkowitz, Sternberg, Boat, *et al.*, 1996; Lamb, Hershkowitz, Sternberg, Esplin, *et al.*, 1996; Sternberg, Lamb, Hershkowitz, Esplin, *et al.*, 1996). Derived from Stern's open versus closed dichotomy, there have been two approaches to the examination of questions (Richardson, Dohrenwend, & Klein, 1965). One approach is by use of broad categories to describe the overall style of the interview in terms as 'active' versus 'passive' (DuBois, 1937), and 'guided' versus 'free' (Kluckhohn, 1945). The other approach is by identifying characteristics, such as the length of answers, and to consider the effect of these characteristics at the level of the individual question. Richardson *et al.* (1965) introduced one of the first classification systems tending to divide the closed questions into three subcategories, denoted: 'identification', 'selection', or 'yes/no'. Based on categories like Richardson *et al.* (1965), several classification systems have been developed, all emphasising the importance of asking questions yielding long and detailed responses through the interviewees' free narratives (Fisher & Geiselman, 1992; Griffiths & Milne, 2006; Poole & Lamb, 1998; Powell & Snow, 2007; Shepherd, 2007; Walker & Hunt, 1998). In the present study, we have used the classification system developed by Richardson *et al.* (1965).

Focusing on the importance of the use of open questions to elicit long narratives, the interviewers' level of competence was introduced as an important variable (Fisher, Geiselman, & Raymond, 1987). The interviewers' competence could be described according to their theoretical and practical level, based on the length of their formal, scientifically based training at the institutional level, and the number of conducted FIIC (Fisher & Geiselman, 1992; Lamb, Sternberg, Orbach, Esplin, & Mitchell, 2002; Lamb, Sternberg, Orbach, Hershkowitz, *et al.*, 2002). The basic implicit assumption is that police officers who have received long and more specialised training generate longer answers from the interviewed children. We hypothesise that interviewers at a high-competence level will achieve longer answers from the interviewed children, compared to interviewers with lower competence.

Several studies of interviewed children have also revealed variables that might affect the length of responses. Studies of children aged 3–12 years all observed an effect of the children's age on the number of words in the children's responses in FIIC, with the vocabularies of young children being more limited and less descriptive than those of older children (Davies, Westcott, & Horan, 2000; Hershkowitz, Orbach, Lamb, Sternberg, & Horowitz, 2002; Korkman, Santtila, Westeråker, & Sandnabba, 2008; Walker, 1999). This is in line with the studies by Hershkowitz, Horowitz, and Lamb (2005; Hershkowitz, Fisher, Lamb, & Horowitz, 2007) examining all child abuse investigations conducted in Israel between 1998 and 2002. They also reported a gender effect with girls providing

more details than boys (Lamb, Hershkowitz, Orbach, & Esplin, 2008), supporting the developmental theory stating that girls being more linguistically mature than boys. Based upon these studies that indicate clear effects of both gender and age, we include those variables in the present study.

To sum up, a review of several studies indicate that factors related to the kind of questions, the competence of the interviewer, and the age and gender of the interviewed child all affect the FIIC. The purpose of the present study is to analyse if the same factors also affect the outcome of the FIIC either as evaluated by the prosecutor as IEP, or by the court judged as acquittals or convictions. The analysis of the FIIC could be done stepwise by first comparing the IEP with those judged by the court, and second, by comparing the convicted with the acquitted FIIC outcome. Methodologically, we argue that each of the one hundred FIIC in this study has been classified independently in one of the three outcome possibilities. Thus, there is no relation between the three different outcomes and we have treated the outcome as one variable.

METHOD

Sample

The study included 100 FIIC conducted by separate police interviewers; all statutorily mandated to conduct investigative interviews of children in Norway. Each interviewer met the following characteristics:

- 1) The child interviewed had to be 6–16 years of age (school age).
- 2) No deviation in the child's linguistic development or skills.
- 3) The interviewers' gender, age, training level (low/high), and the number of practiced FIIC.
- 4) The age and gender of the interviewed child.
- 5) The interview had to be fully video recorded, from the child entering the interview room until leaving the room.
- 6) The interview should be completely transcribed and approved by the judge according to the national Norwegian regulations.
- 7) The FIIC should be the only interview of the child in the case.
- 8) The case should be completed either as 'IEP' or proceeded by the court. The outcome of the proceeding in the court should be known.

The theoretical competence of the Norwegian police officers conducting the interviews could be divided into two groups, 'high' and 'low', respectively. The interviewers with high theoretical competence have substantially more in-depth and broader theoretical training lectures at the Norwegian Police University College compared to the interviewers with low theoretical competence. The two groups are denoted 'theoretical high' (T+) and 'theoretical low' (T-), respectively.

Procedure

A list with the individual case reference numbers in the National Register of Criminal Convictions (STRASAK) was presented together with the corresponding FIIC category for an officer at the National Criminal Investigation Services. The officer, who was

uninformed of the research process validated the FIIC outcome for each of the case reference numbers. No corrections were made to any of the FIIC outcomes initially reported by the interviewers submitting the interviews.

All the transcripts of the interviews were electronically digitised and validated word by word to ensure their accuracy and completeness with the video. Each of the interviewers' utterances were classified in one of the following six categories first developed by Richardson *et al.* (1965) and later modified by Myklebust and Bjørklund (2006):

- 1) *Instruction*: Information given by the interviewer to the child during the interview. The information is not directly related to the case that is investigated but about the interview itself, e.g. date and time the interview is conducted, equipment used during the interview, people present in the monitoring room, the importance of answering the questions loud and truthfully, and information about the context the interview is held.
- 2) *Open question*: A question that requires more than a few words for an adequate response, e.g. 'Tell me about that...', 'Describe for me...', 'Explain...'. The child's focus is not delimited more than in a general way.
- 3) *Identification*: (person, place, group, time, number etc.); who, where, when, how many, or which questions, e.g. 'What was his name?'
- 4) *Selection*: Fixed-alternative question, e.g. 'Was the door open or closed?'
- 5) *Yes-no question* answered by 'yes' or 'no', e.g. 'Do you know his name?', possibly supplemented by a redundant phrase such as 'I think so' or 'I doubt it'.
- 6) *Facilitators*: Utterances like 'okay', restatements of the child's previous utterance, and non-suggestive words of encouragement that are designed to encourage the child to keep talking.

Closed question was defined as category 3, 4, and 5. The closed questions include a limited number of response alternatives and could be answered adequately in a few words. *Open questions* deal with broader topics than closed questions and allow children more flexibility to choose which aspects of an event they will describe, e.g. 'Tell me about the person you met' (open question) and 'What kind of jacket did he wear?' (closed, identification question).

The classification of the interviewers' utterances was validated by randomly choosing 10 of the interviews in each of the two groups of interviewers (T-/T+), providing a total of 6,438 utterances. The results from the 20 validated interviews gave an inter-rater reliability coefficient of 0.96, indicating a high level of agreement between the authors. The validation revealed disagreement in 238 questions classified as 147 open/identification questions, 69 open/facilitator and 22 instructions/yes-no, respectively. The number of words in the children's answers to each of the investigators' utterances was counted in all the interviews.

Data analyses

The mean number of words in the children's answers to the six categories of interviewer utterances was calculated in all interviews and used as cell means for subsequent statistical analysis. The numbers of words in the children's answers were used as the dependent variable. To perform the statistical analysis, SPSS v. 16.0 was used. Given that all parametric assumptions were met (Langdridge, 2004) (data should be at least interval level, normally distributed, with homogeneity of variance), analysis of variance (ANOVA), and post hoc were conducted on the analysis. Five percent was used as the alpha level in all

the statistical tests. Furthermore, in order to investigate any differences between the outcome groups as to characteristics of children and interviewers, chi-square tests were performed.

The study was approved by the Director of the Public Prosecution and conducted according to the standards of the Norwegian regional committee of research ethics.

RESULTS

Table 1 shows the distribution of interviewed children and the outcome of FIIC.

By median split half method, we divided the children below 11 years in the ‘young’ children group (N = 52) and those above 10 years in the ‘old’ children group (N = 48). When we included gender as an independent variable, we got four groups of children. The total numbers of children in the FIIC outcomes are 31, 15, and 54 for the ‘IEP’, ‘acquitted’, and ‘convicted’ outcomes, respectively. Chi-square showed a significant difference between the groups ($\chi^2 = 17.129, p < 0.01$). Inspection of the residuals shows the greatest differences for the old girls with no cases of acquittals and over-representation in the convicted outcome. The old girls group is also the only group that is over-represented as convicted.

Table 2 shows the mean number of words in the children’s responses related to the three FIIC outcomes and the six categories of interviewer utterances.

Table 1. The distribution (residuals in parentheses) of interviewed children related to their gender, age-group and the three FIIC outcomes

Child group	FIIC outcome			Total
	IEP*	Acquitted	Convicted	
Young girls	10 (-0.2)	8 (3.0)	15 (-2.8)	33
Young boys	8 (2.1)	5 (2.2)	6 (-4.3)	19
Old girls	9 (-2.8)	0 (-5.7)	29 (8.5)	38
Old boys	4 (0.9)	2 (0.5)	4 (-1.4)	10
Total	31	15	54	100

*Insufficient evidence to proceed.

Table 2. Mean numbers of words (SD in parentheses) in the children’s responses related to the three FIIC outcomes and six categories of interviewers’ utterances

Interviewer utterances	FIIC outcome			
	IEP* N = 31	Acquitted N = 15	Convicted N = 54	Total N = 100
Instruction	3.1 (2.8)	2.6 (4.7)	2.1 (3.4)	2.5 (3.4)
Yes-no	5.4 (3.2)	4.7 (2.3)	7.1 (5.7)	6.2 (4.7)
Selection	7.0 (5.0)	6.8 (4.3)	9.9 (11.9)	8.6 (9.4)
Identification	9.6 (4.5)	7.7 (2.0)	11.4 (7.6)	10.3 (6.3)
Facilitators	6.6 (5.4)	8.7 (8.3)	7.7 (9.9)	7.5 (8.4)
Open question	26.8 (15.3)	21.6 (13.8)	49.8 (37.2)	38.4 (31.5)
Total	9.8 (10.7)	8.7 (9.3)	14.7 (23.2)	12.2 (18.6)

*Insufficient evidence to proceed.

ANOVA revealed a significant main effect of the three FIIC outcomes ($F(2,582) = 11.570$, $p < 0.001$, partial Eta squared = 0.038). The total mean numbers of words were 9.8, 8.7, and 14.7 for the IEP, acquitted, and convicted outcomes, respectively. Post hoc analysis showed that the number of words in the children's answers in the 'convicted' category was significantly higher than the two other FIIC outcomes (all p 's < 0.001), whilst the mean numbers of words in the acquitted and the IEP cases were not significantly different.

ANOVA showed a significant main effect of the six categories of utterances ($F(5,582) = 50.187$, $p < 0.001$, partial Eta squared = 0.301). The mean numbers of words for the categories 'instruction', 'yes-no', 'selection', 'identification', 'facilitator', and 'open' questions were 2.5, 6.2, 8.6, 10.3, 7.5, and 38.4 words, respectively. Post hoc analysis revealed that the number of words in the children's answers to open questions was significantly different from all the other categories (all p 's < 0.001). In addition, identification was significantly different from instruction ($p < 0.05$). ANOVA demonstrated a significant interaction effect between the three FIIC outcomes and the six categories of utterances ($F(10,582) = 6.491$, $p < 0.001$, partial Eta squared = 0.100). Post hoc analysis showed that the FIIC outcomes only affected the children's number of words on open questions (see Table 2). The number of words to open questions in the convicted outcome (49.8 words) was significantly different from the two other outcomes (IEP = 26.8 words, acquitted = 21.6 words, both p 's < 0.001), whilst the differences between the IEP and acquitted were not significant ($p > 0.05$). The convicted FIIC outcome was 1.7 and 1.5 times longer than the acquitted and IEP outcome, respectively.

Table 3 shows the distribution of the number of interviews related to the interviewers' competence and the outcome of FIIC.

The competence of the interviewers consists of two aspects; one is the theoretical competence and the other is the practical competence. By means of median split half method, the investigators were divided into two groups. Those who have conducted less than 15 interviews were denoted 'practical low' (P-) and interviewers with more or equal to 15 interviews as 'practical high' (P+). We used the information about the interviewers theoretical competence to divide them into 'theoretical low' (T-) or 'theoretical high' (T+). The combination of low/high theory (T-/T+) and low/high practice (P-/P+) gave four competence groups with 36, 13, 12, and 39 interviewers in the (T-/P-), (T-/P+), (T+/P-), and (T+/P+) categories, respectively. Chi-square test showed no significant differences between the groups ($\chi^2 = 7.135$, $p > 0.05$).

Table 3. The distribution (residuals in parentheses) of the number of interviews related to the interviewers' competence and the three FIIC outcomes

Competence level	FIIC outcome			Total
	IEP*	Acquitted	Convicted	
Group 1 (T-/P-)	10 (-1.2)	3 (-2.4)	23 (3.6)	36
Group 2 (T-/P+)	6 (2.0)	1 (-1.0)	6 (-1.0)	13
Group 3 (T+/P-)	2 (-1.7)	2 (0.2)	8 (1.5)	12
Group 4 (T+/P+)	13 (0.9)	9 (3.2)	17 (-4.1)	39
Total	31	15	54	100

*Insufficient evidence to proceed.

Table 4. Mean numbers of words (SD in parenthesis) in the children's responses to open questions related to the four groups of interviewer's competence and three FIIC outcomes

Competence level	FIIC outcome			Total
	IEP*	Acquitted	Convicted	
Group 1 (T-/P-)	31.2 (16.1)	23.7 (4.5)	48.6 (40.6)	41.7 (34.6)
Group 2 (T-/P+)	26.4 (17.5)	39.6 (0)	66.4 (50.0)	45.9 (39.6)
Group 3 (T+/P-)	15.4 (17.5)	9.5 (2.1)	68.3 (33.0)	49.7 (36.9)
Group 4 (T+/P+)	25.2 (14.1)	21.6 (15.7)	36.7 (24.3)	29.4 (20.2)
Total	26.8 (15.3)	21.6 (13.8)	49.7 (37.2)	38.4 (31.5)

*Insufficient evidence to proceed.

Table 4 shows the mean number of words in the children's responses to open questions related to the interviewer's competence and the FIIC outcomes.

ANOVA did not show any main effect of the interviewers' competence on the open responses ($F(3,88) = 0.681, p > 0.05$, partial Eta squared = 0.023). ANOVA showed a main effect of the FIIC outcomes on the open responses, ($F(2,88) = 9.087, p < 0.001$, partial Eta squared = 0.171), as previously described. The interaction between the three FIIC outcomes and the four categories of competence did not reveal any significant effect on the number of words in the children's answer to the open questions ($F(6,88) = 0.981, p > 0.05$, partial Eta squared = 0.063).

Since ANOVA did not reveal any significant differences between 'IEP' and 'acquitted' open responses, these two categories were merged into one category; denoted 'not-convicted'. ANOVA did not reveal any significant main effect of the four groups of children on the open responses ($F(3,92) = 1.729, p > 0.05$, partial Eta squared = 0.053) nor any interaction effect between the groups of children and the two FIIC outcomes (convicted/not convicted) ($F(3,92) = 0.762, p > 0.05$, partial Eta squared = 0.024).

We also divided the FIIC into short (less than 49 minutes, $N = 52$) and long interviews (above 48 minutes, $N = 48$). ANOVA showed no main effect between long or short interviews on open responses ($F(1,96) = 1.083, p > 0.05$, partial Eta squared = 0.011). The interaction between the length of the interviews and the two FIIC outcomes was not significant ($F(1,96) = 2.504, p > 0.05$, partial Eta squared = 0.025).

DISCUSSION

We have used the outcome variable from 100 FIIC in order to identify any differences between IEP (31 cases), acquittals (15 cases), and convictions (54 cases).

We found that none of the old girls (more than 10 years old) were represented in the acquitted outcome and that they were over-represented in the convicted outcome. The fact that the court always judges the cases with old girls as convicted outcomes has never been observed before. We believe that the studies related to verbal competence might be a possible candidate to enlightening this finding. The children's verbal competence have been investigated by Davies *et al.* (2000), Korkman *et al.* (2008), Lamb *et al.* (2008), Lamb, Sternberg, & Esplin (2000), Sternberg *et al.* (1996), and Sternberg *et al.* (1997). Taken together, these studies show that old girls have longer verbal responses compared

to both boys and younger girls. Our finding might indicate that the court decisions are affected by children who use many words in their testimonies. We have called this effect the 'child verbal competence effect in court'. We also found that the children's responses were longer in the convicted outcome compared to the other outcomes. Especially in the categories with open questions, the responses were about twice as long in the convicted outcome. These results support the child verbal competence effect in the court. We do not have any independent information about the one hundred children in the study. Hence, we are not able to evaluate the theory without additional information about verbal competence. The analyses, including the four groups of children, failed to find any differences in the response lengths between them. Thus, the child verbal competence effect in court might not be solely related to old girls.

The finding that responses to open questions are about twice as long in the convicted outcome might be explained in, at least, three ways. Based upon the studies that indicate a high correlation between the length of children's responses and number of details in FIIC (Lamb, Hershkowitz, Sternberg, Boat, & Everson, 1996; Lamb, Hershkowitz, Sternberg, Esplin *et al.*, 1996; Sternberg *et al.*, 1996), one could expect that children who experienced CSA in reality actually have more to talk about than other children. Undeutsch (1967; 1982) demonstrated a causal relationship between the children's personal experiences and the length of their verbal utterances in FIIC, with children who have been exposed to sexual assaults perceiving a large amount of information that generates several details in their statements. This finding is referred to as the 'Undeutsch-hypothesis' and according to the theory, children who have experienced real assaults will use many words in their responses to open questions compared to children without this experience. Our result, with children using more words in their responses in cases resulting in convictions compared to the other outcomes, might fit with the 'Undeutsch-hypothesis'.

Another explanation for why some children have responses nearly twice as long in the convicted outcome, compared to children in the IEP or acquitted outcomes, could be the children's communicative training. By having children practising in providing detailed narrative accounts of experienced events in day-to-day conversations, several children are entrained in their response styles by their parents, teachers and other adults. As a consequence, children who have grown up with high degree of verbal stimulation or with verbal talents might use more words in their responses resulting in longer answers in FIIC. Contrary, other adults entrain their children by frequently interrupting them, using their adult authority to decide how much time may be spent discussing certain topics, and testing whether the children have acquired certain knowledge that the adults already possess. As a consequence, these children are accustomed to adults possibly correcting what they have said, and in their interaction, the adults become the most active conversational partner taking on large responsibility for formulating and interpreting what is being said. Being entrained in this uneven balance of power between children and adults, the latter group of children might give shorter reports in the FIIC situation. The last explanation of our result focuses on the effect of specific stimulation on the children's memory system. Through history several studies and CSA cases (Ceci & Bruck, 1995; Garven, Wood, Malpass, & Shaw, 1998) have demonstrated that errors arise when victims and witnesses are exposed to information that is false or to social pressures that encourage particular types of answers. This is generally referred to as 'suggestibility' and might explain why some children give long statements in their interviews even if the incident under investigation has never taken place. If the children in the present study have been exposed to false information or interviewed in a suggestive way, they might use a larger amount of words from this exposure

compared with children without this specific stimulation (Bruck *et al.*, 1995; Ceci *et al.*, 1994; Leichtman & Ceci, 1995; White *et al.*, 1997). Unfortunately, we do not have any independent and objective evidence that might indicate which of the three possible explanations that best fits to the finding that the responses is twice as long in the convicted outcome compared to the other outcomes.

The duration of the interviews in this study varied between 14 and 104 minutes, but we did not find any difference between the short and long interviews on the number of words in the children's answers. This indicates that the length of the interview do not affect the length of the responses in FIIC.

Based upon the literature (Fisher & Geiselman, 1992; Fisher, Geiselman, & Amador, 1989; George & Clifford, 1992; Lamb, Sternberg, Orbach, Esplin *et al.*, 2002; Lamb, Sternberg, Orbach, Hershkowitz, *et al.*, 2002; Sternberg, Lamb, Orbach, Esplin, & Mitchell, 2001), we assumed that police officers with high level of theoretical competence would generate longer responses amongst the children compared to other interviewers. Our results do not support any such competence theory. One of the authors of the present study has himself conducted a large number of police interviews, including FIIC. It is possible that the training of the officers is good, but that the situation with shortage of time to plan, absence of information about the child and the offence under investigation etc. may counteract good quality in the FIIC.

The present findings could be vulnerable to several factors. One factor is that we have only analysed the interviews of the children, not including any of the other documentations or evidence presented to the court, which might have been essential in the assessment and evaluation by the different courts reaching their final verdicts. Second, another factor is that our results related to the competence of the interviewers could possibly be explained by how the officers understand the term 'open questions'. The Norwegian literature in this area is scarce. Most literature used in the training of the interviewers at NPUC is referring and quote the English literature and it is also expected that the interviewers read some of the literature in English. The literatures on definition of question types are inconsistent and somewhat contradictory of what constitutes an 'open question'. Thus, the questions which begin with '*Tell, Explain, or Describe*' are said to be open by researchers, police training departments and practitioners. There are less agreement within the closed question category, with e.g. the 'wh..' lexis of '*who*', '*what*', '*where*', '*which*', and '*when*', categorised as 'closed-specific' by Loftus (1982), as open by CPTU (1992), on the contrary Shepherd (2007) argue that 'wh' questions could be classified both as open and closed questions. This divergent description of open questions could confuse the police officers who conduct FIIC. The third factor might be that the police officers, themselves, have chosen the interviews that have been included in the study and perhaps the officers with the lowest competence have used more resources to select an outstanding FIIC. The fourth factor could be that in some of the cases the suspect might have pleaded guilty in the initial suspect interview, before the case went to court. However, this would not have an effect on the way the interview of the child has been conducted, all the time the offended child is interviewed before the suspect in the case. At our knowledge all the children were interviewed before the suspect in the present study.

In conclusion, the analysis of one hundred FIIC demonstrated that amongst girls older than 10 years there were no cases of acquittals and an over-representation of the conviction condition. The results indicate that the courts appear to be affected by children who use many words in their testimonies, what we have defined as the 'child verbal competence effect in court'. FIIC that contain longer responses to open questions in court results in

convictions, compared with shorter responses that result in acquittals or with evidence found insufficient to proceed to court.

ACKNOWLEDGEMENTS

The authors wish to thank the Director of Public Prosecution for the interest in the project, and the investigative interviewers and Police Departments that made their interviews available for study.

REFERENCES

- Bruck, M., Ceci, S. J., Francoeur, E., & Barr, R. (1995). 'I hardly cried when I got my shot!' Influencing children's reports about a visit to their pediatrician. *Child Development*, *66*, 193–208.
- 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., Loftus, E., Leichtman, M. D., & Bruck, M. (1994). The role of source misattributions in the creation of false beliefs among preschoolers. *International Journal of Clinical and Experimental Hypnosis*, *42*(4), 304–320.
- CPTU (1992). *Investigative interviewing: A guide to interviewing*. London: Central Planning and Training Unit.
- Davies, G. M., Westcott, H. L., & Horan, N. (2000). The impact of questioning style on the content of investigative interviews with suspected child sexual abuse victims. *Psychology, Crime and Law*, *6*, 81–97.
- DuBois, C. (1937). Some psychological objectives and techniques in ethnography. *Journal of Social Psychology*, *8*, 285–300.
- Fisher, R. P., & Geiselman, R. E. (1992). *Memory-enhancing techniques for investigative interviewing. The cognitive interview*. Springfield, IL: Charles C. Thomas Publisher.
- 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*(5), 722–727.
- Fisher, R. P., Geiselman, R. E., & Raymond, D. S. (1987). Critical analysis of police interview techniques. *Journal of Police Science and Administration*, *15*(3), 177–185.
- Fridriksdóttir, H. (2008). Afhörung af børn i sager om seksuelle overgreb [Investigative interviews of Children in CSA cases]. In A. M. B. (Ed.), *Forhandlingerne ved Det 38. nordiske Juristmøde i København 21.-23. august 2008 Bind I* (pp. 177–187). København: Den danske lokalstyrelse.
- 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*(3), 347–359.
- George, R., & Clifford, B. R. (1992). Making the most of witnesses. *Policing*, *8*, 185–198.
- Griffiths, A., & Milne, B. (2006). Will it all end in tiers? Police interviews with suspects in Britain. In T. Williamson (Ed.), *Investigative interviewing. Rights, research, regulation*. Cullompton, UK: Willian.
- Hershkowitz, I., Fisher, S., Lamb, M. E., & Horowitz, D. (2007). Improving credibility assessment in child sexual abuse allegations: The role of the NICHD investigative interview protocol. *Child Abuse & Neglect*, *31*, 99–110.
- Hershkowitz, I., Horowitz, D., & Lamb, M. E. (2005). Trends in children's disclosure of abuse in Israel: A national study. *Child Abuse & Neglect*, *29*, 1203–1214.
- Hershkowitz, I., Orbach, Y., Lamb, M. E., Sternberg, K. J., & Horowitz, D. (2002). A comparison of mental and physical context reinstatement in forensic interviews with alleged victims of sexual abuse. *Applied Cognitive Psychology*, *16*, 429–441.
- Korkman, J., Santtila, P., Westeråker, M., & Sandnabba, N. K. (2008). Interviewing techniques and follow-up questions in child sexual abuse interviews. *European Journal of Developmental Psychology*, *5*(1), 108–128.

- Kluckhohn, C. (1945). The personal document in anthropological science. In L. Gottschalk, C. Kluckhohn & R. Angell (Eds.), *The use of personal documents in history, anthropology, and sociology* (pp. 79–173). New York: Social Science Research Council.
- Lamb, M. E., Hershkowitz, I., Orbach, Y., & Esplin, P. W. (2008). *Tell me what happened: Structured investigative interviews of child victims and witnesses*. Chichester: John Wiley & Sons, Ltd.
- Lamb, M. E., Hershkowitz, I., Sternberg, K. J., Boat, B. W., & Everson, M. D. (1996). Investigative interviews of alleged sexual abuse victims with and without anatomical dolls. *Child Abuse and Neglect*, 20(12), 1251–1259.
- Lamb, M. E., Hershkowitz, I., Sternberg, K. J., Esplin, P. W., Hovav, M., Manor, T., & Yudilevitch, L. (1996). Effects of investigative utterance types on Israeli children's responses. *International Journal of Behavioral Development*, 19(3), 627–637.
- Lamb, M. E., Sternberg, K. J., & Esplin, P. W. (2000). Effects of age and delay on the amount of information provided by alleged sex abuse victims in investigative interviews. *Child Development*, 71(6), 1586–1596.
- Lamb, M. E., Sternberg, K. J., Orbach, Y., Esplin, P. W., & Mitchell, S. (2002). Is ongoing feedback necessary to maintain the quality of investigative interviews with allegedly abused children? *Applied Developmental Science*, 6(1), 35–41.
- Lamb, M. E., Sternberg, K. J., Orbach, Y., Hershkowitz, I., Horowitz, D., & Esplin, P. W. (2002). The effects of intensive training and ongoing supervision on the quality of investigative interviews with alleged sex abuse victims. *Applied Developmental Science*, 6(3), 114–125.
- Langdridge, D. (2004). *Introduction to research methods and data analysis in psychology*. Harlow: Pearson.
- Leichtman, M. D., & Ceci, S. J. (1995). The effects of stereotypes and suggestions on preschoolers' reports. *Developmental Psychology*, 31(4), 568–578.
- Loftus, E. (1982). Interrogating eyewitnesses—good questions and bad. In R. Hogarth (Ed.), *Question framing and response consistency*. San Francisco: Josey-Bass.
- Myklebust, T., & Bjørklund, R. A. (2006). The effect of long-term training on police officers' use of open and closed questions in field investigative interviews of children (FIIC). *Journal of Investigative Psychology and Offender Profiling*, 3, 165–181.
- Poole, D. A., & Lamb, M. E. (1998). *Investigative interviews of children: A guide for helping professionals*. Washington, DC: American Psychological Association.
- Powell, M. B., & Snow, P. C. (2007). Recommendations for eliciting a disclosure of abuse from a young child. *Australian Police Journal*, 61(2), 76–80.
- Richardson, S. A., Dohrenwend, B. S., & Klein, D. (1965). *Interviewing: Its forms and functions*. London: Basic Books, Inc., Publishers.
- Shepherd, E. (2007). *Investigative interviewing: The conversation management approach*. Oxford: Oxford University Press.
- Stern, W. (1903/1904). *Beiträge zur Psychologie der Aussage*. Leipzig: Verlag von Johann Ambrosius Barth.
- Sternberg, K. J., Lamb, M. E., Hershkowitz, I., Esplin, P. W., Redlich, A. D., & Sunshine, N. (1996). The relation between investigative utterance types and the informativeness of child witnesses. *Journal of Applied Developmental Psychology*, 17, 439–451.
- Sternberg, K. J., Lamb, M. E., Hershkowitz, I., Yudilevitch, L., Orbach, Y., Esplin, P. W., & Hovav, M. (1997). Effects of introductory style on children's abilities to describe experiences of sexual abuse. *Child Abuse and Neglect*, 21(11), 1133–1146.
- Sternberg, K. J., Lamb, M. E., Orbach, Y., Esplin, P. W., & Mitchell, R. (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*, 85(5), 997–1005.
- Undeutsch, U. (1982). Statement reality analysis. In A. Trankell (Ed.), *Reconstructing the past: The role of psychologists in criminal trials* (pp. 27–56). Stockholm: P A Norstedt & Söners förlag.
- Undeutsch, U. (1967). Beurteilung der Glaubhaftigkeit von Aussagen. In U. Undeutsch (Ed.), *Handbuch der Psychologie Bd.11: Forensische Psychologie* (pp. 26–181). Goettingen: Hogrefe.
- Walker, A. G. (1999). *Handbook on questioning children: A linguistic perspective* (2nd ed.). Washington, DC: American Bar Association: Center on Children and the Law.

- Walker, N. E., & Hunt, J. S. (1998). Interviewing child victim-witnesses: How you ask is what you get. In C. P. Thompson, D. J. Herman, J. D. Read, D. Bruce, D. G. Payne & M. P. Toglia (Eds.), *Eyewitness memory: Theoretical and applied perspectives* (pp. 55–87). Mahwah, NJ: Lawrence Erlbaum Associates.
- White, T. L., Leichtman, M. D., & Ceci, S. J. (1997). The good, the bad, and the ugly: Accuracy, inaccuracy, and elaboration in preschoolers' reports about a past event. *Applied Cognitive Psychology*, *11*, 37–54.