

**Preschoolers' true and false reports: Comparing effects of the Sequential Interview and
NICHD protocol**

Mikaela Magnusson^{1*}, Malin Joleby¹, Emelie Ernberg¹, Lucy Akehurst², Julia Korkman^{3,4} &
Sara Landström¹

¹University of Gothenburg² University of Portsmouth, ³Åbo Akademi University

⁴Helsinki University Hospital

In press, Legal and Criminological Psychology

Author Note

Funding: This work was financially supported by the Swedish Crime Victim
Compensation and Support Authority [Grant number 03005/2016].

* Corresponding author: Mikaela Magnusson, Department of Psychology, University of
Gothenburg, P.O. Box 500, SE 405 30 Gothenburg, Sweden, Phone: +4631-7861696, E-mail:
mikaela.magnusson@psy.gu.se

Abstract

Purpose: The current study aimed to examine a Norwegian technique for conducting investigative interviews with preschoolers: –the Sequential Interview (SI). The SI advocates for increased initial rapport building and includes a pre-determined break before the substantive phase. To explore the potential benefits and risks of the SI, the technique was compared with an adapted version of the National Institute of Child Health and Development (NICHD) protocol.

Methods: A total of 129 preschoolers (3–6 years) were interviewed with either the SI or NICHD protocol about a self-experienced (Exp. I) or non-experienced (Exp. II) event.

Result: For Exp. I, no significant difference was observed across interview conditions in the number of reported details about a self-experienced event. However, children interviewed with the SI exhibited a slightly lower accuracy rate compared to those interviewed with the NICHD protocol. For Exp. II, a total of 31.1% of the preschoolers initially assented to remembering a fictive (false) experience and 15.6% gave an account (> 40 details) of the non-experienced event. We found no difference between interviewing conditions in assent rates or number of false accounts.

Conclusions: The study provides valuable insights into the difficulties involved when interviewing young children. The results showed few differences between the novel SI model and the well-established NICHD protocol. While many preschoolers could provide accurate testimony, some embedded worrisome false details in their narratives. Furthermore, a minority of children gave false reports about non-experienced events when interviewed with the two techniques. Methodological limitations and suggestions for future research will be discussed.

Keywords: child interview, children’s testimony, preschooler, NICHD protocol, sequential interview, interview protocol

**Preschoolers' true and false reports: Comparing effects of the Sequential Interview and
NICHD protocol**

Young children can generally be reliable witnesses from around three to four years of age (Brubacher et al., 2019). However, a wide range of factors influence the reliability of their reports. Specifically, preschoolers have been found to be particularly vulnerable to certain forms of social influence including suggestive interviewing techniques (Ceci & Bruck, 1993; but see also Otgaar et al., 2018). Considering the limited cognitive and verbal abilities of preschoolers, they cannot give as complete accounts as older children or adults (Goodman & Reed, 1986). Although their ability to give detailed and coherent narratives develop throughout childhood (Fivush, 2011), many preschoolers can be highly accurate when questioned in a developmentally sensitive manner (Brubacher et al., 2019). During the last decades, concerns have been raised regarding how to interview preschoolers to facilitate their witness accounts without compromising their accuracy (Poole et al., 2014). While a substantial amount of research has focused on establishing risk factors during these difficult interviews (e.g., Ceci & Bruck, 1995), few studies have compared the effects of different investigative interviewing protocols with preschool-aged children (Benia et al., 2015; Saywitz et al., 2018). The present paper aimed to address this limitation by comparing preschoolers' reports about self-experienced and non-experienced events when interviewed with the pre-substantive interview structure of the Sequential Interview model (Langballe & Davik, 2017) or the original National Institute of Child Health and Development (NICHD) protocol (Lamb et al., 2008).

Through laboratory and field studies, researchers have been able to identify both strengths and weaknesses in young children's memory for experienced and non-experienced events. In situations where preschoolers are interviewed in a non-suggestive manner using primarily open-ended questions, their statements can be both detailed and accurate (e.g.,

EFFECTS OF INTERVIEWING TECHNIQUE

Gagnon & Cyr, 2017; Melinder et al., 2006). However, older preschoolers typically outperform younger preschoolers in different eyewitness paradigms (e.g., Alexander et al., 2002; Goodman & Reed, 1986; Melinder et al., 2006). Furthermore, suggestive influences increase the risk of obtaining false accounts from young children (for reviews, see Bruck & Melnyk, 2004; Ceci & Bruck, 1995, Poole et al., 2014). Numerous factors might contaminate preschoolers' testimony, including positive and negative reinforcement, introduction of misinformation, social pressure, compliance to authority figures, invitations to speculate, and induction of stereotypes (Ceci & Bruck, 1995).

Although the risks of suggestive techniques are well known, it is less discussed that a small but significant proportion of children give false reports during interviews following research-based protocols (Brubacher et al., 2019). In a laboratory study examining NICHD protocol interviews with 5–7-year-olds, Brown et al. (2013) reported that 10.9% gave a false report about a non-experienced visit to a fire station. Similarly, research on the Narrative Elaboration Technique with elementary school children found that 4% of the children falsely described a fictive desert trip in response to a free recall request (Camparo et al., 2001). During police investigations, forensic interviewers more often than not need to interview children in situations where it is not known whether abuse has occurred (Korkman, Antfolk, Fagerlund, & Santtila, 2018). It is therefore important to study children's reports about both self-experienced and non-experienced events when examining the potential benefits and risks of different child interviewing techniques (e.g., Powell, Jones, & Campbell, 2003). The two studies described above showed that a minority of school-aged children gave false reports when interviewed with research-based protocols. Considering that preschoolers have been found more vulnerable to suggestive interviewing compared to school-aged children (Ceci & Bruck, 1993), more research is needed to understand young children's susceptibility to false reporting in response to different child interview techniques.

EFFECTS OF INTERVIEWING TECHNIQUE

One of the child interviewing techniques examined in the present study was the Sequential Interview model (SI; see Langballe & Davik, 2017, for an overview). The SI model was created by a team of Norwegian researchers and practitioners who were influenced by the Extended Forensic Interview (Carnes et al., 1999) and the Dialogical Communication Model (Gamst & Langballe, 2004). A SI is segmented into different sequences (each around 15-20 minutes long). Since preschoolers have a limited attention span, there are breaks in between each sequence. The first sequence focuses on rapport building. Interviewers are encouraged to combine the use of different props (e.g., jigsaw puzzles, drawings) with open-ended questions about the props (see Magnusson et al., 2020b). Following ‘best practice’ guidelines (Newlin et al., 2015), this sequence also comprises an introduction, explanation of ‘ground rules’ for the subsequent phases of the interview, questions about personal interests, and an episodic practice narrative. Before transitioning to the information-gathering (henceforth referred to as the substantive) phase of the interview, the SI includes a break after the initial sequence to counter fatigue effects. The substantive phase is thereafter also segmented into shorter sequences with breaks in between, followed by a closure phase that ends on a neutral topic (for more information about the SI model, see Langballe & Davik, 2017).

Following the implementation of the SI model in Norway (October 2015), two field studies have examined the technique with varied results in terms of the question types used by forensic interviewers (Baugerud et al., 2020) and the dialogical communication patterns between child and interviewer (Melinder et al., 2020). However, no study has examined the effects of the SI in a controlled laboratory setting where children’s response accuracy can be assessed. The present paper aimed to address this gap by comparing the pre-substantive phase structure of the SI to an adapted version of the original NICHD protocol (see Lamb et al.,

EFFECTS OF INTERVIEWING TECHNIQUE

2008¹, for an overview). The NICHD protocol is a well renowned child interviewing technique that builds upon field and laboratory research on children's witness abilities (see e.g., Benia et al., 2015; La Rooy et al., 2015). The protocol largely overlaps with the current 'best practice' guidelines outlined by the US Justice Department (Newlin et al., 2015) and other research-based techniques (Faller, 2015). Of relevance for the present study, the NICHD protocol includes a briefer pre-substantive phase compared to the SI model. Specifically, the protocol includes an introduction, ground rules training, questions about personal interests and an episodic practice narrative (see Table 1 for a comparative overview of the NICHD protocol and the SI model). Thus, the two interviewing techniques differ in the content and length of the pre-substantive phase. This raises the question of whether a longer prop-based rapport building phase followed by a break (as recommended in the SI model) would be better suited to meet the needs of preschool-aged witnesses compared to a briefer verbal rapport building phase (as recommended in the original NICHD protocol).

[Insert Table 1 about here]

The inclusion of a rapport building phase during child interviews has been proposed to have a range of benefits, such as decreasing children's anxiety and reluctance, providing an opportunity for children to practice answering questions, and enabling interviewers to adjust their interviewing strategies depending on the children's abilities (e.g., Collins, Doherty-Sneddon, & Doherty, 2014; Hershkowitz, 2011; Saywitz et al., 2015). These benefits are, in turn, assumed to positively affect the completeness and accuracy of children's witness reports (e.g., Brown et al., 2013; Hershkowitz, 2009; Lyon et al., 2014; Roberts, Lamb & Sternberg, 2004; Sternberg et al., 1997; Teoh & Lamb, 2010; Yi & Lamb, 2018, but see also Sauerland

¹ During data collection for the present study, a revised version of the NICHD handbook was published (Lamb et al., 2018) focusing on increased socio-emotional support. Similar to the SI model, the revised NICHD protocol recommends a more extensive rapport building phase compared to the original protocol. As of yet (September 2020), no experimental studies have been published on the potential effects of the revised NICHD protocol on preschoolers' response accuracy.

EFFECTS OF INTERVIEWING TECHNIQUE

et al., 2019). However, few controlled laboratory studies have examined the potential impact of different rapport building strategies during investigative interviews with children (Saywitz et al., 2015). Due to the lack of ground truth in field research, experimental studies are needed to assess the effects on children's recall accuracy. For example, there is a paucity of research on the effects of rapport building on children's reports of non-experienced events. This gap in the literature is important to address considering the implications false reports can have on a criminal investigation. A more extended rapport building phase may make children more comfortable correcting the interviewer, and in turn, better able to resist suggestive influence (Saywitz et al., 2015). On the other hand, extensive rapport building might increase the risk of children providing false information to please the interviewer (Sauerland et al., 2019).

Forensic interviewers have also expressed concerns regarding the length and content of the pre-substantive interview phase with preschool-aged children (see Magnusson et al., 2020a). Specifically, practitioners have described working under time pressure as young children rapidly begin experiencing fatigue during interviews. Potentially, a lengthy pre-substantive phase involving different rapport building activities might have a negative impact on young children's ability to focus on the task at hand (e.g., to monitor their response accuracy and stay on topic). In line with these concerns, Davies, Westcott and Horan (2000), found that rapport building sessions that lasted more than 8 minutes were associated with less informative accounts during the substantive phase of child interviews. Implementing a break after the pre-substantive phase, as prescribed in Sequential Interviews, may help children recharge their energy (Saywitz & Camparo, 2014). In classroom settings, the use of breaks has for example been associated with increased attention levels among preschoolers (Holmes, Pellegrini, & Schmidt, 2007). While most child interviewing protocols recommend employing breaks to some extent throughout the interview, few laboratory studies have

EFFECTS OF INTERVIEWING TECHNIQUE

incorporated breaks as part of the experimental design. Another distinct feature of the SI model concerns the use of different collaborative activities involving rapport building props (e.g., jigsaw puzzles, drawings, describing pictures from story books) during the pre-substantive phase (Langballe & Davik, 2017). Props such as a jigsaw puzzle may serve as a good ‘ice-breaker’ and encourage children to interact with the interviewer by collaborating on a shared task. On the other hand, props may be distracting for young children and might shift their focus from the conversation toward focusing primarily on the prop. A recent experiment of relevance for the present study showed that the use of a prop-based rapport strategy (solving a jigsaw puzzle) did not appear to compromise preschool-aged children’s accuracy when describing a transgression compared to a verbal rapport building strategy (talking about personal interests, Magnusson et al., 2020b). Given that child interviewing protocols differ in whether they employ verbal or prop-based rapport building techniques, and that limited research has investigated these differences, examining the effects of these different techniques is crucial.

To summarize, the present study sought to compare preschoolers’ statements as a function of two different interviewing techniques based on the SI (Langballe & Davik, 2017) and the original NICHD protocol (Lamb et al., 2008). Specifically, we examined effects of the pre-substantive phase of the two techniques. See Table 1 for an overview of the interview structures. Two contrasting hypotheses could be made regarding whether preschoolers would give fuller and more accurate accounts following a prolonged prop-based introduction with a break before the substantive phase, as advocated in the SI, or from a briefer verbal introduction as advocated by the original NICHD protocol. On the one hand, increased rapport building activities may help children feel more at ease, which could have positive effects on their productivity and accuracy (Saywitz et al., 2015). On the other hand, the prolonged interviewing time might lead to fatigue and inattention, which could lead to shorter

EFFECTS OF INTERVIEWING TECHNIQUE

or more erroneous statements by the preschoolers (Magnusson et al., 2020a). Furthermore, extended rapport building may make children more comfortable correcting the interviewer and more resistant to suggestive influence (Saywitz et al., 2015). Alternatively, children may not want to disappoint the interviewer and in turn, be more prone to falsely agreeing to misleading questions after an extended rapport phase (Sauerland et al., 2019). We therefore chose a two-tailed approach for examining possible differences across interviewing protocols. The preschoolers' amount of details and accuracy rates regarding self-experienced (Exp. I) and non-experienced (Exp. II) events functioned as dependent variables. Prior to the data collection, the research project was reviewed and approved by the Regional Ethics Board.

Experiment I

The first experiment was designed to investigate the quantity and accuracy of information included in preschoolers' statements regarding a self-experienced event as a function of the two interviewing techniques (SI vs. NICHD protocol). We hypothesized that; there would be a difference between interview protocols in terms of the amount (H1) and accuracy (H2) of details, and that older preschoolers would report more details (H3) and exhibit higher accuracy rates (H4) compared to younger preschoolers.

Method

Recruitment and participants. All preschools located in the metropolitan, suburban, and rural areas of *[location]*, were invited to participate during an annual Science Fair. Nineteen preschools spread across low-, middle- and high-income areas, chose to take part. Written parental consent was collected prior to the data collection and children were asked for their assent before they were interviewed. Of 168 children scheduled to participate, 88 had parental consent and gave verbal assent on the day of their interview². Four children

² The large attrition in Exp. I and II was mainly due to children being sick on the day of their interview, parental consent forms arriving late in the mail, and preschools forgetting to hand out the consent forms.

EFFECTS OF INTERVIEWING TECHNIQUE

were excluded from the analyses due to interviewers deviating from the interview guide. Thus, the analyses are based on interviews with 84 children (53 girls and 31 boys) aged 42–75 months ($M = 61.7$ months, $SD = 8.75$). All children were randomly allocated to the interviewing conditions, with 45 children ($M = 62.8$ months, $SD = 8.6$, 26 girls, 19 boys) in the NICHD condition and 39 children ($M = 60.3$ months, $SD = 8.87$, 27 girls, 12 boys) in the SI condition. According to a sensitivity analysis using G*Power (Faul et al., 2007), our sample could detect a main effect size of $d = 0.62$ with 80% power ($\alpha = .05$, two-tailed).

Materials and procedure

About one week prior to the children's interviews at the Science Fair, ($M = 6.6$ days, $SD = 1.4$, $Mdn = 7$, range 2-10³), the children took part in a staged 'pirate' event at their preschool (adapted from Brown et al., 2013). Shortly before the children's interviews, they also participated in a second staged event, the visit of detective 'Clever Clara' which was the topic of the practice narrative phase of their interviews, see Figure 1.

[Insert figure 1 about here]

The children were separately interviewed by one of ten RAs following semi-structured interview guides based on the SI or NICHD protocol. All interviewers were naïve to the study hypotheses and had received a two-day training course in child interviewing. At the beginning of all interviews, the interviewers introduced themselves, demonstrated the recording equipment, and asked the children if they wanted to be interviewed. If the children agreed to participate, they were randomly allocated to the NICHD or SI condition.

The NICHD condition was adapted from the original NICHD protocol (Lamb et al., 2008) and translated to Swedish. After the introduction, the interviewers proceeded to a ground rules phase that included instructions explaining that the children should tell if the

³ The retention interval length varied depending on the schedule of the preschools. Retention was therefore statistically controlled for in the analyses.

EFFECTS OF INTERVIEWING TECHNIQUE

interviewer said something they did not understand, and a practice in saying “I don’t know” and correcting the interviewer. This was followed by a rapport building phase where the interviewer asked the children to expand upon personal interests. The children were encouraged to elaborate on details they mentioned through invitations (“*Tell me more?*”), cued recall prompts (“*You said you like dogs, tell me more about that?*”), facilitators (“*uhmh*”, “*ok*”), and directive open-ended questions (“*What does your dog look like?*”). Next, the interviewer conducted an episodic practice narrative asking the children to describe everything that happened with detective ‘Clever Clara’, from beginning to end. The children were asked to elaborate using invitations, cued recall prompts, facilitators, and directive open-ended questions. The interviewers thereafter proceeded to the substantive phase of the interviews. The NICHD protocol introduction took around 5–7 minutes to complete.

The SI condition followed the same interview structure with two key exceptions. First, between the rapport building phase and the practice narrative, the SI also included a prop-based rapport building task. During this stage, the children were asked to work on a jigsaw puzzle and answer open-ended questions about it (see *-reference omitted for peer review*). Second, after the practice narrative, the interviewers in the SI condition stated that they would take a short break (5 min.) before asking more questions. During the break, the children were seated in another area of the interview room and given a set of crayons and paper for drawing. The children were thereafter led back to the interview area for the substantive phase of the interview. The introduction and break in the SI condition took around 15–17 minutes to complete.

For both conditions, the interviewers transitioned to the substantive phase by directing the topic to the pirate visit (“*I heard that you met Pia the Pirate, do you remember that?*”). If the child agreed, the interviewer asked him or her to elaborate (“*Tell me everything that happened, from the beginning to end.*”). If the child did not respond, the interviewer further

EFFECTS OF INTERVIEWING TECHNIQUE

clarified the context (“*Clever Clara said that a pirate visited your preschool, tell me about that*”). If the child started to talk about the event, the interviewer used invitations (“*Tell me more*”), open-ended cued prompts (“*You mentioned a sword, tell me more about that*”), facilitators (“*okay*”) and open-ended directive questions (e.g., “*What did she look like?*”). After exhausting the children’s free recall, all children were asked four scripted questions, see Table 2. The scripted questions included both leading (e.g., “*Did anyone take a picture of you?*”) and misleading (e.g., “*I heard that Pia the pirate showed you a pink rabbit, did you see it?*”) statements about information that may not have been mentioned by the children during the free recall phase. The children’s responses to the scripted questions were therefore analyzed separately. The interviewers thereafter progressed to another interviewing phase, which formed the basis for a separate study on drawing during interviews with children. The children’s responses during this drawing phase were not included in the present analyses. Lastly, the children were thanked for their participation and took part in a debriefing session and pirate treasure hunt.

Coding

The interviews were video-recorded and transcribed verbatim. Trained coders (naïve to the study hypotheses) coded the verbal content into separate details (similar to Brown et al., 2013). All details about actors (e.g., the *pirate*), actions and affective states (e.g. *showed*), objects (e.g. *suitcase*), temporal details (e.g. *before*) and settings (e.g. *preschool*) were given one point each. Additional attributes were also scored separately (e.g. “a *plastic silver sword*” was coded as three details). Repeated details were only counted the first time they occurred. Off-topic details were not coded. For exploratory purposes, the coders also marked all utterances that contained aggressive themes (e.g. “*The sword stabbed me*”). Importantly, there were no aggressive acts during the staged ‘pirate’ event.

EFFECTS OF INTERVIEWING TECHNIQUE

Each detail was assessed as either accurate, false, or non-verifiable (including subjective experiences). Accuracy rates were calculated from the amount of correct details divided by the total amount of accurate and false details. For the scripted questions, the children's responses were coded as correct, incorrect, or non-responsive ("*I don't know*"). If a child changed their answer ("*The bag was black, or no, it was brown*"), the coders only scored the last response. As response changes were rare ($n = 4$), we were unable to conduct separate analyses on this type of response pattern. Inter-rater reliability analyses were performed on 20% of the data. Cohen's $\kappa = .83-.87$ were achieved for all codes, indicating an adequate level of agreement. Disagreements were resolved through discussion.

Results

Preliminary regression analyses indicated that retention interval length (in days) did not have a significant impact on the following statistical tests (p 's $> .21$). Furthermore, as the interviews were carried out by ten different research assistants, we initially controlled for differences between interviewers but found no significant impact of including this control variable in the regression analyses (p 's $> .46$). We have therefore chosen to report the original regression models without retention interval and interviewer difference as covariates in the main result section. Moreover, preliminary analyses (using Welch t -tests) on the content of the substantial phase with regard to the number of interviewer prompts (NICHD condition $M = 33.4$, $SD = 12.3$; SI condition $M = 35.0$, $SD = 9.7$) and time spent talking (NICHD condition $M = 5.2$ minutes, $SD = 1.7$; SI condition $M = 5.5$, $SD = 1.6$) indicated that there was no significant differences between the two interviewing conditions, $t(81.267) = -.644$, $p = .521$, and $t(81.918) = -.799$, $p = .427$ respectively. Lastly, the interviewing conditions were balanced regarding child age, gender, and retention interval length (p 's $> .05$).

Amount of Detail during Free Recall

EFFECTS OF INTERVIEWING TECHNIQUE

First, we examined the effects of the interviewing technique by comparing the amount of details given in the SI ($M_{\text{raw}} = 96.8$, $SD_{\text{raw}} = 59.9$) and NICHD conditions ($M_{\text{raw}} = 102.7$, $SD_{\text{raw}} = 66.8$). A visual inspection of the distributions of data indicated non-normality in both groups (see Figure 2), which was confirmed by a Kolmogorov Smirnov test, $p = .003$. A Log10 transformation (+ 1) was performed to correct for non-normality in the data. A Welch t -test on the log transformed data indicated that there was no significant difference between conditions, $t(81.743) = 0.127$, $p = .899$, Hedges $g = 0.027$, 95% CI [-0.40, 0.46]. A hierarchical multiple linear regression analysis was thereafter performed with child age (in months) and interview structure (0 = NICHD protocol, 1 = SI) entered as predictor variables in Step 1. This model explained around 7% of the variance for the log-transformed total amount of details, $F(2, 81) = 4.45$, $p = .015$, $R^2_{\text{Adj.}} = 0.07$. Specifically, child age was significantly associated with the number of reported details ($b = 0.03$, $\beta = .318$, $p = .004$) indicating that older preschoolers reported more details than younger preschoolers. Interview condition was not associated with the total amount of details when controlling for child age ($b = 0.05$, $\beta = .032$, $p = .925$). To explore potential interactions between child age and interview condition, we entered an interaction term of age (mean centered) and interviewing condition in Step 2 of the regression analysis. The interaction term did not have a significant impact on the predictive value of the model, $F(1, 80) = 1.40$, $p = .240$, $\Delta R^2 = .016$.

[Insert Figure 2 about here]

Report Accuracy during Free Recall

The children's accuracy rates were generally high ($M = .83$, $SD = .13$, range .24-1), see Figure 3. As accuracy rates are bounded between 0 and 1, a logit-transformation (with a constant of .01 added to 0 values and subtracted from 1 values) was carried out before the inferential test. A Welch t -test indicated that the logit transformed accuracy rates differed significantly between interviewing conditions, $t(81.999) = 2.75$, $p = .007$, Hedges' $g = 0.59$,

EFFECTS OF INTERVIEWING TECHNIQUE

95% CI [0.15, 1.03]. Children in the NICHD condition ($M_{\text{raw}} = .86$, $SD = .10$) exhibited slightly higher accuracy rates compared to children in the SI condition ($M_{\text{raw}} = .79$, $SD = .15$). As can be seen in Figure 3, the SI condition contained one extreme outlier (more than two standard deviations from the mean) that could be influencing the results. A second Welch t-test was therefore conducted without the outlier, indicating that the effect was still significant without the extreme value, $t(80.51) = 2.54$, $p = .013$, Hedges' $g = 0.54$, 95% CI [0.10, 0.99]. However, a visual inspection of the data (see Figure 3) may help explain the observed accuracy effect. While both conditions included a small number of children with low accuracy scores, there were a few more children in the SI condition who exhibited accuracy scores on the lower end of the distribution compared to in the NICHD condition. We also carried out a hierarchical linear regression analysis with the children's logit-transformed accuracy rates as dependent variable. First, child age (in months) and interviewing condition (0 = NICHD protocol, 1 = SI) were entered as predictor variables. This model explained around 14% of the variance, $F(2, 81) = 7.66$, $p > .001$, $R^2_{\text{Adj.}} = .138$. In line with our hypothesis, child age was a significant predictor ($b = 0.03$, $\beta = .255$ $p = .008$) indicating that older preschoolers exhibited higher accuracy rates. Interviewing condition was also a significant predictor ($b = -0.53$, $\beta = -.276$ $p = .018$) when controlling for age differences, with children in the NICHD condition exhibiting higher accuracy rates. Including an interaction term in Step 2 between age (mean centered) and interviewing condition did not have a significant impact on the model, $\Delta R^2 = .009$, $F(1, 80) = 0.81$, $p = .375$.

[Insert Figure 3 about here]

[Insert Table 2 about here]

Reponses to the Scripted Questions

Towards the end of the substantive phase, the children were asked four leading and misleading questions about the pirate event. A descriptive overview of the children's

EFFECTS OF INTERVIEWING TECHNIQUE

responses to the scripted questions can be found in Table 2. Chi-square tests of independence and Fisher's Exact test were used to analyse the children's report accuracy. Due to low cell counts for the non-responsive category ("I don't know" responses), these responses were omitted from the main analyses. We found no significant difference between the SI and NICHD condition for any of the scripted questions; "*Did anyone take a picture of you?*" $\chi^2(1, N = 80) = 0.17, p = .679$, "*Was Pia the Pirate's suitcase brown or yellow?*" Fisher's test $> .05$, "*Was Pia the Pirate's hat red or green?*" $\chi^2(1, N = 72) = 1.27, p = .260$, "*I heard that Pia the pirate showed you a pink rabbit, did you see it?*" $\chi^2(1, N = 80) = 0.235, p = .628$.

Details Containing Aggressive Themes

Twelve children (14.3%) mentioned details containing aggressive themes during the substantive phase. The children were equally distributed across the two interviewing protocols (6 children in the NICHD condition; 6 children in the SI condition). Most of these details appear related to stereotypical representations of pirates cutting someone with a sword ($n = 3$, "*The sword stabbed me*", "*She cut someone*", "*A sad story because some died when they cut their throats*"), having guns ($n = 2$, "*She probably had a gun*", "*Pirates can shoot*"), or fighting ($n = 2$, "*They were out going and then fight with stupid pirates*", "*But these were real pirates that we met, real kind also. I have met real that can fight with swords and that are mean and take things*"). One girl (50 months) gave a longer false report about going home to Pia the Pirate with some friends to play games. She said that Pia the Pirate had then punched a hole in the wall with her fist. A boy (56 months) explained to the interviewer that during the pirate visit; "*There was no fighting... there was no hitting ... no pulling in the hair ... no pinching in the ear ... and no pinching in the lip ...*". Another boy (72 months) gave a longer false report about a pirate cutting someone across the throat. The responses containing aggressive themes were preceded by directive wh-questions ($n = 6$, e.g., "*What did she look like?*") and open-ended invitations ($n = 5$, e.g., "*Tell me everything that happened*"). One

EFFECTS OF INTERVIEWING TECHNIQUE

response containing an aggressive theme was elicited after the interviewer had accidentally posed an option-posing question ($n = 1$, “*Was it a funny or sad story?*”).

Performance During the Practice Narrative

Lastly, we also explored whether there were any substantial differences between interviewing conditions on children’s episodic practice narrative reports. Preliminary analyses indicated that there was no significant time difference (in minutes spent talking) during the episodic practice phase in the two conditions, $t(74.43) = -0.21, p = .837$. A Welch t -test indicated that the amount of details during the practice narrative did not differ significantly between children in the NICHD condition ($M = 51.09, SD = 32.34$) and the SI condition ($M = .43.97, SD = 44.80$), $t(68.08) = 0.82, p = .413$, Hedges’ $g = 0.18, 95\% CI [-0.25, 0.61]$. Most children were highly accurate during the practice narrative phase in both the NICHD condition ($M_{raw} = .94, SD = .20$) and the SI condition ($M_{raw} = .91, SD = .23$). A second Welch t -test with logit transformed accuracy rates revealed no significant difference between conditions, $t(72.71) = 0.59, p = .554$, Hedges’ $g = 0.13, 95\% CI [-0.31, 0.58]$.

Conclusions

We found no difference in the amount of reported details between the interviewing conditions. However, we did observe a higher accuracy rate among children interviewed with the NICHD protocol compared to the SI. Age was a significant predictor for children’s amount of detail and accuracy rate. Some children embedded worrisome false details containing aggressive themes in their reports regardless of the interviewing method used.

Experiment II

The second experiment examined whether children’s accounts of a non-experienced event would differ as a function of interviewing technique (NICHD protocol vs. SI). Specifically, we aimed to explore whether there would be a difference in terms of the false assent rate (H1) and amount of false details (H2) between the two interview conditions.

EFFECTS OF INTERVIEWING TECHNIQUE

Method

Participants. Approximately 100 children from twenty local preschools were initially scheduled to participate. On the day of their interview, 46 preschoolers had written parental consent and gave verbal assent to participate. One child was excluded from the analyses because of deviations from the interview guidelines. Thus, the final sample consisted of 45 preschoolers (24 girls and 21 boys, $M_{\text{age}} = 64.8$ months, $SD = 8.2$, age range 47 - 81 months) from 17 preschools. Importantly, none of the children had participated in the ‘pirate’ event staged for Exp. I, and their preschool staff confirmed that the children had not taken part in any similar event at the preschool. However, some children (33.3%) attended the same preschool as children in Exp. I and this was therefore something that was taken into consideration during the analyses.

Materials and Procedure

The same procedure was used as in Exp. I with regard to recruitment and the practice narrative event (interacting with detective ‘Clever Clara’). The children were separately interviewed by one of eleven RAs following the same interviewing guidelines as the previous experiment. If a child denied meeting the pirate after the two misleading introductory questions, they took part in a voluntary drawing session conducted for another study on draw-and-talk techniques (their performance during this part was not included in the present analyses). After the interviews, the children were debriefed and participated in a pirate treasure hunt.

Coding

The interviews were video-recorded and transcribed verbatim. Two RAs (unaware of the study hypotheses) coded the verbal content. False assents were coded as whether the child initially agreed that the event had occurred in response to the introductory questions. An additional code was added for situations where a child first assented but on follow-up

EFFECTS OF INTERVIEWING TECHNIQUE

questioning explained that they had not met the pirate. The amount of false details in their accounts included all details pertaining to meeting ‘Pia the Pirate’ and the responses were quantified following the same coding system as the previous experiment. The false details did not include off-topic comments that were not directly related to meeting the pirate. Inter-rater reliability analyses were conducted on 20% of the data. The coders reached an adequate level of inter-rater reliability, Cohen’s $\kappa = .86-.94$ for all codes. Disagreements were solved through discussion.

Results

Fourteen (31.1%) children initially assented to the false event. However, upon follow up questions, six children withdrew their assent and explained that they had not met the pirate⁴. There was no significant difference in the initial false assent rates between children interviewed with SI (8 of 22 children assented) and NICHD protocol (6 of 23 children assented), $\chi^2(1, N = 45) = 0.554, p = .457$. Five children (SI condition; 3 children, NICHD condition; 2 children) who initially assented to meeting the pirate belonged to a preschool where other children had participated in the pirate event. Thus, there is a possibility that these children had heard their peers talk about the event. Importantly, the inclusion or exclusion of these participants did not change the outcome of the statistical test. Across all interviews, the children’s reports contained an average of 17.5 false details ($SD = 44.6$, range 0–185), with a mean of 21.4 ($SD = 52.8$, range 0–185) in the SI condition and a mean of 13.8 ($SD = 35.7$, range 0–155) in the NICHD condition. However, there were large floor effects in both conditions with most of the children reporting no false details (31 of 45 children). Among the children who did report false details, the distribution was as follows: 3 children reported 1 false detail, 2 children reported between 3–4 false details, 2 children reported between 10–11

⁴ The children typically withdrew their false assent after one ($n = 3$ children) or two ($n = 2$ children) free-recall requests asking the child to describe everything that happened. One child withdrew her false assent after five open-ended prompts.

EFFECTS OF INTERVIEWING TECHNIQUE

false details, and 7 children reported more than 40 false details (41, 53, 64, 117, 143, 155, and 185 false details). Due to the substantial floor effects in the amount of false details, we were unable to carry out any inferential tests on this variable.

Among the children who assented to meeting the pirate, seven (15.6%) gave a longer false report (>40 details) about the non-experienced event with a mean of 108.3 false details ($SD = 56.1$, range 41–185, $Mdn = 117$). Among the false reporters, only one child belonged to a preschool where other children had participated in Exp I. There was no difference in the frequency of the false reports between the SI (3 of 22 children) and NICHD condition (4 of 23 children), Fishers' exact test = 1. Six of the seven children who gave a longer false report were girls. In response to the scripted questions, the children varied in their responses. To the question regarding whether someone had taken a photo during the pirate event, only two out of the seven participants (28.6%) acquiesced. However, six children (85.7%) tried to guess the colour of the suitcase and four children (57.1%) guessed the colour of the pirate's hat. In response to the leading question regarding whether the pirate had a pink rabbit, two children (28.6%) agreed and provided more details about the fictive rabbit.

With regard to the content of the false reports, six out of seven children placed themselves within the narrative of meeting the fictional pirate (e.g., describing what they did or felt) and five children described that other people (e.g., friends, parents) were involved in the event. All seven children reported at least one stereotypical detail about pirates (e.g., that the pirate wore a black hat, sailed a pirate ship, had a sword or a gun, and searched for treasures) Only two children gave longer false reports clearly containing fantastical elements, including for example sailing a pirate ship and hunting for a buried treasure together with the pirate. The other five children reported events containing plausible details. These included rich false reports about meeting an adult dressed up as 'Pia the Pirate' at their preschool ($n = 3$), which both the children's parents and preschool teachers confirmed had not occurred. The

EFFECTS OF INTERVIEWING TECHNIQUE

other two children described events that may reflect source monitoring errors. These included reports about playing with a pirate at a birthday party ($n = 1$) and eating ice-cream with a pirate at a playground ($n = 1$).

Conclusions

About one third of all children falsely agreed to having met a fictional pirate character and 15.6% provided false narratives (more than 40 words) when asked open-ended follow-up questions. The distribution of false assents and longer false reports were equally distributed across the SI and NICHD conditions.

General Discussion

The current study compared effects of the pre-substantive phase in two different interviewing techniques (SI vs. NICHD protocol) when questioning preschool-aged children. A secondary aim was to study age differences among preschoolers. As expected, and in line with past studies (e.g., Alexander et al., 2002; Melinder et al., 2006), older preschoolers provided more details and exhibited higher accuracy rates compared to younger preschoolers. While the children did not differ across interview conditions, in the amount of reported details, preschoolers interviewed with the SI exhibited a slightly lower accuracy rate when describing a self-experienced event. However, it should be noted that this effect size was small, and that the observed difference may be explained by differences among a smaller subset of children. Specifically, a few more children in the SI condition exhibited accuracy rates towards the lower end of the distribution compared to children in the NICHD condition. Possibly, this trend could be explained by fatigue effects. Although children in the SI condition did have a short break before the substantive phase of their interviews, this may not have been enough to mitigate the potential negative effects on young children's attention from the prolonged rapport building phase (see also Davies et al., 2000; Teoh & Lamb, 2010). It is important to note that the SI typically advises using a longer break after the first

EFFECTS OF INTERVIEWING TECHNIQUE

session (around 45–60 minutes; Langballe & Davik, 2017) than was possible to carry out in the present experiments. Hence, it cannot be excluded that the results would be different in interviews with a longer break.

Furthermore, the current design focused on examining differences on children's recall accuracy following the entire pre-substantive phase within each of the two interviewing techniques. It is therefore difficult to draw any firm conclusions about the specific component (or combination of components) that accounted for the accuracy effect observed in the first experiment. Specifically, the SI condition differed in several ways (i.e. longer rapport building phase including props, the use of a break after the pre-substantive phase) that may have impacted the children's response monitoring during the substantive phase. Follow-up studies are needed to study the effects of each component in isolation. Interestingly, our exploratory analyses of the children's responses to the episodic practice narrative phase indicated that there were no significant differences between the two techniques. At that point of the interviews, the only difference between the two techniques included the use of a jigsaw puzzle task in the SI condition (see Table 1). This indicates that the difference between conditions may have emerged after the episodic practice narrative, potentially from fatigue after the longer rapport-building phase in the SI condition.

Concerns regarding fatigue effects were also recently reported by forensic child interviewers, who described a need to abbreviate the pre-substantive phase of the NICHD protocol when interviewing preschoolers (Magnusson et al., 2020a). Since both the SI and the revised NICHD protocol (Lamb et al., 2018) advocates for increased early rapport building, further investigation is needed on the effects of different rapport building strategies during interviews with young children (Saywitz et al., 2015). Extended rapport building during the initial phase of child interviews may be particularly important to overcome reluctance among abused child victims (e.g., Herskowitz et al., 2015). However, participants in the current

EFFECTS OF INTERVIEWING TECHNIQUE

experiment were given no external motivation for concealing the target event during their interviews. Potentially, extending the pre-substantive phase may have positive effects for children who are initially reluctant to disclose information. Future research could take this observation into consideration by, for example, including elements of secrecy into the experimental design (see Magnusson et al., 2020b).

When interviewed about a non-experienced event, about one third of the children incorrectly acquiesced to meeting a fictional pirate and some gave a false report about the event. There was no difference between interviewing conditions in the amount of false assents. These findings are concerning considering that the interview guides were adapted from techniques currently used by practitioners (see also Brown et al., 2013). However, it should be noted that we included two suggestive introductory questions (“*I heard that you met Pia the Pirate, do you remember that?*” “*Clever Clara said that a pirate visited you at your preschool, tell me about that*”) to direct the conversation to the pirate visit. These topic prompts could be considered leading for the children who had experienced the event, and misleading for the children who had not experienced the event. Hence, these findings demonstrate the problematic nature of trying to steer the conversation towards the issue under investigation without being suggestive (for more information on this topic, see Earhart, Dandy, Brubacher, Powell, & Sharman, 2018). Even with relatively minor forms of suggestive influence (followed by open-ended requests to elaborate), some preschoolers who had not met the pirate provided false information that appeared largely based on stereotypes about pirates (see Ceci & Bruck, 1995, for more information about the risks of stereotype induction). This is problematic as real-life forensic interviewers need to address the suspicion under investigation in some way and calls for careful hypothesis-testing during child interviews (Korkman, Pakkanen, & Laajasalo, 2017).

EFFECTS OF INTERVIEWING TECHNIQUE

Furthermore, independent of interviewing technique, some children who had met the pirate reported false details containing aggressive themes including the pirate fighting, stabbing, and hitting children. These critical, false details were embedded in otherwise accurate reports about self-experienced events. Notably, the details containing aggressive themes were mainly elicited from recommended question types including open-ended directive questions and invitations. The phenomenon occurred among a minority of children who were interviewed with either the NICHD protocol or the SI model (with no differences in frequency between interviewing conditions). Although embedded false details could be problematic from an applied perspective, there is a paucity of research on this topic in interviewing and deception literature (Verigin, Meijer, & Vrij, 2020). Embedded false information, such as claims about hitting or fighting, may have severe consequences for the continued direction of an investigation, particularly if other details within the report could be corroborated by other information in the case. Like the findings from Exp II, the currently observed pattern relating to the aggressive details in Exp I may have been related to stereotypes about pirates. Still, we believe future research may benefit from focusing on the content of children's false details.

Some methodological concerns need to be addressed. First, due to recruitment issues the statistical power was limited. However, even with the limited sample, it is concerning that some children provided a false narrative about the event regardless of whether they were interviewed with the SI or NICHD protocol. Second, the external validity is limited as laboratory research cannot capture the complex factors involved in criminal cases with child witnesses. The current paradigm enabled us to examine preschoolers' accuracy when describing an interactive staged event including multiple components (i.e. interacting with unfamiliar adults, reading a storybook, watching different objects hidden inside a suitcase, dressing up as pirates, and being photographed; see also Brown et al., 2013). As the false

EFFECTS OF INTERVIEWING TECHNIQUE

details sometimes comprised stereotypical information about pirates, future studies may benefit from using other encoding paradigms to avoid unintended stereotype induction (Ceci & Bruck, 1995) or source monitoring errors with children recounting other pirate-related experiences (Johnson et al., 1993). Specifically, the imaginary nature of the current paradigm could have affected the children's report accuracy. The use of a fun interactive event may also have limited the children's report monitoring compared to situations with more distressing events.

Third, the interviews were conducted by research assistants who received a two-day training course in child interviewing. In real settings, forensic interviewers are recommended to attend extensive specialist training coupled with continuous feedback and supervision (Powell, 2008). Lastly, we did not measure the level of rapport experienced by the children and interviewers. Future research may benefit from integrating techniques aimed at capturing direct effects of different rapport building techniques across interviews (see e.g., Johnston, Brubacher, Powell, & Fuller-Tyszkiewicz, 2019). As reduced anxiety levels are one of the presumed benefits from extended rapport building, researchers may also want to include state anxiety measures at different points of the interview (see e.g., Almerigogna, Ost, Bull, & Akehurst, 2007).

Notwithstanding these limitations, the present study provides valuable insights into the investigative interviewing of preschool-aged children. When comparing the pre-substantive phase of two child interviewing techniques currently used by Scandinavian practitioners, we found few significant differences on children's recall of a staged event. While many children provided accurate testimony when interviewed with either protocol, some preschoolers gave false reports or embedded worrisome false details in their statements. Taken together, preschoolers interviewed with the novel sequential interviewing approach (used by Norwegian police interviewers) performed similarly to preschoolers interviewed

EFFECTS OF INTERVIEWING TECHNIQUE

with the well-established NICHD protocol. The only observable difference between techniques concerned the slightly lower accuracy rate in the sequential interview condition, which might have derived from fatigue following a prolonged rapport building phase. As both the sequential interview model and the revised NICHD protocol focuses on extended rapport building, future studies may benefit from examining the use of longer breaks after the pre-substantive phase. Furthermore, it is important to note that the current study only manipulated the pre-substantive phase of the two interview protocols. Other aspects of the two techniques were thus not included in the current design. More laboratory research is required to further understand the effectiveness and limitations of different interviewing techniques with young children. Considering that legal practitioners have described cases involving preschoolers as particularly challenging (Ernberg et al., 2020; Magnusson et al., 2020a), greater efforts are needed to ensure that current interview guidelines are applicable to this group of young, vulnerable witnesses.

References

- Alexander, K. W., Goodman, G. S., Schaaf, J. M., Edelstein, R. S., Quas, J. A., & Shaver, P. R. (2002). The role of attachment and cognitive inhibition in children's memory and suggestibility for a stressful event. *Journal of Experimental Child Psychology, 83*, 262-290. [http://doi.org/10.1016/S0022-0965\(02\)00149-2](http://doi.org/10.1016/S0022-0965(02)00149-2)
- Almerigogna, J., Ost, J., Bull, R., & Akehurst, L. (2007). A state of high anxiety: How non-supportive interviewers can increase the suggestibility of child witnesses. *Applied Cognitive Psychology, 21*, 963-974. <http://doi.org/10.1002/acp.1311>
- Baugerud, G-A., Johnson, M. S., Hansen, H. B. G., Magnussen, S., & Lamb, M. E. (2020). Forensic interviews with preschool children: An analysis of extended interviews in Norway (2015-2017). *Applied Cognitive Psychology*, advanced online publication. <http://doi.org/10.1002/acp.3647>
- Brown, D. A., Lamb, M. E., Lewis, C., Pipe, M. E., Orbach, Y., & Wolfman, M. (2013). The NICHD investigative interview protocol: An analogue study. *Journal of Experimental Psychology Applied, 19*, 367-382. <http://doi.org/1037/a0035143>
- Brubacher, S. P., Peterson, C., La Rooy, D., Dickinson, J. J., Poole, D. A. (2019). How children talk about events: Implications for eliciting and analyzing eyewitness reports. *Developmental Review, 51*, 70-89. <http://doi.org/10.1016/j.dr.2018.12.003>
- Bruck, M., & Melnyk, L. (2004). Individual differences in children's suggestibility: A review and synthesis. *Applied Cognitive Psychology, 18*, 947-996. <https://doi.org/10.1002/acp.1070>
- Camparo, L. B., Wagner, J. T., & Saywitz, K. J. (2001). Interviewing children about real and fictitious events: Revisiting the narrative elaboration procedure. *Law & Human Behavior, 25*, 63-80. <http://doi.org/10.1023/A:1005691926064>

EFFECTS OF INTERVIEWING TECHNIQUE

- Carnes, C. N., Wilson, C., & Nelson-Gardell, D. (1999). Extended forensic evaluation when sexual abuse is suspected: A model and preliminary data. *Child Maltreatment, 4*, 242–254. <https://doi.org/10.1177/1077559599004003005>
- Ceci, S. J., & Bruck, M. (1993). Suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin, 113*, 403–439. <http://doi.org/10.1037/0033-2909.113.3.403>
- Ceci, S. J., & Bruck, M. (1995). *Jeopardy in the courtroom: A scientific analysis of children's testimony*. American Psychological Association. <https://doi.org/10.1037/10180-000>
- 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. <https://doi.org/10.1080/10683160008410834>
- Earhart B., Danby M. C., Brubacher S. P., Powell M. B., Sharman S. (2018). A comparison of responses to substantive transition prompts in interviews with children. *Child Maltreat, 23* 221–225. <http://doi.org/10.1177/1077559518756827>
- Ernberg, E., Magnusson, M., & Landström, S. (2020). Prosecutors' experiences investigating alleged sexual abuse of preschoolers. *Psychology, Crime, & Law*. Advanced online publication. <https://doi.org/10.1080/1068316X.2019.1709639>
- Faller, K. C. (2015). Forty years of forensic interviewing of children suspected of sexual abuse, 1974-2014: Historical benchmarks. *Social Sciences, 4*, 34-65. <http://doi.org/10.3390/socsci4010034>
- Faul, F., Erdfelder, E., Lang, A-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioural, and biomedical sciences. *Behavior Research Methods, 39*, 175-191. <http://doi.org/10.3758/BF03193146>
- Fivush, R. (2011). The development of autobiographical memory. *Annual Review of Psychology, 62*, 559-582. <http://doi.org/10.1146/annurev.psych.121208.13>

EFFECTS OF INTERVIEWING TECHNIQUE

- Gagnon, K., & Cyr, M. (2017). Sexual abuse and preschoolers: Forensic details in regard of question types. *Child Abuse & Neglect*, *67*, 109–118.
<http://doi.org/10.1016/j.chiabu.2017.02.022>
- Gamst, K. T., & Langballe, Å. (2004). *Barn som vitner: En empirisk og teoretisk studie av kommunikasjon mellom avører og barn i dommeravhør: Utvikling av en avhørsmetodisk tilnærming*. [Children who testify: An empirical and theoretical study of the communication between interviewer and child during forensic interviews: The development of an interviewing approach] (Doctoral dissertation). University of Oslo.
- Goodman, G. S., & Reed, R. S. (1986). Age differences in eyewitness testimony. *Law & Human Behavior*, *10*, 317-332. <http://doi.org/10.1007/BF01047344>
- Hershkowitz, I., Lamb, M. E., Katz, C., & Malloy, L. C. (2015). Does enhanced rapport-building alter the dynamics of investigative interviews with suspected victims of intra-familial abuse? *Journal of Police and Criminal Psychology*, *30*, 6-14.
<https://doi.org/10.1007/s11896-013-9136-8>
- Johnson, M. K., Hashtroudi, S., Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*, 3-28. <http://doi.org/10.1037/0033-2909.114.1.3>
- Johnston, V., Brubacher, S. P., Powell, M., & Fuller-Tyszewicz, M. (2019). Patterns of nonverbal rapport behaviors across time in investigative interviews with children. *Journal of Nonverbal Behavior*, *42*, 411-434. <http://doi.org/10.1007/s10919-019-00206-1>
- Korkman, J., Antfolk, J., Fagerlund, M., & Santtila, P. (2018). The prevalence of unfounded suspicions of child sexual abuse in Finland. *Nordic Psychology*, *71*, 39-50.
<http://doi.org/10.1080/19012276.2018.1470554>
- Korkman, J., Pakkanen, T., & Laajasalo, T. (2017). Child forensic interviewing in Finland: Investigating suspected child abuse at the forensic psychology unit for children and

EFFECTS OF INTERVIEWING TECHNIQUE

- adolescents. In S. Johansson, K. Stefansen, E. Bakketeig, & A. Kaldal (Eds.), *Collaborating against child abuse: Exploring the Nordic barnahus model* (pp. 145-164). Springer International Publishing. <http://doi.org/10.1007/978-3-319-58388-4>
- Lamb, M. E, Hershkowitz, I., Orbach, Y., & Esplin, P. W. (2008). *Tell me what happened: Structured investigative interviews of child victims and witnesses* (1st ed.). Wiley. <http://doi.org/10.1002/9780470773291>
- Langballe, Å., & Davik, T. (2017). Sequential interviews with preschool children in Norwegian barnahus. In S. Johansson, K. Stefansen, E. Bakketeig, & A. Kaldal (Eds.), *Collaborating against child abuse: Exploring the Nordic barnahus model* (pp. 165-183). Springer International Publishing. <http://doi.org/10.1007/978-3-319-58388-4>
- La Rooy, D., Brubacher, S. P., Aromäki-Stratos, A., Cyr, M., Hershkowitz, I., Korkman, J., ... Lamb, M. E. (2015). The NICHD protocol: A review of an internationally-used evidence-based tool for training child forensic interviewers. *Journal of Criminological Research, Policy and Practice*, 1, 76-89.
- Magnusson, M., Ernberg, E., Landström, S., & Akehrst, L. (2020a). Forensic interviewers' experiences interviewing children of different ages. *Psychology, Crime & Law*. Advanced online publication. <https://doi.org/10.1080/1068316X.2020.1742343>
- Magnusson, M., Ernberg, E., Landström, S., Joleby, M., & Akerhurst, L. (2020b). Can rapport building strategies, age, and question type influence preschoolers' disclosures of adult wrongdoing? *Scandinavian Journal of Psychology*, 61, 393-401. <http://doi.org/10.1111/sjop.12626>
- Melinder, A., Endestad, T., & Magnussen, S. (2006). Relations between episodic memory, suggestibility, theory of mind and cognitive inhibition in the preschool child.

EFFECTS OF INTERVIEWING TECHNIQUE

Scandinavian Journal of Psychology, 47, 485-495. <http://doi.org/10.1111/j.1467-9450.2006.00542>.

Melinder, A., Magnusson, M., & Gilstrap, L. L. (2020). What Is a child-appropriate interview? Interaction between child witnesses and police officers. *International Journal on Child Maltreatment: Research, Policy, and Practice*. Advanced online publication. <http://doi.org/10.1007/s42448-020-00052-8>

Newlin, C., Steele, L. C., Chamberlin, A., Anderson, J., Kenniston, J., Russell, A., ... Vaughan-Eden, V. (2015). Child forensic interviewing: Best practices. *Juvenile Justice Bulletin*. U.S. Department of Justice. Retrieved from <https://www.ojjdp.gov/pubs/248749.pdf>

Otgaar, H., Howe, M. L., Merckelbach, H., & Muris, P. (2018). Who is the better eyewitness? Sometimes adults but at other times children. *Current Directions in Psychological Science*, 27, 378-385. <http://doi.org/10.1177/0963721418770998>

Poole, D. A., Dickinson, J. J., & Brubacher, S. P. (2014). Sources of unreliable testimony from children. *Roger Williams University Law Review*, 19, 382-410.

Powell, M. B. (2008). Designing effective training programs for investigative interviewers of children. *Current Issues in Criminal Justice*, 20, 189-208. <http://doi.org/10.1080/10345329.2008.12035804>

Powell, M. B., Jones, C. H., & Campbell, C. (2003). A comparison of preschoolers recall of experienced versus non-experienced events across multiple interviews. *Applied Cognitive Psychology*, 17, 935-952. <http://doi.org/10.1002/acp.932>

Sauerland, M., Brackmann, N., & Otgaar, H. (2019). Rapport: Little effect on children's, adolescents', and adults' statement quantity, accuracy, and suggestibility. *Journal of Child Custody*, 15, 268-285. <http://doi.org/10.1080/15379418.2018.1509759>

EFFECTS OF INTERVIEWING TECHNIQUE

Saywitz, K. J., Larson, R. P., Hobbs, S., & Wells, C. R. (2015). Developing rapport with children in forensic interviews: Systematic review of experimental research.

Behavioural Sciences and the Law, 33, 372-389. <http://doi.org/10.1002/bsl.2186>

Saywitz, K. J., Lyon, T. D., & Goodman, G. S. (2018). Interviewing Children. In J. Conte & B. Klika (Eds.), *APSAC Handbook on Child Maltreatment*. (4th ed. pp-310-329).

Newbury Park, CA: Sage.

Teoh, Y.-S., & Lamb, M. E. (2010). Preparing children for investigative interviews: Rapport-building, instruction, and evaluation. *Applied Developmental Science*, 14, 154–163.

<https://doi.org/10.1080/10888691.2010.494463>

Verigin, B. L., Meijer, E. H., & Vrij, A. (2020). Embedding lies into truthful stories does not affect their quality. *Applied Cognitive Psychology*, 34, 516-525.

<http://doi.org/10.1002/acp.3642>

EFFECTS OF INTERVIEWING TECHNIQUE

Table 1

Interview guides adapted from the NICHD protocol (Lamb et al., 2008) and the SI model (Langballe & Davik, 2017)

Interview phase	NICHD protocol	SI model
Pre-substantive phase	Introduction (1 min) Ground rules (1-2 min) Questions about interests (1-3 min) Episodic narrative practice (2-3 min)	Introduction (1 min) Ground rules (1-2 min) Questions about interests (1-3 min) Jigsaw puzzle task (3-5 min) Episodic narrative practice (2-3 min) Break (5 min)
Substantive phase	Free recall Specific questions	Free recall Specific questions

EFFECTS OF INTERVIEWING TECHNIQUE

Figure 1

Overview of staged events in Exp. 1

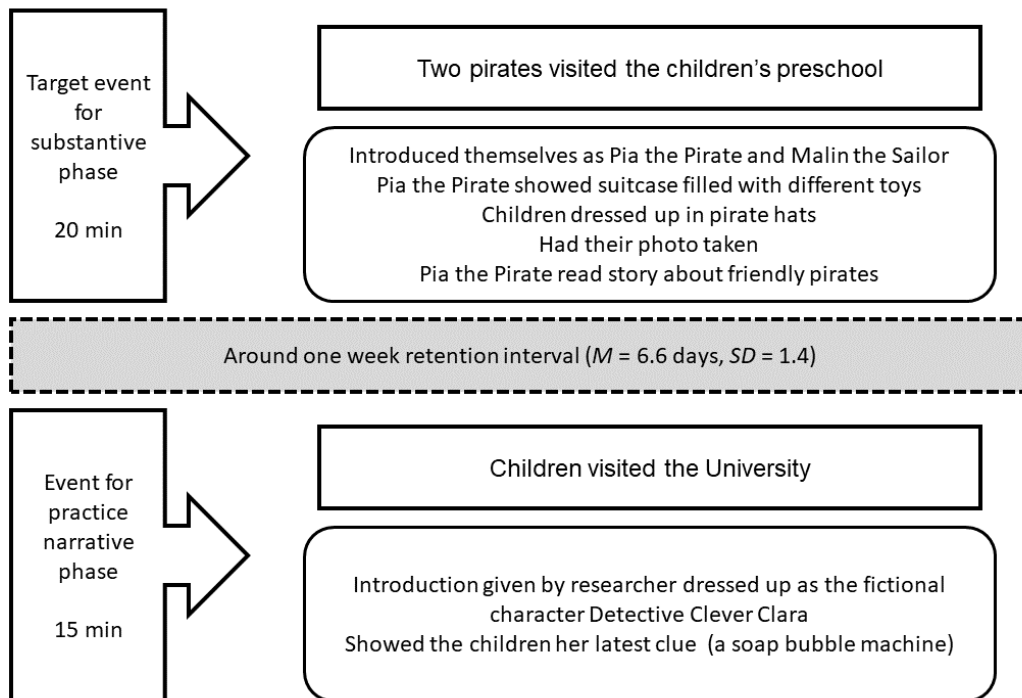
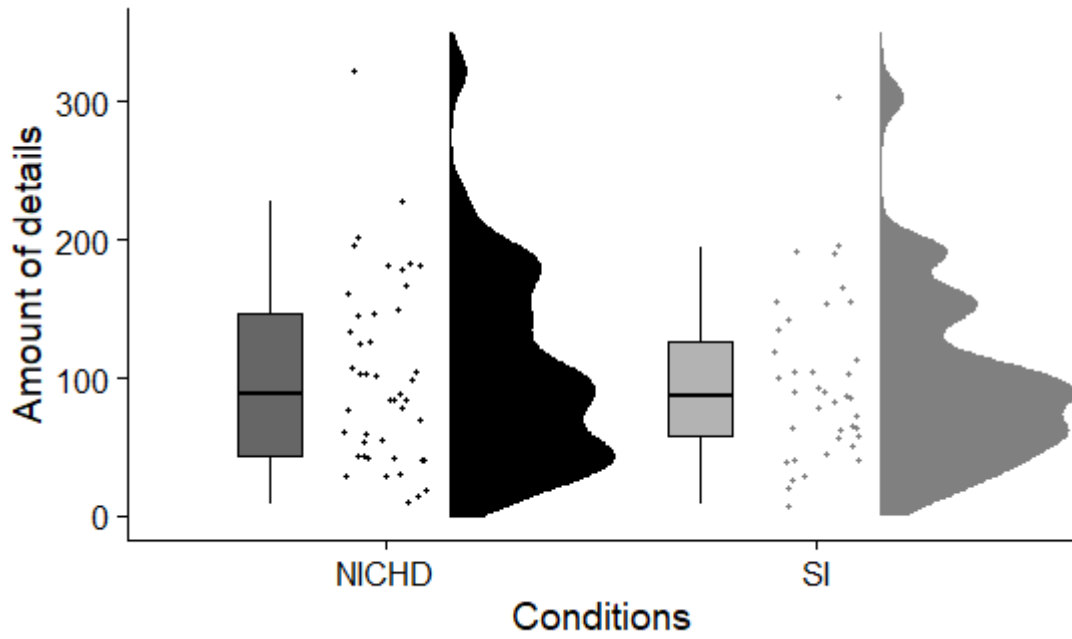


Figure 2

Raincloud Plot of the Distribution of Amount of Detail (Raw Data) in the NICHD and SI Conditions (Exp. 1)

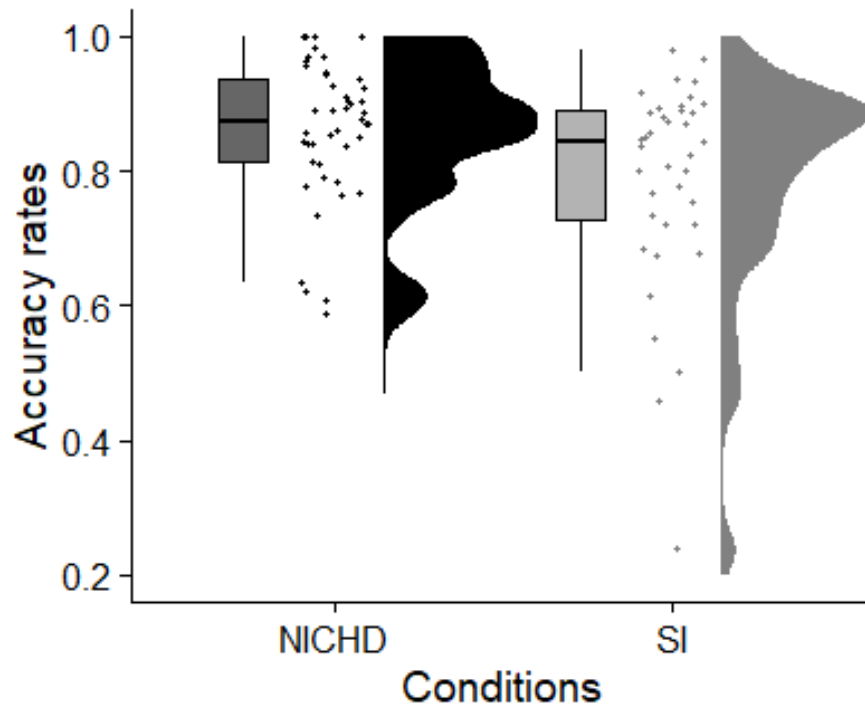


EFFECTS OF INTERVIEWING TECHNIQUE

Figure 3

Raincloud Plot of the Distribution of Accuracy Rates (Raw Data) in the NICHD and SI

Conditions (Exp. I)



EFFECTS OF INTERVIEWING TECHNIQUE

Table 2

Proportions of Response Types to the Scripted Questions Divided by Interview Condition (Exp. I)

Scripted question	NICHD protocol			SI model		
	Correct	Incorrect	Non-responsive	Correct	Incorrect	Non-responsive
<i>“Did anyone take a picture of you?” (leading yes/no question)</i>	82.22%	15.56%	2.22%	76.32%	18.42%	5.26%
<i>“Was Pia the Pirate’s suitcase brown or yellow?” (option-posing question containing one correct and one incorrect option)</i>	86.67%	6.67%	6.67%	92.31%	5.13%	2.56%
<i>“Was Pia the Pirate’s hat red or green?” (option-posing question containing two incorrect options)</i>	61.36%	25%	13.64%	73.68%	15.79%	10.53%
<i>I heard that Pia the pirate showed you a pink rabbit, did you see it?” (misleading yes/no question)</i>	79.55%	20.45%	0%	71.05%	23.68%	5.26%