An analysis of the quality of French investigative interviews with children: Age of the witnesses does matter

Fanny Verkampt[1*], Olivier Dodier[2*], Becky Milne[3], & Magali Ginet[2]

[1]CLLE, Université de Toulouse, CNRS, UT2J, France

[2]Université Clermont Auvergne, Clermont-Ferrand, France

[3]University of Portsmouth, Portsmouth, UK

*Corresponding Author:

Fanny Verkampt, Laboratoire CLLE, Université de Toulouse, 5 Allées Antonio Machado,

31 058 Toulouse Cedex 9, France. E-mail: fanny.verkampt@univ-tlse2.

Abstract

A context-driven analysis was performed to assess the quality of French investigative interviews with three age groups of child witnesses (under 7 years old, 7-10 years old, and 11-17 years old). We measured how age was related to the quality of 24 real-life interviews by evaluating how child-centered the interview was: (i) did it follow recommendations for each phase of the interview and , (ii) was the questioning style appropriately adapted to the child's previous answer. Results showed that the older children gave more detailed responses to open questions. However, investigators did not encourage their recall through the use of appropriate questions, but instead asking closed and leading questions. With the younger children, interviewers also asked predominantly closed and leading questions but after evasive answers. Our findings suggest different strategies underpin the use of inappropriate questions with respect to the age of the witnesses. Implications for training are discussed.

Keywords: investigative interviewing, child witnesses, age, questioning techniques, dynamics of verbal exchanges

Introduction

To give victims the best possible opportunity to provide sufficient and accurate evidence, French investigators are trained to use child-centered interviews that are structured around four main phases (i.e. introduction, free recall, questioning, and closure), and in the use of appropriate questioning techniques (e.g. focus on open questions, see Achieving Best Evidence developed in UK, Ministry of Justice, 2011, 2014). Despite a scientific consensus on the benefits associated to open-ended questions, a large body of research has suggested that, in the real world, practitioners make extensive use of inappropriate closed or suggestive questions during child interviews (e.g. Kask, 2012; Luther, Snook, Barron, & Lamb, 2015). Among the factors which affect the use of appropriate questioning, a child's willingness to disclose and the ability of the interviewer to apply recommendations in an effective way are important to understanding why best practice guidance is not followed. In the field, the challenge faced by the interviewer is to tailor their question use to the particular child's developmental needs and to the specific context of an interview, that is: to consider the interview phase (Griffiths & Milne, 2006) and to respond dynamically to the answers provided by the child (i.e. the dynamics of verbal exchanges; see, Korkman, Santilla, Westeråker, & Sandnabba, 2008; Lindholm, Cederborg, & Alm, 2014).

Prior research has investigated the quality of child interviews by counting the number of appropriate vs. inappropriate questions asked during the interview (e.g. Kask, 2012). Others have considered the context within which questions are asked, by considering either the interview phase (e.g. Westcott & Kynan, 2006) or the dynamics of verbal exchanges (e.g. Wolfman, Brown, & Jose, 2016). Nevertheless, to the best of our knowledge, no analysis of the quality of child interviews has examined the relation between the witness's age and the two context-related elements, specific phase and dynamics of verbal exchanges, in the use of questions.

Therefore, the first aim of the current archival study was to utilise a context-driven analysis of question use by interviewers, by considering the specific phase and dynamics of verbal exchanges, with three age groups of young witnesses: very young children (under 7 years old), young children (7-10 years old), and older witnesses (11-17 years old). Since limited research has examined real-life interviews in this country, the second aim was to explore this issue within French forensic interviews for the first time. This exploration may also broaden our understanding of the quality of child interviews by offering an analysis of interviews conducted in an inquisitory criminal justice system (vs. accusatory system as investigated in most studies; for example, see Abbe & Brandon, 2014; Hughes-Scholes & Powell, 2013).

Background Information

Specificities of child victims

Age differences in memory recollection. Young victims may be uncommunicative during investigative interviews because of socio-emotional factors (e.g. shame, modesty; see Leander, 2010) and memory retrieval difficulties. Memory retrieval is known to vary in accordance with the individual differences of the rememberer (e.g. cognitive development levels; Fivush, 2010) and depending on the to-be-remembered event itself (e.g. trauma factors; Alexander et al., 2005). In addition, research demonstrated that the age of the interviewee affects the quantity of the resultant recall, with younger children tending to recall less information (Jack, Leov, & Zajac, 2014).

The difficulty in retrieving memories may be more pronounced with victims younger than 7 years due to having less advanced retrieval strategies (e.g. Fivush, 1993). As a consequence, younger children rely on prompts, specifically open and probing questions, in order to aid their retrieval (Brown & Lamb, 2015). Furthermore, a less extensive vocabulary

exacerbates the difficulties, due to having less of a resource to help describe the occurrence (Brown, 1973; Walker, 1999).

Although this recall difficulty exists also for children aged from 7 to 10 years, their memory performance nevertheless improves thanks to the development of language and of the repetition of personal experiences. Their parental conversations and school learning (Fivush & Hamond, 1990) contribute to the development of their narratives (i.e. canonical linguistic forms specifying a sequence of actions and the links between them; Fivush, 2010). The development of canonical linguistic forms helps them not only to organize their recall but also to improve their recollection of memory (Fivush, 2010).

Finally, adolescence is marked by the development of the prefrontal cortex involved in perception (Rossi, Pessoa, Desimone, & Ungerleider, 2009), coding and organization of information in memory, and narrative production (Ansado, Chiasson, & Beauchamp, 2014). Although witnesses from 11 to 17 years may be generally considered as more able in retrieval tasks compared to young (from 7 to 10 years) and very young children (younger than 7 years), adolescent's individual variability is important to consider. Thus, assessment is required to tailor such interview to each individual.

Non-linear trend in suggestibility. Children are known to be vulnerable to misinformation suggested prior to or during investigative interviews (e.g. authority figures, co-witnesses; Carol & Schreiber Compo, 2015; Garven, Wood, & Malpass, 2000; Grisso et al., 2003). Very young children (under 7) are also highly vulnerable to suggestive interview techniques, such as: (mis)leading questions (i.e. questions suggesting a particular piece of information), misleading information, misleading props, repeated questions, etc. (see Ceci & Bruck, 1995; Gabbert & Hope, 2017, for a review). However, the link between age and suggestibility is not linear. A key period for vulnerability (i.e. period of utilization deficiency) seems to occur around the age of 8 (Miller, 1990), where children who have recently started

to spontaneously generate their own memory recollection strategies often find it difficult to attend fully to a recall task. Thus, a deficit in attentional resources may explain why 8 year-old children comply more often with information suggested after their free recall (Dewhurst & Robinson, 2004).

(Pre)adolescents (11-17 years old) are also particularly susceptible to suggestion when: (i) The suggestion is semantically similar to the original information; and (ii) they are interviewed with high-pressure tactics (e.g. leading questions and positive or negative reinforcement provided by the investigators; see Gudjonsson & Singh, 1984; Richardson, Gudjonsson, & Kelly, 1995). According to Fuzzy trace theory (Brainerd & Reyna, 2002, 2005) and Associative-activation (Howe, Winner, Gagnon, & Plimpton, 2009) theories, semantic knowledge plays a prominent role in recall, recognition, and false memory. Children of any age are thus more likely to incorporate suggested information in their accounts that are related to (piece of) the original experience (Ceci, Papierno, & Kulkofsky, 2007; Dewhurst & Robinson, 2004), and this detrimental effect increases as knowledge and experience develop through repeated exposure and development. It can be thus seen that children and (pre)adolescents need to be interviewed carefully to obtain reliable information.

Child forensic interviews in the field

Even though international interviewing guidance outlines the risks associated with inappropriate questioning and promotes the use of open-ended questions, a substantial number of field studies have shown that the use of open-ended questions is worryingly low in the field (e.g. Hershkowitz et al., 2006; Powell & Guadagno, 2008; Sternberg, Lamb, Davies, & Westcott, 2001). More recently, and despite a subsequent development of professional training focusing on the nature of questions in child interviews, field analysis still found similar patterns of poor questioning (Kask, 2012; Luther et al., 2014; Wolfman et al., 2016).

Several factors have been identified as potential impediments to the use of appropriate

and child-centered questioning. First, there is a non-adherence to the four-phase structure (Philipps, Oxburgh, Gavin, & Myklebust, 2012). Westcott and Kynan (2006) found that the free recall and closure phases were absent in about half of the interviews they examined, while the introduction[i] and questioning phases were present in the majority of interviews examined. Interviewers tend to rush children into the questioning phase of the interview rather than promoting free recall.

Another factor in the underuse of appropriate questioning with the youngest children (under 7 years vs. 7-9 and 10-12 years) may be the unresponsiveness of the child victims themselves (Westcott & Kynan, 2006). Practitioners may in turn believe they need to ask inappropriate questions to overcome this barrier. If this explanation is correct, non-recommended utterances (e.g. inappropriate closed or suggestive questions) might be used more frequently with very young child victims, whose memory capacity and language abilities could lead to a higher number of non- or evasive responses. Kask (2012), for example, found that suggestive questions were used significantly more often with 4 to 7-year-old children (vs. 8-11 years, 12-14 years). Looking at "Interviewer utterance – Interviewee answer" associations, work on the dynamics of verbal exchanges has found that open questions (invitations) and probing questions (directives) were more likely to elicit both request-confirming (Keselman et al., 2010; Lindholm et al., 2014) and evasive answers (Lindholm et al., 2014) from 13 to 18-year-old alleged victims of human trafficking. Not surprisingly, closed or suggestive questions elicited more agreement and disagreement and, to some degree, more extended task-related answers (Keselman et al., 2010; Lindholm et al., 2014; see also Wolfman et al., 2016). These findings support the argument that the overuse of inappropriate questions might be a result of children's unresponsiveness (Gilstrap & Papierno, 2004).

When examining "Interviewee answer-Interviewer utterance" associations, Korkman

and her colleagues even found that after an extensive and judicially significant answer, interviewers still relied more on suggestive questions than on the recommended open-ended utterances with alleged victims aged 3-8 years (Korkman et al., 2008) and 4-7 years (Korkman, Santilla, & Sandnabba, 2006). The fact that this observation was obtained (i) with interviewers who did not use a structured approach to conduct the interviews and (ii) with children known to be more dependent on interviewers' support (e.g. probing questions) may both explain the findings. Regarding the latter, as hypothesized by Korkman et al. (2008), interviewers might have based their questioning more on their general beliefs and expectations about the recall capacities of young children, rather than on the informative nature of answers previously given by the young victims. In order to determine the relation between the non-responsiveness of children and the quality of interviews, a context-driven analysis interview quality (i.e. with adherence to the four-phase approach and the dynamics of verbal exchanges) with different age groups of young witnesses, including older children (11 years and older), was conducted.

Present Study

The major aim of this archival study was therefore to conduct a context-driven analysis of the quality of real-life child interviews in France. Such analysis was conducted with three age groups of child witnesses (under 7 years, 7-10 years, 11 years and older). Additionally, we examined whether interviews adhered to the four-phase structure *and* the dynamics of verbal exchanges. Given the above-mentioned differences in terms of abilities in recollecting memories and of vulnerability to suggestion, different contexts of question use (i.e. interview phase, dynamics within investigative interviews) were examined as a function of the age of the child.

A second aim was to explore these issues in France because, to our knowledge, no research has assessed the quality of forensic child interviews either before or after a new law

reformed the area of child protection and child investigative interviews in France (Law n°2007-293 of March 5, 2007). Before 2007, interviewer training included information on non-suggestive and child-centred interviews (see Berthet & Monnot, 2007, for a review). From 2007, modifications in interviewer training were made to better meet with the requirement of the new law. This led to inclusion of a module on video-recorded interviews and how to transcribe them, more exercises on evidence-based content, and more role playing. The mandatory video-recording of child interviews, as provided for in the law, encouraged the use of qualified professionals for such interviews. However, this recommendation was not accompanied by mandatory monitoring of implementation. Finally, from 2016, training in the National Institute Child Health and Human Development protocol (NICHD protocol, see Cyr & Lamb, 2009; Lamb, Orbach, Hershowitz, Esplin, & Horowitz, 2007) was also implemented in France. To date, 240 military interviewers have been trained in this protocol and, at the end of 2018, 200 additional military interviewers should have been trained (R. Job, personal communication, February 5, 2018).

Material and Method

Data

The Home Office, the National Police and Military Police, as well as the Court District of Clermont-Ferrand (France) gave the official authorizations for this study. Our interview data consisted of closed cases, in which children under 18-year-old were interviewed as witnesses (and/or victims) of sexual and domestic violence. Interviews conducted with young offenders were thus excluded. We only included interviews conducted after the adoption of the 2007 law. A member of the Court District staff selected the cases that met these criteria. Then our team examined these interviews in a dedicated room, inside the Court District. All descriptive statistics regarding these interviews are reported in Table 1.

Witnesses and interviewers. Our initial sample consisted of 28 children aged from 3 to 17 years. Due to missing data (i.e. age of the children), 4 interviews were excluded from the sample. Chi-square (χ^2) tests of independence showed a non-significant differences between the wearing of a uniform and the age of the witness, χ^2 (2, $N = 23$) = 0.059, $p = .97$, Cramer's V = 0.044.

Interviews and Characteristics of the Allegations.

Chi-square (χ^2) tests of independence showed a non-significant difference between the type of abuse and the child-suspect relationship, χ^2 (9, $N = 24$) = 8.643, $p = .47$, Cramer's V = 0.20. There was a significant age difference regarding the type of sexual abuse (non-penetrative vs. penetrative sexual abuse, $n = 22$), where the witnesses interviewed about non-penetrative sexual abuse being older ($M = 9.53$; $SD = 3.89$) than those interviewed about penetrative sexual abuse ($M = 5.00$; $SD = 2.10$), $t(17) = 2.608$, $p = .014$, Cohen's $d = 1.28$.

[Table 1 here]

Procedure

Transcription. The interview transcripts were produced *in situ* during time slots decided on by the Court District. With the exception of the victims' age and gender, personal details and references to places that might make it possible to identify the child victims were removed from the transcripts.

Coding. The interviewers' and children's utterances were coded by the second author as well as by three research assistants who were unaware of the study's purposes ($n = 4$). In order to evaluate whether the interviewer utilized the appropriate questioning style adapted to the child's previous answer, the codes for each interviewer' and child's utterance were mutually exclusive and exhaustive (see Bakerman & Quera, 2011; Wolfman et al., 2016). This allows the following sequential analyses: Interviewer → Child; Child → Interviewer. As the nature of the interviewer's utterances was dependent upon the interview context (i.e. the

10

interview phase), we coded the content of the interviewer's utterances and the child's answers once the coding of the four-phase structure had been completed.

In order to evaluate whether each phase identified in the Achieving Best Evidence guidance was used, the four phases were first coded as present or absent, as well as their order of presentation. One point was then given when one element of the interview was representative of one specific phase. For example, if one element of the introduction phase was identified (e.g. neutral discussion), it was coded as "1-Introduction". Interviewer-interviewee exchanges made from the presentation of the purpose of the interview (e.g. Do you know why you are here today?), until the time at which the children no longer freely explained in their own words what had happened, were coded as "1-Free recall". Interviewer-interviewee exchanges from then on until the interviewer stated that the interview was ended were coded as "1-Questioning". The final interviewer-interviewee exchanges (e.g. So how are you feeling?) were coded as "1-Closure".

The content of the interviewers' and the children' utterance types were then coded. We coded the interviewers' utterances using the Griffiths Question Map coding scheme (Griffiths & Milne, 2006; see also Dodier & Denault, 2018). Interviewer's utterances were first coded as either (i) appropriate; that is being child-centered and adapted to the individual characteristics of the child (language, active mental images) and thus in turn supporting more complete and accurate recall, or (ii) inappropriate; that is utterances limiting completeness and accuracy of recall (cf. Table 2). Appropriate utterances were sub-classified into one of the three following categories: (i) open-ended question, (ii) probing question, and (iii) closed (yes-no) question. Inappropriate utterances were sub-classified as: (i) open-ended question, (ii) closed (yes-no) question, (iii) multiple question, (iv) forced-choice question, (v) opinion or statement, and (vi) leading question. All the interviewers' utterances were coded, with one point being given for each utterance. For example, the interviewer's question "Tell me

everything about that?*" was coded "1-Appropriate-open question" if it was asked immediately after the child had mentioned that his father had punched him and was done so without interrupting the child's speech. If this question interrupted the child's speech, it was coded "1-Inappropriate-open question".

The coding scheme used for the children's utterances was similar to the one used by Keselman et al. (2010), that is each utterance being whether: (i) extended task-related, (ii) request-confirming, (iii) agreement, (iv) disagreement, (v) no answer, and (vi) evasive response. All the children's answers were coded and one point was given for each answer. For example, if a child answered "I was in my mom's car" to the question "Where were you?", this was coded "1–Disclosure-request-confirming". If s/he answered "I was in my mom's car with my little sister who was crying because dad hurt mummy", her/is answer was coded "1-Disclosure–extended task-related". An example of interviewers' utterances and child's responses coding is presented in Appendix.

[Table 2 here]

Training and inter-coder reliability

The four main coders and the one involved in the inter-coder reliability measures (named "inter-coder reliability" coder) were trained in the use of the GQM and of Keselman et al's coding scheme in two training sessions. The first session consisted of lectures explaining the different phases of a child-centered interview and the utterances recommended to the interviewers in international guidelines. The second training session consisted of several coding practices, with immediate and individual feedback. The "inter-coder reliability" coder, who was unaware of the purposes of the current study, coded a random sample ($n = 10; \approx 42\%$) of the interviews. This sample was composed of two interviews coded by coder 1, two interviews coded by coder 2, three interviews coded by coder 3, and

three interviews coded by coder 4. Therefore, inter-coder reliability was assessed across two coders, using Cohen's Kappa coefficient. The results for the interviewers' utterances (Cohen's Kappa$_{appropriate\ questions}$= .91, p < .001, Cohen's Kappa$_{inappropriate\ questions}$ = .93, p < .001, Cohen's Kappa$_{answers}$ = .95, p < .001) suggested excellent inter-rater reliability.

Results

Total Base Rates

The duration of the interviews ranged from 10 to 60 min (M = 28 min; SD = 16 min). In total, the interviewers and child victims produced 4046 sequences of question/answer pairings and 4023 answer/question pairings. Investigators asked an average of 154.42 questions per interview (SD = 75.37), ranging from 51 to 324 questions. Among them, 29.8% were appropriate questions (with 2.2% open-ended questions, 26.2% probing questions, 1.4% appropriate closed questions), and 70.2% were inappropriate questions (with 0.3% inappropriate open-ended questions, 32.7% inappropriate closed questions, 7.1% multiple questions, 3.7% forced choice questions, 8.6% opinions or statements, and 17.8% leading questions). Regarding the witnesses' answers, 83.6% were disclosure answers (with 26% request-confirming, 24.7% agreement, 16.6% disagreement, and 16.3% extended task-related), and 16.4% were non-disclosure answers – including 11.2% no answer, and 5.2% evasive responses.

Preliminary Analyses

Using a series of separate t-test, we first analyzed the influence of (i) interviewers' gender, (ii) interviewers' rank, and (iii) the wearing of a uniform during interviews on the total number of questions asked and on the total disclosure answers provided by the child witnesses. While it may influence the child's willingness to talk, in particular with regard to potentially traumatic events such as sexual abuse (Lamb & Garretson, 2003), the results

showed a non-significant influence of interviewer gender on the total number of questions asked ($t(22) = - .742$, $p = .47$, Cohen's $d = .32$) and on the total number of disclosure answers provided by the witnesses ($t(22) = - .817$, $p = .42$, Cohen's $d = .35$). Concerning the interviewer's rank, analyses were performed with the two most represented ranks in our sample—ranks that also reflect a real hierarchical difference in the organization of the French military police, that is: Staff sergeant ($n = 17$) and Warrant officers ($n = 4$). Results showed a non-significant influence of interviewer's rank on the total number of questions asked ($t(19) = 1.302$, $p = .209$, $d = 0.723$) and on the total number of disclosure answers provided by the witnesses ($t(19) = 1.541$, $p = .140$, $d = 0.856$). The same analysis was made on the wearing of a uniform during the interviews since previous research has showed a negative impact of this factor on the amount of information (accurate and inaccurate) reported by 6- to 8-year-old children (Powell & Croft, 2000). The results did not show any significant influence of uniform on the total number of questions asked by the interviewers ($t(21) = 1.289$, $p = .21$, Cohen's $d = .55$) or on the total number of disclosure answers provided by the witnesses($t(22) = .567$, $p = .32$, Cohen's $d = .44$). Additionally, two "2(Uniform: yes, not) $\times$ 3(Age group: under 7, 7-10, 11 years and older)" Kruskal-Wallis tests were performed and showed a non-significant difference on the total number of questions asked ($\chi^2 (5, N = 23) = 7.238$, $p = .20$, Cramer's V $= 0.25$) and on the total number of disclosures ($\chi^2 (5, N = 23) = 7.061$, $p = .22$, Cramer's, V $= 0.25$).Therefore, the "Interviewer gender", "Interviewer rank" and the "Uniform" variables were not included in any of the subsequent reported statistical tests. All the variables entered in the non-parametric analyses were normally distributed. However, the variance of the *multiple questions* variable was significantly different. This was therefore transformed (by computing its square root) to make its variance was homogenous ($p = .188$).

Statistical Analyses

Three statistical analyses were run in order to better examine the quality of investigative interviews with child witnesses. For the three age groups (under 7 years, from 7 to 10 years, and older than 11 years), we analyzed the use of the four-phase interview structure, the questioning techniques as a function of the age of the witness, and the association between the interviewers' utterances and the children's answers.

First, adherence to the four-phase structure was analyzed with Friedman tests. Wilcoxon tests were computed on significant main effects. The relationship between the use of the four interview phases and the age of the witnesses was examined using χ^2 tests. Second, the influence of the witnesses' age on the nature of the questions asked by the interviewers was examined with a series of ANOVAs. LSD post hoc tests were then computed on significant main effects (using an adjusted alpha of .017). Third, possible associations between the interviewers' utterances and the children's answers were examined using χ^2 tests for all three age groups of witnesses. Specific results for all the question types are presented in Tables 3 to 6. However, in the interest of brevity and clarity, the analyses presented below focuses on appropriate open-ended and probing questions, as well as on inappropriate closed and leading questions.

Use of the four-phase structure as a function of the age of the victim

The analysis revealed significant differences in the use of the four interview phases (χ^2 (3, $N = 24$) = 20.250, $p < .001$, $r = 0.92$), with *free recall* being significantly less used than *introduction* ($Z = -3.000$, $p = .003$, $r = -0.61$), *questioning* ($Z = -3.000$, $p = .003$, $r = -0.61$), and *closure* ($Z = -2.121$, $p = .034$, $r = -0.43$). There was a non-significant difference between the use of the other three phases, $ps \geq .076$. When we looked at the association between the age of the witnesses and the use of the four-phase structure, we did not find any significant difference in the use of the introduction (used in all of the studied interviews), free recall (χ^2 (2, $N = 24$) = 1.705, $p = .43$, Cramer's, V = 0.19) and the questioning phase (used in all of the

studied interviews). However, while the closure phase was systematically used with children in the "7-10 years" and the "11 years and older" groups, it was used to a significantly lesser extent with the "under 7 years" group, χ^2 (2, $N = 24$) = 8.327, $p = .032$, Cramer's, V = 0.42. To summarize, whatever the age of the witness, the free recall was absent in about 30% of the interviews examined. This suggests that evidence was mainly provided by the witnesses' responses to the interviewers' questions.

Which question types for which age of child?

The question types used throughout the interviews differed significantly from one another not only across each interview as a whole, but also within each specific interview phase: introduction, free recall, questioning, and closure (see Table 3 for observed frequencies, confidence intervals, degrees of freedom, Ns, χ^2, p-values, and effect sizes). From the opening of the interviews, the interviewers used more inappropriate closed questions, probing questions, and leading questions. Open questions as recommended in international guidance were underused by the interviewers. However, when we considered the interviews in which free recall was present ($n = 17$), we observed that the interviewers mostly used appropriate questions, and more specifically open questions and probing questions, to gather evidence from the child victims.

[Table 3 here]

We then examined whether or not the use of the abovementioned questions was linked to the age of the child. We found for example that closed, probing, and leading questions were over-represented in interviews with very young children, young children, and 11 years and older witnesses (see Table 4 for average observed frequencies, degrees of freedom, F values, p values and effect sizes). Are there different dynamics behind this common observation for all witnesses? If so, do they differ regarding the age of the witnesses *per se* or their abilities to provide detailed responses? The two next sub-sections will attempt to answer these questions.

[Table 4 here]

How did the child witnesses answer the interviewers' questions?

We examined a simple two-code chain (Interviewer → Child) in order to analyze the association between the interviewer's question and the child's answer. This analysis was performed for the three age groups independently (see Table 5 for confidence intervals, degrees of freedom, Ns, χ^2s, p-values, and effect sizes).

Children in the "under 7 years" group. In line with the rule holding that absolute values of standardized residuals greater than two yield significant χ^2 values (Hinkle, Wiersman, & Jurs, 1998; for an application, see Lindholm et al., 2014), the results showed that both "no answer" and "extended task-related" responses were significantly more frequent answers to *appropriate open-ended* questions, while "agreements" were less frequent due to the nature of the question. Additionally, while "request-confirming" answers were significantly more frequent with *probing* questions, "agreements" and "disagreements" were less frequently elicited by these question types. The opposite pattern was observed for *inappropriate closed* questions. These elicited significantly more "agreements" and "disagreements" but significantly fewer "request-confirming" answers, as well as fewer "extended task-related" answers and "no answers". Finally, "agreements" and "disagreements" were both significantly more frequent in response to *leading* questions. At the same time, "request-confirming" and "extended task-related" answers were significantly less frequent in response to this question type. In summary, children younger than 7 years formulated their answers only in response to appropriate open questions. However, this question type was also associated with more silences. Probing questions allowed these children to describe the event, albeit with fewer details than open questions. Finally, inappropriate closed and leading questions only elicited yes-no answers.

Children in the "7-10 years" group. The results indicated that significantly more "extended task-related" as well as "no answer" and "evasive answers" were elicited with *appropriate open-ended* questions. In turn, this question type was significantly less associated with "request-confirming" answers, "agreements" and "disagreements". The children produced significantly more "request-confirming" and, to an even greater extent, "extended task-related" answers in response to *probing* questions. At the same time, they produced significantly fewer "agreements", "disagreements" or "evasive answers" in response to these questions. *Inappropriate closed* questions led to a significant increase in "agreements" and "disagreements" and a significant decrease in "request-confirming", "no answer" and "evasive" answers. As previously observed with children in the "under 7 years" group, the children aged between 7 and 10 years produced significantly more "agreements" and "disagreements" and significantly fewer "request-confirming" and "extended task-related" responses in response to *leading* questions. In summary, the children aged from 7 to 10 years produced more extensive answers when answering appropriate open-ended and probing questions. Unlike the open questions, probing questions were not associated with evasive answers or with a failure to answer. Finally, these children produced more "yes" or "no" answers in response to inappropriate closed and leading questions.

Children in the "11 years and older" group. Interviewers' *appropriate open-ended* questions elicited significantly more "extended task-related" answers, as did the *probing* questions. The latter were also significantly more highly associated with "request-confirming", and "evasive" answers and were also associated with fewer "agreements" and "disagreements". *Inappropriate closed* questions elicited significantly more "agreements" and "disagreements", and fewer "request-confirming" and "extended task-related" answers. Finally, as observed in the other two age groups, "agreements" and "disagreements" were significantly more frequent in response to *leading* questions, while "request-confirming" and

"evasive" answers were significantly less frequent in response to this question type. In summary, open questions gave the 11 years and older witnesses the possibility of producing a more detailed description of the offence. The same benefit was observed with probing questions, even if they were also associated with a larger number of evasive answers. Finally, the 11 years and older witnesses produced more "yes" or "no" answers in response to inappropriate closed and leading questions.

To summarize, subsequent extended task-related answers to open questions were more often observed with the "7-10 years" and the "11 years and older" groups. The overuse of closed and leading questions (cf. p. 16) might thus reflect an attempt to help the youngest to provide information they could not give on their own. The final part of this section will expand on this. We will also try to identify why such non recommended utterances were used with 11 years and older witnesses, who are able to spontaneously give detailed responses.

[Table 5 here]

How did the interviewers adapt their questions to the children's answers?

We examined the simple two-code chain Child → Interviewer. This analysis was performed for the three age groups independently (see Table 6 for confidence intervals, degrees of freedom, Ns, χ^2s, p-values, and effect sizes).

Children in the "under 7 years" group. No significant associations were observed between the children's *extended task-related* answers and the interviewers' subsequent question type, despite a significant difference in the use of questions following this kind of answer. The *request-confirming* answers were significantly more associated with subsequent "inappropriate closed" and "probing" questions, and less so by "closed" or "leading" questions. *Agreements* were significantly more likely to be followed by "appropriate open" and "probing" questions. *Disagreements* were significantly more likely to be followed by "appropriate closed-ended" questions. *No answers* were significantly more associated with

subsequent "leading" questions, and less so with "inappropriate closed" and "probing" questions. Finally, the interviewers used significantly more "appropriate closed" questions immediately after *evasive* responses. In summary, interviewers used more appropriate (open-ended, probing) questions after a request-confirming answer or an agreement. Additionally, they more often used appropriate closed questions after a disagreement. Nevertheless, inappropriate question types (closed or leading questions) were also the most frequent utterances to follow a request-confirming answer or an absence of answer.

Children in the "7-10 years" group. No significant associations emerged between the children's *extended task-related* answers and the interviewers' subsequent question types, despite a significant difference in the use of questions following this kind of answer. *Request-confirming* answers were significantly more likely to be followed by "inappropriate closed" and "probing" questions. No significant associations emerged between the children's *agreements* and the interviewers' subsequent question types, despite a significant difference in the use of questions following this kind of answer. *Disagreements* were significantly more associated with "leading" and "inappropriate closed" questions, and less so with "probing" questions. The interviewers used significantly more "appropriate open-ended" questions after an absence of response from the child witnesses (*no answers*). Finally, *evasive* answers were significantly more likely to be followed by "leading" questions. In summary, the interviewers used more subsequent appropriate (open-ended, probing) questions after a silence (no answer) or a request-confirming answer. However, these answers also more often led the interviewers to use inappropriate closed questions. Disagreement and evasive answers by the children were more likely to be followed by an inappropriate question type (closed or leading questions).

Children in the "11 years and older" group. The interviewers used significantly more "leading" questions and fewer "inappropriate closed" questions after the children had answered the previous question with an *extended task-related* response. *Request-confirming*

20

answers were significantly more often associated with "opinions and statements". *Agreements* were significantly less likely to be followed by "leading" questions. However, *disagreements* were significantly more associated with "inappropriate" and "appropriate closed" questions and less associated with "probing" questions. While a significant decrease in "probing" questions was observed after an absence of answer, a significant increase in interviewers' "opinions and statements" was observed immediately after this kind of non-disclosure answer *(no answer)*. Finally, *evasive* responses elicited significantly more "forced choice" questions . In summary, the interviewers used few subsequent appropriate (open-ended, probing) questions to explore the statements of 11 years and older witnesses in more detail. Indeed, even when the children provided an extended task-related answer or a request-confirming answer, the interviewers pursued the interview with closed questions or comments.

To summarize, whether witnesses gave a potentially informative response or not, we observed the same subsequent strategies among the interviewers: asking a leading or a closed question.

[Table 6 here]

Discussion

In this archival study, we explored for the first time in France the context in which interviewers used appropriate (vs. inappropriate) questions for interviewing three groups of young witnesses, that is: under 7 years old, 7 to 10-year-old children, and 11-year-old and older children. Two important findings emerged. First, the free recall phase and associated use of appropriate open-ended questions were underused in the interviews overall, and this occurred whatever the age of the witness. Secondly, the tendency to rely on closed and leading questions occurred whatever the age of witnesses. The two dynamics underlying this use may be associated more with the age of the witnesses *per se* than with their non-

responsiveness. These findings and their implication for current child interview training in France will be reviewed below.

The overall quality of child interviews in France

The results concerning the adherence to the four-phase structure and the appropriate questioning of child witnesses were consistent with prior research. In particular, the underuse of the free recall compared to the introduction and the questioning phases was similar to the majority of international research examining real-life forensic interviews with children (e.g. Canada: Luther et al., 2014; Estonia: Kask, 2012; New Zealand: Wolfman et al., 2016; Sweden: Lindholm et al., 2014). In addition, the interviewers relied heavily on inappropriate questions to elicit and support the statements. Specifically, in total, only 2.2% of all the interviewers' utterances were appropriate open-ended questions (vs. more than 50% of inappropriate closed and leading questions).

The free recall phase was mainly initiated with appropriate open-ended and probing questions (which made up 54% of the interviewers' questions). However, almost 30% of the inappropriate closed and leading questions were also used in this phase. Moreover, the use of inappropriate questions, including leading questions, was found from the introduction of the interview right through until their closure (see Sternberg et al., 2001).

The predominant use of inappropriate (closed and leading) questions could reflect everyday conversational roles on the part of interviewers who have not been appropriately trained in the conduct of child interviews (Korkman et al., 2008). The child interview-related recommendations might also be difficult to apply and require interviewers with high-level skills and knowledge (Griffiths, Milne, & Cherryman, 2011) on which current French child interview training needs to focus. Finally, the non-responsiveness of the young witnesses might have increased the perceived difficulty of applying some of the recommendations. Our data enabled us to explore only the third hypothesis. To do this, we examined the outputs of

the analyses conducted on: (i) The relationship between the age of the witness and the question types used, and; (ii) the dynamics of verbal exchanges during the investigative interviews.

First of all, it seems important to note that the non-compliance with the international recommendations during the application of the four-phase structure, and in particular the weakness of the introduction phase, may also have played a key role in the non-responsiveness of the children. Indeed, when looking deeper at the opening phase, we observed that: (i) Interviewers mostly used it to introduce the reason of the interview, but; (ii) they did not transfer the control of the interview to the witnesses, did not explain to them the "I don't know/understand" instruction, and did not determine their witnesses' understanding of the distinction between truth and lies for example. These components of the introduction were absent from the 24 studied interviews. The omission of important information during the introductory phase could be explained by the failure of interviewers to acknowledge the differences that exist between an every-day life conversation and a formal investigative interview with its specific requirements (Davies, Bull, & Milne, 2016). Another explanation is the possibility that they took place off-camera. Nevertheless, even if it was the case, their benefits would have been greatly limited by the parallel use of inappropriate closed and leading questions. These types of question might have negatively impacted upon the dynamic of the interviews by reinforcing the idea that the interviewer controlled the interview and that short answers were required. Such beliefs are known not only to reduce the quantity of the reported details (e.g. Verkampt, Ginet, & Colomb, 2014), but also their accuracy and the children's resistance to misleading questions (e.g. Mulder & Vrij, 1996).

Children's non-responsiveness and the use of inappropriate closed and leading questions

The majority of our analyses showed no significant relationship between the age of the witness and the question types used. An exception was observed for the use of multiple

questions (more often used with witnesses under 7 years) and the giving of

opinions/statements (more often used with 11 years and older witnesses, cf. Table 4).

Therefore, contrary to Sternberg et al.'s (2001) findings, the questions did not become less

specific as the age of the witnesses increased along with their ability to retrieve and report

experiences. For instance, leading questions were asked as frequently with the very young

witnesses as they were with the 11 years and older witnesses (i.e. 19.3% and 16.9% of all the

interviewers' utterances, respectively; see Kask, 2012, for contrasting results). Indeed, open-

ended questions did not account for more than about 3% of all the interviewers' utterances,

whatever the age of the witness.

Despite this, the analyses performed on the "Interviewers' utterances → Children's

answers" sequences indicated that the witnesses were generally very responsive to the

interviewers' questions (between 84.3% and 90.8% of responses across the age groups; see

Wolfman et al., 2016, for similar results). However, in line with the developmental literature,

their answers questions were associated with their age (Brown & Lamb, 2015). Indeed,

although some of the witnesses under 7 years of age and those aged 7 to 10 years were able to

provide a detailed answer to an open-ended question, some of them provided only an evasive

answer or did not respond at all. With these two groups of children, more answers were thus

given after probing questions (with 84% and 90.7% of responses, respectively) than after

open-ended questions (66.7% and 74.4% of responses, respectively). By contrast, the 11 years

and older witnesses were equally responsive to open-ended questions (91.7% of responses vs.

8.3% of non-responses) and probing questions (91.1% of responses vs. 8.9% of non-

responses). However, a reliance on probing questions inhibited the possibility for older

children to give more extensive answers (see also Keselman et al., 2008). Finally,

inappropriate closed and leading questions were associated with more "yes" and "no" answers

irrespective of the age of the witnesses; producing few details of importance to the

investigation. To better analyze the reliance on these two question types, we next turn to the question: Did inappropriate closed and leading questions follow a silence or any other indication of the witnesses' difficulty in providing information?

When we looked at the information provided by the witnesses prior to the interviewers' questions, two distinct patterns of results emerge, namely one for the two youngest groups of witnesses and one for the 11 years and older witnesses. The non-responsiveness of the witnesses in the two youngest groups (silence, evasive answer) was associated with the giving of a subsequent leading question. However, our results also showed that request-confirming answers were also more often associated with a subsequent leading question in the witnesses under 7 years and with a closed question in the witnesses aged 7 to 10 years. Equally, disagreements were more often associated with a subsequent closed question in the 7 to 10 year-old witnesses. Taken together, these findings partially support the idea that non-responsiveness on the part of very young and young children leads interviewers to use inappropriate (closed and leading) questions. Indeed, interviewers also rely on this type of inappropriate question with responsive young and very young witnesses who attempt to engage with the question asked, even if this takes the form of request-confirming responses (see Korkman et al., 2008). One explanation could be that interviewers might perceive closed and leading questions as being equally efficient as probing questions in helping children to formulate their answers. Alternatively, the distinction between appropriate probing, closed, and leading questions might not be sufficiently clear to them. Including more tests on question types and role-playing on this issue in training events *early* in professionals' careers would help them to improve their skill in eliciting more judicially relevant information in child interviews without relying on very specific questions, which sometimes risk contaminating the child's responses (see Powell et al., 2015).

With 11 years and older witnesses, our analyses revealed different dynamics. For example, interviewers asked inappropriate closed questions immediately after a disagreement (about 39% of the subsequent utterances) or an extended task-related answer (25.1% of the interviewers' subsequent utterances). Similarly, leading questions were more often associated with a prior extended task-related answer (about 21% of the subsequent utterances) or an agreement (10.3% of the interviewers' subsequent utterances). Interestingly, an absence of answer (silence) on the part of these witnesses more often resulted in a subsequent probing question or an opinion or statement (making up 48.6% of the interviewers' subsequent utterances). Therefore, even when the 11 years and older witnesses were informative, the interviewers failed to pursue the positive dynamic of the interview with neutral and recommended utterances. Instead, they relied on inappropriate questions known to elicit shorter answers that may be less accurate and which risked contaminating the witnesses' statements. One weakness of our study was that neither the quality of the witnesses' verbal behaviours nor the non-verbal behaviours of either the witness or the interviewer were analyzed. In consequence, some specific difficulties associated with older witnesses might have influenced the interviewers' questioning practices. Further research is therefore needed to explore the interviewers' strategies that underlie the use of inappropriate closed questions, and above all of leading questions, with responsive (pre)adolescents. In addition, because this study was conducted on a small number of interviews conducted by interviewers for whom the training-related information were not available, it must also be acknowledged that training-related factors might have influenced the conduct of the interviews (see Powell, Fisher, & Wright, 2005; Powell, Hughes-Scholes, Smith, & Sharman, 2015). Research using a larger interview sample and conducted by trained interviewers is thus to be encouraged.

Two other limitations to our study can be discussed. First, while we performed our analysis on the question types used by the investigators, no qualitative analysis was done in

order to order to determine whether the vocabulary used by the investigators or the length of each question, for example, were appropriated to the age of the witnesses. As such analysis might require a joint analysis of child non-verbal and verbal behaviours, further real-life forensic interviews analyses should also focus on content analyses to shed new light on interviewing practices with child and adolescent witness. Second, we used the Griffiths Question Map to code the type of questions asked by interviewers. It must be acknowledge that such an approach may be the subject of debate (see Oxburgh, Myklebust & Grant, 2010), the GQM more accounting for the classification of questions than their function within an interview. Although our inter-rater reliability values were satisfactory and our results on the proportions of questions asked are consistent with the literature using other tools (e.g. Korkman et al., 2006), one cannot fully rule out the possibility that an analysis of our interviews sample with other coding tools might put in light different results.

Conclusion

Our findings are consistent with past and current child interview practices of trained interviewers, as assessed in countries other than France. Results of this context-driven analysis (i.e. study of the four-phase structure and of the dynamics of the verbal exchanges) can thus improve our understanding of the overuse of closed and leading questions in child interviews.

The non-compliance with the recommendations formulated for each of the interview phases might have dramatically reduced the likelihood that the witnesses would feel at ease, as well as their readiness to tell an unknown adult what had happened to them. The results also enable us to identify two prevailing dynamics behind the inappropriate questioning style of interviewers, which may be associated to the age of the witnesses (under 10 years old and above 10 years old). Indeed, the use of inappropriate closed and leading questions was not systematically associated with the failure of the youngest witnesses to respond to an earlier

question but was, more worryingly, associated with informative prior answers on the part of the oldest witnesses. Further research should thus explore other factors that may be associated with the quality of child-centered interviews, such as the degree of control interviewers like to exercise when conducting interviews with child witnesses (Griffiths et al., 2011).

In terms of child interview training, these findings tend to support the main factors within the Framework of Investigative Transformation for investigative interviewing (Griffith & Milne, 2019). For example, more practical tests and role-playing within training is recommended in order to help trainees to better discriminate closed, specific and (mis)leading questions. The personal perceptions (bias) trainees could have about child vs. adolescent witnesses could also be discussed. Indeed, wrong beliefs could be associated with interviewers' expectancies that might be implicitly or explicitly communicated via the questions type they used to interview young witnesses. Therefore, focus group discussions allowing trainees to deconstruct these beliefs would be beneficial.

References

Abbe, A., & Brandon, S.E. (2014) Building and maintaining rapport in investigative interviews. *Police Practice and Research: An International Journal, 15(3)*, 207-220. Doi: 10.1080/15614263.2013.827835

Alexander, K. W, Quas, J.A., Goodman, G.S, Ghetti, S., Edelstein, R.S, Redlich, A.D, Gordon, I. M., & Jones, D.P.H. (2005). Traumatic impact predicts long-term memory for documented child sexual abuse. *Psychological science, 16(1)*, 33-40.

Ansado, J., Chiasson, V. & Beauchamp, M. H. (2014). Croissance cérébrale et Neurodéveloppement à l'adolescence. Dans M. Claes, & L. Lannegrand-Willems (Éds.), *La psychologie de l'adolescence* (pp. 45-71). Les Presses de l'Université de Montréal, Montreal, QC.

Bakeman, R., & Quera, V. (2011). *Sequential Analysis and Observational Methods for the Behavioral Sciences*. Cambridge, UK: Cambridge University

Berthet, G., & Monnot, C. (2007). L'audition du mineur victime. [Child investigative interview]. *Enfances & Psy, 3(36)*, 80-92. Doi: 10.3917/ep.036.0080

Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science, 11*, 164-168.

Brainerd, C. J., & Reyna, V. F. (2005). *The Science of False Memory*. New York : Oxford University Press.

Brown, R. (1973). *A first Language*, Harvard University Press, Cambridge, MA

Brown, D. A.,& Lamb, M. E. (2015). Can children be useful witnesses? It depends how they are questioned. *Child Development Perspectives, 9,* 250-255. Doi:10.1111/cdep.12142

Carol, R.N. & Schreiber Compo, N. (2015). Other People: A Child's Age Predicts a Source's Effect on Memory. *Legal and Criminological Psychology, 22*(1), 74-87. Doi: 10.1111/lcrp.12078

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., Papierno, P. B., & Kulkofksy, S. (2007). Representational constraints on children's suggestibility. *Psychological Science, 18(6)*, 503-509. Doi: 10.1111/j.1467-9280.2007.01930.x

Cyr, M., & Lamb, M.E. (2009). Assessing the effectiveness of the NICHD investigative interview protocol when interviewing French-speaking alleged victims of child sexual abuse in Quebec. *Child Abuse & Neglect, 33*, 257-268. Doi: 10.1016/j.chiabu.2008.04.002

Davies, G., Bull, R., & Milne, R. (2016). Analyzing and improving the testimony of vulnerable witnesses interviewed under the 'Achieving Best Evidence' protocol. In R. Radcliffe, G.H. Gudjonsson, A. Heaton-Armonstrong, & D. Wolchover (Eds.), *Witness testimony in sexual cases* (pp. 207-219). Oxford: Oxford University Press.

Dewhurst, S. A., & Robinson, C. A. (2004). False memories in children: Evidence for a shift from phonological to semantic associations. *Psychological Science*, *15*, 782–786.

Dodier, O., & Denault, V. (2018). The Griffiths Question Map: A forensic tool for expert witnesses' assessments of witnesses and victims' statements. *Journal of Forensic Sciences, 63(1)*, 266-274. doi. 10.1111/1556-4029.13477

Fivush, R. (1993). Developmental perspectives on autobiographical recall. In G. S. Goodman & B.L. Bottoms (Eds.), *Child Victims, Child Witnesses* (pp. 1-24). New York: The Guilford Press

Fivush, R. (2010). The development of autobiographical memory. *Annual Review of Psychology, 62*, 559-582.doi: 10.1146/annurev.psych.121208.131702.

Fivush, R., & Hamond, N.R. (1990). Autobiographical memory across the preschool years: Towards reconceptualizing in young childhood amnesia. In R. Fivush, & J.A. Hudson

(Eds.), *Knowing and remembering in young children* (pp. 223-248). New York: Cambridge University Press.

Gabbert, F., & Hope, L. (2017). Suggestibility in the courtroom: How memory can be distorted during the investigative and legal process. In H. Otgaar, & M. Howe (Eds.), *Finding the truth in the courtroom: Dealing with deception, lies, and memories* (pp. 31-58). New York: Oxford University Press.

Garven, S., Wood, J.M., & Malpass, R.S. (2000). Allegations of wrongdoing: The effects of reinforcement on children's mundane and fantastic claims. *Journal of Applied Psychology, 85*(1), 38-49.

Gilstrap, L. L., & Papierno, P. B. (2004). Is the cart pushing the horse ? The effects of child characteristics on children's and adults' interview behaviours. *Applied Cognitive Psychology, 18(8)*, 1059-1078. Doi: 10/1002/acp.1072

Griffiths, A., & Milne, R. (2006). Will it all end in tiers? In T. Williamson (Ed.), *Investigative Interviewing; Research rights and regulation* (pp167-189). Cullompton: Willan.

Griffiths, A., & Milne, R. (2019). *The Psychology of Criminal Investigation: From Theory to Practice*. Issues in Forensic Psychology. Oxford: Routledge.

Griffiths, A., Milne, R., & Cherryman, J. (2011). A question of control? The formulation of suspect and witness interview question strategies by advanced interviewers. *International Journal of Police Science & Management, 13(3)*, 255-267. Doi:10.1350/ijps.2011.13.3.219

Grisso, T., Steinberg, L., Cauffman, E., Scott, E., Graham, S., …Schwartz, R. (2003). Juveniles' competence to stand trial: A comparison of adolescents' and adults' capacities as trial defendants. *Law and Human Behavior, 27*(4), 333-363.

Gudjonsson, G. H., & Singh, K. (1984). Interrogative suggestibility and delinquent boys: An empirical validation study. *Personality and Individual Differences*, *5*, 425-430.

Hershkowitz, I., Orbach, Y., Lamb, M. E., Sternberg, K. J., & Horowitz, D. (2006). Dynamics of forensic interviews with suspected abuse victims who do not disclose abuse. *Child Abuse & Neglect, 30,* 753-769. Doi: 10.1016/j.chiabu.2005.10.016

Hinkle, D. E., Wiersman, W., & Jurs, S. G. (1998). *Applied statistics for the behavioral sciences* (4th ed.). Boston, MA: Houghton Mifflin.

Home Office. (2011). *Achieving best evidence in criminal proceedings: Guidance for vulnerable and intimidated witnesses, including children*. London, UK: Home Office Communications Directorate.

Howe, M.L., Wimmer, M. C., Gagnon, N., & Plumpton, S. (2009). An associative-activation theory of children's and adults' memory illusions. *Journal of Memory and Language, 60*, 229-251. Doi: 10.1016/j.jml.2008.10.002

Hughes-Scholes, C.H., & Powell, M.B. (2013). Techniques used by investigative interviewers to elicit disclosures of abuse from child witnesses: A critique. *Police Practice and Research: An International Journal, 14*(1), 45-52. Doi: 10.1080/15614263.2012.68071

Jack, F., Leov, J., & Zajac. R. (2014). Age-related differences in the free-recall accounts of child, adolescent, and adult witnesses. *Applied Cognitive Psychology, 28*(1), 30-38. Doi: 10.1002/acp.2951

Kask, K. (2012). Dynamics in using different question types in Estonian police interviews of children. *Applied Cognitive Psychology*, *26*, 324-329.

Keselman, O., Cederborg, A-C., Lamb, M. E., & Dahlström, O. (2010). Asylum-seeking minors in interpreter-mediated interviews : What do they say and what happens to their responses? *Child and Family Social work*, *15*, 325-334.

Keselman, O., Cederborg, A. C., Lamb, M. E., & Dahlström, Ö. (2008) Mediated communication with minors in asylum-seeking hearings. *Journal of Refugee Studies*, *21(1)*, 104–115. Doi: 10.1093/jrs/fem051

Koriat, A., Goldsmith, M., Schneider, W., & Nakash-Dura, M. (2001). The credibility of children's testimony: Can children control the accuracy of their memory reports? *Journal of Experimental Child Psychology, 79*, 405–437.

Korkman J., Santilla 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., Santilla P., Westeraker, M. & Sandnabba, N. K. (2008). Interviewing techniques and follow-up questions in child sexual abuse interview. *European Journal of Developmental Psychology*, *5(1)*, 108-128.

Lamb, M. E., & Garretson, M. E. (2003). The effects of interviewer gender and child gender of the informativeness of alleged child sexual abuse victims in forensic interviews. *Law and Human Behavior, 27*, 157-171. doi: 10.1023/A:1022595129689.

Lamb, M.E., Orbach, Y., Hershowitz, I., Esplin, P.W., & Horowitz, D. (2007). Structured forensic interview protocols improve the quality and informativeness of investigative interviews with children: A review of research using the NICHD Investigative Interview Protocol. *Child Abuse & Neglect, 31*, 1201-1231. Doi: 10.1016/j.chiabu.2007.03.021.

Leander, L. (2010). Police interviews with child sexual abuse victims: patterns of reporting, avoidance and denial. *Child Abuse & Neglect, 34(3)*, 192-205. Doi: 10.1016/j.chiabu.2009.09.11.

Lindholm, J., Cederborg, A-C., & Alm, C. (2014). Adolescent girls exploited in the sex trade: informativeness and evasiveness in investigative interviews. *Police Practice and Research: An International Journal, 16(3)*, 197-210. Doi: 10.1080/15614263.2014.880839

Luther, K., Snook, B., Barron, T., & Lamb, M. E. (2015). Child interviewing practices in Canada: A box Score from field observations. *Journal of Police and Criminal Psychology, 30,* 204-212. Doi: 10.1007/s11896-014-9149-y

Miller, P. H. (1990). The development of strategies of selective attention. In D.F. Bjorklund (Ed.), *Children's strategies: Contemporary views of cognitive development* (pp. 157-184). Hillsdale, NJ : Erlbaum.

Mulder, M. R., & Vrij, A. (1996). Explaining conversation rules to children: An intervention study to facilitate children's accurate responses. *Child Abuse & Neglect, 20*, 623–631.

Myklebust, T. (2017). The Nordic model of handling children's testimonies. In S. Johansson, K. Stefansen, E. Baketeig & A. Kaldal (Eds.), *Collaborating against child abuse: Exploring the Nordic Barnahus Model* (pp. 97-120). Palgrave Macmillan, Cham. Doi: 10.1007/978-3-319-58388-4_5

Oxburgh, G.E., Myklebust, T., & Grant, T. (2010). 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, 17(1)*, 45-66. Doi: 10.1558/ijsll.v17i1.45

Philipps, E., Oxburgh, G. E., Gavin, A., & Myklebust, T. (2012). Investigative interviews with victims of child sexual abuse: The relationship between question type and investigation relevant information. *Journal of Police and Criminal Psychology, 27(1)*, 45-54. doi/10.1007/s11896-011-9093-z

Powell, M.B., & Croft, C. (2000). The effect of uniform and prior knowledge on children's event reports and disclosure secrets. *Journal of Police and Criminal Psychology, 15(1)*, 27-40. Doi: 10.1007/BF02802655.

Powell, M., Fisher, R., & Wright, R. (2005). Investigative Interviewing. In N. Brewer, & K. Williams (Eds.), *Psychology and law: An empirical perspective* (pp. 11-42). New York: Guilford Press.

Powell, M. B., Hughes-Scholes, C. H., Smith, R., & Sharman, S. J. (2015). The relationship between investigative interviewing experience and open-ended question usage. *Police*

Practice and Research: An international Journal, 15(4), 283-292. Doi: 10.1080/15614263.2012.704170

Powell, M.B., & Guadagno, B. (2008). An examination of the limitations in investigative interviewer's use of open-ended questions. *Psychiatry, Psychology, and Law, 15*(3), 382-395. Doi: 10.1080/13218710802101621

Richardson, G., Gudjonsson, G. H., & Kelly, T. P. (1995). Interrogative suggestibility in an adolescent forensic population. *Journal of Adolescence*, *18*, 211–216.

Rossi, A. F., Pessoa, L., Desimone, R., & Ungerleider, L. G. (2009). The prefrontal cortex and the executive control of attention. *Experimental Brain Research*, *192*(3), 489-497.

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.

Tickle-Degen, L., & Rosenthal, R. (1990). The nature of rapport and its nonverbal correlates. *Psychological Inquiry, 1*, 285-293. Doi: 10.1207/s15327965pli0104_1

Verkampt, F., Ginet, M., & Colomb, C. (2014). The influence of social instructions on the effectiveness of a cognitive interview used with very young child witnesses. *European Review of Applied Psychology*, *64*, 323-333.

Walker, A.G. (1999). *Handbook on questioning children. A linguistic perspective* (2Ed.), American Bar Association : Center on Children and the Law, Washington, DC.

Westcott, H. & Kynan, S. (2006). Interviewer practice in investigative interviews for suspect child sexual abuse. *Psychology, Crime & Law*, *12*, 367-382.

Wolfman, M., Brown, D., & Jose, P. (2016). Talking past each other: Interview and child verbal exchanges in forensic interviews. *Law and Human Behavior*, *40(2)*, 107-117. Doi: 10.1037/lhb0000171.

Table 1. Description of the sample of interviews

Variables	Frequency	%	CI 95%	Rank	Mean (SD)
Interviewers					
Chief warrant officer	1	4.2	[0.1, 21.1]		
Warrant officers	4	16.7	[4.7, 37.4]		
Staff sergeants	17	70.8	[48.9, 87.4]		
Officers	2	8.3	[1.0, 27.0]		
Uniform	16	66.7	[44.7, 84.4]		
Without uniform	7	29.7	[12.6, 51.1]		
Clothes unspecified	1	4.2	[0.1, 21.1]		
Witnesses	24	100			
Age				3 - 17 years old	8.38 (3.85)
Male	14	58.3	[36.6, 77.9]		
Female	10	41.7	[22.1, 63.4]		
< 7 years old	8	33.3	[15.6, 55.3]		
7 - 10 years old	9	37.5	[18.8, 59.4]		
> 10 years old	7	29.2	[12.6, 51.1]		
Interview dates					
2007	1	4.2	[0.1, 21.1]		
2009	4	16.7	[4.7, 37.4]		
2010	14	58.3	[36.6, 77.9]		
2011	5	20.8	[7.1, 42.2]		
Lenght of the interview				10 - 60 min.	28 min. (16 min.)
Presence of the interviewer					
Video room	14	58.3	[36.6, 77.9]		
Another room	9	37.5	[18.8, 59.4]		
Interview room	1	4.2	[0.1, 21.1]		
Number of interview per child					
Once	21	87.5	[67.6, 97.3]		
Twice	3	12.5	[2.7, 32.4]		
Disclosure of abuse					
Yes	24	100.0	[85.8, 100.0]*		
Case follow-up					
No follow-up (lack of evidence)	24	100.0	[85.8, 100.0]*		
Type of abuse					
non-penetrative sexual abuse	13	54.2	[32.8, 74.4]		
penetrative sexual abuse	6	25.0	[9.8, 46.7]		
physical violence	2	8.3	[1.0, 27.0]		
sexual abuse	3	12.5	[2.7, 32.4]		
Status of the interviewee					
Victim	23	95.8	[78.9, 99.9]		
Eyewitness	1	4.2	[0.1, 21.1]		

Perpatrators' characteristics			
Father	9	37.5	[18.8, 59.4]
Member of the family	4	16.7	[4.7, 37.4]
Known but unrelated to the family	10	41.7	[22.1, 63.4]
Non-identified	1	4.2	[0.1, 21.1]

*One-sided 97.5% CI.

Table 2. The question types according to the Griffiths Question Map.

Group	Question types	Definition	Example
Appropriate	Open-ended	Allow a full range of response	Tell me everything about that?
	Probing/specific	Defined as more intrusive and requiring a more specific answer, usually commencing with the active words "who", "what", "why", "where", "when", "which", or "how"	When did that happen?
	Closed	Used at the conclusion of a topic where open and probing questions have been exhausted	Can you tell me where it was?
Inappropriate	Open-ended	When the question interrupts the speech of the witness	… wait wait… What happened at this point?
	Closed	Used at the wrong point in the interview	Do you know this man?
	Multiple	Constitute a number of subquestions asked at once	Did he do it several times? Where were you? How did you feel?
	Forced choice	Only offer the interviewee a limited number of possible responses	This woman, was she red-headed or blonde?
	Opinion/ Statement	Defined as posing an opinion or putting statements to an interviewee as opposed to asking a question	I think that this man is not nice to you… You did not tell me that, the first time…
	Leading	Suggest an answer in formal content to an interviewee	Were they an alcoholic?

Table 3. Frequency (%) of interviewers' utterances during the interviews and within each phase (confidence intervals are indicated in brackets)

Utterance types	Appropriate questions				Inappropriate questions							df	N	χ^2	p	r
	Open-ended	Probing	Closed	Total	Open-ended	Closed	Multiple	Forced choice	Opinion/statement	Leading	Total					
Overall interviews	2.2 [1.7, 2.7]	26.2 [24.9, 27.7]	1.4 [1.1, 1.9]	29.8 [28.5, 31.3]	0.3 [0.1, 0.5]	32.7 [31.4, 34.3]	7.1 [6.4, 8.0]	3.7 [3.1, 4.3]	8.6 [8.1, 9.9]	17.8 [16.1, 18.4]	70.2 [68.8, 71.6]	8	24	158.844	< .001	2.57
Introduction	0.7 [0.2, 1.7]	32.5 [30.7, 37.9]	1.1 [0.5, 2.3]	34.3 [32.5, 39.9]	0.1 [0.0, 0.8]	29.1 [27.2, 34.2]	6.6 [5.2, 9.2]	4.4 [3.2, 6.5]	10.1 [8.4, 13.2]	15.2 [13.4, 19.0]	65.7 [60.1, 67.5]	8	24	108.423	< .001	2.13
Free recall	31.5 [22.6, 43.1]	22.5 [14.6, 33.2]	0.0 [0.0, 4.2]*	53.9 [44.1, 65.9]	0.0 [0.0, 4.2]*	15.7 [9.1, 25.5]	5.6 [1.9, 12.9]	2.2 [0.3, 8.1]	9.0 [4.1, 17.3]	13.5 [7.3, 22.9]	46.1 [34.1, 55.9]	8	24	61.393	< .001	1.60
Questioning	1.7 [1.3, 2.3]	23.7 [22.2, 25.3]	1.1 [0.8, 1.6]	26.6 [25.0, 28.2]	0.4 [0.2, 0.7]	35.1 [33.4, 36.9]	7.7 [6.7, 8.7]	3.7 [3.0, 4.4]	8.1 [7.2, 9.2}	18.4 [17.1, 19.9]	73.4 [71.8, 75.0]	8	24	156.570	< .001	2.55
Closure	1.4 [0.2, 5.1]	5.0 [2.0, 10.0]	11.4 [6.7, 17.9]	17.9 [11.9, 25.2]	0.0 [0.0, 2.6]*	32.1 [24.5, 40.6]	2.1 [0.4, 6.1]	5.7 [2.5, 10.9]	27.1 [20.0, 35.3]	15.0 [0.5, 22.0]	82.1 [74.8, 88.1]	8	24	50.830	< .001	1.46

*One-sided 97.5 % CI

Table 4. Mean proportions (%) and standard deviations of question types used across the interviews as a function of the age of the witnesses

		Age-based groups						df	F	p	η^2_p
		Younger than 7 years		7-10 years		11 years and older					
		(n = 8)	CI 95%	(n = 9)	CI 95%	(n = 7)	CI 95%				
Appropriate	Open-ended	2.0 (2.0)	[0.3, 3.6]	2.5 (1.9)	[1.1, 4.0]	2.7 (2.4)	[0.5, 4.9]	2, 21	0.234	.793	.022
	Probing	25.1 (6.0)	[20.1, 30.1]	24.1 (7.0)	[18.7, 29.4]	26.4 (8.2)	[18.8, 34.0]	2, 21	0.227	.799	.021
	Closed	1.2 (1.4)	[0.1, 2.3]	1.7 (0.8)	[1.0, 2.3]	1.8 (2.4)	[0.0, 4.0]	2, 21	0.286	.754	.027
Inappropriate	Open-ended	0.1 (0.2)	[0.0, 0.3]	0.52 (0.6)	[0.1, 1.0]	0.2 (0.6)	[0.0, 0.8]	2, 21	1.215	.317	.104
	Closed	29.9 (14.8)	[17.5, 42.3]	31.8 (10.8)	[23.5, 40.1]	30.3 (4.7)	[26.0, 34.6]	2, 21	0.072	.930	.007
	Multiple	13.1 (7.3)	[6.9, 19.2]	9.1 (3.1)	[6.6, 11.4]	5.6 (4.8)	[1.2, 10.1]	2, 21	3.739	.041	.263
	Forced Choice	3.0 (2.1)	[1.3, 4.8]	4.7 (3.6)	[2.0, 7.5]	2.6 (1.9)	[0.8, 4.4]	2, 21	1.415	.265	.119
	Opinion/statement	6.3 (5.4)	[1.8, 10.8]	5.8 (3.9)	[2.8, 8.8]	13.6 (6.8)	[7.3, 19.9]	2, 21	5.186	.015	.331
	Leading	19.3 (13.4)	[8.1, 30.5]	19.9 (9.7)	[12.5, 27.3]	16.8 (5.5)	[11.8, 21.9]	2, 21	0.203	.818	.019

Table 5. Associations between the interviewers' utterances and the young witnesses' responses in the overall interviews (%) for each age group (confidence intervals are indicated in brackets)

		Response categories										
	Response types (%)	Request-confirming	Agreement	Disagreement	Extended Task-related	No answer	Evasive Response	df	N	χ^2	p	r
		27.5	27.3	16.0	9.1	15.7	4.4					
Under 7 years old	Appropriate open-ended questions	25.0 [9.8, 46.7]	**4.2 [0.1, 21.1]**	4.2 [0.1, 21.1]	**25.0 [9.8, 46.7]**	**33.3 [15.6, 55.3]**	8.3 [1.0, 27.0]	5	24	11.500	.042	0.69
	Probing questions	**62.9 [57.2, 68.3]**	**2.0 [0.7, 4.2]**	**1.3 [0.4, 3.3]**	11.7 [8.3, 15.9]	16.0 [12.0, 20.5]	6.2 [3.8, 9.5]	5	307	501.319	< .001	1.28
	Appropriate closed questions	14.3 [3.0, 36.3]	23.8 [8.2, 47.2]	**42.9 [21.8, 66.0]**	9.5 [1.2, 30.4]	4.8 [0.1, 23.8]	4.8 [0.1, 23.8]	5	21	13.571	.019	0.80
	Inappropriate open-ended questions	-	-	-	-	-	-	-	-	-	-	-
	Inappropriate closed questions	**11.1 [8.3, 14.4]**	**47.3 [42.6, 52.2]**	**20.8 [17.1, 24.9]**	**4.9 [3.0, 7.3]**	**12.2 [9.3, 15.7]**	3.7 [2.1, 5.9]	5	433	342.081	< .001	0.89
	Multiple questions	27.4 [18.7, 37.5]	**12.6 [6.7, 21.0]**	22.1 [14.2, 31.8]	12.6 [6.7, 21.0]	**23.2 [15.1, 32.9]**	2.1 [0.3, 7.4]	5	95	24.558	< .001	0.51
	Forced choice questions	**74.4 [57.9, 87.0]**	**7.7 [1.6, 20.9]**	7.7 [1.6, 20.9]	5.1 [0.6, 17.3]	**0.0 [0.0, 9.0]***	5.1 [0.6, 17.3]	5	39	72.154	< .001	1.36
	Opinion/statement	**05.6 [1.5, 13.6]**	18.1 [10.0, 28.9]	11.1 [4.9, 20.7]	**16.7 [8.9, 27.3]**	**44.4 [32.7, 56.6]**	4.2 [0.9, 11.7]	5	72	46.833	< .001	0.81
	Leading questions	**10.4 [6.7, 15.2]**	**39.4 [32.9, 46.1]**	**26.2 [20.6, 32.6]**	9.1 [5.6, 13.6]	11.3 [7.5, 16.2]	3.6 [1.6, 7.0]	5	221	119.751	< .001	0.74
		Response categories										
	Response types (%)	Request-confirming	Agreement	Disagreement	Extended Task-related	No answer	Evasive Response					
		26.2	22.2	18.9	17.4	9.2	6.1					
7-10 years old	Appropriate open-ended questions	**10.3 [2.9, 24.2]**	**0.0 [0.0, 9.0]***	**0.0 [0.0, 9.0]***	**51.3 [34.8, 67.6]**	25.6 [13.0, 42.1]	**12.8 [4.3, 27.4]**	5	39	16.487	.001	0.65
	Probing questions	**54.8 [49.9, 59.6]**	**1.2 [0.4, 2.8]**	**2.9 [1.5, 5.0]**	23.2 [19.2, 27.6]	9.3 [6.7, 12.5]	**8.6 [6.1, 11.7]**	5	418	512.660	< .001	1.11
	Appropriate closed questions	12.0 [2.5, 31.2]	32.0 [14.9, 53.5]	**40.0 [21.1, 61.3]**	8.0 [1.0, 26.0]	4.0 [0.1, 20.4]	4.0 [0.1, 20.4]	5	25	17.960	.003	0.85

	Inappropriate open-ended questions	50.0 [15.7, 84.3]	0.0 [0.0, 36.9]*	12.5 [0.3, 52.7]	12.5 [0.3, 52.7]	12.5 [0.3, 52.7]	12.5 [0.3, 52.7]	4	8	4.500	.343	0.75
	Inappropriate closed questions	**10.2 [7.7, 13.1]**	**39.7 [35.5, 44.0]**	**27.4 [23.6, 31.4]**	15.2 [12.2, 18.6]	**5.2 [3.5, 7.5]**	**2.3 [1.2, 4.0]**	5	519	319.416	<.001	0.79
	Multiple questions	**36.0 [28.0, 44.7]**	**10.3 [5.7, 16.7]**	19.1 [12.9, 26.7]	20.6 [14.1, 28.4]	7.4 [3.6, 13.1]	6.6 [3.1, 12.2]	5	136	50.971	<.001	0.61
	Forced choice questions	**44.4 [33.4, 55.9]**	**8.6 [3.5, 17.0]**	12.4 [6.1, 21.5]	18.5 [10.8, 28.7]	7.4 [2.8, 15.4]	8.6 [3.5, 17.0]	5	81	49.000	<.001	0.78
	Opinion/statement	**16.7 [9.8, 25.6]**	26.0 [17.6, 36.0]	**7.3 [3.0, 14.4]**	**9.4 [4.4, 17.1]**	**35.4 [25.9, 45.8]**	5.2 [1.7, 11.7]	5	96	41.000	<.001	0.65
	Leading questions	**10.7 [7.5, 14.7]**	**31.5 [26.3, 37.0]**	**32.5 [27.3, 38.0]**	**10.7 [7.5, 14.7]**	7.1 [4.5, 10.6]	7.5 [4.8, 11.0]	5	308	132.260	<.001	0.66

		Response categories										
	Response types (%)	Request-confirming	Agreement	Disagreement	Extended Task-related	No answer	Evasive Response					
		24.3	25.3	14.1	22.0	9.4	5.0					
11 years old and more	Appropriate open-ended questions	16.7 54.5, 37.4]	0.00 [0.0, 14.2]*	0.0 [0.0, 14.2]*	**66.7 [44.7, 84.4]**	8.3 [1.0, 27.0]	08.3 [1.0, 27.0]	5	24	22.667	<.001	0.97
	Probing questions	**52.2 [46.7, 57.7]**	**1.5 [0.5, 3.4]**	**0.6 [0.1, 2.1]**	27.9 [23.2, 33.0]	8.9 [6.1, 12.5]	**8.9 [6.1, 12.5]**	5	337	404.383	<.001	1.10
	Appropriate closed questions	**0.0 [0.0, 24.7]***	**53.9 [25.1, 80.8]**	15.4 [1.9, 45.4]	23.1 [5.0, 53.8]	7.7 [0.2, 36.0]	0.0 [0.0, 24.7]*	3	13	6.385	.094	0.70
	Inappropriate open-ended questions	-	-	-	-	-	-	-	-	-	-	-
	Inappropriate closed questions	**12.5 [9.3, 16.3]**	**39.4 [34.4, 44.6]**	**22.6 [18.4, 27.2]**	**14.4 [11.0, 18.4]**	8.2 [5.6, 11.4]	3.0 [1.5, 5.3]	5	368	184.065	<.001	0.71
	Multiple questions	20.7 [11.2, 33.4]	20.7 [11.2, 33.4]	08.6 [2.9, 19.0]	27.6 [16.7, 40.9]	13.8 [6.1, 25.4]	8.6 [2.9, 19.0]	5	58	10.069	.073	0.42
	Forced choice questions	**50.0 [30.6, 69.4]**	**7.1 [0.9, 23.5]**	14.3 [4.0, 32.7]	17.9 [6.1, 36.9]	3.6 [0.1, 18.3]	7.1 [0.9, 23.5]	5	28	24.714	<.001	0.94
	Opinion/statement	14.8 [10.0, 20.8]	**35.2 [28.2, 42.6]**	**11.0 [6.8, 16.5]**	19.2 [13.8, 25.7]	**15.4 [10.5, 21.5]**	4.4 [1.9, 8.5]	5	182	58.593	<.001	0.57
	Leading questions	**6.8 [3.7, 11.4]**	**36.1 [29.3, 43.4]**	**27.8 [21.5, 34.7]**	21.5 [15.9, 28.0]	6.8 [3.7, 11.4]	**1.1 [0.1, 3.7]**	5	191	110.351	<.001	0.76

Note: the values in bold contribute to the significant relation between the investigators' utterances and the response types given by the children and adolescents (measured by the standardized residuals worth ± 2.0). *One-sided 97.5% CI.

Table 6. Associations between the young witnesses' responses and the interviewers' subsequent utterances in the overall interviews (%) for each age group (confidence intervals are indicated in brackets)

	Utterance types (%)	Appropriate open-ended questions	Probing questions	Appropriate closed questions	Inappropriate open-ended questions	Inappropriate closed questions	Multiple questions	Forced choice questions	Opinion / statement	Leading questions	df	N	χ^2	p	r
		2.0	25.1	1.2	0.1	29.9	13.1	3.0	6.3	19.3					
Under 7 years old	Request-confirming	1.8 [0.7, 3.9]	**29.9 [25.0, 35.1]**	**0.0 [0.0, 1.1]***	0.3 [0.0, 1.7]	**43.3 [37.9, 48.8]**	6.9 [4.4, 10.1]	2.1 [0.8, 4.3]	**3.3 [1.7, 5.8]**	**12.5 [9.2, 16.6]**	7	355	465.597	<.001	1.15
	Agreement	**3.3 [1.7, 5.9]**	**29.7 [24.8, 34.9]**	1.2 [0.3, 3.1]	**0.0 [0.0, 1.1]***	31.5 [26.5, 36.8]	**5.2 [3.0, 8.1]**	3.0 [1.5, 5.5]	7.3 [4.7, 10.6]	18.8 [14.7, 23.4]	7	330	284.933	<.001	0.93
	Disagreement	2.6 [0.8, 5.9]	20.7 [15.2, 27.1]	**3.6 [1.5, 7.3]**	0.0 [0.0, 1.9]*	38.9 [31.9, 46.1]	9.3 [5.6, 14.3]	3.1 [1.1, 6.6]	5.2 [2.5, 9.3]	16.6 [11.5, 22.6]	7	193	171.062	<.001	0.94
	Extended Task-related	0.0 [0.0, 3.3]*	26.1 [18.2, 35.3]	0.9 [0.0, 4.9]	0.0 [0.0, 3.3]*	35.1 [26.3, 44.8]	11.7 [6.4, 19.2]	2.7 [0.6, 7.7]	7.2 [3.2, 13.7]	16.2 [9.9, 24.4]	6	111	73.712	<.001	0.81
	No answer	0.5 [0.0, 3.0]	**15.7 [10.8, 21.7]**	0.0 [0.0, 2.0]*	3.2 [1.2, 6.9]	**27.0 [20.8, 34.0]**	9.2 [5.4, 14.3]	5.4 [2.6, 9.7]	8.7 [5.0, 13.7]	**30.3 [23.7, 37.4]**	7	185	124.578	<.001	0.82
	Evasive Response	1.9 [0.0, 10.3]	17.3 [8.2, 30.3]	**7.7 [2.1, 18.5]**	0.0 [0.0, 6.8]*	34.6 [22.0, 49.1]	11.5 [4.4, 23.4]	5.8 [1.2, 15.9]	3.9 [0.5, 13.2]	17.31 [8.2, 30.3]	7	52	32.923	<.001	0.80

	Utterance types (%)	Appropriate open-ended questions	Probing questions	Appropriate closed questions	Inappropriate open-ended questions	Inappropriate closed questions	Multiple questions	Forced choice questions	Opinion / statement	Leading questions					
		2.5	24.1	1.7	0.5	31.8	9.0	4.7	5.8	19.9					
7 – 10 years old	Request-confirming	2.4 [1.1, 4.3]	**32.6 [28.2, 37.3]**	1.9 [0.8, 3.7]	0.5 [0.1, 1.7]	**27.9 [23.7, 32.5]**	8.9 [6.4, 12.0]	4.5 [2.7, 6.9]	5.4 [3.5, 8.0]	16.0 [12.6, 19.8]	8	335	431.915	<.001	1.10
	Agreement	2.2 [1.0, 4.3]	28.3 [23.7, 33.3]	1.4 [0.5, 3.2]	0.0 [0.0, 1.0]*	32.5 [27.7, 37.6]	6.9 [4.5, 10.1]	5.00 [3.0, 7.8]	6.4 [4.1, 9.4]	17.2 [13.5, 21.5]	7	360	295.644	<.001	0.91
	Disagreement	2.3 [0.9, 4.7]	**14.4 [10.6, 18.8]**	1.6 [0.5, 3.8]	0.3 [0.0, 1.8]	**37.9 [32.4, 43.6]**	8.8 [5.9, 12.6]	6.7 [4.3, 10.3]	4.3 [2.3, 7.2]	**23.5 [18.9, 28.7]**	8	306	340.765	<.001	1.06
	Extended Task-related	1.4 [0.4, 3.6]	29.9 [24.7, 35.6]	1.4 [0.4, 3.6]	1.1 [0.2, 3.1]	30.3 [25.0, 36.0]	7.4 [4.6, 11.1]	3.5 [1.7, 6.4]	7.6 [4.9, 11.5]	17.3 [13.0, 22.2]	8	284	289.211	<.001	1.01
	No answer	**6.2 [2.9, 11.5]**	18.6 [12.6, 25.9]	0.7 [0.0, 3.8]	0.7 [0.0, 3.8]	33.1 [25.5, 41.4]	10.3 [5.9, 16.5]	6.2 [2.9, 11.5]	4.1 [1.5, 8.8]	20.0 [13.8, 27.4]	8	145	121.834	<.001	0.92
	Evasive Response	1.0 [0.0, 5.5]	18.2 [11.1, 27.2]	0.0 [0.0, 3.7]*	1.0 [0.0, 5.5]	31.3 522.4, 41.4]	8.1 [3.6, 15.3]	4.0 [1.1, 10.0]	8.1 [3.6, 15.3]	**28.3 [19.7, 38.2]**	7	99	79.990	<.001	0.90

	Utterance types (%)	Appropriate open-ended questions	Probing questions	Appropriate closed questions	Inappropriate open-ended questions	Inappropriate closed questions	Multiple questions	Forced choice questions	Opinion / statement	Leading questions

		2.7	26.4	1.8	0.2	30.3	5.6	2.6	13.6	16.9					
11 years old and more	Request-confirming	1.4 [0.4, 3.5]	32.5 [29.1, 40.4]	1.0 [0.2, 3.0]	0.3 [0.0, 1.9]	34.6 [29.1, 40.4]	3.2 [1.7, 6.2]	2.7 [1.2, 5.3]	**9.6 [6.5, 13.6]**	14.4 [10.6, 18.9]	8	292	384.973	< .001	1.15
	Agreement	3.3 [1.6, 6.0]	30.2 [25.1, 35.8]	0.7 [0.1, 2.4]	0.0 [0.0, 1.2]*	30.9 [25.7, 36.5]	6.0 [3.6, 9.3]	2.0 [0.7, 4.3]	16.6 [12.6, 21.3]	**10.3 [7.1, 14.3]**	7	301	253.286	< .001	0.92
	Disagreement	1.8 [0.4, 5.1]	**18.8 [13.2, 25.5]**	**2.9 [1.0, 6.7]**	0.0 [0.0, 2.1]*	**38.8 [31.5, 46.6]**	2.4 [0.6, 5.9]	1.é [0.1, 4.2]	13.5 [8.8, 19.6]	20.6 [14.8, 27.5]	7	170	168.259	< .001	0.99
	Extended Task-related	1.1 [0.2, 3.3]	31.2 [25.6, 37.2]	0.4 [0.0, 2.1]	0.0 [0.0, 1.4]*	**25.1 [20.0, 30.8]**	7.2 [4.4, 11.1]	1.1 [0.2, 3.3]	12.9 [9.1, 17.6]	**20.9 [16.2, 26.3]**	7	263	212.772	< .001	0.90
	No answer	2.7 [0.6, 7.7]	**18.9 [12.1, 27.5]**	0.9 [0.0, 4.9]	0.0 [0.0, 3.3]*	26.1 [18.2, 35.3]	3.6 [1.0, 9.0]	1.8 [0.2, 6.4]	**29.7 [21.4, 39.1]**	16.2 [9.9, 24.4]	7	111	85.396	< .001	0.88
	Evasive Response	1.7 [0.0, 8.9]	23.3 [13.4, 36.0]	1.7 [0.0, 8.9]	0.0 [0.0, 6.0]*	21.7 [12.1, 34.2]	5.0 [1.0, 13.9]	**10.0 [3.8, 20.5]**	20.0 [10.8, 32.3]	16.7 [8.3, 28.5]	7	60	27.467	< .001	0.68

Note: the values in bold contribute to the significant relation between the response types given by the children and adolescents and the investigators' subsequent utterances (measured by the standardized residuals worth ± 2.0). *One-sided 97.5% CI.

Appendix. Coding example based on a 13-year-old girl interview

Utterances	Phase	Utterance type	Appropriateness of the question (Q) / Level of disclosure of the response (A)	Rationale
(Q) And then, what happened?	Questions	Open question	Appropriate	
(A) He took my pant off. . . and he *thinged* (sic) me.	Questions	Extended task-related	Disclosure	
(Q) What does "to *thing*" mean?	Questions	Probing question	Appropriate	It elaborates the previous answer provided by the child and it is not suggestive.
(A) Well… he *thinged* me.	Questions	Evasive answer	Non-disclosure	It provides an answer but it is not informative or remains unclear.
(Q) So. . . you told me he took your pants off, right?	Questions	Closed question	Appropriate	It allows a yes or no response. It is appropriate as the interviewer uses it to summarize previous collected information and to refocus the interview on the perpetrator's actions. The expected answer to this question will then be used by the interviewer as a starting point to the following perpetrators' actions.
(A) Yes.	Questions	Agreement	Disclosure	
(Q) So, what did he do right after that?	Questions	Probing question	Appropriate	It is asked to elaborate on the perpetrators' actions.
(A) He. . . He *thinged* (sic) me.	Questions	Request conforming	Disclosure	
(Q) To thing. . . does that mean having sex?	Questions	Suggestive question	Inappropriate	The child did not say before that 'to thing' means 'having sex'.
(A) Yes.	Questions	Agreement	Disclosure	
(Q) Was it long?	Questions	Suggestive question	Inappropriate	The child never provided any information related to the length of the abuse. The interviewer suggests only one response modality (i.e. it was long)
(A) No. I don't think so. I don't know.	Questions	Evasive answer	Non-disclosure	
(Q) So you say that he did enter there below?	Questions	Suggestive question	Inappropriate	The child only agreed with the fact that 'to thing' means having sex, but did not specify before the type of sexual abuse she was victim of.
(A) Yes.	Questions	Agreement	Disclosure	
(Q) Many times?	Questions	Suggestive question	Inappropriate	The child did not provide any information related to the frequency of abuses before. The interviewer suggests only one response modality (i.e. many times)
(A) Twice.	Questions	Request Conforming	Disclosure	

Endnotes

[i] In their paper, Westcott and Kynan (2006) referred to the rapport building phase rather than the introduction of the interview. The rapport building is part of the introduction phase (see Luther et al., 2015). Specifically, rapport building consists of mutual attentiveness, positivity and co-ordination (Tickle-Degen & Rosenthal, 1990) in order to create a more comfortable interview atmosphere. The rapport building phase described by Westcott and Kynan, with components such as truth-lies ceremony, discussion about non-abuse topics, explain the reasons of the interview and the presentation of the ground rules for example, is therefore more closed to the introduction phase.