The productivity of wh- prompts when children testify

Samantha J. Andrews[1], Elizabeth C. Ahern[1], Stacia N. Stolzenberg[2], Thomas D. Lyon[3]

University of Cambridge[1]

Arizona State University[2]

University of Southern California[3]

IN PRESS: APPLIED COGNITIVE PSYCHOLOGY

Author Note

Corresponding author: Samantha J. Andrews, Department of Psychology, University of Cambridge, Free School Lane, Cambridge CB2 3RQ, U.K. Email: sja57@cam.ac.uk

Abstract

Wh- prompts (what, how, why, who, when, where) vary widely in their specificity and accuracy, but differences among them have largely been ignored in research examining the productivity of different question-types in child testimony. We examined 120 6- to 12-year-olds' criminal court testimony in child sexual abuse cases to compare the productivity of various wh- prompts. We distinguished among what/how prompts, most notably: what/how-happen prompts focusing generally on events, what/how-dynamic prompts focusing on actions or unfolding processes/events, what/how-causality prompts focusing on causes and reasons, and what/how-static prompts focusing on non-action contextual information regarding location, objects, and time. Consistent with predictions, what/how-happen prompts were the most productive, and both what/how-dynamic prompts and wh- prompts about causality were more productive than other wh- prompts. Prosecutors asked proportionally more what/how-dynamic prompts and fewer what/how-static prompts than defense attorneys. Future research and interviewer training may benefit from finer discrimination among wh- prompts.

Keywords: Wh- prompts, directive questions, child sexual abuse, defense cross-examination, prosecution direct-examination.

The productivity of wh- prompts when children testify

Best-practice guidelines for child interviewers universally recommend that interviewers ask open-ended questions and avoid closed-ended and suggestive questions (American Professional Society on the Abuse of Children [APSAC], 2012; Home Office, 2011; Lamb, Hershkowitz, Orbach, & Esplin, 2008), because open-ended questions are more productive and elicit fewer errors. However, there is substantial uncertainty regarding what constitutes open questions, particularly whether wh- prompts (what, how, why, who, when, where) are open-ended. Interviewers typically ask large numbers of wh- prompts (Lamb et al., 2008), and they include a very diverse group of utterances that vary widely in their specificity, including prompts about actions, causality, people, places, and objects. Distinctions among these prompts have largely been ignored in previous research examining the productivity of different question types in child testimony but are potentially valuable both for researchers and practitioners. Because of the legal significance of children's actual performance in court, the present study investigated the prevalence and productivity differences between wh- prompts when prosecutors and defense attorneys questioned children about sexual abuse.

Defining Open-ended Questions

The definition of open-ended varies widely among different researchers and practitioners (Oxburgh, Myklebust, & Grant, 2010). The present study conceptualizes open-ended questions in accordance with the NICHD Investigative Interview Protocol coding scheme. This coding scheme classifies questions on the basis of the memory type accessed (i.e., recall vs. recognition), rather than the number of words the question typically elicits. According to the NICHD Protocol, there are two types of open-ended prompts. The first type is invitations, which elicit free-recall responses from children. General invitations are characterized by the absence of a specific memory cue (e.g., "Tell me everything that

happened from the very beginning to the very end."), and cued invitations refocus children's attention on previously mentioned details (e.g., "You mentioned [content previously mentioned by the child]. Tell me more about that."). In the NICHD scheme, a second type of open-ended prompt is directives, which refocus children on previously mentioned details of the allegation, and are phrased as wh- prompts, including what, how, why, who, when, and where. In some classification schemes, wh- prompts are not regarded as open-ended questions, unless they request free recall, such as "what happened" or "what happened next?" (e.g., Benson & Powell, 2015; Milne & Bull, 1999; see Oxburgh et al., 2010, for a review). They are treated as open-ended by the NICHD scheme because they tap cued-recall memory. Although directives are not as effective as invitations in encouraging children to provide spontaneous and elaborative accounts, directives are more productive than option-posing questions, which include yes/no and forced-choice questions (Lamb et al., 2008). Option-posing questions tap recognition memory and tend to elicit brief responses from children as well as increase the risk of error (Bruck, Ceci, & Hembrooke, 1998, Lamb et al., 2008; Lyon, 2014).

Some very open-ended wh- prompts may be considered invitations (e.g., "What happened?"), but even when interviewers follow the NICHD protocol, they are likely to use a large number of wh- prompts that are more specific than invitations and classified as directives (Lamb, Sternberg, & Esplin, 2000). Directive wh- prompts are a means of eliciting information that children might not produce in response to invitations and cued invitations. For example, children often fail to spontaneously mention their subjective reactions to abuse (Snow, Powell, & Murfett, 2009), and questions such as "how did you feel when…" have been effective in eliciting that information (Lyon, Scurich, Choi, Handmaker, & Blank, 2012). If children are motivated to conceal information, they are less likely to do so if interviewers move beyond free recall and ask more direct questions (Pipe & Wilson 1994).

Furthermore, directive wh- prompts may be more effective than invitations at eliciting informative responses from younger children because they make specific requests that demand less retrieval effort (Kulkofsky, Wang, & Ceci, 2008; Schneider & Bjorklund, 1998). There are large developmental improvements in children's ability to self-generate cues that enable them to recall information (Bjorklund & Muir 1988). Indeed, in interviews conducted using the NICHD protocol, directives elicited more informative responses from preschoolers (3- to 4-year-olds) than invitations, whereas the reverse occurred for older children (Hershkowitz, Lamb, Orbach, Katz, & Horowitz, 2012).

Differentiating Among Wh- Prompts

Previous research has not distinguished among various types of directives. As a result, little is known about productivity differences among wh- prompts. Some researchers have distinguished between wh- prompts focusing on static contextual information (e.g., "What did he wear?") and dynamic based wh- prompts focusing on actions or events (e.g. "How did you get hurt?), but without examining their differences in productivity (Peterson & McCabe, 1992; Price & Roberts, 2011). Lamb and colleagues found that cued invitations that referenced actions were more productive than invitations about appearances and locations (Lamb, Sternberg, Orbach, Esplin, Stewart, & Mitchell, 2003). The same may be true of directives. Wh- prompts that focus on actions may be especially productive because children are likely to remember actions better than descriptions (Goodman, Hirschman, Hepps, & Rudy, 1991; Peterson, Dowden, & Tobin, 1999). For example, Peterson and colleagues (1999) questioned 3- to 5-year-olds one week after a play interaction with an adult. Whereas children's responses to wh- questions about actions were quite accurate (84% correct, 5% error), their answers to questions about clothing (43% correct, 29% error) and the room (14% correct, 24% error) were quite inaccurate.

Actions are fundamental in sexual abuse cases since the abuse is itself a series of actions performed by the perpetrator, as is grooming in preparation for the abuse, and attempts at concealment. Furthermore, because most sexual abuse prosecutions involve familiar perpetrators and repeated abuse (Stolzenberg & Lyon, 2014), it is unlikely that descriptions of people and places in those cases play a substantial role in determining whether abuse occurred (Ceci, Ross, & Toglia, 1987).

Wh- prompts are classified as suggestive according to the NICHD Protocol if they presuppose information that the child has not provided (e.g., "What other things did you wear?" if the child had not mentioned additional clothing, see Brown, Lamb, Lewis, Pipe, Orbach, & Wolfman, 2013). Wh- suggestive questions, however, may be less error-prone than suggestive questions that are narrower in form, such as tag questions (e.g., "He touched you, didn't he?"), which are known to elicit higher rates of error among younger children (Krackow & Lynn, 2003). Although we know of no research that has directly compared wh-suggestive questions with other types of suggestive questions, it is consistent with the findings of Brown and colleagues (2013), who conducted an analog study utilizing the NICHD Protocol. They found that children's accuracy in response to suggestive prompts was comparable to their responses to cued invitations (both approximately 80% averaged across conditions). Although the researchers did not note the percentage of suggestive prompts that were worded as wh- questions, this was likely very high, because interviewers were trained in the NICHD Protocol and did not ask any option-posing questions.

We are aware of only one study that has examined productivity differences among different types of wh- questions. Examining rapport-building in the courtroom, Ahern, Stolzenberg, & Lyon (2015) calculated the number of words elicited by different types of wh-prompts asked by prosecutors before sexual abuse was disclosed at trial. They classified what and how questions that asked about actions as what/how-dynamic prompts, questions that

asked about causality as what/how-causality prompts, and questions that asked about descriptions of persons, places or things as what/how-static prompts. What/how-dynamic and causality prompts elicited the highest number of words per prompt. Because the study focused on rapport-building, (i.e., the number of words uttered before abuse was disclosed), the relation between wh- prompts and the number of new abuse-related details could not be determined. Furthermore, they did not distinguish between wh- prompts that would be classified as directives and those that would be classified as invitations or as suggestive prompts according the NICHD Protocol coding scheme.

Attorney Differences in Questioning

When questioning children in court about alleged sexual abuse, both prosecutors and defense attorneys have been found to use more closed-ended questions than open-ended questions, and defense attorneys ask more suggestive questions (e.g., Andrews, Lamb, & Lyon, 2015; Hanna, Davies, Crothers, & Henderson, 2012; Klemfuss, Quas, & Lyon, 2014; Stolzenberg & Lyon, 2014; Zajac & Cannan, 2009). Furthermore, prosecutors ask very few invitations (3%), and defense attorneys virtually never do so (Andrews et al., 2015).

Prosecutors might benefit from the use of invitations, insofar as they could increase the level of detail in the child's testimony and improve the child's credibility. However, prosecutors might deliberately avoid invitations for two reasons. First, they are likely to structure their direct examination quite carefully, based on a child's prior disclosures, the charges in the case, and their opening argument, during which they inform the jury what they expect the child to say. Therefore, they are unlikely to want to elicit new details, or details of abusive acts not previously described, which is more likely if they ask children invitations. Second, a common (but mistaken) belief among many attorneys is that it is clearly objectionable to ask questions that "call for a narrative" (Lyon, 2013). The basis for the objection is that a very open-ended question may elicit inadmissible evidence. However,

virtually everything a child reports about interactions with the defendant are likely to be admissible (the actions will be relevant, and the statements will be admissible as statements by a party opponent), and anything extraneous is subject to a motion to strike.

Defense attorneys understandably avoid invitations in order to focus and limit witnesses' accounts. They are allowed to ask leading questions in cross-examination (Federal Rules of Evidence 611, 2015), and are advised to do so exclusively (Pozner & Dodd, 2004). Indeed Pozner and Dodd argued that "the adept cross-examiner never uses [wh-] questions," explaining that "[t]hese words create the polar opposite of closed-ended questions. These words invite uncontrolled, unpredictable, and perhaps unending answers" (p. 8-14).

Given the virtual non-existence of invitations in court testimony, it is especially profitable to examine the productivity of different types of wh- prompts. Certainly, prosecutors and defense attorneys have very different goals, and are therefore likely to use wh- prompts in different ways. Because children's memory for actions tends to be superior, prosecutors are likely to ask more wh- questions about actions, and defense attorneys are likely to ask more wh- questions seeking descriptions. Furthermore, because defense attorneys are advised to avoid allowing witnesses to explain their answers (Pozner & Dodd, 2004), prosecutors are likely to ask more wh- questions about causality.

Current Study

The current study investigated patterns of productivity among different types of wh- prompts in children's testimony about sexual abuse. We extended previous research (Andrews et al., 2015; Ahern et al., 2015) by supplementing the NICHD Protocol question type coding to further distinguish among different wh- prompts, coding the number of details in children's responses, and by investigating attorney differences in their use of wh- prompts. We predicted that (a) consistent with prior research, invitations would be more productive than directives, and directives more productive than option-posing (yes/no and forced-choice)

questions; (b) what/how-happen prompts (what and how questions that use the word "happened") would be the most productive, and would be classified as invitations; (c) what/how-dynamic prompts (about actions) would be more productive than what/how-static prompts (asking for descriptions); (d) what/how causality prompts would be more productive than what/how-static prompts; (e) prosecutors would ask a higher proportion of what/how-dynamic prompts than defense attorneys, and (f) defense attorneys would ask a higher proportion of what/how-static prompts than prosecutors.

Method

Sample

Transcripts of 106 trials involving a total of 120 alleged victims of child sexual abuse were included in the study. These were selected from a larger sample of 223 trials (309 children) involving felony charges of child sexual abuse that went to trial in Los Angeles County between 1997 and 2001. The children were selected for the present study if they 1) were victims of abuse (as opposed to non-victim witnesses); 2) were aged 12 and under at the time of trial; 3) did not have the assistance of a translator while testifying, and 4) did not fully recant the alleged abuse while testifying. The trials included in the present study involved 68 different prosecutors and 88 different defense attorneys. These trials are the source of a number of previously published papers (e.g. Ahern et al., 2015; Andrews et al., 2015; Stolzenberg & Lyon, 2014), but the coding and research questions in this study are novel.

Children reported single ($n = 43$) or multiple ($n = 77$) sexually abusive experiences involving penetration ($n = 53$), touching under clothes ($n = 37$), touching over clothes ($n = 21$) and indecent exposure ($n = 9$). The final sample included 98 girls and 22 boys from 6 to 12 years of age ($M = 9.6$ years). Children were categorized on the basis of age at the time of trial into 3 groups: 6- to 8-year-olds ($n = 34$), 9- to 10-year-olds ($n = 45$), and 11- to 12-year-

olds ($n = 41$). No information was available concerning the children's socioeconomic and ethnic backgrounds.

All defendants were male. In 90% ($n = 108$) of the cases, children knew the alleged abusers. The suspects were biological parents ($n = 10$), step-fathers/mothers' boyfriends ($n = 23$), other family members ($n = 24$), family friends ($n = 23$), acquaintances ($n = 28$) and strangers ($n = 12$). Defendants were either convicted ($n = 89$) or acquitted ($n = 25$). The remaining 6 cases resulted in mistrials.

Coding of Transcripts

The transcripts contained direct and often redirect examinations, in which prosecutors questioned the children, and cross and often recross examinations, in which defense attorneys questioned the children. Only the substantive prompts were coded. Substantive utterances were defined as those designed to elicit information about what happened during the alleged incidents, what immediately preceded the alleged incidents, within-incident interventions (e.g., unexpected interruptions exposing the abuse), and other features of the abuse (e.g., how long the incidents lasted, where they happened) or surrounding circumstances used to challenge witness character and testimony (e.g., events providing motives for false reports). Children's substantive responses contained incident-related information. Non-substantive prompts that aimed to inform child witnesses about the purpose of the court proceedings, provide details about the examination procedure, and build rapport were not included. By definition, children's non-substantive responses did not contain incident-related information and were also not included.

NICHD utterance types. First, all attorney utterances were categorized into commonly used categories to classify questioners' utterances: invitations, directive, option-posing and suggestive prompts (e.g., Lamb et al., 2008). When a single turn in the dialogue included two or more statements or questions that could be coded differently, the question was coded using

the riskier question type classification (e.g., suggestive over option-posing, option-posing over directive). Definitions and examples of each NICHD Protocol utterance type are provided in Table 1.

Wh- prompts. Wh- prompts were then identified and categorized into general wh-categories: what, how, why, who, when, and where.

What/how prompts. What and how prompts were then further coded using the coding scheme defined in Ahern et al. (2015), with the additional category of what/how-happen. Thus, they were coded as what/how happen, dynamic, causality, evaluative, and static. Definitions and examples of each are provided in Table 2. Because why questions ask about causality and are therefore similar to what/how causality questions, we combined them after assessing their frequency and productivity.

Children's productivity. Responses related to the investigated incident(s) were identified as substantive and coders then tabulated the number of new details conveyed by the child in each utterance using a procedure described by Lamb, Hershkowitz, Sternberg, Esplin, Hovav, Manor, and Yudilevitch (1996). Details were the smallest unit for analyzing information provided by children pertaining to the alleged incidents. Details consisted of the naming, identification, or description of individuals, objects, events, places, actions, emotions, thoughts, and sensations, that are part of an alleged incident, as well as any of their features (e.g., appearances, locations, times, durations, temporal orders, sounds, smells, and textures). Repeated words or details between and within utterances were counted once across the child's testimony, unless the repetition appeared intentional (e.g., for emphasis). Details were only counted when they added to the understanding of the target incident(s), therefore false starts (e.g., "I – they went…"; "Um, well…"), statements that expressed the child's present mental or emotional state (e.g., "I am scared"), phrases that suggested the level of confidence of the interviewee during the interview (e.g. "I know"; "I think"; "Maybe"), and

claims of lack of knowledge/ignorance (e.g., "I don't know"; "I don't remember") are examples of what were *not* counted as substantive details.

Inter-Rater Reliability

Inter-rater reliability in the identification of substantive prompts was 100%. Reliability for the NICHD Protocol question types and details had been previously calculated by two independent raters on a random selection of 20% of the transcripts. Reliability in the classification of attorney question types was reported by Andrews et al. (2015) to be high, *K* = .91. For the what/how prompt coding, one of the authors was trained to code reliably with an experienced coder by practicing on two transcripts. They achieved above 90% agreement on each (number of agreements divided by the number of disagreements). To ensure that coding was consistent over time, 20% of the transcripts were randomly selected and coded by the experienced coder throughout the course of coding (after 25%, 50%, 75%, and 100% completed); Kappas = .87, .91, .95, and .92, respectively. Reliability in the identification and counting of details was also reported by Andrews et al. (2015) to be high, *K* = .81.

<div align="center">

Results

</div>

NICHD Protocol Utterance Types

Frequency tabulations of prompt types by NICHD Protocol utterance types are shown in Table 3. There was a large degree of overlap between what/how happen prompts and invitations: 88% of what/how-happen prompts were invitations, and 86% of invitations were what/how-happen prompts. There was also a large degree of overlap between the other wh-prompts (excluding the what/how happen prompts) and directives: 88% of other wh- prompts were directives, and 93% of the directives were other wh- prompts.

Table 4 displays the prevalence and productivity of NICHD Protocol utterance types. We first examined the productivity differences among NICHD Protocol utterance types in order to determine if the differences observed in prior research examining forensic interviews

would replicate when examining courtroom testimony. A mixed model analysis of variance was conducted (ANOVA) on the number of details provided by children per turn. The within-subject factor was NICHD Protocol utterance type (invitation, directive, option-posing, suggestive) and the between subject factor was age category (6-8, 9-10, 11-12 year olds). Attorney type was not included in the analysis because of the infrequency with which defense attorneys asked invitations, which would have created a serious missing data problem. Main effects emerged for NICHD Protocol utterance type, $F(3, 315) = 91.24$, $p < .001$, $\eta^2_p = .47$ and age group, $F(2, 105) = 4.56$, $p = .013$, $\eta^2_p = .08$.

Invitations elicited more details than directives, directives elicited more details than option-posing questions, and suggestive questions elicited more details than option-posing questions, $ps < .001$. Tukey comparisons revealed that the oldest age group produced more details ($M = 2.25$, $SD = .70$) than the middle ($M = 1.88$, $SD = .47$), $p = .001$, and youngest age groups ($M = 1.62$, $SD = .49$), $ps <.001$; and that the middle and youngest age group produced comparable amounts of details, $p = .20$.

Attorney Differences in Types of Wh- Prompts Asked

The mean number and percentage of wh- prompts asked by each attorney is displayed in Table 5. Why and what/how-causality prompts were used at similar rates and were combined (what/how-causality, $M = .06$, $SD = .05$; why, $M = .05$, $SD = .03$; $t(119) = 1.68$, $p = .10$).

A mixed model ANOVA on the average proportions of questions asked was conducted, with wh- prompts (what/how-happen, what/how-dynamic, what/how-causality or why, what/how-evaluative, what/how-static, when, where, who) and attorney type (prosecutor, defense) entered as within-subject factors and age category (6-8, 9-10, 11-12 year olds) entered as between subjects factors. A main effect for wh- prompt emerged, $F(7, 117) = 16.40$, $p < .001$, $\eta^2_p = .12$, which was qualified by an interaction between wh- prompt

and attorney type, $F(7, 117) = 33.31$, $p < .001$, $\eta^2_p = .22$. T-tests comparing wh- prompts between attorney type revealed that prosecutors asked proportionally more happen, dynamic, evaluative, and where [$t(119)$s = 9.46, 3.28, 4.32, 3.68, ps <.005, respectively] whereas defense attorneys asked proportionally more static and when prompts [$t(119)$s = 5.92, 4.69, ps < .001, respectively].

Wh- Prompt Productivity

The prevalence and productivity (i.e., the number of details they elicited from children) of wh- prompts are shown in Table 6. Why and what/how-causality prompts elicited comparable amounts of details and were combined (what/how-causality, $M = 4.91$, $SD = 3.37$; why, $M = 4.78$, $SD = 3.30$; $t(95) = .27$, $p = .79$). To examine productivity differences among wh- prompts, a mixed model ANOVA was performed. The within-subject factor was wh- prompt (what/how-happen, what/how-dynamic, what/how-causality or why, what/how-evaluative, what/how-static, when, where, who) and the between subject factor was age category (6-8, 9-10, 11-12 year olds). Attorney type was not included in the analysis because of the infrequency with which defense attorneys asked what/how-happen and what/how-evaluative prompts.

Only a main effect due to wh- prompt emerged, $F(7, 455) = 32.82$, $p < .001$, $\eta^2_p = .34$. Paired t-tests comparing the productivity of each wh- prompt revealed that what/how-happen was more productive than other wh- prompt (ps < .001), what/how-dynamic and what/how-causality or why prompts were comparable in productivity and each was more productive than what/how-evaluative, what/how-static, when, where, and who prompts (ps < .001). What/how-static prompts were less productive than what/how-evaluative ($p < .001$), when ($p < .001$), where ($p < .001$), and who ($p = .004$) prompts. When, where, and who prompts elicited similar numbers of details.

Because of the low productivity of some of the wh- prompts, we conducted exploratory analyses comparing the what/how-static, what/how-evaluative, when, where and who prompts to option-posing prompts. Each of these wh- prompts elicited more details in the average child's turn than option posing prompts ($ps < .001$), $t(119) = 11.26$, $t(87) = 7.24$, $t(99) = 6.13$, $t(118) = 7.29$, $t(116) = 7.65$, respectively.

Discussion

With the exception of very open-ended wh- questions (such as "what happened?"), productivity differences between various wh- prompts have been overlooked in prior research despite the fact that these questions are extremely common when children are questioned about abuse. Through examining the substantive components of children's testimony in sexual abuse trials, this study showed that different wh- prompts exhibited substantial differences in children's productivity, consistent with our predictions. Furthermore, prosecutors and defense attorneys showed different patterns in their use of wh-prompts, likely attributable to different motivations for questioning.

Consistent with previous work, children reported more details in response to invitations than to directives, and reported more details in response to directives than to option-posing prompts. Moreover, children received virtually no invitations. Mapping the NICHD Protocol classifications of utterance types onto wh- prompts, we showed that most invitations were what/how-happen prompts (and conversely most what/how-happen prompts were invitations). The other wh- prompts were almost always directives, highlighting the potential value of further categorization. As predicted, clear productivity differences emerged among the wh- prompts. Of course, what/how-happen prompts were the most productive, which is consistent with the finding that invitations are more productive than directives and other question types. But among questions that would typically be grouped together as directive, what/how prompts that asked about actions and what/how and why prompts that

asked about causality were more productive then other types of wh- prompts. What/how prompts that asked for descriptions of people, places, and things were least productive.

The results support the value of distinguishing among different kinds of questions that are termed directive in the NICHD Protocol classification of utterance types. Children's memories for actions tend to be better than their memories for less salient contextual descriptions (Goodman, Hirschman, Hepps, & Rudy, 1991; Peterson, Dowden, & Tobin, 1999). In accordance with research on children's memory, this study showed that child witnesses' responses to wh- prompts about events (what/how-happen), actions (what/how-dynamic prompts) and descriptions (what/how-static prompts) also differed in the number of new details they produce about alleged abuse. Children's responses to attorney questions about causality, either framed as what/how or why questions, were also quite productive, highlighting their ability to provide explanations (Peterson & McCabe, 1992). In line with previous work (Ahern et al., 2015), no productivity differences between what/how-causality and why prompts emerged. This result stands in contrast to recommendations in the clinical literature to avoid asking children "why" questions on the basis that they are accusatory and in response children may become less productive (Anderson et al. 2010; Simmons, 1985). This might be true for specific topics, but our results do not support the recommendation as a general rule.

The results also suggest that it might be valuable to distinguish among different types of suggestive questions. Sixteen percent of what/how dynamic prompts were suggestive (and non-trivial percentages of other what/how prompts). This may be cause for concern, because these questions may be eliciting productive but inaccurate responses. On the other hand, as noted in the introduction, research examining the inaccuracies elicited by suggestive questions has largely focused on suggestive option-posing questions, whereas suggestive wh-questions may be less error-prone (Brown et al., 2013). Suggestive wh- questions require the

child to generate information on his or her own, and they are often easily answered by "nothing" or similar expressions (e.g., "no one", "nowhere"). This suggests a promising avenue for future research on the types of suggestive questions that are most detrimental to accuracy.

Turning to attorney differences, prosecutors asked proportionally more what/how-happen, dynamic, evaluative, and where prompts than defense attorneys, whereas defense attorneys asked proportionally more what/how-static and when prompts than prosecutors. Notably, prosecutors were more inclined than defense attorneys to ask the more productive types of wh- prompts, and with respect to what/how-dynamic and what/how-static prompts, asking about details that children are more likely to recall well. This finding is consistent with research showing that defense attorneys are more likely to focus on peripheral aspects of the abuse when cross-examining children (Ceci & Bruck, 1995). Notably, however, prosecutors asked a large number of wh- prompts calling for descriptions, and this was indeed the most common type of wh- prompt. Compared to prosecutors, defense attorneys also asked a higher proportion of when prompts suggesting that they may focus more on temporal details, often a difficult topic for children (Stewart, Katz, & La Rooy, 2011). However, temporal prompts can be framed in a number of different ways (e.g., "When did it happen?"; "How old were you?"), and future research may benefit from examining productivity differences in temporal questions specifically.

Our findings have potential implications for optimal training on questioning children in forensic settings. Prosecutors (and child investigative interviewers in general) might benefit from training programs that encourage and teach greater use of action-focused wh-prompts. At the same time, some relatively unproductive wh- prompts may be important to ask if they elicit details that the child does not otherwise produce. In these cases, prosecutors could be advised to pair those prompts (such as evaluative questions, e.g., "How did you

feel?", "Bad") with more productive follow-ups, such as a what/how causality prompt ("How did it make you feel bad?") (Ahern & Lyon, 2011). Furthermore, from a training perspective, discussing the productivity of different types of wh- prompts may be easier for interviewees to understand than the invitation/directive distinction. Anecdotally, we have observed that interviewers often believe that they are asking invitations when they use the "tell me" construction coupled with a narrowly focused wh- prompt (e.g. "Tell me who was there") or even a yes/no question (e.g. "Tell me if it was dark"). Greater focus on wh- prompts enables trainers to note that the most productive and open-ended questions make reference to what happened ("What happened?"; "What happened next?"; "How did it happen?"), and the actions involved.

Limitations

In the present sample, all of the trials were tried in a single county 12-17 years ago. Of course, the results may not be representative of attorneys' behavior in other counties and at other times. However, Los Angeles County is the largest and most populous county in the United States, as well as highly diverse, socioeconomically and ethnically, and the cases used in the present study were from courts in 11 different branches throughout the county. Furthermore, there is little evidence that attorneys' questioning techniques have improved over time. For example, Hanna et al. (2012), who analyzed courtroom transcripts in New Zealand from 2008, noted that their results were similar to those reported by Davies and Seymour (1998), who examined transcripts from cases tried in 1994. Nevertheless, it would be fruitful for future research to examine a more recent sample of cases from Los Angeles County to determine whether questioning practices have changed over the years. Future research should also seek to examine trials conducted in other parts of the United States, and indeed other countries, to help determine generalizability. Moreover, the samples to date examining the utility of various wh- prompts have focused on children's criminal court

testimony. It would be valuable to examine children's responsiveness to such prompts in other forensic settings, such as best-practice forensic interviews, when questioners are trained in the use of open-ended questions and are not seeking a specific outcome.

Conclusion

This study shows that there are important productivity differences among wh-prompts asked of child witnesses, with what/how-happen prompts eliciting the greatest number of details. Notably, what/how questions about actions and wh- questions about causality were also very productive, whereas other types of wh- questions were less so, in particular what/how questions asking for descriptions. Furthermore, prosecutors and defense attorneys differed in their use of wh- questions, with prosecutors focusing more on actions and defense attorneys focusing more on descriptions. Future research and training may benefit from finer grained discrimination among wh- questions.

References

Ahern, E. C., & Lyon, T. D. (2011). *Supplemental investigative interview questions.* http://works.bepress.com/thomaslyon/74 (last accessed 15/06/15).

Ahern, E. C., Stolzenberg, S., & Lyon, T. D. (2015). Do attorneys use interview instructions or build rapport with child witnesses? *Behavioral Sciences and the Law, 33*, 476-492. doi: 10.1002/bsl.2183.

Anderson, J., Ellefson, J., Lashley, J., Miller, A. L., Olinger, A., Russell, A., Stauffer, J., & Weigman, J. (2010). The CornerHouse Forensic Interview protocol: RATAC®. *Thomas M. Cooley Journal of Practical and Clinical Law, 12,* 193–331.

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, 28,* 253-261. doi: 10.1002/acp.3103.

American Professional Society on the Abuse of Children, APSAC, (2012). *Practice guidelines: investigative interviewing in cases of alleged child abuse.* http:// www.apsac.org/ (last accessed 3/12/14).

Benson, M.S. & Powell, M.B. (2015). Evaluation of a comprehensive interactive training system for investigative interviews of children. *Psychology, Public Policy, and Law, 21,* 309-322. doi: 10.1037/law0000052.

Bjorklund, D. F. & Muir, J. E. (1988). Children's development of free recall memory: Remembering on their own. In R. Vasta (Ed.), *Annals of child development*, (Vol. 5, pp. 79-123). Greenwich CT: JAI.

Brown, D. A., Lamb, M. E., Lewis, C., Pipe, M-E., Orbach, Y., & Wolfman, M. (2013). The NICHD Investigative Interview Protocol: An analogue study. *Journal of Experimental Psychology: Applied, 19,* 367-382. doi: 10.1037/a0035143.

Bruck, M., Ceci, S. J., & Hembrooke, H. (1998). Reliability and credibility of young children's reports: From research to policy and practice. *American Psychologist, 53,* 136-151. doi: 10.1037/0003-066X.53.2.136.

Ceci, S. J. & Bruck, M. (1995). *Jeopardy in the courtroom: a scientific analysis of children's testimony.* Washington DC: American Psychological Association.

Ceci, S. J., Toglia, M. P., & Ross, D. F. (1987). *Children's Eyewitness Memory*. Springer-Verlag: New York. doi: 10.1007/978-1-4684-6338-5.

Davies, E., & Seymour, F. (1998). Questioning children complainants of sexual abuse: Analysis of criminal court transcripts in New Zealand. *Psychiatry, Psychology, & Law, 5,* 47-61. doi: 10.1080/13218719809524919.

Federal Rules of Evidence. (2014). Retrieved from: http://federalevidence.com/ downloads/rules.of.evidence.pdf (last accessed 3/12/14).

Goodman, G. S., Hirschman, J. E., Hepps, D., & Rudy, L. (1991). Invitational Issue: The Socialization of Children's Cognition and Emotion: Toward an Integrative View *Merrill-Palmer Quarterly, 37*, 109-157.

Hanna, K., Davies, E., Crothers, C., & Henderson, E. (2012). Questioning child witnesses in New Zealand's criminal justice system: Is cross-examination fair? *Psychiatry, Psychology and Law, 19*, 530-546. doi:10.1080/13218719.2011.615813.

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,* 611-622. doi: 10.1111/j.14678624.2011.01704.x.

Home Office (2011). *Achieving the best evidence in criminal proceedings: guidance on interviewing victims and witnesses, and guidance on using special measures.* http://www.cps.gov.uk/publications/docs/best_evidence_in

criminal_proceedings.pdf (last accessed 3/12/14).

Klemfuss J. Z. Quas J. A. and Lyon T. D. (2014), Attorneys' questions and children's productivity in child sexual abuse criminal trials. *Applied Cognitive Psychology, 28,* 780–788. doi: 10.1002/acp.3048.

Krackow, E., & Lynn, S. J. (2003). Is There Touch in the Game of Twister®? The Effects of Innocuous Touch and Suggestive Questions on Children's Eyewitness Memory. *Law and Human Behavior*, *27*, 589-605. doi: 10.1023/B:LAHU.0000004888.52210.35.

Kulkofsky, S., Wang, Q., & Ceci, S. J. (2008). Do better stories make better memories? Narrative quality and memory accuracy in preschool children. *Applied Cognitive Psychology, 22,* 21–38. doi: 10.1002/acp.1326.

Lamb, M. E., Hershkowitz, I., Orbach, Y., & Esplin, P. W. (2008). *Tell me what happened: structured investigative interviews of child victims and witnesses*. Chichester, UK: Wiley.

Lamb, M. E., Hershkowitz, I., Sternberg, K. J., Esplin, P. W., Hovav, M., Manor, T., & Yudilevitch, L. (1996). Effects of investigative utterance types of Israeli children's responses. *International Journal of Behavioral Development, 19,* 627-637. doi: 10.1177/016502549601900310.

Lamb, M. E., Sternberg, K. J., & Esplin, P. W. (2000). Effects of age and delay on the amount of information provided by alleged sex abuse victims in investigative interviews. *Child Development, 71,* 1586-1596. doi: 10.1111/1467-8624.00250.

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,* 926-934. doi: 10.1037/0022-006X.71.5.926.

Lyon, T. D. (2013). Child witnesses and imagination: Lying, hypothetical reasoning, and

    referential ambiguity. In M. Taylor (Ed.), *The Oxford handbook of the development of*

    imagination (pp. 126-136). New York, NY: Oxford.

Lyon, T. D. (2014). Interviewing children. *Annual Review of Law and Social Science, 10,*

    73-89. doi: 10.1146/annurev-lawsocsci-110413-030913.

Lyon, T. D., Scurich, N., Choi, K., Handmaker, S., & Blank, K. (2012). "How did you feel?":

    Increasing child sexual abuse witnesses' production of evaluative information. *Law*

    & Human Behavior, 36, 448-457. doi: 10.1037/h0093986.

Milne, R. & Bull, R. (1999). *Investigative interviewing: psychology and practice.* Chichester:

    Wiley.

Oxburgh, G.E., Myklebust, T. & Grant, T. (2010). The question of question types in police

    interviews: A review of the literature from a psychological and linguistic perspective.

    International Journal of Speech, Language and the Law, 17, 45-66. doi:

    10.1558/ijsll.v17i1.45.

Peterson, C., Dowden, C. & Tobin, J. (1999). Interviewing preschoolers: Comparisons of

    yes/no and wh-questions. *Law and Human Behavior*, *23*, 539-555. doi:

    10.1023/A:1022396112719.

Peterson, C. & McCabe, A. (1992). Parental styles of narrative elicitation: Effect on

    children's narrative structure and content. *First Language, 12(36),* 299-321. doi:

    10.1177/014272379201203606.

Pipe, M. E. & Wilson, J. C. (1994). Cues and secrets: Influences on children's event

    reports. *Developmental Psychology, 30,* 515–525. doi: 10.1037/0012-1649.30.4.515.

Pozner, L. S., & Dodd, R. J. (2004). *Cross-examination: Science and techniques (2d ed.).*

    Newark (NJ): LexisNexis.

Price, H. L. & Roberts, K. P. (2011). The effects of an intensive training and feedback program on police and social workers' investigative interviews of children. *Canadian Journal of Behavioral Science, 43*, 235 – 244. doi:10.1037/a0022541.

Schneider, W. & Bjorklund, D. F. (1998). Memory. In W. Damon, D. Kuhn, & R. S. Siegler (eds.), *Handbook of child psychology: Cognition, perception, and language* (5th edn., Vol. 2, pp. 467-521). New York: Wiley.

Simmons, J.E. (1985). *Psychiatric examination of children (4th ed.).* Philadelphia, PA: Lea & Febiger.

Snow, P., Powell, M., & Murfett, R. (2009). Getting the story from child witnesses: Exploring the application of a story grammar framework. *Psychology, Crime, and Law, 15,* 555–568. doi:10.1080/ 10683160802409347

Stewart, H., Katz, K., & La Rooy, D.J. (2011). Training forensic interviewers. In M.E. Lamb, D.J. La Rooy, L.C. Malloy, and C. Katz (Eds.), *Children's Testimony: A Handbook of Psychological Research and Forensic Practice* (pp. 199-216). West Sussex: Wiley-Blackwell.

Stolzenberg, S. & Lyon, T.D. (2014). How attorneys question children about the dynamics of sexual abuse and disclosure in criminal trials. *Psychology, Public Policy, & Law, 20,* 19-30. doi:10.1037/a0035000.

Zajac, R. & Cannan, P. (2009). Cross-examination of sexual assault complainants: A developmental comparison. *Psychiatry, Psychology and Law, 16,* 36-54. doi: 10.1080/13218710802620448.

Table 1

Attorney Question Types (NICHD Protocol Prompts)

Prompt	Definition	Examples
Invitation	Open-ended, input-free utterances used to elicit free-recall responses from children. Such questions, statements, imperatives, or contextual cues do not restrict the child's focus except in a general sense. Invitations can also follow-up on information just mentioned, or cue for additional free-recall elaboration about details previously mentioned.	"Tell me what happened?" "You mentioned he came into your room. What happened after that?"
Directive	Open-ended cued-recall questions that refocus the child on aspects or details of the allegation that they have previously mentioned, mostly using 'WH' utterances to request further information.	"When did that happen?" "What did he do with his hands?"
Option-posing	These focus the child's attention on details or aspects of the alleged incident that the child has not previously mentioned, asking the child to affirm, negate or select an investigator-given option, thus using recognition memory processes. Option-posing questions do not imply that a particular response is expected.	"Was it over or under your clothes" "Did he touch you?"
Suggestive	Statements or questions that communicates the expected response. They may introduce information not mentioned by the child but assumed by the attorney, or query the truthfulness of the child's response.	"How long did he touch you for?" [when touch was not mentioned by the child]

Table 2

What/how Prompts

Prompt	Definition	Examples
What/how-happen	Questions that included the word "happen" (which refers to a sequence of events that take place over time).	"What happened?"
What/how-dynamic	Questions that asked the child to elaborate on a more specific kind of action or process, often included the word "do" (which refers to actions).	"What did he do?"
What/how-static	Questions that asked for non-action-related descriptions, usually asking for contextual information such as location, time, or objects.	"What color was it?"
What/how causality	Questions that asked for a cause or reason.	"Why did he go?" "How do you remember that?"
What/how-evaluative	Questions that asked the child to make an evaluation about judgments, emotions, thoughts or physical sensations.	"How did you feel?"

Table 3

Prompt Type by NICHD Protocol Utterance Frequency

Prompt Type	NICHD Protocol Utterance				
	Invitation	Directive	Option-posing	Suggestive	Total
What/how-happen	725	54	0	43	822
What/how-dynamic	12	2132	0	402	2546
What/how-causality or why	2	1226	0	127	1355
What/how-evaluative	1	398	0	14	413
What/how-static	19	3290	0	290	3599
When	0	554	0	83	637
Where	0	1872	0	151	2023
Who	0	839	0	256	1095
Not a wh- prompt	81	664	22814	12611	36170
Total	840	11029	22814	13977	48660

Table 4

Prevalence and Productivity of NICHD Protocol Utterance Types

	# NICHD utterances	% NICHD utterances	# Details per Turn
	Mean (SD)	%	Mean (SD)
Invitation	7.00 (9.54)	2	8.97 (7.54)
Directive	91.91 (85.56)	23	3.23 (1.26)
Option posing	190.12 (148.49)	47	1.33 (0.46)
Suggestive	116.48 (112.51)	29	1.53 (0.57)

Note. Analyses conducted at the level of the child ($n = 120$).

Table 5

Mean Number and Percentage of Wh- Prompts by Attorney

	Prosecutors			Defense Attorneys			Total %
Wh- Prompt	Mean	SD	%	Mean	SD	%	%
What/how-Happen	6.27	9.1	8%	0.58	1.4	2%	6%
What/how-Dynamic	16.8	18.33	21%	4.42	5.68	16%	20%
What/how-Causality or Why	8.57	9.17	11%	2.73	5.54	10%	11%
What/how-Evaluative	3.05	4.29	3%	0.39	0.77	1%	3%
What/how-Static	20.47	22.15	27%	9.54	10.52	40%	29%
When	2.89	3.2	4%	2.42	3.13	10%	17%
Where	12.74	11.89	17%	4.12	5.42	13%	5%
Who	6.59	10.76	9%	2.53	4.19	9%	9%

Note. Analyses conducted at the level of the child ($n = 120$).

Table 6

Prevalence and Productivity of Wh- Prompts

	# Wh- Questions		# Details per Turn	
	Mean	SD	Mean	SD
What/how-Happen	6.85	9.43	7.98	5.97
What/how-Dynamic	21.22	21.14	5.00	2.58
What/how-Causality or Why	11.29	12.56	4.76	2.15
What/how-Evaluative	3.44	4.44	3.23	2.59
What/how-Static	30.01	29.05	2.10	0.92
When	5.31	5.16	3.30	3.30
Where	16.86	15.35	3.08	2.59
Who	9.13	14.15	2.63	1.87

Note. Analyses conducted at the level of the child ($n = 120$).