

# University of Southern California Law

---

From the Selected Works of Thomas D. Lyon

---

Spring April 28, 2023

103. Child witnesses productively respond to “How” questions about evaluations but struggle with other “How” questions.

Hayden Henderson, *Stanford University*  
Colleen E Sullivan, *Arizona State University*  
Breanne E. Wylie, *University of Southern California*  
Stacia N Stolzenberg, *Arizona State University*  
Angela D. Evans, *Brock University*, et al.




Available at: <https://works.bepress.com/thomaslyon/201/>

# Child Witnesses Productively Respond to “How” Questions About Evaluations but Struggle With Other “How” Questions

Child Maltreatment  
2023, Vol. 28(3) 417–426  
© The Author(s) 2023  
Article reuse guidelines:  
[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)  
DOI: 10.1177/10775595231175913  
[journals.sagepub.com/home/cm](https://journals.sagepub.com/home/cm)



Hayden Henderson<sup>1</sup>, Colleen E. Sullivan<sup>2</sup> , Breanne E. Wylie<sup>3</sup> , Stacia N. Stolzenberg<sup>2</sup> , Angela D. Evans<sup>4</sup> , and Thomas D. Lyon<sup>3</sup> 

## Abstract

Child interviewers are often advised to avoid asking “How” questions, particularly with young children. However, children tend to answer “How” evaluative questions productively (e.g., “How did you feel?”). “How” evaluative questions are phrased as a “How” followed by an auxiliary verb (e.g., “did” or “was”), but so are “How” questions requesting information about method or manner (e.g., “How did he touch you?”), and “How” method/manner questions might be more difficult for children to answer. We examined 458 5- to 17-year-old children questioned about sexual abuse, identified 2485 “How” questions with an auxiliary verb, and classified them as “How” evaluative ( $n = 886$ ) or “How” method/manner ( $n = 1599$ ). Across age, children gave more productive answers to “How” evaluative questions than “How” method/manner questions. Although even young children responded appropriately to “How” method/manner questions over 80% of the time, specific types of “How” method/manner questions were particularly difficult, including questions regarding clothing, body positioning, and the nature of touch. Children’s difficulties lie in specific combinations of “How” questions and topics, rather than “How” questions in general.

## Keywords

childhood sexual abuse, child witnesses, court testimony, forensic interviews

## Introduction

Child interviewers are taught to maximize their use of broad open-ended requests for recall (also known as invitations), avoid suggestive questions, and minimize their use of yes-no and forced-choice questions (Lamb et al., 2018). Directives, which include most *wh*-questions, occupy a middle ground. They are less favored than invitations because they elicit less information per question, at least for children above four years of age (HersHKowitz et al., 2012). However, they often lead to types of information that invitations fail to elicit (Lyon & Henderson, 2021). This study examined a particular type of directive, “How” questions, because of differences of opinion regarding their utility. Specifically, we examined “How” Auxiliary questions, questions in which “How” is followed by an auxiliary verb (such as “did” or “was”), and compared “How” questions about method or manner (e.g., “How did he touch you?”) to “How” questions in which children are asked for an evaluation (e.g., “How did you feel?”). In what follows, we discuss research examining the development of children’s understanding of “How” questions and why “How” method/manner questions may be more difficult for children than “How” evaluative questions.

## The Difficulty of “How” Questions

A classic finding in language development is that children start asking “How” questions later than “Who,” “What,” and “Where” questions (Brown, 1968; Smith, 1933). The explanations for this have varied, including the possibility that the concepts expressed by “How” questions are more complex for children (Bloom et al., 1982), and the fact that “How” questions appear less often in child-directed speech and are thus less familiar to children (Rowland et al., 2003).

These findings have fueled recommendations that forensic interviewers should avoid asking young children “How” questions. Steward and colleagues (1993) concluded children “cannot

<sup>1</sup>Stanford University, Stanford, CA, USA

<sup>2</sup>Arizona State University, Phoenix, AZ, USA

<sup>3</sup>University of Southern California, Los Angeles, CA, USA

<sup>4</sup>Brock University, St. Catharines, ON, Canada

## Corresponding Author:

Colleen E. Sullivan, Arizona State University, 411 N. Central Ave., Suite 600, Phoenix, AZ 85004, USA.

Email: [Colleen.E.Sullivan@asu.edu](mailto:Colleen.E.Sullivan@asu.edu)

consistently answer” “How” questions until they are 5 or 6 years of age (p. 28; see also Aldridge & Wood, 1998). In the only study examining “How” questions in forensic interviews, Malloy and colleagues (2017) analyzed 49 police interviews with 3- to 5-year-old children, and combining “How” and “Why” questions, found that only 20% of the questions were productive. They concluded that “How/Why prompts may not be effective at eliciting the intended information from preschool children” (p. 64), a view echoed by subsequent researchers (Melinder et al., 2021).

The most vigorous criticism of “How” questions was probably made by Walker and Kenniston (2013) in their practice guide for professionals published by the American Bar Association. They argued that “[m]ost ‘How’ questions in interviews with children... can be difficult even for older children (and some adults); they are problematic, if not impossible for children ages 4-6.” They conclude that “‘How’ questions... should be handled with care with children of any age” (p. 72).

Walker and Kenniston (2013) are correct regarding “How” questions about particular concepts, such as number and time. Children find it difficult to enumerate events and tend to guess in response to “How many times” questions (Wandrey et al., 2012). Children also often struggle to answer “How old” questions about their age at the time of prior events (Wandrey et al., 2012). Furthermore, “How do you know” questions are likely difficult for young children because they require source monitoring abilities, which develop rapidly as children approach school age (Roberts, 2000). In these cases, however, changing the form of the question would probably not reduce children’s difficulties, because the queried concepts are inherently difficult. There is no evidence that children find it easier to answer “What was your age?” than “How old were you?” Similarly, children with source monitoring difficulties often err in specifying whether they saw, heard, or inferred information, regardless of whether they are asked a “How do you know?” question (O’Neill & Chong, 2001).

Blanket recommendations against “How” questions are sure to concern practitioners, who often ask “How” questions in order to avoid asking yes-no questions or forced-choice questions. For example, questions about clothing placement during abuse are often asked as yes-no questions (“Were your clothes on?”) or as forced-choice questions (“Were your clothes on or off?”), but these have been shown to elicit misleading answers from young children when the clothes were neither totally on nor off (Stolzenberg et al., 2017; Wylie et al., 2021). The alternative usually cited by experienced interviewers (Anderson et al., 2010), researchers (London et al., 2017; Ruddock, 2006), and expert witnesses (Idaho v. Herod, 2016) is the question “How were your clothes?” But if “How” questions are themselves difficult, this creates a dilemma.

### **“How” Method-Manner Questions versus “How” Evaluative Questions**

An important aspect of “How” questions is that different “How” questions request very different types of information.

In developmental work examining children’s language acquisition, the focus is on “How” questions asking about method or manner (Cairns & Hsu, 1978; Ervin-Tripp, 1970; Gullo, 1981, 1982; Tyack & Ingram, 1977), which can be defined as “by what means, by what way, in what way, and with what style” (Tsai, 1999, p. 155). For example, if one asks, “How do you ride a bicycle?” a productive answer is to describe mounting the bike, balancing, and pushing the pedals. “How” questions about method or manner are usually structured such that “How” is immediately followed by an auxiliary verb, such as “do,” “did,” “was,” or “were.” We will call “How” questions of this sort “How” method/manner questions.

Interviewers may ask “How” method/manner questions about a number of different aspects of sexual abuse. As noted above, interviewers are likely to ask about clothing placement and removal by asking “How were your clothes?” and “How were his clothes?” Sullivan et al., (2022a) identified 130 cases in which miscommunication occurred between attorneys and children in child sexual abuse trials when discussing the body mechanics of abuse, and identified “How” questions as a problem in 16% of the cases. The authors noted that “how-manner questions about touching seemed to be particularly difficult,” and cited examples such as “How was he touching you?” and “How were his fingers?” (p. NP12388).

Another type of “How” question asks for evaluation, which in child witness research has been defined as “[q]uestions that asked the child to make an evaluation about judgments, emotions, thoughts, or physical sensations” (Andrews et al., 2016, p. 344; see also Ahern et al., 2015, 2018). Similar to “How” method/manner questions, “How” evaluative questions are worded as “How” auxiliary questions. A typical example of a “How” evaluative question is “How did you feel?” Several studies have found that although children rarely spontaneously mention their emotional or bodily reactions to child sexual abuse when questioned with broad, open-ended requests for recall (also known as invitations) or option-posing questions (yes-no and forced-choice questions), they usually do so when asked “How” evaluative questions (Lyon et al., 2012; Stolzenberg et al., 2021). Malloy and colleagues (2017) found that “How” and “Why” questions were usually unproductive with 3- to 5-year-olds, but acknowledged that “‘How’ questions may be particularly useful in eliciting evaluative content from children” (p. 64). However, they did not separately assess “How” evaluative questions in their sample.

One disadvantage of “How” evaluative questions is that children, particularly younger children, often give terse responses (e.g., Q: “How did you feel?”/A: “Bad”) (Fangstrom et al., 2017), making responses to “How” evaluative questions less productive than wh- questions (including “What” and “How”) that ask about causality and actions (Ahern et al., 2015, 2018; Andrews et al., 2016). However, interviewers can use pairing, in which one follows up a child’s response with an invitation. For example, if a child responds “bad” to “How did you feel?” one would follow up with “You said bad. Tell me

more about that.” Pairing “How” evaluative questions with invitations leads to productive responses, with children often explaining the reasons for their evaluations (Ahern & Lyon, 2013; Stolzenberg et al., 2021).

Though children’s acquisition of “How” evaluative questions has not received the same attention in the literature as their acquisition of “How” method/manner questions, their understanding of “How” evaluative questions appears to emerge early. For example, Rowland and colleagues (2003) found that “How are you?” and “How do you do?” were among the earliest appearing “How” questions in children’s speech. However, children’s use of these polite expressions, as well as their appropriate responses, are often characterized as merely “ritualized” (Gullo, 1982; Hamdan & Hamdan, 2020) or “memorized” (Walker & Kenniston, 2013), and thus dismissed as not evincing true understanding of “How” questions.

It is notable that polite questions (including “How are you?” “How are you doing?” “How have you been?” “How are things?”) are worded as “How” evaluative questions (“How do I look?” “How did it go?”), including the forensically relevant “How did you feel?” It might be the case that children initially learn to ask and answer “How are you?” questions with little understanding, but are ultimately able to generalize from these questions to “How” evaluative questions that seek information. We have not been able to identify any research examining the emergence of English-speaking children’s understanding of “How” evaluative questions. However, one acquisition study with two Jordanian children (speaking Arabic) noted that although one child’s first appropriate response to a “How” question was to “the ritualized and routine question” “How are you?” (the child answering with the customary greeting “Thanks to Allah”), soon thereafter the child was capable of engaging in short exchanges beyond the ritualized “How” question, answering “Excellent” to “How was the trip?” (Hamdan & Hamdan, 2020; p. 40).

Children’s early experience with “How” evaluative questions may generate ambiguities and miscommunications when they are asked “How” method/manner questions. As noted above, “How” method/manner questions are phrased similarly to “How” evaluative questions, with an auxiliary verb (such as “did” or “was”) immediately following “How.” This often results in ambiguity, which can be illustrated by the use of “How” questions in humor. On the one hand, a “How” evaluative question can be answered as if it were a “How” method/manner question: “How did you find business conditions abroad, Mr. Kane?”/“With great difficulty!” (Kael et al., 1971, p. 109). On the other hand, a “How” method/manner question can be answered as if it were a “How” evaluative question: “My dog’s got no nose.”/“How does he smell?”/“Awful!” (Chapman et al., 1989; p. 10).

### The Current Study

We tested children’s difficulty in responding to “How” method/manner questions and “How” evaluative questions.

We coded several hundred court and forensic interview transcripts and identified “How” questions in which “How” was immediately followed by an auxiliary verb. We then classified these questions as “How” method/manner questions unless they requested an evaluative judgment, in which case we classified them as “How” evaluative questions. We excluded “How do you know” questions because they require source monitoring (identifying the sources of one’s memory), which is known to be difficult for young children. Following Walker and Kennistons’ (2013) suggestion that “How” questions may be difficult even for older children, we selected a large age range (5–17 years old), though we oversampled younger children from the court samples to ensure sufficient numbers of children at different ages.

We predicted that children would produce unproductive responses more often in response to “How” method/manner questions than “How” evaluative questions. We based our prediction on language development research suggesting that children understand “How” evaluative questions earlier than “How” method/manner questions (Rowland et al., 2003), and observational research examining abuse disclosures finding that children are responsive to “How” evaluative questions (Lyon et al., 2012; Stolzenberg et al., 2021) but struggle with some “How” method/manner questions (Sullivan et al., 2022a). We identified three types of unproductive responses: inappropriate, uninformative, and non-responsive. We classified responses as inappropriate when they clearly failed to answer the question. These are particularly interesting because they may reflect misinterpretation of the intent of the question. We suspected that children might misinterpret “How” method/manner questions as “How” evaluative (because of their similar form) and thus provide inappropriate responses more often in response to “How” method/manner questions. Responses were coded as uninformative when they were clearly implied or obvious based on the question or the child’s previous response. Responses were coded as non-responsive when children gave “don’t know” answers, “don’t understand” answers, or otherwise failed to answer. We also conducted exploratory analyses of “How” method/manner questions that in our experience (and in some research; Sullivan et al., 2022a) are often asked in sexual abuse cases but may be especially difficult, namely questions about clothing placement, body positioning, and touching.

## Method

### Sample

The initial sample included 458 5- to 17-year-old children questioned about sexual abuse, either in court ( $n = 237$ ) or by forensic interviewers ( $n = 221$ ). The sample was drawn from three sources: 1) Maricopa County, Arizona criminal cases, 1/2005-12/2015: 134 5- to 17-year-olds ( $M = 12.48$ ,  $SD = 3.34$ ); 2) Los Angeles County, California criminal cases, 1/1997-11/2001: 103 5- to 9-year-olds ( $M = 7.61$ ,

$SD = 1.29$ ); 3) Los Angeles County, California forensic interviews: 221 5- to 7-year-olds, 2004-2013 ( $M = 5.27$ ,  $SD = 1.13$ ). The criminal case samples were compiled from public records (for more information see [Denne et al., 2020](#); [Szojka et al., 2023](#)). The Arizona transcripts were the original sample, which we supplemented with the California transcripts and interviews in order to increase the number of younger children. The forensic interviews were conducted at one of five different Child Advocacy Center sites in Southern California. The interviews were transcribed from video recordings and anonymized for training purposes, with the consent of the parent or legal guardian. Because the transcripts were anonymized, use of the transcripts for research was approved by the Institutional Review Board as exempt (45 CFR Section 46.014(d)(4)(ii)). An unspecified (but very small) percentage of the forensic interviews focused on physical abuse allegations.

### Coding

We began by identifying all “How” questions in the samples through machine coding, double-checked manually. “How” Auxiliary questions (“How” questions in which “How” was immediately followed by an auxiliary verb) were machine identified. The auxiliary verbs identified included “be” (“am,” “are,” “is,” “was,” “were,” “being,” “been”); “can” (“could”); “do” (“does,” “did”); “have” (“has”); and “will” (“would”). We included questions asked by either the prosecutor or defense (in the court transcripts) or the interviewer (in the forensic interviews) and excluded the small number of questions in the court transcripts asked by the judge ( $n = 23$ ) or clerk/court reporter ( $n = 8$ ). Because they implicate source monitoring abilities, we excluded “How do you know?”/“How did you know?” questions. We manually coded the questions and classified them as “How” method/manner questions unless they asked about the child’s current or prior subjective state (e.g., “How are you doing?”, “How did you feel?”), in which case we classified them as “How” evaluative.

Three types of unproductive responses were identified: inappropriate, uninformative, and non-responsive. Responses were coded as inappropriate when they reflected misunderstanding because they clearly did not answer the question (e.g., Q: “How do you start your day in second grade?”/A: “August,” Q: “How did he do that?”/A: “Easy”). Responses were coded as uninformative when they were clearly implied or obvious based on the question or the child’s previous response (e.g., Q: “How did he touch your butt?”/A: “He touched it”). Responses were coded as non-responsive when children answered “I don’t know,” sought clarification, expressed uncertainty (without providing any information, e.g., “I’m not sure”), were off-topic, or silent.

Two coders each coded half of the sample, and reliability was analyzed throughout the development of the coding guide as well as during the final coding of the transcripts. We report

the Prevalence-Adjusted Bias-Adjusted Kappa (PABAK). Because some of the response types had high prevalence indexes (due to their infrequency), Cohen’s Kappa is difficult to interpret meaningfully, and PABAK has been recommended as an alternative measure ([Brennan & Silman, 1992](#); [Byrt et al., 1993](#)). We have also reported the percent agreement as it is a more familiar and intuitive inter-rater agreement index. PABAK exceeded 0.93 for all variables, and percent agreement exceeded 96% for all variables. Specifically, for inappropriate responses, PABAK = .93, 96% agreement; for uninformative responses, PABAK = .95, 98%; and for non-responsive responses, PABAK = .99, 99%. All disagreements were resolved by discussion.

### Analysis Plan

We conducted generalized mixed models (GLMMs) to determine whether “How” question type (method/manner, evaluative) and child’s age (entered as a continuous variable and tested for linear effects) were associated with the likelihood of unproductive responses overall as well as individually (inappropriate, uninformative, or non-responsive). The models were cross-validated to identify the best fit model, which was determined by the Akaike Information Criteria (AIC), an estimator of the relative quality of a model for a given set of data ([Vrieze, 2012](#)). Because the analyses used a question and answer pair as the unit of analysis, and questions were nested within-child, a random effect for child was included to account for differences among children. Analyses were performed using the *glmer* function in the R package lme4 with the bobyqa optimizer ([Bates et al., 2015](#)) and estimated marginal means were calculated using the *emmeans* function in the R package emmeans ([Lenth, 2020](#)).

GLMMs combine the properties of linear mixed models (which incorporate random effects) and generalized linear models (which handle non-normal data) and are preferable to traditional analysis of variance (ANOVA) models because they have fewer assumptions, handle response variables from different distributions (e.g., binary), and maximize power while simultaneously estimating between-subject variance ([Bates et al., 2015](#)). Estimated marginal means correct for unbalanced data (in which the distribution of questions is unequal among question type and age), because they correct the design’s imbalance by giving equal weight to each group ([Lenth, 2010](#); [Mangiafico, 2016](#)). The best fit models are reported below accompanied by the unstandardized fixed effect estimates ( $\beta$ ), standard errors of the estimates (SE), and estimates of significance (Z and p values).

## Results

### Preliminary Analyses and Descriptives

Of the 458 children within the age range, 80% of children ( $n = 368$ ) were asked at least one “How” method/manner question



( $M = 5$ ,  $SD = 5.83$ , range = 1-49). The sample included 1599 (65%) “How” method/manner questions. Sixty-two percent of children ( $n = 286$ ) within the age range were asked at least one “How” evaluative question ( $M = 3$ ,  $SD = 3.26$ , range = 1-38). The sample included 886 (35%) “How” evaluative questions. The final sample thus consisted of 368 5- to 17-year-old children ( $M = 9.93$ ,  $SD = 4.18$ ).

Preliminary analyses confirmed that the proportion of “How” method/manner to “How” evaluative questions and the proportions of unproductive to productive responses were not significantly different based on interview setting (i.e., California trial, Arizona trial, California forensic interview) when examining the same age range (i.e., under 8 years old), so the three samples were collapsed for subsequent analyses (see Table S1 in Supplemental Materials).

### Is “How” Type and Age Associated With Unproductive Responses?

The best fit model included main effects of “How” Type (method/manner, evaluative) and age, with no interaction. “How” method/manner questions were more likely to elicit unproductive responses (16%,  $SE = 1.4$ ) than “How” evaluative questions (11%,  $SE = 1.3$ ;  $B = 0.43$ ,  $SE = 0.13$ ,  $Z = 3.36$ ,  $p < .001$ ). Younger children were more likely to produce unproductive responses ( $B = -0.18$ ,  $SE = 0.02$ ,  $Z = -8.08$ ,  $p < .001$ ). The percentages of unproductive responses among different age groups by “How” type is presented in Table 1. We then examined the different types of unproductive responses separately. Note that because the models for inappropriate, uninformative, and non-responsive responses controlled for imbalances within each response, the sum of their estimated means was not equal to the estimated means for unproductive responses.

**Inappropriate Responses.** The best fit model included “How” type and age, with no interaction. “How” method/manner questions were significantly associated with more inappropriate responses (2%,  $SE = 0.5$ ) than “How” evaluative questions (1%,  $SE = 0.3$ ;  $B = 0.74$ ,  $SE = 0.26$ ,  $Z = 2.83$ ,  $p = .005$ ). As children got older, they were significantly less likely to give inappropriate responses ( $B = -0.24$ ,  $SE = 0.05$ ,  $Z = -5.12$ ,  $p < .001$ ).

**Uninformative Responses.** The best fit model only included “How” type. “How” method/manner questions were associated with more uninformative responses (3%,  $SE = 0.4$ ) than “How” evaluative questions (1%,  $SE = 0.1$ ;  $B = 3.22$ ,  $SE = 1.01$ ,  $Z = 3.19$ ,  $p = .001$ ).

**Non-Responsive Responses.** The best fit model only included age. As children got older, they were significantly less likely to be non-responsive ( $B = -0.13$ ,  $SE = 0.03$ ,  $Z = -5.26$ ,  $p < .001$ ). Children were equally likely to be non-responsive in response to “How” method/manner questions (8.7%,  $SE = 1.2$ ) and “How” evaluative questions (8.5%,  $SE = 1.1$ ).

### Exploratory Examination of Unproductive Responses

Surprised by children’s success in providing productive responses (even the 5-year-olds were productive in response to 83% of the “How” method/manner questions), we reviewed children’s unproductive responses to determine if particular “How” questions might appear more difficult. Because this approach was exploratory, we did not analyze these percentages using inferential statistics. Three types of questions were notably unproductive. Almost half (44%) of “How” questions about clothing ( $n = 25$ ) were unproductive. For example, Q: “Ok Joseph how were Angel’s clothes?”/A: “Huh?”; Q: “How were your clothes when he did that?”/A: “Fine [3 second pause] fine.” “How” questions about body positioning ( $n = 103$ ) were unproductive 28% of the time. For example, Q: “How was your body when your dad was checking your private?”/A: “What?”; Q: “How was your body?”/A: “Not too good.” Finally, 22% of “How” questions about touching ( $n = 77$ ) were unproductive. For example, Q: “How did he touch you with his hand?”/A: “I don’t get it;” Q: “And how did he touch you?”/A: “He touched me, he touched me, regular touch with [child trails off].”

We informally examined whether inappropriate responses might constitute misinterpretation of “How” method/manner questions as “How” evaluative questions. We considered whether children’s responses were graded (e.g., slow or fast) or valenced (e.g., good or bad), because “How” evaluative questions often elicit a graded or valenced response. Out of 107 inappropriate responses, 18 were graded or valenced (17%). For example, Q: “How did he move your head?” A:

**Table 1.** Percentage of Unproductive Responses by “How” Type and Age.

	“How” Method/Manner Questions, %	“How” Evaluative Questions, %
5-year-olds	17	8
6- to 7-year-olds	10	2
8- to 10-year-olds	4	1
11- to 13-year-olds	4	1
14- to 17-year-olds	3	1

“Like he moved it--like medium;” Q: “How did that conversation come up?” A: “Fine;” Q: “How were you shown them?” A: “Bored.” In sum, our exploratory examination suggested that although children’s overall performance was good, there were specific types of “How” method/manner questions that presented difficulties, and there was some evidence that their understanding of “How” evaluative questions interfered with their understanding of “How” method/manner questions.

## Discussion

Because practitioners are often advised that “How” questions are unproductive, particularly with young children, we examined how well 5- to 17-year-old children answered “How” questions in court and forensic interviews. We compared two different kinds of “How” auxiliary questions (in which the “How” is followed by an auxiliary verb, such as “did” or “were”): “How” method/manner questions, such as “How did he touch you?” or “How were your clothes?”, and “How” evaluative questions, which ask the child for their subjective judgment, such as “How did you feel?” Based on developmental research and research examining children’s responses to questions about sexual abuse, which suggests that children acquire an early understanding of “How” evaluative questions, we predicted that “How” method/manner questions would lead to more unproductive responses than “How” evaluative questions, defining unproductive as responses that were inappropriate, uninformative, or non-responsive. Our prediction was supported. Furthermore, although we predicted that younger children would have special difficulty with “How” method/manner questions, the relative difficulty of “How” method/manner questions did not vary with age, though older children were more productive generally.

Children’s overall productivity was high. Even the youngest children (5-year-olds) gave productive answers to 83% of the “How” method/manner questions. Nevertheless, we identified specific types of “How” method/manner questions that appeared particularly difficult, namely “How” questions about clothing (e.g., “How were your clothes?”), body positioning (e.g., “How was your body?”), and touching (e.g., “How did he touch you?”). Furthermore, consistent with our supposition that children’s understanding of “How” evaluative questions is superior to their understanding of “How” method/manner questions, we found a fair proportion of children’s inappropriate answers to “How” method/manner questions were evaluative. That is, children were answering the “How” method/manner questions as if they were “How” evaluative questions.

## Limitations and Future Research

This study had several limitations. First, because we examined forensic interviews of children in the field, we could not assess the accuracy of children’s responses. Second, although we

interpreted the results as due to children’s difficulty in understanding “How” method/manner questions, we could not control for other potential explanations, including question topic, complexities in question wording, memory failure, and reluctance. The strongest evidence of difficulty in understanding came from children’s inappropriate responses, which were significantly higher in response to “How” method/manner questions, but rare.

Future experimental work, which can assess accuracy and hold constant other factors, will allow researchers to systematically compare “How” method/manner with “How” evaluative questions. Furthermore, of most interest to practitioners, future work can compare potentially difficult “How” method/manner questions with alternative means of eliciting the same information. Questions about clothing, body positioning, and touching are particularly relevant. Researchers have found that “where” questions are superior to yes-no and forced-choice questions in several respects: They are more likely to elicit accurate descriptions of intermediate clothing placement (Stolzenberg et al., 2017; Wylie et al., 2021), and they are generally more productive (Stolzenberg & Lyon, 2017). However, researchers have not compared “Where” questions (e.g., “Where were your clothes?”) to “How” questions (e.g., “How were your clothes?”). Similarly, in the future, researchers could compare “Where” questions about body positioning (e.g., “Where were your legs?”) to “How” questions (e.g., “How were your legs?”). With respect to touching, prior research has found that “What” questions about the use of the hands (e.g., “What did he do with his hands?”) are more productive than yes-no questions (e.g., “Did he do anything with his hands?”) (Henderson et al., 2023). Future work could compare “What” questions to “How” questions (e.g., “How did he touch you?”).

We predict that “Where” and “What did” questions are more productive than “How” method/manner questions. On the other hand, it might be the case that any kind of question about the method and manner of abuse is inherently difficult for children, particularly young children. Adequately describing clothing placement, body positioning, and the nature of touch requires an ability to describe spatial relationships (including a good understanding of prepositions) (Stolzenberg et al., 2017), anatomical awareness, including awareness of one’s genitalia (Milam & Nugent, 2017), and an understanding of the word “touch” (Sullivan et al., 2022b).

Another limitation is that we did not track children’s responses across the interviews (or testimony), and instead treated questions and responses as the unit of analysis. It is possible that interviewers can elicit information either before or after asking difficult “How” questions. In the future, researchers can examine whether persistent use of invitations (broad open-ended requests for recall) obviate the need to ask “How” method/manner questions, or whether follow-up questions after “How” method/manner questions might successfully resolve children’s difficulties. In forensic interviews, invitations are always a preferable means of eliciting

information, because they elicit more information per question than *wh*-questions (which include “*How*” questions) (Lamb et al., 2018). Although observational work has demonstrated that “*How*” evaluation questions are needed to elicit emotional information from children, because they tend not to spontaneously mention their emotional reactions when questioned using invitations (Stolzenberg et al., 2021), similar work has not assessed whether specific questions about clothing, body positioning, and the nature of touching are necessary after interviewers have elicited information about the abuse with invitations. Our sample does not provide a good test of the potential productivity of invitations, because it includes a large number of court transcripts, and attorneys very rarely ask invitations (Andrews et al., 2016).

### *Implications for Practitioners*

Rather than avoiding “*How*” questions as a class, practitioners should recognize that different topics call for different types of questions. “*How*” method/manner questions should be used with caution, particularly with younger children, and particularly when asking about clothing, body positioning, and touch. When children respond inappropriately, this is clear evidence of confusion. Furthermore, younger children often respond to questions they don’t understand with “don’t know” responses (Henderson & Lyon, 2021). Therefore, when children provide uninformative responses to “*How*” method/manner questions, interviewers should consider confusion as a possible explanation. Interviewers should try rewording their questions, including asking “*where*” questions about clothing placement and body positioning. As we noted in the introduction, other types of “*How*” questions are hazardous because of the underlying concepts. Children’s difficulties in estimating number and time counsel caution in asking “*How many*” and “*How old*” questions (Wandrey et al., 2012), and preschool children’s difficulty with source monitoring makes “*How do you know*” questions difficult (Roberts, 2000).

“*How*” evaluation questions, on the other hand, are unlikely to elicit confusion, and can be productive in a number of different contexts. Interviewers can ask children to describe their feelings during the abuse, after abuse, and when they went to the bathroom, the latter question screening for dysuria and other physical and psychological consequences of abuse. Children are adept at describing their emotional and physical reactions to abuse (Lyon et al., 2012; Stolzenberg et al., 2021), and their ability improves with age (Szojka et al., 2023). Questions referring to feelings in general usually lead to descriptions of emotional reactions, whereas questions specifying the child’s body usually lead to descriptions of physical reactions (Stolzenberg et al., 2021). Children’s description of their emotional reactions can lend credibility to their reports as well as assist factfinders in appreciating the impacts of abuse (Stolzenberg et al., 2021). This is particularly important given children’s tendency to exhibit little affect when describing abuse (Castelli & Goodman, 2014), which

leads to skepticism among jurors (Golding et al., 2003). Children’s descriptions of physical reactions similarly facilitate assessment of their credibility, and descriptions of pain can support the legally relevant inference that penetration occurred (Myers, 2022).

We reviewed the “*How*” evaluative questions in this sample, and in addition to questions about children’s emotional and physical reactions to abuse, we identified other types of questions of interest to practitioners. “*How*” evaluation questions spanned the entire history of the child’s case and asked the child how they felt about specific people or circumstances. Interviewers asked children about the suspect before and after the abuse first occurred, the suspect’s preparatory behaviors, the suspect’s threats and inducements to keep the abuse a secret, the mother’s witnessing of the abuse or presence in the house during abuse, the child’s disclosure, the mother’s reactions to disclosure, and the consequences of disclosure, including the arrival of the police, the forensic medical exam, the suspect leaving the home, the child’s placement in foster care, and the child’s separation from siblings.

These questions appeared to serve several purposes, including understanding the effects of abuse; assessing potential bias; explaining delays, inconsistencies, and reluctance; assessing maternal protectiveness; and assessing the child’s needs for additional services. For example, a child’s positive feelings for the suspect before the abuse started (including affection for or trust in the suspect as a father-figure) helps to explain the child’s obedience to the suspect’s actions (supporting a finding of duress), as well as the child’s subsequent reluctance to disclose the abuse. It also illustrates the damage that abuse inflicts on the child’s important relationships, particularly when coupled with the child’s description of the emotional consequences of abuse. Conversely, from a defense standpoint, negative feelings toward the suspect before the alleged abuse occurred could suggest bias.

At trial, there were also evaluative questions about testifying, about needing to testify in front of the perpetrator and the jury, and seeing the perpetrator after testifying. These questions could help to explain a child’s difficulties in providing a complete and consistent report at trial. They can also demonstrate the need for accommodations, such as breaks, a support person, or even remote testimony. Finally, there were general questions about the child’s current feelings in both the forensic interviews and at trial, including the stereotypical “*How are you?*” greetings. These questions may build rapport, and help determine if the child might need additional support.

In conclusion, this study examined children’s potential difficulty in answering “*How*” questions, specifically “*How*” questions that are immediately followed by an auxiliary verb such as “*did*” or “*was*.” We found that when “*How*” auxiliary questions were about evaluations, including “*How did you feel*” questions, children had little difficulty. “*How*” questions about method or manner were more difficult, particularly when they were about clothing, body positioning, and touching.



## Acknowledgments

Special thanks to Hannah Siepmann, Jina Hur, Nika Moussavi, Selin Yalcinkaya, Owen Friend, and Scarlet Cho.

## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported in part by the National Institute of Child Health and Human Development Grants HD087685 and HD101617.

## ORCID iDs

Colleen E. Sullivan  <https://orcid.org/0000-0001-7881-6673>

Breanne E. Wylie  <https://orcid.org/0000-0003-2526-4443>

Stacia N. Stolzenberg  <https://orcid.org/0000-0001-8082-631X>

Angela D. Evans  <https://orcid.org/0000-0003-3409-5979>

Thomas D. Lyon  <https://orcid.org/0000-0001-8179-759X>

## Supplemental Material

Supplemental material for this article is available online.

## References

- Ahern, E. C., Andrews, S. J., Stolzenberg, S. N., & Lyon, T. D. (2018). The productivity of wh- prompts in child forensic interviews. *Journal of Interpersonal Violence*, 33(13), 2007–2015. <https://doi.org/10.1177/0886260515621084>
- Ahern, E. C., & Lyon, T. D. (2013). Facilitating maltreated children's use of emotional language. *Journal of Forensic Social Work*, 3(2), 176–203. <https://doi.org/10.1080/1936928X.2013.854124>
- Ahern, E. C., Stolzenberg, S. N., & Lyon, T. D. (2015). Do prosecutors use interview instructions or build rapport with child witnesses? *Behavioral Sciences & the Law*, 33(4), 476–492. <https://doi.org/10.1002/bsl.2183>
- Aldridge, M., & Wood, J. (1998). *Interviewing children: A guide for child care and forensic practitioners*. John Wiley & Sons.
- Anderson, J., Ellefson, J., Lashley, J., & Miller, A. (2010). The cornerstone forensic interview protocol: RATA. *Thomas M. Cooley Journal of Practical and Clinical Law*, 12(2), 193–332.
- Andrews, S. J., Ahern, E. C., Stolzenberg, S. N., & Lyon, T. D. (2016). The productivity of wh-prompts when children testify. *Applied Cognitive Psychology*, 30(3), 341–349. <https://doi.org/10.1002/acp.3204>
- Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). *Parsimonious mixed models*. arXiv preprint arXiv:1506.04967. <https://doi.org/10.48550/arXiv.1506.04967>
- Bloom, L., Merkin, S., & Wootten, J. (1982). Wh-questions: Linguistic factors that contribute to the sequence of acquisition. *Child Development*, 53(4), 1084–1092. <https://doi.org/10.1111/j.1467-8624.1982.tb01374.x>
- Brennan, P., & Silman, A. (1992). Statistical methods for assessing observer variability in clinical measures. *BMJ (Clinical Research ed.)*, 304(6840), 1491–1494. <https://doi.org/10.1136/bmj.304.6840.1491>
- Brown, R. (1968). The development of wh-questions in child speech. *Journal of Verbal Learning and Verbal Behavior*, 7(2), 279–290. [https://doi.org/10.1016/S0022-5371\(68\)80002-7](https://doi.org/10.1016/S0022-5371(68)80002-7)
- Byrt, T., Bishop, J., & Carlin, J. B. (1993). Bias, prevalence and kappa. *Journal of Clinical Epidemiology*, 46(5), 423–429. [https://doi.org/10.1016/0895-4356\(93\)90018-V](https://doi.org/10.1016/0895-4356(93)90018-V)
- Cairns, H. S., & Hsu, J. R. (1978). Who, why, when, and how: A development study. *Journal of Child Language*, 5(3), 477–488. <https://doi.org/10.1017/S0305000900002105>
- Castelli, P., & Goodman, G. S. (2014). Children's perceived emotional behavior at disclosure and prosecutors' evaluations. *Child Abuse & Neglect*, 38(9), 1521–1532. <https://doi.org/10.1016/j.chiabu.2014.02.010>
- Chapman, G., Cleese, J., Gilliam, T., Idle, E., Jones, T., Palin, M., & Wilmot, R. (1989). In R. Wilmot (Ed.), *The complete Monty Python's Flying Circus: All the words, volume one*. Pantheon Books.
- Denne, E., Sullivan, C., Ernest, K., & Stolzenberg, S. N. (2020). Assessing children's credibility in courtroom investigations of alleged child sexual abuse: Suggestibility, plausibility, and consistency. *Child Maltreatment*, 25(2), 224–232. <https://doi.org/10.1177/1077559519872825>
- Ervin-Tripp, S. (1970). Discourse agreement: How children answer questions. In J. R. Hayes (Ed.), *Cognition and the development of language*. John Wiley & Sons.
- Fangstrom, K., Sarkadi, A., Lucas, S., Calam, R., & Eriksson, M. (2017). "And they gave me a shot, it really hurt": Evaluative content in investigative interviews with young children. *Children and Youth Services Review*, 82(1), 434–443. <https://doi.org/10.1016/j.childyouth.2017.10.017>
- Golding, J. M., Fryman, H. M., Marsil, D. F., & Yozwiak, J. A. (2003). Big girls don't cry: The effect of child witness demeanor on juror decisions in a child sexual abuse trial. *Child Abuse & Neglect*, 27(11), 1311–1321. <https://doi.org/10.1016/j.chiabu.2003.03.001>
- Gullo, D. F. (1981). Social class differences in preschool children's comprehension of "Wh"-questions. *Child Development*, 52(2), 736–740. <https://doi.org/10.2307/1129202>
- Gullo, D. F. (1982). A developmental study of low-and middle-class children's responses to wh-questions. *First Language*, 3(9), 211–221. <https://doi.org/10.1177/014272378200300904>
- Hamdan, J. M., & Hamdan, H. J. (2020). Emergence and development of wh-questions in Jordanian Arabic-speaking children: A longitudinal study. *East European Journal of Psycholinguistics*, 7(2), 28–55. <https://doi.org/10.29038/eejpl.2020.7.2.ham>
- Henderson, H. M., Lundon, G. M., & Lyon, T. D. (2023). Suppositional wh-questions about perceptions, conversations, and actions are more productive than paired yes-no questions when questioning maltreated children. *Child Maltreatment*, 28(1), 55–65. <https://doi.org/10.1177/10775595211067208>
- Henderson, H. M., & Lyon, T. D. (2021). Children's signaling of incomprehension: The diagnosticity of practice questions during

- interview instructions. *Child Maltreatment*, 26(1), 95–104. <https://doi.org/10.1177/1077559520971350>
- Hershkowitz, I., Lamb, M. E., Orbach, Y., Katz, C., & Horowitz, D. (2012). The development of communicative and narrative skills among preschoolers: Lessons from forensic interviews about child abuse. *Child Development*, 83(2), 611–622. <https://doi.org/10.1111/j.1467-8624.2011.01704.x>
- Idaho v. Herod (2016). *WL 4072211* (Ct. App. 2016).
- Kael, P., Mankiewicz, H. J., & Welles, O. (1971). *The Citizen Kane Book*. Limelight editions.
- Lamb, M., Brown, D., Hershkowitz, I., Orbach, Y., & Esplin, P. W. (2018). *Tell me what happened: Questioning children about abuse* (2nd ed.). John Wiley & Sons. <https://doi.org/10.1002/9781118881248>
- Lenth, R. (2020). *Emmeans: Estimated marginal means, aka least-squares means*. R package version 1.5.0. Retrieved from. <https://github.com/rvleth/emmeans>
- Lenth, R. V. (2010). Response-surface methods in R, using rsm. *Journal of Statistical Software*, 32(7), 1–17. <https://doi.org/10.18637/jss.v032.i07>
- London, K., Hall, A. K., & Lytle, N. E. (2017). Does it help, hurt, or something else? The effect of a something else response alternative on children's performance on forced-choice questions. *Psychology, Public Policy, and Law*, 23(3), 281–289. <https://doi.org/10.1037/law0000129>
- Lyon, T. D., & Henderson, H. (2021). Increasing true reports without increasing false reports: Best practice interviewing methods and open-ended wh-questions. *American Professional Society on the Abuse of Children Advisor*, 33(1), 29–39. Available at SSRN: <https://ssrn.com/abstract=3759908>
- Lyon, T. D., Scurich, N., Choi, K., Handmaker, S., & Blank, R. (2012). "How did you feel?": Increasing child sexual abuse witnesses' production of evaluative information. *Law and Human Behavior*, 36(5), 448–457. <https://doi.org/10.1037/h0093986>
- Malloy, L. C., Orbach, Y., Lamb, M. E., & Walker, A. G. (2017). "How" and "why" prompts in forensic investigative interviews with preschool children. *Applied Developmental Science*, 21(1), 58–66. <https://doi.org/10.1080/10888691.2016.1158652>
- Mangiafico, S. S. (2016). *Summary and analysis of extension program evaluation in R, version 1.20.01*. <https://rcompanion.org/documents/RHandbookProgramEvaluation.pdf>
- 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>
- Milam, L. J., & Nugent, W. R. (2017). Children's knowledge of genital anatomy and its relationship with children's use of the word "inside" during questioning about possible sexual abuse. *Journal of Child Sexual Abuse*, 26(1), 23–39. <https://doi.org/10.1080/10538712.2016.1269863>
- Myers, J. E. B. (2022). *Myers on evidence of interpersonal violence, child maltreatment, intimate partner violence, rape, stalking, and elder abuse* (6th ed., cumulative supplement). Aspen Publishers.
- O'Neill, D. K., & Chong, S. (2001). Preschool children's difficulty understanding the types of information obtained through the five senses. *Child Development*, 72(3), 803–815. <https://doi.org/10.1111/1467-8624.00316>
- Roberts, K. P. (2000). An overview of theory and research on children's source monitoring. In K. P. Roberts & M. Blades (Eds.), *Children's source monitoring* (pp. 21–68). Laurence Erlbaum Associates.
- Rowland, C. F., Pine, J. M., Lieven, E. V., & Theakston, A. L. (2003). Determinants of acquisition order in wh-questions: Re-Evaluating the role of caregiver speech. *Journal of Child Language*, 30(3), 609–635. <https://doi.org/10.1017/S0305000903005695>
- Ruddock, A. C. (2006). *The relationship of interviewer rapport behaviors to the amount and type of disclosure from children during child abuse investigations*. Unpublished dissertation. Alliant International University.
- Smith, M. (1933). The influence of age, sex and situation on the frequency, form and function of questions asked by preschool children. *Child Development*, 4(3), 201–381. <https://doi.org/10.2307/1125682>
- Steward, M. S., Bussey, K., Goodman, G. S., & Saywitz, K. J. (1993). Implications of developmental research for interviewing children. *Child Abuse & Neglect*, 17(1), 25–37. [https://doi.org/10.1016/0145-2134\(93\)90005-p](https://doi.org/10.1016/0145-2134(93)90005-p)
- Stolzenberg, S. N., & Lyon, T. D. (2017). "Where were your clothes?" Eliciting descriptions of clothing placement from children alleging sexual abuse in criminal trials and forensic interviews. *Legal and Criminological Psychology*, 22(2), 197–212. <https://doi.org/10.1111/lcrp.12094>
- Stolzenberg, S. N., McWilliams, K., & Lyon, T. D. (2017). Spatial language, question type, and young children's ability to describe clothing: Legal and developmental implications. *Law and Human Behavior*, 41(4), 398–409. <https://doi.org/10.1037/lhb0000237>
- Stolzenberg, S. N., Williams, S., McWilliams, K., Liang, C., & Lyon, T. D. (2021). The utility of direct questions in eliciting subjective content from children disclosing sexual abuse. *Child Abuse & Neglect*, 116(Pt 1), 103964. <https://doi.org/10.1016/j.chiabu.2019.02.014>
- Sullivan, C., George, S. S., Stolzenberg, S. N., Williams, S., & Lyon, T. D. (2022a). Imprecision about body mechanics when child witnesses are questioned about sexual abuse. *Journal of Interpersonal Violence*, 37(13–14), NP12375–NP12397. <https://doi.org/10.1177/0886260521997941>
- Sullivan, C. E., Stolzenberg, S. N., Williams, S., & Lyon, T. D. (2022b). Children's underextended understanding of touch. *Psychology, Public Policy, & Law*. <https://doi.org/10.1037/law0000374>
- Szajka, Z. A., Moussavi, N., Burditt, C., & Lyon, T. D. (2023). Attorneys' questions and children's responses referring to the nature of sexual touch in child sexual abuse trials. *Child Maltreatment*. <https://doi.org/10.1177/10775595231161033>
- Tsai, W. D. (1999). The hows of why and the whys of how. *UCI Working Papers in Linguistics*, 5(1), 155–184.

- Tyack, D., & Ingram, D. (1977). Children's production and comprehension of questions. *Journal of Child Language*, 4(2), 211–224. <https://doi.org/10.1017/S0305000900001616>
- Vrieze, S. I. (2012). Model selection and psychological theory: A discussion of the differences between the Akaike information criterion (AIC) and the bayesian information criterion (BIC). *Psychological Methods*, 17(2), 228–243. <https://doi.org/10.1037/a0027127>
- Walker, A. G., & Kenniston, J. (2013). *Handbook on questioning children: A linguistic perspective* (3d. ed.). American Bar Association.
- Wandrey, L., Lyon, T. D., Quas, J. A., & Friedman, W. J. (2012). Maltreated children's ability to estimate temporal location and numerosity of placement changes and court visits. *Psychology, Public Policy, and Law*, 18(1), 79–104. <https://doi.org/10.1037/a0024812>
- Wylie, B., Stolzenberg, S. N., McWilliams, K., Evans, A. D., & Lyon, T. D. (2021). Young children's ability to describe intermediate clothing placement. *Child Maltreatment*, 26(1), 87–94. <https://doi.org/10.1177/1077559520930825>