



Beyond Field Studies: An analogue comparison of NICHD and MoGP forensic interview guidelines in children's testimony

Mario Pérez-Cano^{1*}, Ángel Prieto Redín¹, Manuel García Sánchez², and Judith Meseguer Aragón¹

¹ Universidad de Murcia, Spain

² UCAM Universidad Católica de Murcia, Spain

Título: Más allá de los estudios de campo: una comparación análoga de las directrices de entrevista forense NICHD y MoGP en el testimonio infantil.

Resumen: Este estudio análogo comparó la eficacia del Protocolo NICHD y el Memorandum de Buenas Prácticas (MoGP) en la obtención de información precisa y relevante desde el punto de vista forense en entrevistas con menores. Participaron 123 niños (60 de 6–7 años y 63 de 10–11 años) que asistieron a un evento simulado, seguido de una entrevista retrasada realizada conforme a uno de los dos protocolos. Los entrevistadores, estudiantes de posgrado, recibieron formación específica según el protocolo asignado. Las entrevistas se codificaron según el tipo de pregunta y la cantidad, precisión y relevancia de los detalles. Los modelos lineales mixtos revelaron que las entrevistas con el protocolo NICHD generaron más detalles totales, correctos, centrales y periféricos, especialmente mediante invitaciones e invitaciones guiadas. Las preguntas cerradas fueron las menos eficaces. Las diferencias de edad en precisión y productividad disminuyeron al controlar la informatividad de las preguntas. Sin embargo, las preguntas cerradas se usaron con frecuencia en ambos protocolos. Los hallazgos destacan la importancia de la formulación de las preguntas, las ventajas del protocolo NICHD y la necesidad de mejorar la formación de los entrevistadores para reducir el uso de formatos menos eficaces y promover buenas prácticas en entrevistas investigativas.

Palabras clave: Entrevista forense. Abuso sexual infantil. Protocolo NICHD. Tipo de pregunta. Precisión de la entrevista.

Abstract: This analogue study compares the effectiveness of two forensic interview protocols—the NICHD Protocol and the Memorandum of Good Practice (MoGP)—in eliciting accurate and forensically relevant information from children. A total of 123 children (60 aged 6–7 and 63 aged 10–11) participated in a staged event, followed by a delayed interview conducted using either the NICHD or MoGP guidelines. Interviewers were postgraduate students who received protocol-specific training. Transcripts were coded for prompt type and the amount, accuracy, and relevance of elicited details. Generalized linear mixed models revealed that interviews conducted using the NICHD Protocol elicited more total, correct, central, and peripheral details, particularly through invitations and cued invitations. Option-posing prompts were the least effective, yielding the fewest correct and peripheral details. Age-related differences in productivity and accuracy were attenuated after controlling for the informativeness of prompt types. Option-posing questions were frequently used in both conditions, including NICHD-guided interviews. These findings highlight the critical role of prompt formulation in shaping children's narrative responses and underscore the relative advantages of the NICHD Protocol. They further suggest that enhanced interviewer training is needed to reduce reliance on less effective prompt types and to promote best practices in investigative interviewing.

Keywords: Forensic interview. Child sexual abuse. NICHD Protocol. Question type. Interview accuracy.

Introduction

According to the World Health Organization, one in five women and one in seven men report having experienced sexual abuse during childhood (World Health Organization, 2024). In Spain, the prevalence of child sexual abuse (CSA) before the age of 18 among university students is 17.9% (Pereda & Forns, 2007). When such cases are brought before the courts, the absence of objective evidence often makes children's statements central to legal proceedings (London et al., 2005, 2007, 2008; Principe & London, 2022). This poses a significant challenge for professionals, as children can provide forensically relevant testimony (Benia et al., 2015; Hershkowitz et al., 2014; Orbach & Lamb, 2000) but they may also make errors in their reports (Lamb et al., 2011; Poole & Lindsay, 1995). Therefore, the quality of the forensic interview and its capacity to elicit reliable information is critical for the integrity of criminal investigations (Herman, 2010; Hershkowitz & Lamb, 2020; Yi & Lamb, 2018).

Collecting Testimony from Child Sexual Abuse Victims

Most professional guidelines for collecting testimony from CSA victims emphasize the importance of eliciting information through open-ended prompts (Bull, 2014; Fisher & Geiselman, 1992; Home Office, 1992; Lyon, 2014; Saywitz & Camparo, 2014). A recent review (Fernandes et al., 2024) found that most forensic interviewing protocols for alleged abuse follow a similar structure, typically comprising three components: (1) free narrative recall of the alleged incidents, (2) a questioning phase to elicit additional details, and (3) a closure phase, during which the child is invited to ask questions and the interview concludes with a neutral topic. An example is the Memorandum of Good Practice on Video-Recorded Interviews with Child Witnesses for Criminal Proceedings (Home Office, 1992). The MoGP provides comprehensive guidelines for conducting child abuse interviews consistent with evidentiary standards in England and Wales, emphasizing the exhaustive use of open-ended questions before resorting to more focused prompts, initiating interviews broadly and narrowing gradually, and minimizing the use of closed or leading questions.

Extensive experimental and field research supports these recommendations. Numerous studies have shown that free-recall prompts elicit more information than closed-ended questions (Lamb et al., 1996; Sternberg et al., 1996; Stern-

* Correspondence address [Dirección para correspondencia]:

Mario Pérez-Cano. E-mail: mario.perez.1204@gmail.com

(Article received: 17-04-2025; reviewed: 27-05-2025, accepted: 16-07-2025)

berg, Lamb, Orbach, et al., 2001) and that the information obtained through such prompts is more likely to be accurate (Brown et al., 2013; Dent, 1986; Dent & Stephenson, 1979; Hutcheson et al., 1995; Lamb, Orbach, Hershkowitz, Horowitz, et al., 2007; Lamb & Fauchier, 2001; Oates & Shrimpton, 1991; Orbach & Lamb, 2001). However, forensic practitioners frequently deviate from these guidelines, relying instead on closed or suggestive questions (Aldridge & Cameron, 1999; Job et al., 2025; Korkman et al., 2006; Powell & Hughes-Scholes, 2009). Sternberg, Lamb, Davies, et al. (2001), for example, observed that practitioners trained under the MoGP framework used open-ended prompts infrequently, comprising only 6% of their utterances, with most information elicited through closed questions.

NICHD Protocol

The National Institute of Child Health and Human Development (NICHD) interview protocol (Hershkowitz et al., 2015; La Rooy et al., 2015; Lamb, Orbach, Hershkowitz, Esplin, et al., 2007; Orbach et al., 2000) is the most widely cited guideline for interviewing CSA victims (Fernandes et al., 2024). This structured protocol employs memory-retrieval and communication techniques designed to minimize suggestibility and enhance the reliability of children's accounts. Detailed guidelines are provided for each interview stage, specifying appropriate question types. In the substantive phase, the interviewer is expected to begin with an open invitation (e.g., "Tell me everything that happened from the beginning to the end as best you can remember") and other free-recall prompts or invitations (e.g., "Tell me more about that"); (2) elicit incident-specific information using follow-up and cued invitations that reference details previously mentioned by the child (e.g., "You said you wore a costume. Tell me more about that"); (3) after exhausting recall-based prompts, address previously mentioned details and gather additional information using directive *Wh-* questions (e.g., about time, appearance); and (4) ask limited option-posing questions regarding new but crucial information that has not yet been reported (Lamb, Orbach, Hershkowitz, Esplin, et al., 2007). Suggestive utterances are strongly discouraged, and interviewers must return to invitations after using focused question types. NICHD Protocol also includes examples and detailed explanations of how to implement each step appropriately (Lamb et al., 2018).

Studies comparing the NICHD Protocol with other interviews have primarily been field studies analyzing investigative interviews from real cases involving alleged child sexual abuse victims (Ahern & Lamb, 2017; Cyr & Lamb, 2009; Lamb et al., 2009). In these types of studies, the NICHD Protocol has consistently been shown to improve the quality of forensic interviews (Job et al., 2025; Lamb, Orbach, Hershkowitz, Horowitz, et al., 2007). A meta-analytic review concluded that interviewers trained in the NICHD Protocol ask more invitations, and that children provide more central details in response to such prompts (Benia et al., 2015).

When the NICHD Protocol has been directly compared to interviews conducted using other evidence-based guidelines, such as the MoGP, findings have shown that NICHD interviews elicit more information through free-recall invitations and less through directive, option-posing, and suggestive questions than MoGP interviews (Lamb et al., 2009; Ahern & Lamb, 2017).

Although analogue studies with children have examined the effectiveness of different question types (Dent & Stephenson, 1979; Henry & Gudjonsson, 1999, 2003) or mnemonic techniques (Milne & Bull, 2002; Verkamp & Ginet, 2010), there is currently limited research on the accuracy of information elicited using the NICHD Protocol. To date, to our knowledge, only Brown et al. (2013) have evaluated the NICHD Protocol in an analogue setting. In their study, 128 children aged 5 to 7 were interviewed about a staged event that occurred 4 to 6 weeks earlier. All children were interviewed using the NICHD Protocol, with the only variation occurring in the rapport. Results showed that invitations and cued invitations elicited more details than directive questions, and that errors were more likely to occur in response to directive questions.

The Present Study

Given the limited analogue research directly comparing structured forensic interview protocols, the current study aims to compare the NICHD Protocol with MoGP. Although MoGP has been examined in field comparisons with NICHD (Ahern & Lamb, 2017; Lamb et al., 2009), to our knowledge, no analogue study has directly contrasted these protocols. Additionally, most NICHD studies have been conducted in field settings and with samples of children undergoing evaluation due to suspected child sexual abuse (for a meta-analysis, see Benia et al., 2015). Consequently, this type of study presents two main weaknesses. First, it cannot be established with certainty that all the interviewed children had abusive experiences to disclose, and second, there is no possibility of verifying the accuracy of the information they provide. In contrast, analogue methodologies offer a higher degree of experimental control. The present design allowed us to compare protocols under standardized conditions and not only to examine the amount of information elicited but also to assess error rates and levels of accuracy with precision, as the ground truth of the staged events was fully documented. This study assesses both the quantity and accuracy of children's responses to various prompt types: invitations, cued invitations, directives, option-posing, and suggestive prompts.

Following Brown et al. (2013), in our study, interviewers were provided with a list of specific topics to explore if not spontaneously mentioned by the child. These included games played, actions and interactions with the facilitator, the facilitator's identity, and relevant objects. Interviewers were instructed to elicit this information primarily through open-ended questions. This practice aligns with NICHD

recommendations for the minimal and strategic use of directive and option-posing questions.

Additionally, two design features were implemented with the aim of mitigating potential carry-over effects associated with interviewers' increasing familiarity with the event content. First, a relatively large pool of interviewers was employed (six per condition). Second, each child individually experienced a unique event consisting of three out of 14 possible activities, counterbalanced across participants. Accordingly, and to preserve the independence of each interview, interviewers were explicitly instructed to disregard narratives from prior sessions when conducting subsequent ones. Within our experimental setting, it was difficult to fully ensure the independence of each child's account, apart from instructing participants not to discuss their experiences with peers. Nevertheless, we found no indication that this factor differentially affected the two experimental conditions.

It was hypothesized that NICHD-based interviews would employ more invitations and fewer directive, option-posing, and suggestive questions than MoGP interviews. It was further anticipated that NICHD would yield greater overall information, particularly through open-ended prompts. Regarding accuracy, invitations were expected to generate the most accurate responses, followed by cued invitations, directive questions, option-posing questions, and suggestive prompts. Given NICHD's emphasis on open-ended strategies, this protocol was predicted to elicit more central information and greater accuracy in retrieving forensically relevant details.

Method

Participants

Children in 1st grade (6–7 years old) and 6th grade (11–12 years old) were recruited from primary schools in the Region of Murcia (Spain). Educational institutions were selected from both urban and peripheral areas and included public and semi-private schools, with the aim of obtaining a sample diverse in terms of sociocultural background. Informed consent was obtained from the parents of 137 children; however, only data from 60 children in 1st grade ($M = 6.28$ years, $SD = 0.45$; 45% female [$n = 27$], 55% male [$n = 33$]) and 63 children in 6th grade ($M = 11.27$ years, $SD = 0.45$; 49.21% female [$n = 31$], 50.79% male [$n = 32$]) were included in the analyses. Most participants were born in Spain (77.24%), while 22.76% were born in other countries, including Morocco (13.82%), several Latin American countries (7.32%), and South Africa (1.63%). Exclusion reasons included technical recording problems ($n = 3$), refusal to participate ($n = 2$), denial of the event/no disclosure ($n = 4$), and failure to attend the interview session ($n = 5$).

An a priori power analysis using G*Power 3.1 (Faul et al., 2009) for a repeated measures design with between-subjects factors, assuming a medium effect size of $f = .25$ ($a = .05$), four groups, four repeated measures, and a moderate

correlation between measures (.3 to .5), indicated that a total sample size between 88 and 116 participants would be required to achieve a power of .80 (Cohen, 1992). Alternative design options resulted in smaller sample sizes; therefore, the most conservative approach was selected.

To account for potential attrition, a recruitment target of at least 137 participants was set. Following exclusions ($n = 14$), data from 123 children were analyzed, exceeding the suggested sample size by 6.3%. No data on family income or other socioeconomic variables were collected. The study was approved by the Ethics Committee of the University of Murcia and all procedures were conducted in accordance with the Declaration of Helsinki. This study was not preregistered.

Procedure

Staged event

Each child individually participated in a 15-minute staged event at their school. A research assistant escorted the participant to a room separate from their classroom, where a trained "game facilitator" (one of six postgraduate psychology students, four female) conducted the session, which was video recorded. The session included a rapport-building phase (scripted questions about child's interests) followed by three play activities selected from a pool of 14 culturally familiar games. The order and selection of games were counterbalanced across participants, but each child was guaranteed to participate in at least two physically active games (e.g., Put on a children's costume, taking a photo of the monitor, building shapes, sculptor, or guessing letters written on the back) and one predominantly verbal game (e.g., hangman, word chains).

Facilitators received specialized training and were blind to the study's hypotheses. Although teachers and parents were instructed not to discuss the event with participants, control over children's discussions among peers was not possible.

Interview

Interviews occurred between 12 and 17 days after the staged event, conducted in a different school room from the event location. A 2 (protocol: NICHD vs. MoGP) \times 2 (age: 6–7 vs. 11–12 years) ANOVA confirmed no significant condition differences between interview protocols in delay, $F(1, 119) = 1.91$, $p = .17$ ($M_{\text{NICHD}} = 14.34$ days, $SD = 1.81$; $M_{\text{MoGP}} = 14.8$ days, $SD = 1.81$).

The children were interviewed once by 12 postgraduate psychology students (six per condition), who completed extensive training led by trainers (authors). Training workshops included theoretical instruction on each protocol, as well as role-play exercises with real-time feedback. Interviewers remained blind to hypotheses and event specifics, receiving ongoing supervision and feedback.

The presubstantive interview phase was standardized across conditions, encompassing introductions, rapport building, task clarification, ground rules (e.g., "if you don't know the answer to a question, you should say 'I don't know'"), and a practice narrative phase. Following this, the interview about the staged event began with a series of scripted prompts designed to help orient the child to the play session with the game facilitator. These prompts progressed from general (e.g., "Have you done any special activities at school lately?") to more specific cues (e.g., "I heard that a few weeks ago someone took you from your class to play with a monitor."). Once the presubstantive and transition phases were completed, the substantive phase of the interview began.

Interviewers following the NICHD Protocol received instruction on identifying each type of prompt, understanding when and how to use them, and engaging in supervised role-play interviews, with emphasis on applied practice. They were trained in the strategy to follow during the investigative phase of the interview. Specifically, they were instructed to begin with a main invitation (e.g., "Tell me everything that happened") and to continue with as many free-recall prompts as possible. The training emphasized the importance of exhausting and prioritizing invitations and cued invitations before progressing to more focused forms of questioning. Directive questions were to be used only after all free-recall prompts had been exhausted, and limited option-posing questions were reserved for situations in which crucial details were still missing. Following any focused prompt, interviewers were instructed to return to the most open style of questioning, in accordance with the protocol's recommendations (Lamb et al., 2018).

MoGP interviewers received specific training in the protocol's guiding principles and were instructed to conduct the substantive phase in four stages: free narrative, open-ended questions; closed questions; and closure. Interviewers were trained to prioritize open-ended prompts and to begin the interview using them, gradually transitioning to focused

questions only once open-ended approaches had been exhausted.

In both protocols, emphasis was placed on avoiding the use of option-posing and suggestive questions. Additionally, interviewers were provided with a checklist of topics to address in each interview, including all games played, the child's actions and interactions with the game facilitator, the facilitator's identity, and objects relevant to the activities.

Coding

Interview transcripts were independently coded by two trained research assistants blind to conditions and hypotheses. Interviewer prompts were categorized (see Table 1) following NICHD guidelines (Orbach et al., 2000). As no suggestive questions were identified in the coded interviews, they were not included in the results. Child responses were coded for new details. Total details were defined as the sum of four categories: spatial details, temporal details, perceptual details, and action details. This analysis was based on the reality monitoring framework (Johnson & Raye, 1981), which has been applied in studies of children's testimony (Alonso-Quecuty & Hernández-Fernaud, 1997; Stromwall & Granhag, 2005; Vrij et al., 2004). Spatial details referred to information about locations, object arrangements, or physical surroundings (e.g., "The pencil was on the table"). Temporal details included information about when events occurred or the sequence of events (e.g., "First I drew a picture and then I dressed up as a pirate"). Perceptual details described what was seen, heard, felt, tasted, or smelled (e.g., "The ball made a high-pitched sound every time it bounced"). Action details described actions or the process of actions performed by the interviewee (e.g., "I threw the ball into the basket"). All details were categorized according to their accuracy as either correct if present during the play session or incorrect if present but inaccurately described, or not present at all. Vague, subjective (e.g., "I had fun"), unverifiable (e.g., inaudible or invisible), or irrelevant information (e.g., being picked up from the classroom) was not coded.

Table 1
Coding Definitions for Interviewer Utterances.

Utterance Type	Definition	Example
Invitations	Open-ended prompts for free-recall responses that did not focus on specific content	Tell me everything that happened Tell me more
Cued Invitations	Open-ended prompts that invited elaboration on content already disclosed by the child	"You said you were drawing, tell me more about that"
Directives	Wh- questions seeking details about previously reported information	If the child mentioned a pencil, "What color was the pencil?"
Option-posing	Ask the child to affirm, negate, or choose from alternatives provided by the interviewer	"Did he give you the pencil?"
Suggestive	Questions implied an expected response or introduced information not previously mentioned by the child.	If the child had not mentioned a pencil, "What color was the pencil?"

The accuracy of each coded detail was verified by reviewing the video recording of the play session. Additionally, each detail (whether correct or incorrect) was coded as either cen-

tral or peripheral. This distinction followed definitions used in previous research employing the NICHD Protocol (Dion & Cyr, 2008; Orbach & Lamb, 2000, 2001). Central details

referred to elements that outlined or defined the core aspects of the event and whose omission or modification would change the plot (e.g., actions and interactions with game facilitator, identification of key activities and objects, event location). Peripheral details described features that were not integral to the event and whose omission or alteration would not affect its overall meaning (e.g., color of the balloon, room description).

To assess coding reliability, two research assistants blind to the experimental conditions and study hypotheses independently coded the question types and the identification of details for a subsample of 40 participants (31.71%). Intra-class correlation coefficients (ICCs) were calculated using a two-way random-effects model, based on single ratings and absolute agreement. Interrater reliability was excellent (Koo & Li, 2016) for all question types: invitations (ICC = .96, CI [.93, .98]), cued invitations (ICC = .94, CI [.90, .97]), directives (ICC = .93, CI [.87, .96]), and option-posing questions (ICC = .90, CI [.83, .95]). Regarding the coding of details, ICCs also indicated excellent reliability: correct details (ICC = .99, CI [.98, .99]), central details (ICC = .98, CI [.96, .99]), peripheral details (ICC = .95, CI [.91, .97]), and incorrect details (ICC = .94, CI [.88, .96]).

Data analysis

Analyses examined differences between interview protocols in the amount of total and correct information elicited, and whether these effects interacted with age. The role of different prompt types in shaping the amount, accuracy, and forensic relevance of the elicited information was also explored. The study employed a $4 \times 2 \times 2$ mixed factorial design: question type (invitations, cued invitations, directives, option-posing) as a within-subjects factor, and interview protocol (NICHD vs. MoGP) and age group (7 vs. 11 years) as between-subjects factors (see Table 2).

Measures were obtained for the information elicited, as well as for the amount (correct and incorrect) and type (central and peripheral) of details. In addition, indices of informativeness (number of details elicited per question asked) and accuracy rates (proportion of correct details) were computed for each type of prompt and interview condition.

Given the count nature of dependent variables and overdispersion, Generalized Linear Mixed Models (GLMMs) with negative binomial distributions were used. This approach handles non-normal distributions while modeling fixed and random effects. Participant-level random intercepts accounted for repeated measures and individual variability.

A diagonal residual covariance structure was specified based on low intercorrelations among question types. Pearson residual plots confirmed model adequacy, showing no systematic heteroscedasticity or autocorrelation.

Model estimation used safeguards to ensure robust inference: Satterthwaite's approximation for degrees of freedom, robust (Huber-White) standard errors, and family-wise error

corrections (Bonferroni or Holm-Bonferroni) for planned comparisons. For interpretability, raw means and standard deviations are reported to contextualize model-based findings. All statistical analyses were performed using SPSS version 28.

Results

How Were the Interviews Constructed?

To assess differences in prompt use, a GLMM with a negative binomial distribution was fitted due to overdispersion. A diagonal residual covariance structure was specified given the low intercorrelations among question types ($r_s < .36$), and Pearson residual plots showed no systematic heteroscedasticity. Model fit was acceptable ($AIC_c = 1159.49$, $BIC = 1180.19$), and the participant-level random effect was significant ($\sigma^2 = .09$, $SE = 0.04$, $z = 2.29$, $p = .022$).

Significant fixed effects emerged for interview protocol, $F(1, 118) = 22.51$, $p < .001$; question type, $F(3, 126) = 28.61$, $p < .001$; their interaction, $F(3, 126) = 10.80$, $p < .001$; and for protocol by age group, $F(1, 118) = 3.96$, $p = .05$.

Table 2
Mean (SD) Number of Prompts Posed During Interview.

Type of prompt	Protocol Interview			
	MoGP		NICHD	
	6-7 Years	11-12 Years	6-7 Years	11-12 Years
Invitations	2.83 (2.63)	2.56 (2.01)	4.80 (3.89)	6.32 (4.57)
Cued	2.77 (3.30)	2.78 (2.86)	4.70 (3.93)	9.35 (6.68)
Invitations	4.57 (3.76)	4.41 (3.89)	5.20 (4.94)	4.48 (3.67)
Directives	7.63 (4.95)	8.06 (6.69)	6.07 (4.11)	9.32 (5.48)
Option-posing	17.80 (9.54)	17.81 (11.66)	20.77 (10.71)	29.48 (10.82)
Total				

Pairwise comparisons revealed that NICHD interviews included more invitations and cued invitations than MoGP, with no protocol differences for directives or option-posing prompts. MoGP used more option-posing questions than other types, whereas NICHD favored cued invitations and option-posing over directives. The protocol-by-age interaction indicated that 11-year-olds interviewed with NICHD received more prompts than any other condition.

Amount of Information Elicited by Prompt Type

To examine whether the amount of information elicited varied as a function of prompt type, age, and protocol, a GLMM with a negative binomial distribution was fitted. The outcome variable was the number of elicited details (Table 3). Question type was modeled as a within-subjects factor; protocol and age as between-subjects factors. Overdisper-

sion across levels of the within-subject factor justified the use of the negative binomial distribution. A diagonal covariance structure was used (all $r < .26$), and a random intercept accounted for individual variability ($\Delta AIC = 1553.88$, $BIC = 1554.58$).

Significant fixed effects emerged for age group, $F(1, 120) = 9.14$, $p = .003$; interview protocol, $F(1, 120) = 14.29$, $p < .001$; question type, $F(3, 151) = 40.46$, $p < .001$; and the protocol \times question type interaction, $F(3, 151) = 3.98$, $p = .009$. The random effect was significant ($\sigma^2 = .17$, $SE = .06$, $\zeta = 2.81$, $p = .005$).

Pairwise comparisons showed that 11-year-olds reported more details than 7-year-olds. NICHD elicited more information than MoGP only with invitations and cued invitations; no differences appeared for directives or option-posing prompts.

Within MoGP, option-posing prompts yielded fewer details than all other types. In contrast, all question types in NICHD differed significantly: cued invitations elicited the most information, followed by invitations, directives, and option-posing.

To assess the average informativeness of each prompt type, a relative information index was computed by dividing the number of details elicited by the number of prompts posed, following Brown et al. (2013). A GLMM was fitted with age, protocol, and question type as predictors. Given the right-skewed distribution and the proportional nature of the index, a gamma distribution with a log-link function was used. Low intercorrelations among prompt types justified a diagonal covariance structure. The participant-level random effect was non-significant and excluded.

Table 3
Mean Number (SD) of Details Elicited per Prompt.

Type of Prompt	Interview Protocol			
	MoGP		NICHD	
	6-7 Years	11-12 Years	6-7 Years	11-12 Years
Invitations	7.13 (5.62)	8.78 (12.77)	12.10 (10.25)	22.06 (19.47)
Cued	8.93 (16.17)	12.19 (16.48)	21.07 (21.61)	26.35 (28.32)
Invitations	5.30 (6.86)	9.75 (16.72)	7.93 (10.02)	11.29 (10.91)
Directives	3.23 (4.40)	3.03 (2.75)	2.27 (3.19)	4.32 (6.49)
Option-posing	24.23 (17.73)	33.66 (27.81)	43.37 (31.85)	64.48 (42.65)
Total				

A significant main effect of question type emerged, $F(3, 136) = 26.41$, $p < .001$. Pairwise comparisons indicated that cued invitations ($M = 4.96$, $SE = .76$) and invitations ($M = 3.47$, $SE = .24$) were the most informative prompt types, yielding significantly more details per question than directives ($M = 2.87$, $SE = .50$) and option-posing prompts ($M = 0.94$, $SE = .14$). Directive prompts were also more informative than option-posing. No interaction effects were found,

indicating that this pattern was consistent across age groups and protocols.

Covariate Models: Prompt Frequency, Informativeness, and Interview Length

Total prompt count was added as a covariate to the main model. Model fit did not improve ($\Delta AIC = 1542.85$, $BIC = 1563.16$), but the covariate had a significant effect, $F(1, 93) = 4.47$, $p = .037$. All other fixed effects remained significant, indicating that protocol effects on elicited details could not be attributed solely to the number of prompts. Interview length was also included as a covariate in all models but did not yield significant effects, suggesting that differences in information elicitation were not driven by the number of questions formulated in each interview.

A second model included the number of times each prompt type was used per-interview as covariates. This model also failed to improve fit ($\Delta AIC = 1568.84$, $BIC = 1589.57$). Only the number of invitations had a significant effect, $F(1, 38) = 24.63$, $p < .001$, and cued invitations showed a marginal trend, $F(1, 98) = 3.72$, $p = .06$. After adjusting for these covariates, the effect of protocol was no longer significant ($p = .23$), while the effects of age group ($p = .02$) and the protocol \times question type interaction ($p = .03$) remained significant. These results suggest that the protocol advantage was primarily driven by greater use of invitations and, to a lesser extent, cued invitations.

A third model added the relative information index as a predictor. Model fit improved ($\Delta AIC = 1505.51$, $BIC = 1522.11$), and all four informativeness indices were significantly associated with the number of elicited details: invitations, $F(1, 479) = 51.39$, $p < .001$; cued invitations, $F(1, 479) = 72.27$, $p < .001$; directives, $F(1, 479) = 36.50$, $p < .001$; and option-posing, $F(1, 479) = 10.17$, $p = .002$. The age effect became marginal ($F(1, 479) = 3.51$, $p = .06$), suggesting that age-related differences were partially mediated by prompt informativeness.

Finally, a combined model included both the number and informativeness of prompts. Model fit remained comparable ($\Delta AIC = 1553.87$, $BIC = 1570.44$). The number of invitations, $F(1, 435) = 206.32$, $p < .001$ and their informativeness $F(1, 479) = 152.90$, $p < .001$, had the strongest effects. Informativeness of other prompt types and the number of cued invitations also showed significant effects. In contrast, the frequency of directives and option-posing prompts was unrelated to outcome. Under this model, age group ($p = .63$), protocol ($p = .98$), and their interaction with question type ($p = .07$) were no longer significant, though the main effect of question type remained ($p = .001$). These results indicate that differences in productivity across conditions are better explained by jointly considering the frequency and informativeness of each prompt type. In contrast, they suggest that once prompt effectiveness is accounted for, age-related differences in productivity are substantially reduced.

How Accurate Was the Information Elicited Using Each Type of Prompt?

The primary objective of the study was to determine whether the NICHD protocol facilitated the collection of a greater quantity of correct information without increasing the amount of incorrect information. Consequently, accuracy rates were calculated for each question type (see Table 4). A GLMM with binomial distribution and logit link function was fitted to examine accuracy. Number of correct details served as the outcome variable, with total number of trials specified as the binomial denominator. A random intercept was included for participants to account for the repeated-measures structure of the data.

A significant main effect was found for age group, $F(1, 224) = 5.83, p = .02$. Comparison of estimated means revealed that 11-years old showed an accuracy rate ($M = .97, SE = .01$) significantly higher than 7-years-old. No other effect reached statistical significance, perhaps due to a ceiling effect.

Table 4

Proportion (SD) of Reported Information That Was Correct (Accuracy) by Group Age, Protocol Interview and Prompt Type.

Type of Prompt	Protocol Interview			
	MoGP		NICHD	
	6-7 Years	11-12 Years	6-7 Years	11-12 Years
Invitations	.88 (.31)	.99 (.02)	.87 (.28)	.99 (.02)
Cued Invitations	.95 (.07)	.97 (.06)	.94 (.18)	.96 (.09)
Directives	.93 (.15)	.95 (.17)	.93 (.16)	.92 (.15)
Option-posing	.90 (.30)	.89 (.27)	.91 (.19)	.93 (.23)

Subsequently, the next step was to assess accuracy of information through quantity of correct and incorrect details (see Table 5). A GLMM with a negative binomial distribution was fitted to analyze the number of correct details elicited. The model included question type as a within-subject factor, and age group and interview protocol as between-subject factors. Given the count nature of the data and evidence of overdispersion, the negative binomial family was used. A diagonal residual covariance structure was applied based on low intercorrelations among question types. Model fit was acceptable ($AIC = 1554.00, BIC = 1574.70$), and the random effect was significant ($\sigma^2 = 0.19, SE = 0.06, \chi^2 = 2.88, p = .004$).

Table 5

Mean Number of Correct and Incorrect Details Elicited, by Type of Prompt, Age and Protocol Interview Condition (SD).

Type of Prompt	Protocol Interview			
	MoGP		NICHD	
	6-7 Years	11-12 Years	6-7 Years	11-12 Years
Correct Details				
Invitations	6.70 (5.53)	8.56 (12.81)	11.10 (9.86)	21.90 (19.47)
Cued Invitations	8.10 (13.79)	11.72 (16.00)	20.10 (21.39)	25.71 (28.09)
Directives	4.93 (6.61)	9.53 (16.76)	7.47 (9.72)	10.52 (12.18)
Option-posing	3.10 (4.26)	2.78 (2.62)	2.13 (3.14)	4.13 (6.47)
Incorrect Details				
Invitations	0.43 (1.16)	0.22 (0.49)	1.00 (1.68)	0.16 (0.45)
Cued Invitations	0.83 (1.84)	0.47 (1.05)	0.97 (2.27)	0.66 (2.03)
Directives	0.37 (0.85)	0.22 (0.55)	0.47 (1.20)	0.77 (1.52)
Option-posing	0.13 (0.35)	0.25 (0.67)	0.13 (0.43)	0.19 (0.75)

Significant main effects were found for age group, $F(1, 121) = 9.63, p = .002$; interview protocol, $F(1, 121) = 13.67, p < .001$; and question type, $F(3, 147) = 40.16, p < .001$. A significant interaction emerged between protocol and question type, $F(3, 147) = 4.15, p = .007$.

Pairwise comparisons revealed that 11-year-olds provided more accurate details than 7-year-olds. NICHD interviews elicited more correct details than MoGP, particularly through invitations and cued invitations. No differences emerged for directives or option-posing questions. Under MoGP, option-posing prompts generated fewer correct details than other formats. NICHD interviews showed hierarchical effectiveness: cued invitations were most effective, followed by invitations, directives, and option-posing prompts. Controlling for total number of questions posed did not

substantially improve model fit, though the covariate was significant, $F(1, 92) = 3.99, p = .05$. All other fixed effects remained significant.

When number of each prompt type was included as covariates, only invitations ($F(1, 44) = 21.34, p < .001$) and marginally cued invitations ($p = .06$) predicted correct details. Under this model, the effect of protocol became non-significant ($p = .24$), while the remaining effects persisted. These findings suggest that protocol differences in eliciting correct information are largely driven by the use of open-ended prompts.

In parallel with the informativeness analysis, a relative accuracy index was computed for each type of question by dividing the number of correct details by the number of questions asked. A GLMM was fitted to examine the effects

of age, protocol, and question type on this index, with question type as a within-subject factor. A gamma distribution with a log-link function was used given the proportional nature of the data. As the participant-level random effect was not significant, it was excluded. Model fit was acceptable ($AIC = 1356.41$, $BIC = 1372.12$). Only question type had a significant effect, $F(3, 133) = 25.99$, $p < .001$. Option-posing prompts yielded significantly lower relative accuracy ($M = 0.93$, $SE = .15$) than directives ($M = 2.74$, $SE = .49$), invitations ($M = 3.44$, $SE = .24$), and cued invitations ($M = 4.83$, $SE = .74$).

When relative accuracy was added as a covariate in the model predicting correct details, model fit improved ($AIC = 1531.51$, $BIC = 1548.11$). Relative accuracy was a significant predictor, and the age effect became marginal, $F(1, 467) = 3.55$, $p = .06$, suggesting that age differences in correct details may stem from variation in relative accuracy. The effects of protocol and question type remained significant or increased.

Finally, when both relative accuracy and prompt frequency were included, model fit declined slightly ($AIC = 1577.83$, $BIC = 1594.40$). Under this model, only the effect of question type remained significant, $F(3, 142) = 23.87$, $p < .001$. Age group ($p = .73$), protocol ($p = .87$), and their interaction with question type ($p = .06$) were no longer significant.

Number of Incorrect Details

To further examine the nature of the information elicited, a GLMM was fitted to predict the number of incorrect details. The model included question type as a within-subject factor and interview protocol and age group as between-subjects factors. Due to overdispersion in the count data, a negative binomial distribution and a diagonal residual covari-

ance structure were specified. This choice was supported by low intercorrelations among question types (all $r < .32$) and residual diagnostics. A participant-level random effect was included. Model fit was acceptable ($AIC = 2217.11$, $BIC = 2237.81$), and the random effect was significant ($\sigma^2 = 1.779$, $\chi = 4.750$, $p < .001$).

Only the fixed effect of question type reached statistical significance, $F(3, 48) = 8.11$, $p < .001$. Pairwise comparisons showed that in comparison with cued invitations, option-posing questions elicited fewer incorrect details. No other comparisons were statistically significant.

What Kind of Information Was Elicited?

To analyze the forensic relevance of the elicited information, two GLMM were fitted, separately examining the number of correct central and peripheral details (see Table 6). Both models included question type as a within-subject factor, and interview protocol and age group as between-subjects factors. Descriptive analyses showed substantial overdispersion; therefore, a negative binomial distribution and diagonal covariance structure were specified. For central details, significant fixed effects were observed for age group, $F(1, 117) = 9.09$, $p = .003$; interview protocol, $F(1, 111) = 9.46$, $p = .003$; and question type, $F(3, 136) = 37.52$, $p < .001$. A significant interaction between protocol and question type also emerged, $F(3, 136) = 4.62$, $p = .005$. Pairwise comparisons revealed that 11-year-olds provided more central details than 7-year-olds, and NICHD interviews elicited more central details than MoGP—but only for invitations and cued invitations. Under MoGP, option-posing prompts generated fewer correct details than other formats. NICHD interviews showed hierarchical effectiveness: cued invitations and invitations were most effective, followed by directives, and option-posing prompts.

Table 6

Mean (SD) Number of Central and Peripheral Details Reported by Children to Each Prompt Type, by Age and Protocol Interview Condition.

Type of Prompt	Interview Protocol			
	MoGP		NICHD	
	6-7 Years	11-12 Years	6-7 Years	11-12 Years
Central				
Invitations	6.33 (5.55)	7.60 (13.10)	10.70 (9.66)	20.29 (18.54)
Cued Invitations	6.83 (10.88)	9.81 (12.16)	18.03 (19.31)	23.65 (27.55)
Directives	4.07 (5.02)	8.31 (13.64)	6.40 (9.03)	9.68 (11.29)
Option-posing	2.37 (3.37)	2.31 (2.62)	1.73 (3.01)	3.16 (6.31)
Peripheral				
Invitations	0.37 (1.13)	0.94 (1.88)	0.40 (1.71)	1.61 (3.63)
Cued Invitations	1.27 (3.75)	1.92 (5.23)	2.07 (3.64)	2.06 (4.00)
Directives	0.87 (2.10)	1.22 (3.40)	1.07 (1.96)	0.84 (1.95)
Option-posing	0.73 (1.72)	0.47 (0.84)	0.40 (0.85)	0.97 (2.10)

When frequencies of use of each question type were included as covariates, the effect of protocol was no longer significant. Likewise, the effect of age group was no longer

significant after controlling for the relative accuracy index. When both covariates were entered simultaneously, none of the fixed effects or interactions reached significance.

A second GLMM was fitted for peripheral details using the same structure. Overdispersion was present, justifying the use of a negative binomial distribution. Model fit was acceptable ($AIC = 2138.38$, $BIC = 2159.15$), and only the fixed effect of question type was significant, $F(3, 120) = 7.00$, $p < .001$. The participant-level random effect was significant ($\sigma^2 = 0.76$, $\chi = 3.31$, $p < .001$). Pairwise comparisons showed that cued invitations elicited more peripheral details than both invitations and option-posing prompts.

Including question frequency as covariates yielded no significant effects. When relative accuracy was added to the model, fit declined slightly ($AIC = 2197.79$, $BIC = 2214.35$), and the random effect was removed to achieve convergence. Question type remained a significant predictor ($p = .022$), although with reduced strength. All relative accuracy indices except for specific prompts ($p = .14$) were significant predictors of peripheral details ($ps < .001$).

Data on incorrect central and peripheral details were not included in the analyses because their distributions exhibited an excessive inflation of zeros.

Discussion

This study represents the first controlled analogue comparison between interviewers trained using the NICHD Protocol and those trained with the MoGP guidelines. To our knowledge, it is also the first time that the NICHD Protocol has been directly compared to another forensic interview protocol in an analogue context. Although previous research has demonstrated the advantages of the NICHD Protocol over non-structured or minimally guided interviews, most of these studies have been conducted in naturalistic settings and compared NICHD-guided interviews to baseline practices (Cyr & Lamb, 2009; Orbach et al., 2000; Sternberg, Lamb, Orbach, et al., 2001; Yi et al., 2016).

Importantly, this investigation employed ecologically sophisticated analogue methodology. In contrast to previous analogue studies that often relied on a limited number of interviewers or exposed them to statements of identical events (Dent, 1986; Milne & Bull, 2002), the present study involved a relatively large number of interviewers (12 in total) thereby reducing the potential for expectancy effects. Furthermore, each child participated in a unique staged event that involved a different combination of activities, selected for their analogy with those typically present in CSA cases (e.g., putting on and taking off clothes, physical contact, and having conversations).

Present findings provide compelling evidence for the relative effectiveness of the NICHD investigative interview protocol in eliciting more detailed, accurate, and forensically relevant information from children. Across models examining prompt frequency, information yield, accuracy, and content forensic relevance, NICHD interviews consistently outperformed those conducted using the MoGP protocol, with the greatest advantages observed for invitations and cued invitations. Regarding forensic relevance of information re-

ported, NICHD interviews elicited more central details than MoGP, again driven by the use of invitations and cued invitations. Peripheral details were also more frequently elicited by cued invitations, suggesting that open-ended prompts facilitate a more comprehensive recall of the event. However, once question frequency and accuracy indices were controlled, the effects of age and protocol were no longer significant, reinforcing the conclusion that question quality and formulation are the primary drivers of informational productivity.

These findings are aligned with prior literature underscoring the importance of open-ended questioning strategies in promoting children's narrative productivity and minimizing suggestive influence (Danby & Sharman, 2023; Gagnon & Cyr, 2017; Hershkowitz, 2001; Lamb et al., 2003; Lee & Kim, 2021; Melinder et al., 2021). This pattern may be explained by the fact that the NICHD Protocol offers a structured framework for the interview, including clearly defined phases, recommended question stems, and concrete examples to guide interviewer practice.

Age-related differences were also evident, with older children providing more correct and central details and demonstrating higher accuracy rates. However, the inclusion of prompt informativeness and accuracy indices in the models attenuated these age effects, suggesting that age differences in performance were partially mediated by the cognitive demands and communicative affordances of the questions asked. This result accords with developmental accounts positing that younger children require more support to access and communicate episodic memories and that age-related disparities can be mitigated through interviewer scaffolding strategies such as open-ended prompts (Pipe et al., 2007).

Our findings contribute to the existing literature on question types (Andrews et al., 2015; Gagnon & Cyr, 2017; Lamb & Fauchier, 2001). In our study, prompt type consistently emerged as a critical determinant of interview outcomes. Invitations and cued invitations were the most productive and informative across all measures, eliciting more total, correct, central, and peripheral details than other question formats. These prompts also exhibited the highest informativeness and relative accuracy indices. Interviews conducted under the NICHD protocol, a vast majority of correct details (82.27% of all reported details) and correct forensically relevant information (77.61%) were obtained through invitations and cued invitations. In contrast, option-posing prompts consistently underperformed, yielding the fewest correct and peripheral details and the lowest informativeness and relative accuracy. These patterns held across both protocols and age groups, underscoring the generalizability of the observed prompt hierarchies.

An unexpected finding was the high frequency with which option-posing questions were asked in both protocols, and—contrary to expectations—were highly prevalent even within interviews guided by the NICHD protocol. This observation is inconsistent with the protocol's guidelines and

with its foundational logic, which prioritizes open-ended prompts (Lamb et al., 2018). Previous research has documented similar deviations from protocol in applied settings, where interviewers may default to closed-ended formats under conditions of uncertainty or when children provide minimal responses (Cederborg et al., 2013; Kask, 2012; Lamb, Sternberg, Orbach, Esplin, et al., 2002; Lamb, Sternberg, Orbach, Hershkowitz, et al., 2002). The overreliance on option-posing prompts may also reflect difficulties in maintaining the open-ended style across the entire interview (Aldridge & Cameron, 1999; Korkman et al., 2006; Powell & Guadagno, 2008; Powell & Hughes-Scholes, 2009; Sternberg, Lamb, Orbach, et al., 2001). Furthermore, interviewers may underestimate the potential costs of these question types, particularly when used as follow-ups rather than as initial prompts (Powell & Guadagno, 2008). Alternatively, this pattern may have been influenced by the specific instructions given to interviewers. In an effort to simulate real-world forensic contexts, where professionals often need to obtain highly specific and relevant details, interviewers in both protocols were instructed to use more focused or option-posing questions when narrative approaches failed to elicit forensically relevant information. It is likely that the frequent use of option-posing questions reflected attempts to compensate for the limited success of open-ended strategies in eliciting essential details.

The interaction between protocol and prompt type further elucidated the mechanisms underlying protocol effectiveness. NICHD interviews not only included a higher frequency of open-ended prompts but also demonstrated greater effectiveness (informativeness and relative accuracy) when such prompts were used. In contrast, the MoGP protocol exhibited limited use of invitations and cued invitations, and when employed, these prompts elicited fewer details. These results suggest that adherence to structured protocols emphasizing open-ended questioning may enhance both the frequency and effectiveness of these strategies, likely through improved interviewer training and monitoring (Lamb, Sternberg, Orbach, Hershkowitz, et al., 2002). Notably, despite being the most frequently used question type, option-posing questions elicited fewer total details, fewer forensically relevant details compared to other prompt types, and showed the lowest relative accuracy of all question types. Interestingly, its accuracy rate is not different from that of other question types. These findings raise concerns regarding their use and reinforce the importance of prioritizing recall-based strategies over recognition-based prompts (Lamb et al., 2018).

The absence of significant differences between protocols in number of incorrect details elicited supports the claim that the NICHD Protocol enhances the quantity of accurate information without increasing the risk of error. However, our findings showed that cued invitations elicited slightly more incorrect details than option-posing questions. This parallels the results of Brown et al. (2013), who observed that while invitations and cued invitations substantially increased the

amount of information provided by children, they also slightly increased the probability of errors. However, this effect was minor compared to their substantial benefits in eliciting accurate and forensically central information.

Analogue studies necessarily differ from real forensic contexts, particularly when children are interviewed about potentially traumatic or abusive experiences. In our study, as in previous research, responses to close-ended prompts were generally less detailed and, in many cases, more error-prone (Lamb et al., 2018). From a forensic perspective, adherence to and training in the NICHD Protocol may be beneficial for child victims, as it can facilitate the elicitation of a detailed, non-directive narrative while minimizing the use of potentially leading or suggestive questions and reducing the need for repeated interviews to clarify the child's account.

Limitations

Despite the strengths of this study, several limitations should be acknowledged. First, findings may not generalize to forensic interviews involving CSA or other potentially traumatic or distressing events. Second, the high overall accuracy rates limited the analysis of between-group differences in the amount of incorrect information provided. Third, our sample included only older children (7- and 11-year-olds), which may have reduced the likelihood of detecting developmental differences, particularly those observed in preschool-aged populations, who often pose greater challenges for eliciting forensically relevant information. Finally, the relatively short delay between the event and the interview (12–17 days) might restrict the generalizability of the findings to situations involving longer retention intervals. Future research should include younger children—especially those aged 4 to 5—and examine how longer delays affect both the quantity and accuracy of information elicited. Such studies would help determine whether the advantages of structured protocols like NICHD persist under more ecologically valid conditions that resemble real-world forensic contexts. Additionally, in our study, we did not control for the order in which the interviewers conducted the interviews or the order in which the children participated in the activities relative to other children. Nor did we control for participants' socioeconomic status. Future studies could address these variables to examine how interviewers' experience and contextual factors affect outcomes and to reduce potential biases.

Conclusions

The present study offers a novel analogue comparison between the NICHD Protocol and the MoGP guidelines within a controlled setting, focusing on the impact of question type and structure on the quality of children's forensic testimony. Overall, interviews conducted using the NICHD Protocol elicited more correct and forensically relevant details than those conducted using the MoGP. This advantage appears to be primarily driven by the strategic use of recall-

based prompts—particularly invitations and cued invitations. These findings extend previous research based on forensic interview transcripts by demonstrating that NICHD training supports the elicitation of accurate and central information under standardized conditions. As such, the present results provide strong empirical support for the effectiveness of protocol-based interviewing and underscore the critical role of question structure in shaping the quality of children's testimony.

References

- Ahern, E. C., & Lamb, M. E. (2017). Children's Reports of Disclosure Recipient Reactions in Forensic Interviews: Comparing the NICHD and MoGP Protocols. *Journal of Police and Criminal Psychology, 32*(2), 85–93. <https://doi.org/10.1007/s11896-016-9205-x>
- Aldridge, J., & Cameron, S. (1999). Interviewing Child Witnesses: Questioning Techniques and the Role of Training. *Applied Developmental Science, 3*(2), 136–147. https://doi.org/10.1207/s1532480xads0302_7
- Alonso-Quecuty, M. L., & Hernández-Fernaudo, E. (1997). Tócala otra vez Sam: repitiendo las mentiras [Play it again, Sam: Repeating lies]. *Estudios de Psicología, 18*(57), 29–37. <https://doi.org/10.1174/021093997320972025>
- Andrews, S. J., Lamb, M. E., & Lyon, T. D. (2015). Question Types, Responsiveness and Self-contradictions when Prosecutors and Defense Attorneys Question Alleged Victims of Child Sexual Abuse. *Applied Cognitive Psychology, 29*(2), 253–261. <https://doi.org/10.1002/acp.3103>
- Benia, L. R., Hauck-Filho, N., Dillenburg, M., & Stein, L. M. (2015). The NICHD Investigative Interview Protocol: A Meta-Analytic Review. *Journal of Child Sexual Abuse, 24*(3), 259–279. <https://doi.org/10.1080/10538712.2015.1006749>
- 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*(4), 367–382. <https://doi.org/10.1037/a0035143>
- Bull, R. (2014). *Investigative interviewing*. Springer.
- Cederborg, A. C., Alm, C., Lima da Silva Nises, D., & Lamb, M. E. (2013). Investigative interviewing of alleged child abuse victims: An evaluation of a new training programme for investigative interviewers. *Police Practice and Research, 14*(3), 242–254. <https://doi.org/10.1080/15614263.2012.712292>
- Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155–159. <https://doi.org/10.1037/0033-2909.112.1.155>
- 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*(5), 257–268. <https://doi.org/10.1016/j.chiabu.2008.04.002>
- Danby, M. C., & Sharman, S. J. (2023). Open-ended initial invitations are particularly helpful in eliciting forensically relevant information from child witnesses. *Child Abuse & Neglect, 146*, 1–13. <https://doi.org/10.1016/j.chiabu.2023.106505>
- Dent, H. R. (1986). An experimental study of the effectiveness of different techniques of questioning mentally handicapped child witnesses. *British Journal of Clinical Psychology, 25*(1), 13–17. <https://doi.org/10.1111/j.2044-8260.1986.tb00666.x>
- Dent, H. R., & Stephenson, G. M. (1979). An experimental study of the effectiveness of different techniques of questioning child witnesses. *British Journal of Social and Clinical Psychology, 18*(1), 41–51. <https://doi.org/10.1111/j.2044-8260.1979.tb00302.x>
- Dion, J., & Cyr, M. (2008). The Use of the NICHD Protocol to Enhance the Quantity of Details Obtained from Children with Low Verbal Abilities in Investigative Interviews: A Pilot Study. *Journal of Child Sexual Abuse, 17*(2), 144–162. <https://doi.org/10.1080/10538710801916564>
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160. <https://doi.org/10.3758/BRM.41.4.1149>
- Fernandes, D., Gomes, J. P., Albuquerque, P. B., & Matos, M. (2024). Forensic Interview Techniques in Child Sexual Abuse Cases: A Scoping Review. *Trauma, Violence, & Abuse, 25*(2), 1382–1396. <https://doi.org/10.1177/15248380231177317>
- Fisher, R. P., & Geiselman, R. E. (1992). *Memory-Enhancing Techniques for Investigative Interviewing: The Cognitive Interview*. Charles C. Thomas.
- Gagnon, K., & Cyr, M. (2017). Sexual abuse and preschoolers: Forensic details in regard of question types. *Child Abuse & Neglect, 67*, 109–118. <https://doi.org/10.1016/j.chiabu.2017.02.022>
- Henry, L. A., & Gudjonsson, G. H. (1999). Eyewitness Memory and suggestibility in Children With Mental Retardation. *American Journal on Mental Retardation, 104*(6), 491–508. [https://doi.org/10.1352/0895-8017\(1999\)104<0491:EMASIC>2.0.CO;2](https://doi.org/10.1352/0895-8017(1999)104<0491:EMASIC>2.0.CO;2)
- Henry, L. A., & Gudjonsson, G. H. (2003). Eyewitness memory, suggestibility, and repeated recall sessions in children with mild and moderate intellectual disabilities. *Law and Human Behavior, 27*(5), 481–505. <https://doi.org/10.1023/A:1025434022699>
- Herman, S. (2010). The role of corroborative evidence in child sexual abuse evaluations. *Journal of Investigative Psychology and Offender Profiling, 7*(3), 189–212. <https://doi.org/10.1002/jip.122>
- Hershkowitz, I. (2001). Children's responses to open-ended utterances in investigative interviews. *Legal and Criminological Psychology, 6*(1), 49–63. <https://doi.org/10.1348/135532501168190>
- Hershkowitz, I., & Lamb, M. E. (2020). Allegation rates and credibility assessment in forensic interviews of alleged child abuse victims: Comparing the revised and standard NICHD protocols. *Psychology, Public Policy, and Law, 26*(2), 176–184. <https://doi.org/10.1037/law0000230>
- Hershkowitz, I., Lamb, M. E., & Katz, C. (2014). Allegation rates in forensic child abuse investigations: Comparing the revised and standard NICHD protocols. *Psychology, Public Policy, and Law, 20*(3), 336–344. <https://doi.org/10.1037/a0037391>
- 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*(1), 6–14. <https://doi.org/10.1007/s11896-013-9136-8>
- Home Office. (1992). *Memorandum of Good Practice on Video Recorded Interviews with Child Witnesses for Criminal Proceedings*. Her Majesty's Stationery Office.
- Hutcheson, G. D., Baxter, J. S., Telfer, K., & Warden, D. (1995). Child witness statement quality: Question type and errors of omission. *Law and Human Behavior, 19*(6), 631–648. <https://doi.org/10.1007/BF01499378>
- Job, R., Brunel, M., & Przygodzki-Lionet, N. (2025). Child Interviewing Practices: Effects of the Implementation of the NICHD Protocol Training in France. *Journal of Police and Criminal Psychology, 40*(1), 103–115. <https://doi.org/10.1007/s11896-024-09707-5>
- Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88*(1), 67–85. <https://doi.org/10.1037/0033-295X.88.1.67>
- Kask, K. (2012). Dynamics in Using Different Question Types in Estonian Police Interviews of Children. *Applied Cognitive Psychology, 26*(2), 324–329. <https://doi.org/10.1002/acp.1831>
- Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine, 15*(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Korkman, J., Santtila, P., & Sandnabba, N. K. (2006). Dynamics of verbal interaction between interviewer and child in interviews with alleged vic-

Complementary information

Conflict of interest.- The authors declare no conflict of interest.

Financial support.- No funding.

Data Research Availability Statement.- The data used for this study are available upon reasonable request to the corresponding author

- tims of child sexual abuse. *Scandinavian Journal of Psychology*, 47(2), 109–119. <https://doi.org/10.1111/j.1467-9450.2006.00498.x>
- La Rooy, D., Brubacher, S. P., Aromäki-Stratos, A., Cyr, M., Hershkowitz, I., Korkman, J., Myklebust, T., Naka, M., Peixoto, C. E., Roberts, K. P., Stewart, H., & 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(2), 76–89. <https://doi.org/10.1108/JCRPP-01-2015-0001>
- Lamb, M. E., Brown, D. A., Hershkowitz, I., Orbach, Y., & Esplin, P. W. (2018). *Tell Me What Happened. Questioning Children About Abuse* (2nd ed.). John Wiley & Sons.
- Lamb, M. E., & Fauchier, A. (2001). The effects of question type on self-contradictions by children in the course of forensic interviews. *Applied Cognitive Psychology*, 15(5), 483–491. <https://doi.org/10.1002/acp.726>
- Lamb, M. E., Hershkowitz, I., Sternberg, K. J., Esplin, P. W., Hovav, M., Manor, T., & Yudilevitch, L. (1996). Effects of Investigative Utterance Types on Israeli Children's Responses. *International Journal of Behavioral Development*, 19(3), 627–637. <https://doi.org/10.1177/016502549601900310>
- Lamb, M. E., La Rooy, D., Malloy, L. C., & Katz, C. (2011). *Children's testimony: a handbook of psychological research and forensic practice*. John Wiley & Sons.
- Lamb, M. E., Orbach, Y., Hershkowitz, I., Esplin, P. W., & Horowitz, D. (2007). A structured forensic interview protocol improves the quality and informativeness of investigative interviews with children: A review of research using the NICHD Investigative Interview Protocol. *Child Abuse and Neglect*, 31(11–12), 1201–1231. <https://doi.org/10.1016/j.chiabu.2007.03.021>
- Lamb, M. E., Orbach, Y., Hershkowitz, I., Horowitz, D., & Abbott, C. B. (2007). Does the type of prompt affect the accuracy of information provided by alleged victims of abuse in forensic interviews? *Applied Cognitive Psychology*, 21(9), 1117–1130. <https://doi.org/10.1002/acp.1318>
- Lamb, M. E., Orbach, Y., Sternberg, K. J., Aldridge, J., Pearson, S., Stewart, H. L., Esplin, P. W., & Bowler, L. (2009). Use of a structured investigative protocol enhances the quality of investigative interviews with alleged victims of child sexual abuse in Britain. *Applied Cognitive Psychology*, 23(4), 449–467. <https://doi.org/10.1002/acp.1489>
- Lamb, M. E., Sternberg, K. J., Orbach, Y., Esplin, P. W., & Mitchell, S. (2002). Is Ongoing Feedback Necessary to Maintain the Quality of Investigative Interviews with Allegedly Abused Children? *Applied Developmental Science*, 6(1), 35–41. https://doi.org/10.1207/S1532480XADS0601_04
- Lamb, M. E., Sternberg, K. J., Orbach, Y., Esplin, P. W., Stewart, H., & Mitchell, S. (2003). Age differences in young children's responses to open-ended invitations in the course of forensic interviews. *Journal of Consulting and Clinical Psychology*, 71(5), 926–934. <https://doi.org/10.1037/0022-006X.71.5.926>
- Lamb, M. E., Sternberg, K. J., Orbach, Y., Hershkowitz, I., Horowitz, D., & Esplin, P. W. (2002). The Effects of Intensive Training and Ongoing Supervision on the Quality of Investigative Interviews With Alleged Sex Abuse Victims. *Applied Developmental Science*, 6(3), 114–125. https://doi.org/10.1207/S1532480XADS0603_2
- Lee, S., & Kim, J. (2021). Rapport quality in investigative interviews: effects on open-ended questions and free recall responses. *Police Practice and Research*, 22(1), 996–1008. <https://doi.org/10.1080/15614263.2020.1786691>
- London, K., Bruck, M., Ceci, S. J., & Shuman, D. W. (2005). Disclosure of Child Sexual Abuse: What Does the Research Tell Us About the Ways That Children Tell? *Psychology, Public Policy, and Law*, 11(1), 194–226. <https://doi.org/10.1037/1076-8971.11.1.194>
- London, K., Bruck, M., Ceci, S. J., & Shuman, D. W. (2007). Disclosure of Child Sexual Abuse: A Review of the Contemporary Empirical Literature. In M. E. Pipe, M. E. Lamb, Y. Orbach, & A.-C. Cederborg (Eds.), *Child Sexual Abuse: Disclosure, Delay, and Denial* (pp. 159–174). Psychology Press. <https://doi.org/10.4324/9780203936832>
- London, K., Bruck, M., Wright, D. B., & Ceci, S. J. (2008). Review of the contemporary literature on how children report sexual abuse to others: Findings, methodological issues, and implications for forensic interviewers. *Memory*, 16(1), 29–47. <https://doi.org/10.1080/09658210701725732>
- Lyon, T. D. (2014). Interviewing Children. *Annual Review of Law and Social Science*, 10(1), 73–89. <https://doi.org/10.1146/annurev-lawsocsci-110413-030913>
- Melinder, A., Magnusson, M., & Gilstrap, L. L. (2021). What Is a Child-Appropriate Interview? Interaction Between Child Witnesses and Police Officers. *International Journal on Child Maltreatment: Research, Policy and Practice*, 3(4), 369–392. <https://doi.org/10.1007/s42448-020-00052-8>
- Milne, R., & Bull, R. (2002). Back to basics: a componential analysis of the original cognitive interview mnemonics with three age groups. *Applied Cognitive Psychology*, 16(7), 743–753. <https://doi.org/10.1002/acp.825>
- Oates, K., & Shrimpton, S. (1991). Children's Memories for Stressful and Non-Stressful Events. *Medicine, Science and the Law*, 31(1), 4–10. <https://doi.org/10.1177/002580249103100102>
- Orbach, Y., Hershkowitz, I., Lamb, M. E., Sternberg, K. J., Esplin, P. W., & Horowitz, D. (2000). Assessing the value of structured protocols for forensic interviews of alleged child abuse victims. *Child Abuse and Neglect*, 24(6), 733–752. [https://doi.org/10.1016/S0145-2134\(00\)00137-X](https://doi.org/10.1016/S0145-2134(00)00137-X)
- Orbach, Y., & Lamb, M. E. (2000). Enhancing children's narratives in investigative interviews. *Child Abuse & Neglect*, 24(12), 1631–1648. [https://doi.org/10.1016/S0145-2134\(00\)00207-6](https://doi.org/10.1016/S0145-2134(00)00207-6)
- Orbach, Y., & Lamb, M. E. (2001). The relationship between within-interview contradictions and eliciting interviewer utterances. *Child Abuse & Neglect*, 25(3), 323–333. [https://doi.org/10.1016/S0145-2134\(00\)00254-4](https://doi.org/10.1016/S0145-2134(00)00254-4)
- Pereda, N., & Forn, M. (2007). Prevalencia y características del abuso sexual infantil en estudiantes universitarios españoles [Prevalence and characteristics of child sexual abuse in Spanish university students]. *Child Abuse & Neglect*, 31(4), 417–426. <https://doi.org/10.1016/j.chiabu.2006.08.010>
- Pipe, M. E., Lamb, M. E., Orbach, Y., & Cederborg, A. C. (2007). Child Sexual Abuse: Disclosure, Delay and Denial. In M. E. Pipe, M. E. Lamb, Y. Orbach, & A. C. Cederborg (Eds.), *Child Sexual Abuse: Disclosure, Delay, and Denial*. Psychology Press. <https://doi.org/10.4324/9780203936832>
- Poole, D. A., & Lindsay, D. S. (1995). Interviewing Preschoolers: Effects of Nonsuggestive Techniques, Parental Coaching, and Leading Questions on Reports of Nonexperienced Events. *Journal of Experimental Child Psychology*, 60(1), 129–154. <https://doi.org/10.1006/jecp.1995.1035>
- Powell, M. B., & Guadagno, B. (2008). An Examination of the Limitations in Investigative Interviewers' Use of Open-Ended Questions. *Psychiatry, Psychology and Law*, 15(3), 382–395. <https://doi.org/10.1080/13218710802101621>
- Powell, M. B., & Hughes-Scholes, C. H. (2009). Evaluation of the Questions Used to Elicit Evidence About Abuse from Child Witnesses: Australian Study. *Psychiatry, Psychology and Law*, 16(3), 369–378. <https://doi.org/10.1080/13218710902930325>
- Principe, G. F., & London, K. (2022). How parents can shape what children remember: Implications for the testimony of young witnesses. *Journal of Applied Research in Memory and Cognition*, 11(3), 289–302. <https://doi.org/10.1037/mac0000059>
- Saywitz, K. J., & Camparo, L. B. (2014). Evidence-based child forensic interviewing: The developmental narrative elaboration interview: Interviewer guide. In *Evidence-based child forensic interviewing: The developmental narrative elaboration interview: Interviewer guide*. Oxford University Press.
- Sternberg, K. J., Lamb, M. E., Davies, G. M., & Westcott, H. L. (2001). The Memorandum of Good Practice: theory versus application. *Child Abuse & Neglect*, 25(5), 669–681. [https://doi.org/10.1016/S0145-2134\(01\)00232-0](https://doi.org/10.1016/S0145-2134(01)00232-0)
- Sternberg, K. J., Lamb, M. E., & Hershkowitz, I. (1996). Child Sexual Abuse Investigations in Israel. *Criminal Justice and Behavior*, 23(2), 322–337. <https://doi.org/10.1177/0803973233>
- Sternberg, K. J., Lamb, M. E., Orbach, Y., Esplin, P. W., & Mitchell, S. (2001). Use of a structured investigative protocol enhances young children's responses to free-recall prompts in the course of forensic interviews. *Journal of Applied Psychology*, 86(5), 997–1005. <https://doi.org/10.1037/0021-9010.86.5.997>
- Stromwall, L. A., & Granhag, P. A. (2005). Children's Repeated Lies and Truths: Effects on Adults' Judgments and Reality Monitoring Scores. *Psychiatry, Psychology and Law*, 12(2), 345–356. <https://doi.org/10.1375/pplt.12.2.345>

- Verkampt, F., & Ginet, M. (2010). Variations of the cognitive interview: Which one is the most effective in enhancing children's testimonies? *Applied Cognitive Psychology*, 24(9), 1279–1296. <https://doi.org/10.1002/acp.1631>
- Vrij, A., Akehurst, L., Soukara, S., & Bull, R. (2004). Let me inform you how to tell a convincing story: CBCA and reality monitoring scores as a function of age, coaching, and deception. *Canadian Journal of Behavioural Science / Revue Canadienne Des Sciences Du Comportement*, 36(2), 113–126. <https://doi.org/10.1037/h0087222>
- World Health Organization. (2024). *Child Maltreatment*. <https://www.who.int/news-room/fact-sheets/detail/child-maltreatment>
- Yi, M., Jo, E., & Lamb, M. E. (2016). Effects of the NICHD Protocol Training on Child Investigative Interview Quality in Korean Police Officers. *Journal of Police and Criminal Psychology*, 31(2), 155–163. <https://doi.org/10.1007/s11896-015-9170-9>
- Yi, M., & Lamb, M. E. (2018). The effects of narrative practice on children's testimony and disclosure of secrets. *Applied Cognitive Psychology*, 32(3), 326–336. <https://doi.org/10.1002/acp.3385>