

University of Southern California Law

From the Selected Works of Thomas D. Lyon

Winter February 10, 2021

84. Don't know responding in young maltreated children: The effects of wh- questions type and enhanced interview instructions.

Kelly McWilliams, *CUNY John Jay College*

Shanna Williams, *McGill University*

Stacia N Stolzenberg, *Arizona State University*

Angela D. Evans, *Brock University*

Thomas D. Lyon

Don't Know Responding in Young Maltreated Children: The Effects of Wh- Questions Type and Enhanced Interview Instructions

Kelly McWilliams¹, Shanna Williams², Stacia N. Stolzenberg³, Angela D. Evans⁴, and Thomas D. Lyon⁵

¹ Department of Psychology, John Jay College of Criminal Justice and the Graduate Center, City University of New York

² Department of Educational and Counseling Psychology, McGill University

³ School of Criminology and Criminal Justice, Arizona State University

⁴ Psychology Department, Brock University

⁵ Gould School of Law, University of Southern California

Objective: Two studies examined 4–7-year-old maltreated children's "I don't know" (IDK) responses to wh- questions after receiving various interview instructions. **Hypotheses:** We predicted (H1) children would be less inclined to give IDK responses and more inclined to guess to color/number questions compared to other wh- questions; (H2) IDK instructions would increase children's IDK responding compared to no instructions, with an increase in accuracy; but (H3) instructions would be less effective in reducing guessing for color/number questions than other wh- questions. In Study 1, we predicted that (H4) verbalizing a commitment to answer IDK would be particularly effective. In Study 2, we predicted that (H5) IDK instructions would reduce children's accurate corrective responses, but that (H6) the negative effect of IDK instructions on corrective responses would be alleviated by a "correct the interviewer" instruction. **Method:** Across 2 studies, 301 four- to seven-year-old ($M = 5.60$, $SD = 1.09$) maltreated children viewed videos and answered wh- questions about true and false details. Both studies included a within-subjects manipulation of wh- types (color/number & wh- detail) and a between-subjects manipulation of instructions (Study 1: IDK practice, IDK practice/verbalize, control; Study 2: IDK, correct me, IDK + correct me, control). **Results:** In both studies, (a) color/number questions elicited more guessing than wh- detail questions, (b) IDK instructions decreased inaccurate responses, but they also decreased accurate responses, including accurate corrective responses, and (c) IDK instructions had a larger effect on wh- detail questions, reducing accurate corrective responses. In Study 1, verbalization failed to enhance the effect of instructions. In Study 2, the negative effect of IDK instructions on accurate corrective responses was not alleviated by instructions to correct the interviewer. **Conclusions:** Among young maltreated children, color/number questions elicit higher rates of guessing than other wh- questions. IDK instructions reduced inaccurate responses, but also reduced accurate responses.

Public Significance Statement

Guessing is a problem when young maltreated children are asked questions about color and number, probably because of the ease with which children can generate a response to such questions. Instructions designed to decrease guessing and increase don't know responses have some positive effects, but also have a tendency to reduce young maltreated children's correct responses, including responses that correct interviewers' assumptions about what occurred. Researchers and interviewers should be attentive to the tradeoffs in encouraging children to answer color/number questions, and the difficulties in improving children's performance by encouraging don't know responses.

Keywords: child maltreatment, forensic interviewing, ground rules, child memory, interview instructions

Supplemental materials: <https://doi.org/10.1037/lhb0000404.supp>

Bradley D. McAuliff served as Action Editor.

Kelly McWilliams  <https://orcid.org/0000-0001-8732-3790>

Shanna Williams  <https://orcid.org/0000-0002-5049-2598>

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

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

This research was supported by National Institute of Health Grant HD087685. For their help with interviewing children and data coding, thanks to Alicia Burrell, Scarlet Cho, Leah Chueh, Olivia De Santis, Rayna

Enriquez, Cate Fischer, Erica Kelley, Jordan Sargent, and Tianyu Wang. A special thank you to Clementine and Indiana Hawke and family for our stimuli videos.

Portions of this data were presented at the 2018 and the 2019 meetings of the American Psychology-Law Society.

Correspondence concerning this article should be addressed to Kelly McWilliams, Department of Psychology, John Jay College of Criminal Justice and the Graduate Center, City University of New York, 524 West 59th Street, New York, NY 10019, United States. Email: kmcwilliams@jjay.cuny.edu

Young children's excessive tendency to guess when asked questions, and their concomitant failure to give "I don't know" (IDK) responses, is a well-documented phenomenon (Lamb & Brown, 2006). Guessing is an obvious problem for forensic interviews, because it is likely to increase inaccuracies and inconsistencies in children's reports. In response, practice guidelines commonly recommend that forensic interviewers questioning children about abuse include an instruction on the appropriateness of answering IDK (American Professional Society on the Abuse of Children Taskforce, 2012; Lamb et al., 2018; Ministry of Justice, 2011). The factors that influence young children's tendency to guess, including question-type and the nature of instructions, are of obvious applied importance, and also provide insight into children's cognitive and social development. In two studies, we examined 4- to 7-year-old maltreated children's tendency to answer IDK varying the types of wh- questions asked and the types of IDK instructions provided.

Response Availability and IDK Responding

A number of studies have found that children are less likely to answer IDK to yes-no and forced-choice questions than wh- questions, both in laboratory studies (Geddie et al., 2001; Gee et al., 1999; Waterman et al., 2000, 2001, 2004) and in observational research (Andrews et al., 2017 [testimony]; Earhart et al., 2014 [forensic interviews]). Respondents can answer yes-no questions with a yes or no; forced-choice questions provide options with the conjunction "or," and wh- questions include what, how, who, where, when, and why questions. In part, children's disinclination to provide IDK responses to yes-no and forced-choice questions is attributable to the fact that yes-no and forced-choice questions are recognition questions, and thus more sensitive to memory than wh- questions, which are recall questions. However, children also show this tendency when asked nonsensical questions, which demonstrates that yes-no and forced-choice questions encourage guessing. For example, Waterman et al. (2000) asked 5- to 8-year-old children either yes-no (e.g., "Is a box louder than a knee?") or wh- (e.g., "What do bricks eat?") nonsensical questions. Children were more inclined to answer IDK to the nonsensical wh- questions than the nonsensical yes-no questions.

Waterman et al. (2000) theorized that children's greater tendency to guess in response to yes-no questions is due to the ease with which children can generate a response to yes-no questions compared to wh- questions (see also Earhart et al., 2014; and Marquis et al., 1972, p. 184 [in adults "willingness to guess . . . may be stronger when questions are forced-choice rather than open-ended"]). To answer a yes-no question children need only respond "yes" or "no" (or even simply nod or shake their head), whereas wh- questions require children to generate information. Similarly, to answer a forced-choice question, children need only choose one of the proffered responses, whereas again wh- questions require children to provide more information. Question types thus vary with respect to the ease with which children can generate a response. We will refer to this characteristic of questions as response availability.

Younger children are likely particularly susceptible to the effects of response availability on IDK responding. In developmental research on the origins of children's ability to reflect on their own ignorance, researchers have argued that children initially

base their judgments solely on whether they are able to generate a response. If a response is unavailable, the youngest children will acknowledge their ignorance, but if they are capable of generating any response, even if based on nothing more than pure speculation, they will claim to know (Kloo et al., 2017; Rohwer et al., 2012). Response availability emerges at a very young age for yes-no questions. Children as young as 15 months of age exhibit an awareness of the need to provide yes or no responses to yes-no questions and will do so despite their ignorance (Horgan, 1978).

An unexplored possibility is that response availability varies not only between wh- questions and option-posing questions (which encompass both yes-no and forced-choice questions), but among different types of wh- questions. At an early stage in language development, children develop "lexical domains" for abstract concepts such as color and number (Tare et al., 2008; Wagner et al., 2013). That is, before they have a good understanding of the specific meanings of color and number terms, they recognize color and number words as such. Researchers have theorized that children use lexical characteristics and contextual cues from parent-child conversations and educational contexts to create links among words in the lexical domains based on repeated and consistent exposure (Bloom & Wynn, 1997; Durkin et al., 1986; Tare et al., 2008). For example, hearing utterances such as "this color is red" or "that is the color green" allows children to link the terms "red" and "green" to the repeated term "color." As a result, when asked questions such as "What color is X?" or "How many X's are there?", very young children (as young as 20 months) provide answers that are within the relevant lexical domain (e.g., they will respond to a question about color with the name of a color) but usually inaccurate (Sandhofer & Smith, 1999; Wynn, 1992). Hence, we predicted that children would be especially prone to guess in response to wh- questions about color and number compared to other types of wh- questions.

If children exhibit a tendency to guess in response to wh- questions about color and number, this has important practical implications for forensic interviewing. Color questions are likely to occur with some frequency in interviews, such as when interviewers ask children descriptive questions about suspects, places, and abusive acts (McWilliams et al., 2013), and number questions are quite common, given the likelihood that abuse victims have experienced multiple episodes of abuse and given legal investigators' concern with establishing the number of times that abuse occurred (Brubacher et al., 2013). Researchers advise interviewers to minimize their use of recognition questions and to increase their use of recall questions to reduce guessing and error (Lamb et al., 2018), but if young children exhibit high rates of guessing in response to some types of recall questions—namely those regarding color and number—this would counsel against their use as well and suggest that interviewers use caution when interpreting children's answers to such questions.

The Effects of Enhanced Instructions on Children's IDK Responding

Although practice guidelines commonly recommend that interviewers instruct children on the acceptability of IDK responses, research support for instructions is somewhat limited (Brubacher et al., 2015). There are a number of potential problems with IDK instructions. First, it may be necessary to utilize practice and

reminders to make IDK instructions clear, sufficiently encourage children to provide IDK answers, and increase the likelihood that children retain and use the instruction throughout the interview. Danby and colleagues showed that practice with the IDK instruction increased the likelihood that children gave IDK responses (Danby et al., 2015 [5- to 9-year-olds]), in contrast to studies failing to include practice that found the instruction ineffective (Ellis et al., 2003 [4- to 5-year-olds]; Peterson & Grant, 2001 [3- to 5-year-olds]). However, practice should not be too time-consuming such that it is impractical in actual forensic interviews. Some researchers supporting instructions used rather extensive training sessions (Nesbitt & Markham, 1999 [3- to 5-year-olds]; Saywitz & Moan-Hardie, 1994 [7-year-olds]). The trick is to give children enough practice and reminders to be effective without being impractical.

Second, IDK instructions may be less effective with younger children, who, as noted earlier, may not know when they know the answer to a question, and, because of their immature executive functioning skills, may have more difficulty keeping the instruction in mind and in using it in formulating their responses. Some researchers examining the utility of IDK instructions with preschool children have found them ineffective (Ellis et al., 2003 [4- to 5-year-olds]; Geddie et al., 2001 [3- to 6-year-olds]; Peterson & Grant, 2001 [3- to 5-year-olds]). Furthermore, younger children and children with developmental delays have more difficulty understanding and practicing the instruction (Brown et al., 2019 [4- to 12-year-olds]; Dickinson et al., 2015 [4- to 12-year-olds]; Fessinger et al., 2021 [4- to 12-year-olds]).

It is particularly important to assess the effects of instructions on young maltreated children involved in court proceedings. In legal contexts, interviewers most commonly question children in cases of suspected maltreatment (Goodman et al., 1999). Maltreated children often exhibit delays in executive functioning (Williams et al., 2020) and verbal development (Lyon & Saywitz, 1999), which are likely to negatively affect their ability both to understand the instructions and to utilize the instructions in regulating their responses. If researchers are to recommend instructions to forensic interviewers, it is helpful to test their efficacy among young maltreated children.

A third potential problem with IDK instructions is that they may encourage children to be excessively cautious in responding, or to respond "IDK" without thinking (Gee et al., 1999 [Study 1: 9- to 13-year-olds]; Nesbitt & Markham, 1999 [3- to 5-year-olds]; Saywitz & Moan-Hardie, 1994 [Study 1: 7-year-olds]). This can lead to two types of error. It can reduce true positives: children may substitute accurate responses with IDK responses. Researchers who only assessed questions about false details overlooked this type of error (Endres et al., 1999 [4- to 7-year-olds]; Mulder & Vrij, 1994 [4- to 10-year-olds]).

Excessive use of the IDK response can also reduce true negatives: children may fail to deny when they should deny and fail to correct false details presupposed by questions. In the adult eyewitness literature, Scoboria and colleagues (Scoboria & Fisico, 2013; Scoboria et al., 2008) discussed how at least two distinct types of responses are commonly classified as IDK responses: those in which the participant has insufficient information to answer the question, and responses that "result from recalling that information was not present" (Scoboria et al., 2008, p. 256). The latter type of response can be called "corrective," because it corrects the

question's presupposition. For example, if an interviewer asks, "What color was the man's hat?" a corrective response would be "the man didn't have a hat." IDK instructions might lead children to provide simple IDK responses instead of giving corrective responses. If the child's corrective responses are accurate, substituting them with simple IDK responses would reduce true negatives. To counter the potentially negative effects of IDK instructions, it may be necessary to add instructions (and practice) regarding the value of responding when one *does* know the answer (Gee et al., 1999 [Study 2: 9- to 11-year-olds]; Saywitz & Moan-Hardie, 1994 [Study 2: 7-year-olds]).

In addition to examining the value of instructions with practice and feedback, we suspected that children might more appropriately use IDK responding if prompted to verbalize the IDK rule (cf. Brown et al., 2019, who recommend such an approach), rather than simply asking them to answer practice questions with "IDK." Two studies have found that when researchers asked young children (3- to 7-year-olds) to verbalize a commitment (a promise not to cheat), compared to simply affirming a request, the children were more likely to adhere to the rule (Evans et al., 2018; Heyman et al., 2015). Evans and colleagues theorized that the verbalization was a more concrete and explicit declaration of a commitment. As such, the verbalization increased children's sense of obligation and they felt more motivated to keep the promise. Analogously, we predicted that asking children to verbalize the IDK rule would improve performance.

The Current Studies

The present investigation examined young maltreated children's tendency to provide IDK responses to different types of wh- questions and whether interview instructions would influence their responses. In both studies children viewed a series of videos depicting sibling conflicts and answered wh- questions about the videos. Questions were either about false details (details not in the videos) or true details (details in the videos). We compared children's tendency to give IDK responses to wh- questions about color/number to wh- detail questions and predicted that children would be less inclined to give IDK responses to color/number questions compared to wh- detail questions (H1).

In Study 1, we examined the effects of different kinds of IDK instructions. We predicted that instructions with practice would increase children's IDK responding, compared to no instructions (H2). We anticipated, however, that the conditions with instructions would be less effective in reducing children's tendency to answer color/number questions than wh- detail questions (H3). We also included an instructions condition in which children verbalized their commitment to provide IDK responses and anticipated that this would enhance the effect of instructions (H4).

Study 2 followed up on the findings of Study 1. First, we again compared wh- color/number questions to wh- detail questions, this time more closely matching content across question types. Second, we tested a hypothesis derived from unexpected results of Study 1, which was that IDK instructions would lead children to give simpler IDK responses and fewer accurate corrective responses, in which they accurately corrected erroneous presuppositions (H5). Third, we added an instruction condition in which we combined IDK instructions with "correct the interviewer" instructions to determine if this would ensure children would distinguish between

IDK responses and accurate corrective responses. We predicted that “correct the interviewer” instructions would reduce children’s tendency to give simple IDK answers rather than accurate corrective answers (H6).

Study 1

Method

Participants

The sample included 105 four- to seven-year-old maltreated children ($M = 5.64$ years old, $SD = 1.08$; 4- to 5-year-olds $n = 45$, 6- to 7-year-olds, $n = 60$). Fifty percent ($n = 53$) of the children were female. Sixty-two percent ($n = 65$) of the children were Latino, 23% ($n = 24$) were African American, 12% were Caucasian ($n = 13$), and 3% ($n = 3$) were Asian. Based on social service records, we classified children as (a) sexual abuse with or without other maltreatment (8%), (b) physical abuse or exposure to violence with or without other maltreatment (66%), or (c) neglect without other maltreatment (30%). The numbers add to more than 100% because some children had experienced both sexual and physical harm.

Because the children included in the sample were not in the legal custody of their parents due to substantiated child maltreatment, the Presiding Judge of the Juvenile Court provided consent for their participation. Children were ineligible if they were awaiting adjudication or contested disposition hearing on the date of testing (because they might be called to testify) or if they were incapable of communicating with the researchers in English. In addition, all child participants provided verbal and written assent prior to participation. In addition to the sample reported above, four children were approached but declined to participate while nine began the study but did not complete the session because they had to leave to meet with their attorneys or attend a court proceeding.

Procedure

We tested participants individually in a private testing room in the shelter care area of dependency court. Shelter care is a large facility equipped with crafts, games, and a large outdoor play area. A research assistant identified and approached eligible children in shelter care and asked if they would like to play. If the child agreed, the research assistant escorted the child to the private testing room. Once in the room, the research assistant informed children that she had five videos for the children to watch. She explained that each video was about two sisters, she showed the children the pictures of each girl and said, “This is the big sister [points to picture]” and “This is the little sister [points to picture].” The research assistant described the videos as involving “sisters playing together.” The research assistant feigned ignorance to the content of the videos, stating that they were new to her. The research assistant said, “I need a child to watch each video for me and tell me everything that happens” and asked, “Will you help me?” When the participants agreed to help, the research assistant placed headphones on the children and turned on a laptop computer to display the videos. The research assistant turned the screen of the laptop away from herself so only the child could see

the screen. The research assistant then provided the following instruction at the start of the first video,

Ok, I will show you a video. When the video is over. Raise your hand, like this [research assistant raised hand to demonstrate]. Then, I will ask you some questions so I can find out everything that happened in the video. Remember, I don’t know what happens in the video. I won’t be able to tell you the answers to my questions.

The research assistant then started the first video. Each video depicted two actors, a “big sister” (8 years old) and a “little sister” (5 years old), engaging in daily activities that resulted in a conflict. The first video (*sweatshirt*) depicted the sisters getting ready for school, with an argument between the sisters over the little sister borrowing the big sister’s sweatshirt without asking. In the second video (*phone*) the big sister took their mother’s phone from her purse, the sisters fight about playing with the phone, and after dropping the phone, it broke. In the third video (*cleaning*) the sisters argued about cleaning their room, and when the little sister refused to help the big sister threw away the little sister’s toys. The fourth video (*breakfast*) depicted the older sister accidentally spilling milk on the little sister’s drawing while making breakfast. The final video (*TV*) portrayed an argument over what the sisters wanted to watch on TV. Each video was approximately one minute in length ($M = 51$ s, range = 30 to 84 s) and the order of videos was consistent across participants.

Materials

Wh- Types. The interviews consisted of free recall questions and wh- questions; the free recall question simply asked the child to describe what happened in the video, and we did not analyze these responses. The key questions were 11 wh- questions. The first two were wh- summary questions regarding central elements of each video (e.g., “In the video, what did the little sister take from her big sister?”). The next nine questions were wh- questions that were difficult and specific: they varied by whether they were asking about true or false details and wh- type (color/number vs. wh- detail). The true detail questions asked about actions/events in the video, and the false detail questions asked about actions/events not in the video (and therefore unanswerable; a correct answer is some form of “that did not happen”). The color/number questions asked about the color or number of objects, and the wh- detail questions asked about noncolor/number details. Both true/false detail and wh- type were varied, resulting in nine questions: four true color/number questions (two color, two number), two false color/number questions (one color, one number), two true wh-detail questions, and one false wh-detail question for each video (see [online supplemental materials](#) for a full list of questions). The order of questions was held constant across participants.

Instruction Conditions. The research assistant questioned children immediately following each video. Prior to being asked questions, we randomly assigned children (stratified by age and gender) to one of three IDK instruction conditions: (a) control ($n = 35$), (b) IDK practice ($n = 35$), or (c) IDK practice/verbalization ($n = 35$). In the control condition the children did not receive any instructions. In the IDK practice condition, we gave children an instruction to say “IDK” as well as the opportunity to practice saying IDK to an unanswerable question and correctly answering an answerable question (Lyon, 2014):

Ok [child's name], if I ask you a question and you don't know the answer, then just say, "I don't know." Let's practice: so, if I ask you "What's my dog's name?" what do you say? (wait for response). That's right, because you don't know. But what if I ask you, "Do you have a dog?" What do you say? (wait for response). Good, because you do know.

For each subsequent video, we gave children the following reminder instruction before the direct questions: "Ok [child's name], remember, if I ask you a question and you I don't know the answer, then just say 'I don't know.' But if you do know the answer, tell me."

In the IDK practice/verbalization condition, we gave children the same instructions as in the IDK practice condition, except that we asked children to verbalize the IDK rule both initially and in the reminder instructions. In the initial instruction, the interviewer said:

Ok [child's name], if I ask you a question and you don't know the answer, what should you say? (waits for response). Let's practice: so if I ask you "What's my dog's name?" what do you say? (waits for response). That's right, because you don't know. But what if I ask you, "Do you have a dog?" What do you say? (waits for response). Good, because you do know.

We reminded children after each subsequent video, before questioning: "Ok [child's name], if I ask you a question and you do not know the answer, what should you say? (waits for response). But if you do know the answer tell me." At the end of the procedure, we thanked children and let them choose a small toy for their participation.

Coding

We coded children's responses as inaccurate, IDK, or accurate. For all responses that included conflicting information we coded children's ultimate answer. For example, if the child said "don't know" but then gave a substantive response we coded the substantive response. If a child gave a substantive response but then ended with a IDK, we coded the response as don't know. When children provided corrective responses (denying that a queried detail appeared in the video, e.g., "That didn't happen") the code depended on the type of question. For questions about false details a corrective response was coded as accurate. However, in response to questions about true details a corrective response was considered inaccurate. For all questions, IDK responses included any type of response that explicitly indicated ignorance, such as "IDK," "I do not remember," and "I have no clue." We created proportion scores for each of the three response types across all five videos by occurrence and wh- type. For reliability, two coders independently coded 100% of the transcripts. Interrater reliability was high ($\kappa = .96$); the two coders discussed and resolved all discrepancies.

Results

Preliminary Analyses

We began by examining whether children's responses differed by ethnicity or sex. We conducted a series of 2 (ethnicity: Latinx, Non-Latinx) $\times$ 2 (sex: male, female) univariate analyses of

variance (ANOVAs) on children's inaccurate, IDK, and accurate responses to wh- summary, false detail, and true detail questions. There were no significant differences in children responses based on ethnicity or sex (see full results from statistical analyses in [online supplemental materials](#)), thus these variables were not included in our main analyses. In addition, in response to a reviewer's request we conducted independent sample *t*-tests to examine the effects of any IDK instruction (i.e., interview instruction conditions collapsed) on our dependent variables. There were no significant effects of instruction, beyond what is reported below (see full results from statistical analyses in [online supplemental materials](#)).

Analysis Plan

Our primary analyses concerned the effects of instruction condition, wh- type, and age on children's responses across wh- summary questions, false details, and true details. We conducted a series of mixed model ANOVAs on the proportion of children's inaccurate, IDK, and accurate responses with instruction condition (*control*, *IDK practice*, *IDK practice/verbalization*) and age group (4- to 5-year-olds, 6- to 7-year-olds) entered as between-subjects variables and wh- type (color/number, wh- detail) entered as a within-subject variable. We conducted separate ANOVAs for wh- summary, false details, and true details. We report only significant statistical tests in the text; for a full reporting see [online supplemental materials](#). For all analyses, we examined significant instruction condition differences with post hoc Fisher's least significant difference (LSD) tests. We examined significant interactions by conducting *t*-tests or univariate ANOVAs to examine the simple effects.

Wh- Summary Questions

First, we investigated children's responses to wh- summary questions (see [Table 1](#)). For children's IDK, $F(1, 99) = 7.21$, $p = .01$, $\eta_p^2 = .06$, 95% CI [.01, .17], and accurate, $F(1, 99) = 11.33$, $p = .001$, $\eta_p^2 = .10$, 95% CI [.02, .21] responses there was an effect of age, whereby older children provided significantly fewer IDK responses (older: $M = .08$, $SD = .08$; younger: $M = .14$, $SD = .17$;) and more accurate responses (older: $M = .78$, $SD = .15$; younger: $M = .66$, $SD = .20$) than younger children. Instruction condition had no effect on children's inaccurate, IDK, or accurate responses. Age did not significantly influence children's inaccurate responses.

Questions About False Details

For children's inaccurate responses (i.e., guesses), (see [Table 2](#)) instruction condition had no effect on children's inaccurate responses to questions about false details. Children in the IDK practice, IDK practice/verbalization, and control conditions provided inaccurate guesses to approximately half of the questions. There were significant main effects of wh- type, $F(1, 99) = 50.68$, $p < .001$, $\eta_p^2 = .34$, 95% CI [.19, .46], and age, $F(1, 99) = 8.53$, $p = .004$, $\eta_p^2 = .08$, 95% CI [.01, .19], as well as an Wh- Type $\times$ Age interaction, $F(1, 99) = 4.13$, $p = .045$, $\eta_p^2 = .04$, 95% CI [.00, .14]. Paired samples *t*-tests revealed an effect of wh- type for both older, $t(58) = 7.13$, $p < .01$, $d = .93$, 95% CI [.62, 1.23], and younger children, $t(45) = 3.27$, $p = .002$, $d = .48$, 95% CI [.17, .79]. Children provided significantly more inaccurate responses to color/

Table 1*Study 1 Means and Standard Deviations for Children's Responses to Wh- Summary Questions*

Age	Inaccurate			I don't know			Accurate		
	Control (n = 35)	Practice (n = 35)	Practice/verbalize (n = 35)	Control (n = 35)	Practice (n = 35)	Practice/verbalize (n = 35)	Control (n = 35)	Practice (n = 35)	Practice/verbalize (n = 35)
Younger (n = 45)	.22 (.35)	.14 (.10)	.21 (.16)	.11 (.12)	.15 (.20)	.17 (.21)	.66 (.20)	.71 (.18)	.61 (.21)
Older (n = 60)	.17 (.09)	.10 (.10)	.17 (.16)	.07 (.09)	.08 (.08)	.09 (.08)	.76 (.13)	.82 (.14)	.75 (.17)
Total	.20 (.12)	.12 (.10)	.19 (.16)	.09 (.11)	.11 (.14)	.12 (.15)	.71 (.18)	.78 (.17)	.69 (.19)

number questions (older: $M = .55$, $SD = .33$; younger: $M = .66$, $SD = .33$) than to wh- detail questions (older: $M = .28$, $SD = .32$; younger: $M = .51$, $SD = .35$). The interaction is attributable to a larger difference among older children.

When we examined children's IDK responses there were no significant effects of instruction condition, wh- type, or age. Examining the means, children in the IDK practice and IDK practice/verbalization conditions gave more IDK responses than children in the control condition (see Table 2). Children gave slightly fewer IDK responses to color/number questions ($M = .33$, $SD = .31$) than wh- detail questions ($M = .37$, $SD = .30$). Older children ($M = .37$, $SD = .28$) gave slightly more IDK responses than younger children ($M = .30$, $SD = .28$). However, none of these differences reached statistical significance.

Finally, we examined children's accurate responses to questions about false details, which are corrective responses in which children explicitly stated the question presupposed something false (e.g., "That didn't happen"). Results revealed main effects of instruction condition, $F(2, 99) = 4.73$, $p = .01$, $\eta_p^2 = .09$, 95% CI [.01, .19]; wh- type, $F(1, 99) = 46.34$, $p < .001$, $\eta_p^2 = .32$, 95% CI [.17, .44]; and age, $F(1, 99) = 8.03$, $p = .01$, $\eta_p^2 = .07$, 95% CI [.01, .19]; as well as significant interactions of Instruction Condition $\times$ Wh- Type, $F(2, 99) = 3.82$, $p = .03$, $\eta_p^2 = .03$, 95% CI [.00, .17]; and Wh- Type $\times$ Age, $F(1, 99) = 4.35$, $p = .04$, $\eta_p^2 = .04$, 95% CI [.00, .14]. For the of Instruction Condition $\times$ Wh- Type interaction, univariate ANOVAs revealed that instructions failed to affect children's accuracy in response to color/number questions, but significantly *reduced* children's accuracy in response to the wh- detail questions, $F(2, 104) = 4.51$, $p = .01$, $\eta_p^2 =$

.08, 95% C [.00, .18]. For wh- detail questions, Tukey's HSD post hoc tests revealed the both the IDK practice ($p = .04$, $d = .60$, 95% CI [.10, 1.05]) and the IDK practice/verbalization ($p = .02$, $d = .63$, 95% CI [.14, 1.10]) resulted in significantly fewer accurate responses compared to the control condition (see Table 2).

For the Wh- Type $\times$ Age interaction, paired samples t -tests showed an effect of wh- type for both older children, $t(58) = 5.79$, $p < .001$, $d = .75$, 95% CI [.46, 1.04], and younger children, $t(45) = 3.70$, $p = .001$, $d = .55$, 95% CI [.23, .85], whereby children gave more accurate responses to wh- detail (older: $M = .31$, $SD = .33$; younger: $M = .17$, $SD = .26$) questions than color/number (older: $M = .10$, $SD = .16$; younger: $M = .05$, $SD = .11$) questions. The interaction is attributable to a larger difference among older children.

Questions About True Details

Next, we investigated children's responses to questions about true details (see Table 3). First, we examined inaccurate responses to questions about true details, which includes both guesses (97%) and corrective responses (3%). There were no significant effects of instruction condition, wh- type or age. Rates of inaccurate responses were equally high across instruction condition (see Table 3), wh- type (color/number: $M = .52$, $SD = .24$; wh- detail: $M = .53$, $SD = .28$), and age (older: $M = .50$, $SD = .21$; younger: $M = .57$, $SD = .26$).

For children's IDK responses about true details a significant effect of wh- type emerged, $F(1, 99) = 19.69$, $p < .001$, $\eta_p^2 = .17$, 95% CI [.05, .29], such that children were significantly less likely

Table 2*Study 1 Means and Standard Deviations for Children's Responses to Questions About False Details by Age, Condition, and Wh- Type*

Wh- type and age	False details								
	Inaccurate			I don't know			Accurate		
	Control (n = 35)	Practice (n = 35)	Practice/verbalize (n = 35)	Control (n = 35)	Practice (n = 35)	Practice/verbalize (n = 35)	Control (n = 35)	Practice (n = 35)	Practice/verbalize (n = 35)
Color/number									
Young (n = 45)	.69 (.34)	.71 (.29)	.57 (.36)	.24 (.31)	.28 (.28)	.37 (.36)	.07 (.16)	.01 (.03)	.06 (.09)
Old (n = 60)	.52 (.34)	.59 (.33)	.53 (.32)	.33 (.338)	.34 (.29)	.39 (.32)	.15 (.23)	.07 (.11)	.09 (.14)
Total	.60 (.35)	.64 (.31)	.55 (.33)	.29 (.32)	.31 (.28)	.38 (.34)	.11 (.20)	.05 (.09)	.07 (.12)
Wh- detail									
Young (n = 45)	.46 (.39)	.50 (.29)	.57 (.38)	.31 (.32)	.37 (.26)	.31 (.29)	.24 (.33)	.13 (.20)	.12 (.22)
Old (n = 60)	.24 (.33)	.30 (.32)	.29 (.31)	.26 (.25)	.46 (.32)	.49 (.34)	.50 (.31)	.24 (.32)	.22 (.30)
Total	.35 (.37)	.38 (.32)	.41 (.36)	.28 (.28)	.42 (.29)	.41 (.33)	.37 (.34)	.19 (.28)	.18 (.27)
Wh- combined									
Young (n = 45)	.61 (.32)	.64 (.27)	.57 (.33)	.26 (.29)	.31 (.25)	.35 (.31)	.13 (.20)	.05 (.08)	.08 (.11)
Old (n = 60)	.43 (.31)	.50 (.29)	.45 (.27)	.31 (.28)	.38 (.24)	.42 (.32)	.27 (.22)	.13 (.15)	.13 (.17)
Total	.52 (.32)	.55 (.29)	.50 (.31)	.29 (.28)	.35 (.24)	.39 (.31)	.20 (.22)	.10 (.13)	.11 (.15)

Table 3*Study 1 Means and Standard Deviations for Children's Responses to Questions About True Details by Age, Condition, and Wh- Type*

Wh- type and age	True details								
	Inaccurate			I don't know			Accurate		
	Control (n = 35)	Practice (n = 35)	Practice/verbalize (n = 35)	Control (n = 35)	Practice (n = 35)	Practice/verbalize (n = 35)	Control (n = 35)	Practice (n = 35)	Practice/verbalize (n = 35)
Color/number									
Young (n = 45)	.63 (.26)	.57 (.23)	.49 (.28)	.26 (.27)	.32 (.25)	.36 (.34)	.10 (.08)	.10 (.07)	.14 (.10)
Old (n = 60)	.53 (.23)	.50 (.21)	.46 (.22)	.30 (.21)	.35 (.226)	.38 (.24)	.17 (.12)	.14 (.09)	.16 (.09)
Total	.57 (.25)	.53 (.22)	.47 (.25)	.28 (.23)	.34 (.23)	.37 (.28)	.14 (.11)	.13 (.08)	.15 (.09)
Wh- detail									
Young (n = 45)	.62 (.28)	.52 (.14)	.56 (.32)	.35 (.30)	.43 (.29)	.40 (.34)	.04 (.08)	.05 (.07)	.04 (.07)
Old (n = 60)	.58 (.21)	.48 (.26)	.47 (.30)	.36 (.19)	.45 (.26)	.46 (.30)	.06 (.07)	.07 (.07)	.07 (.08)
Total	.60 (.24)	.50 (.27)	.51 (.31)	.35 (.25)	.44 (.27)	.43 (.31)	.05 (.07)	.06 (.07)	.06 (.08)
Wh- combined									
Young (n = 45)	.62 (.26)	.55 (.23)	.51 (.28)	.29 (.27)	.36 (.25)	.38 (.33)	.08 (.08)	.09 (.06)	.11 (.07)
Old (n = 60)	.54 (.21)	.49 (.21)	.46 (.23)	.32 (.18)	.38 (.22)	.41 (.24)	.14 (.10)	.12 (.06)	.14 (.07)
Total	.58 (.23)	.52 (.22)	.48 (.25)	.31 (.23)	.37 (.23)	.39 (.28)	.11 (.09)	.11 (.06)	.12 (.07)

to say IDK in response to color/number questions ($M = .33$, $SD = .25$) than wh- detail questions ($M = .41$, $SD = .28$). The IDK instruction did not influence children's responding. Children in the IDK practice and IDK practice/verbalization said IDK slightly more often than children in the control; however, the difference did not reach significance (see Table 3). Age was also unrelated to children's willingness to admit ignorance to questions about true details (older: $M = .37$, $SD = .22$; younger: $M = .34$, $SD = .28$).

Finally, we examined children's accurate responses to true details. There were no significant effects of instruction condition, but there were significant effects of wh- type, $F(1, 99) = 68.16$, $p < .001$, $\eta_p^2 = .41$, 95% CI [.26, .52], and age, $F(1, 99) = 6.96$, $p = .01$, $\eta_p^2 = .07$, 95% CI [.00, .17]. Children were more accurate when responding to color/number questions ($M = .14$, $SD = .09$) compared to wh- detail questions ($M = .06$, $SD = .07$). In addition, older children ($M = .13$, $SD = .08$) provided significantly more accurate responses than younger children ($M = .09$, $SD = .07$).

Discussion

We found some support for our prediction that color/number questions would increase guessing by decreasing children's tendency to give IDK responses. With respect to false details, children gave more inaccurate responses to color/number questions, consistent with guessing, though this was due to a reduced tendency to give corrective responses rather than simple IDK responses. With respect to true details, children were less likely to give IDK responses to color/number questions, but this didn't lead to higher rates of inaccuracy, because their answers to color/number questions were more accurate. One might take the finding regarding true details as evidence that children's memory for color/number was superior but, coupled with their higher error rate in response to the false detail questions, this could also mean that children were simply better at guessing color/number. In Study 2, we matched content across color/number and wh- detail questions.

Instructions had few effects, and when they did affect performance they decreased accuracy. They failed to affect accuracy in

response to the wh- summary questions or in response to color/number questions. Their one effect was to reduce accuracy in response to wh- detail questions about false details. Instructions reduced children's tendency to provide accurate corrective responses. We had predicted that instructions would affect wh- detail questions more than color/number questions, but we did not anticipate that those effects would be negative. Furthermore, children in the verbalization instruction condition performed similarly to children in the instruction condition without a verbalization component.

As we noted in the introduction, after excluding studies that obtained their effects by increasing indiscriminate IDK responding, there are only a few studies that have clearly demonstrated positive effects of IDK instructions. Notably, studies finding a positive effect combined the IDK instructions with correct the interview instructions (Gee et al., 1999 [Study 2]; Saywitz & Moan-Hardie, 1994 [Study 2]), which might mitigate the IDK instruction's tendency to reduce corrective responses. In Study 2, we added a correct the interviewer instruction condition (i.e., correct me), both alone and in combination with the IDK instruction, to determine whether the correct the interviewer instruction would mitigate children's tendency to move away from corrective responses after receiving the IDK instruction, which many protocols recommend in conjunction with an IDK instruction.

Study 2 therefore had several goals. First, we created color/number and wh- detail questions with matched content to provide a clearer test of the effects of question type on guessing. Second, we sought to replicate the unexpected finding that IDK instructions decreased children's tendency to provide accurate corrective responses. Third, we added "correct the Interviewer" instruction conditions to determine if these might offset any negative effect of IDK instructions on corrective responses. To provide a more powerful test of the effects of IDK instructions, we tested a larger group of children, with equal numbers of children receiving and not receiving IDK instructions.

Study 2

Method

One hundred and eighty-nine maltreated 4- to 7-year-olds ($M = 5.55$ years old, $SD = 1.10$; 4- to 5-year-olds [$n = 93$], 6- to 7-year-olds [$n = 96$]) participated in Study 2. Forty-nine percent of the children were female and 59% ($n = 114$) of the children were Latino, 30% ($n = 58$) were African American, 8% ($n = 15$) were Caucasian, 1% ($n = 1$) were Asian and ethnicity was unavailable for 3% ($n = 4$) of participants.

As in Study 1, we examined social service records and classified children as (a) sexual abuse with or without other maltreatment (8%), (b) physical abuse or exposure to violence with or without other maltreatment (65%; 22% of the total sample had specifically suffered from physical abuse), or (c) solely neglect (50%). The recruitment, consent and ethics procedures were identical to Study 1, except for the additional requirement that no child who participated in Study 1 was eligible to participate in Study 2. For Study 2, 14 children were approached but declined to participate either before or immediately following assent. An additional 15 children began the study but did not complete the session because they either withdrew assent, had to meet with their attorney, or attend a court proceeding.

Materials

The majority of study materials, including the five stimulus videos and child photographs were identical to those used in Study 1. However, we changed the interview questions asked in Study 2. The free recall questions were identical, but we modified the wh- questions. First, we removed the wh- summary questions, so that we could increase the number of color/number and wh- detail questions without risking fatigue. Second, to correct for potential confounding effects of question topic (discussed earlier), we constructed a color, number and wh- detail version for each question. To accomplish this, we created 12 questions (six true details, six false details) for each video. Then, for each question we developed three versions: color, number, and wh- detail. For instance, for the topic "stepping stool" the versions were "What color was the stepping stool?" "How many steps were on the stepping stool?", or "What did the sister stand on?" We asked each child one version of each question. We determined the administration of questions by a Latin square, so that the resulting interview per child included 12 questions per video: four color, four number, and four wh- detail questions (with equal numbers of where and what questions) that we split equally across true detail and false detail questions (i.e., two true detail color, two true detail number, two true detail wh- detail, two false detail color, two false detail number, and two false detail wh- detail questions). The order of video presentation and question topic were the same for all children. A full list of questions is presented in the [online supplemental materials](#).

Procedure

The procedure for Study 2 largely mirrored that of Study 1, with the exception of the instruction conditions. In Study 2, we randomly assigned children (stratified by age and gender) to one of four instructions conditions: a) control ($n = 47$), (b) IDK ($n = 48$), (c) correct me ($n = 47$), or (d) IDK + correct me ($n = 47$)

condition. In the control condition the children did not receive any instruction or reminder. For the IDK condition the instruction was identical to the IDK practice condition in Study 1. There was no verbalization in the IDK instruction because this added element did not provide any significant benefit in Study 1. We administered the IDK instruction in full for the first video, with reminders for each subsequent video.

For the correct me condition, after the first video's free recall, the research assistant gave the child a full correct me instruction:

Sometimes I make mistakes or say the wrong thing. When I do, you can tell me that I am wrong. If I say, "You are thirty years old," what do you say? (waits for response) Ok, so how old are you?

For each subsequent video, before each set of questions, we reminded children: "Ok [child's name], remember, sometimes I make mistakes or say the wrong thing. When I do, you can tell me that I am wrong."

In the IDK + correct me condition, we gave children the full IDK and correct me instructions for the first video and both reminders for each of the additional videos. We always administered the IDK instruction prior to the correct me instruction (we applied the same order for the reminders).

Coding

We coded children's overall accuracy into codes of inaccurate, IDK, and accurate (see [Table 1](#) for full code descriptions and examples). As in Study 1, we created proportion scores for children's accuracy type across all five videos by true details/false details and wh- type questions. Two coders independently coded 100% of the transcripts and obtained interrater reliability of $\kappa = .93$. The two coders discussed and resolved all discrepancies.

Results

Preliminary Analyses

We conducted preliminary analyses to determine whether children's responses differed by gender or ethnicity. We conducted a series of 2 (ethnicity: Latinx, Non-Latinx) $\times$ 2 (sex: male, female) univariate ANOVAs on the proportion of children's inaccurate, IDK, and accurate responses for questions about both false and true details. Results revealed no significant differences in response pattern by ethnicity or sex (see [online supplemental materials](#) for full analyses); these variables are not considered further. Similar to Study 1, in response to a reviewer's request, we conducted a series of independent sample t -tests to compare children's responses following a IDK instruction (i.e., IDK instruction and IDK + correct me instruction conditions collapsed) with their responses when no IDK instruction was given (i.e., control condition and correct me instruction conditions collapsed). There were no significant effects of instruction, beyond what is reported below (see full results from statistical analyses in [online supplemental materials](#)).

Analysis Plan

Similarly to Study 1 we examined the effects of instruction condition, wh- type, and age on the proportion of children's inaccurate, IDK, and accurate responses across true detail and false detail

wh- questions using mixed model ANOVAs on children's responses with instruction condition (control, IDK, correct me, IDK + correct me) and age group (4- to 5-year-olds, $n = 93$; 6- to 7-year-olds, $n = 96$) entered as between-subjects variables and wh-type (color/number, wh- detail) entered as a within-subjects variable. We entered the proportion of children's inaccurate, IDK, and accurate responses as dependent variables each in separate analyses. For all analyses, we first present findings for questions about false details followed by findings for questions about true details. We report only significant statistical tests in the text; for a full reporting, see [online supplemental materials](#). We interpreted all significant instruction condition differences using post hoc Tukey's honestly significant difference (HSD) tests (to account for the multiple comparisons resulting from the four instruction conditions). To interpret the significant interactions, we examined the simple effects by conducting t -tests or univariate ANOVAs.

Questions About False Details

First, we examined children's responses to questions about false details (see [Table 4](#)). For children's inaccurate responses there was no significant main effect of instruction condition. However, there were main effects of both wh- type, $F(1, 180) = 8.75, p = .004, \eta_p^2 = .05, 95\% \text{ CI } [.01, .12]$, and age, $F(1, 180) = 21.13, p < .001, \eta_p^2 = .11, 95\% \text{ CI } [.04, .19]$, as well as a significant interaction of Instruction Condition $\times$ Wh- Type, $F(3, 180) = 3.61, p = .01, \eta_p^2 = .06, 95\% \text{ CI } [.002, .12]$. The main effect of age revealed that older children ($M = .41, SD = .34$) gave fewer inaccurate responses than younger children ($M = .64, SD = .32$).

With respect to the Instruction Condition $\times$ Wh- Type interaction, paired samples t -tests revealed that a significant effect of wh-type only occurred in the correct me, $t(46) = 3.21, p = .002, d = .47, 95\% \text{ CI } [.16, .77]$, and control, $t(46) = 2.50, p = .02, d = .36, 95\% \text{ CI } [.07, .66]$ instruction conditions. In these instruction conditions, children gave more inaccurate responses to color/number questions than to wh- detail questions. There were no significant differences between children's color/number and wh- detail responses in the IDK, $t(46) = .66, p = .51, d = .10, 95\% \text{ CI } [-.19, .38]$ or the IDK + correct me, $t(46) = .62, p = .54, d = .09, 95\% \text{ CI } [-.20, .38]$ instruction conditions (see [Table 4](#)).

Next, we assessed children's IDK responses to questions about false details. There were significant main effects of instruction condition, $F(3, 180) = 9.35, p < .001, \eta_p^2 = .14, 95\% \text{ CI } [.05, .22]$ and wh- type, $F(1, 180) = 6.25, p = .01, \eta_p^2 = .03, 95\% \text{ CI } [.00, .10]$, but no significant effect of age. For the main effect of instruction condition, post hoc Tukey's HSD tests revealed that an IDK instruction resulted in significantly more IDK responses than no IDK instruction. Specifically, children in the IDK ($p < .001, d = .89, 95\% \text{ CI } [.51, 1.36]$) and IDK + correct me ($p = .003, d = .71, 95\% \text{ CI } [.33, 1.17]$) conditions provided significantly more IDK responses than children in the correct me condition. Children in the IDK ($p = .001, d = .79, 95\% \text{ CI } [.35, 1.19]$) and IDK + correct me ($p = .01, d = .61, 95\% \text{ CI } [.18, 1.01]$) conditions also each provided significantly more IDK responses than children in the control condition (see [Table 4](#)). For wh- type, children said IDK less often in response to color/number questions ($M = .24, SD = .30$) than wh- detail questions ($M = .27, SD = .31$).

We then examined children's accurate responses to false detail questions, which required children to make corrective responses.

Significant main effects of instruction condition, $F(3, 180) = 5.18, p = .002, \eta_p^2 = .08, 95\% \text{ CI } [.01, .15]$, and age, $F(1, 180) = 18.50, p < .001, \eta_p^2 = .09, 95\% \text{ CI } [.03, .18]$, emerged, as well as a significant interaction of Instruction Condition $\times$ Wh- Type, $F(3, 180) = 3.82, p = .01, \eta_p^2 = .06, 95\% \text{ CI } [.004, .13]$. The main effect of age reflected older children's ($M = .31, SD = .32$) tendency to provide significantly more accurate responses than younger children ($M = .14, SD = .23$). With respect to the significant Instruction Condition $\times$ Wh- Type interaction, univariate ANOVAs revealed significant effects for both color/number questions, $F(3, 188) = 3.20, p = .03, \eta_p^2 = .05, 95\% \text{ CI } [.00, .11]$ and wh- detail questions, $F(3, 188) = 6.31, p < .001, \eta_p^2 = .09, 95\% \text{ CI } [.02, .17]$. Post hoc Tukey's HSD tests revealed that for both color/number and wh- detail questions, children were significantly less likely to provide accurate responses in the IDK (color/number: $p = .04, d = .54, 95\% \text{ CI } [.12, .94]$; wh-detail: $p = .002, d = .66, 95\% \text{ CI } [.23, 1.07]$) or IDK + correct me conditions (color/number: $p = .02, d = .57, 95\% \text{ CI } [.11, .94]$, wh-detail: $p = .001, d = .77, 95\% \text{ CI } [.29, 1.12]$) than in the correct me condition (see [Table 4](#)). The interaction reflects a smaller instruction condition difference for the color/number questions than for the wh-detail questions.

Questions About True Details

For questions about true details (see [Table 5](#)), we first examined children's inaccurate responses, which include incorrect guesses as well as inaccurate corrective responses. As in Study 1, incorrect guesses (83%) accounted for the majority of children's inaccurate responses, compared to false corrective responses (17%). There were significant main effects of instruction condition, $F(3, 180) = 4.60, p = .004, \eta_p^2 = .07, 95\% \text{ CI } [.01, .14]$; wh- type, $F(1, 180) = 8.23, p = .005, \eta_p^2 = .04, 95\% \text{ CI } [.00, .11]$; and age, $F(1, 180) = 8.12, p = .01, \eta_p^2 = .04, 95\% \text{ CI } [.00, .11]$; and a Significant Wh-Type $\times$ Age interaction, $F(1, 180) = 4.34, p = .04, \eta_p^2 = .02, 95\% \text{ CI } [.00, .08]$. For the main effect of instruction condition, post hoc Tukey's HSD tests revealed that the IDK instruction decreased errors compared to the correct me ($p = .02, d = .67, 95\% \text{ CI } [.22, 1.06]$) and control conditions ($p = .02, d = .67, 95\% \text{ CI } [.24, 1.07]$). The IDK + correct me condition did not significantly differ from any other instruction condition (see [Table 5](#)). With respect to the interaction between wh- type and age, paired sample t -tests revealed a significant effect of wh- type only for older children, $t(95) = 3.20, p = .002, d = .33, 95\% \text{ CI } [.12, .53]$, whereby older children provided more inaccurate responses to color/number questions ($M = .55, SD = .25$) than they did to wh- detail questions ($M = .49, SD = .25$). The wh- type effect for younger children was not significant, $t(91) = .62, p = .54, d = .06, 95\% \text{ CI } [-.14, .27]$.

Next, we investigated children's IDK responses to questions about true details. There was a significant main effect of instruction condition, $F(3, 180) = 7.22, p < .001, \eta_p^2 = .11, 95\% \text{ CI } [.03, .19]$. An IDK instruction resulted in a significantly higher rate of IDK responses than no IDK instruction. Specifically, post hoc Tukey's HSD tests revealed that children in the IDK ($p = .004, d = .69, 95\% \text{ CI } [.26, 1.09]$) and IDK + correct me ($p = .04, d = .54, 95\% \text{ CI } [.18, 1.00]$) instruction conditions provided significantly more IDK responses than those in the correct me condition. Furthermore, children in the IDK ($p = .001, d = .81, 95\% \text{ CI } [.32, 1.16]$) and IDK + correct me ($p = .01, d = .65, 95\% \text{ CI } [.25, 1.09]$) conditions provided more IDK responses than those in the control

Table 4
Study 2 Means and Standard Deviations for Children's Responses to Questions About False Details by Age, Condition, and Wh-Type

	False details											
	Inaccurate				IDK				Accurate			
	Control (<i>n</i> = 47)	IDK (<i>n</i> = 48)	Correct me (<i>n</i> = 47)	IDK + Correct me (<i>n</i> = 47)	Control (<i>n</i> = 47)	IDK (<i>n</i> = 48)	Correct me (<i>n</i> = 47)	IDK + Correct me (<i>n</i> = 47)	Control (<i>n</i> = 47)	IDK (<i>n</i> = 48)	Correct me (<i>n</i> = 47)	IDK + Correct me (<i>n</i> = 47)
Color/number												
Young (<i>n</i> = 93)	.68 (.37)	.57 (.37)	.68 (.33)	.66 (.33)	.15 (.28)	.35 (.35)	.12 (.22)	.28 (.33)	.09 (.12)	.04 (.07)	.10 (.15)	.03 (.08)
Old (<i>n</i> = 96)	.56 (.34)	.31 (.29)	.43 (.39)	.37 (.35)	.16 (.20)	.41 (.31)	.13 (.15)	.36 (.34)	.14 (.15)	.13 (.14)	.22 (.18)	.13 (.14)
Total	.62 (.36)	.44 (.36)	.55 (.38)	.51 (.37)	.15 (.24)	.38 (.33)	.12 (.19)	.32 (.33)	.11 (.14)	.09 (.12)	.16 (.17)	.08 (.12)
Wh- detail												
Young (<i>n</i> = 93)	.60 (.37)	.62 (.34)	.59 (.36)	.65 (.35)	.16 (.28)	.31 (.32)	.17 (.24)	.30 (.34)	.24 (.29)	.07 (.14)	.24 (.34)	.05 (.13)
Old (<i>n</i> = 96)	.52 (.40)	.30 (.29)	.38 (.37)	.36 (.34)	.19 (.21)	.47 (.36)	.14 (.19)	.42 (.34)	.30 (.32)	.23 (.31)	.48 (.39)	.22 (.26)
Total	.56 (.36)	.46 (.35)	.48 (.38)	.50 (.37)	.17 (.24)	.39 (.35)	.16 (.21)	.36 (.34)	.27 (.30)	.15 (.25)	.36 (.38)	.14 (.22)
Wh- combined												
Young (<i>n</i> = 93)	.65 (.36)	.59 (.35)	.65 (.33)	.66 (.33)	.15 (.27)	.33 (.33)	.13 (.21)	.29 (.33)	.20 (.25)	.08 (.13)	.22 (.30)	.06 (.15)
Old (<i>n</i> = 96)	.54 (.33)	.31 (.28)	.41 (.37)	.371 (.34)	.17 (.19)	.43 (.32)	.13 (.14)	.38 (.33)	.28 (.31)	.25 (.28)	.46 (.37)	.25 (.26)
Total	.59 (.35)	.45 (.35)	.53 (.37)	.51 (.36)	.16 (.23)	.38 (.33)	.13 (.18)	.33 (.33)	.24 (.28)	.17 (.24)	.34 (.35)	.16 (.23)

Note. IDK = I don't know.

Table 5
Means and Standard Deviations for Children's Responses to Questions About True Details by Age, Condition, and Wh-Type

[illegible]

Note. IDK = I don't know.

condition (see Table 5). There were no significant effects of wh- type or age.

Finally, we examined children's accurate responses to questions about true details. There were significant main effects of instruction condition, $F(3, 180) = 5.19, p = .002, \eta_p^2 = .08, 95\% \text{ CI } [.01, .15]$; wh- type, $F(3, 180) = 5.65, p = .02, \eta_p^2 = .03, 95\% \text{ CI } [.00, .09]$; and age, $F(1, 180) = 35.37, p < .001, \eta_p^2 = .16, 95\% \text{ CI } [.08, .26]$; and a significant interaction of Wh- Type $\times$ Age, $F(1, 180) = 9.19, p = .003, \eta_p^2 = .05, 95\% \text{ CI } [.01, .12]$. For the main effect of instruction condition, post hoc Tukey's HSD tests revealed significant instruction condition differences between both the IDK ($p = .01, d = .54, 95\% \text{ CI } [.17, 1.00]$) and IDK + correct me ($p = .01, d = .45, 95\% \text{ CI } [.06, .88]$) instructions compared to the control condition. Children provided fewer accurate responses in both the IDK ($M = .16, SD = .08$) and the IDK + correct me ($M = .17, SD = .08$) conditions than in the control ($M = .21, SD = .09$) or correct me ($M = .20, SD = .09$) conditions (see Table 5).

For the Wh- Type $\times$ Age interaction, paired samples t -tests revealed that there was a significant effect of wh- type only for older children, $t(95) = 3.51, p = .001, d = .36, 95\% \text{ CI } [.15, .56]$. Older children provided significantly fewer accurate responses to color/number questions ($M = .19, SD = .12$) than to wh- detail questions ($M = .26, SD = .14$). There was no significant effect of wh- type for younger children, $t(91) = .52, p = .61, d = .05, 95\% \text{ CI } [-.15, .26]$.

General Discussion

We examined young maltreated children's tendency to give IDK responses to different types of wh- questions. In both studies there was evidence that children guessed more often in response to color/number questions than in response to wh- detail questions, though this manifested itself in different ways. The effects were clearest with respect to questions about false details, because attempts to answer such questions necessarily constituted guessing. In Study 1, we found that color/number questions about false details led to increased error. Children were less likely to give true corrective responses in which they denied that the details appeared in the videos. In Study 2, we matched color/number and wh- detail questions for content and found that in response to questions about false details, children were less likely to respond IDK to color/number questions than wh- detail questions. In the absence of IDK instructions, children provided more inaccurate responses to color/number questions than to wh- detail questions, likely driven by fewer IDK answers.

We also tested whether various types of interview instructions affect children's tendency to give IDK answers and improve accuracy. The studies found few positive effects. In both studies, IDK instructions, either alone in Study 1, or combined with correct the interview instructions in Study 2, reduced children's tendency to provide corrective responses to questions about false details. That is, the instructions led children to refrain from correcting the interviewer's inaccurate presuppositions. With respect to questions about true details, we found mixed effects. Study 1 found a non-significant decrease in errors with no decrease in accurate responses. With a larger sample, Study 2 found a significant decrease in errors but with a corresponding decrease in accurate responses. The effect sizes for the positive and negative effects were comparable, suggesting that children's increased tendency to

give IDK responses reflected a general tendency to respond IDK more often when instructed to do so, which thus reduced the number of inaccurate responses but also the number of accurate responses.

We predicted that instructions would have a larger effect on wh-detail questions than color/number questions. In both studies, this occurred, but not in the way we had anticipated. Rather than improve performance, instructions were more likely to reduce true corrective responses to wh- detail questions than to color/number questions. With respect to specific instruction manipulations, Study 1 found no support for asking children to verbalize their intent to provide IDK answers, and Study 2 found no benefit to combining IDK instructions with instructions to correct the interview.

Guessing and Response Availability

Even very young children understand that yes-no questions call for a yes or a no (Horgan, 1978) and that forced-choice questions ask one to choose one of the proffered options (Sumner et al., 2019), and this understanding likely contributes to high rates of guessing in response to these kinds of questions (Waterman et al., 2000). By the same token, the results suggest that young children's recognition that questions about color and number call for a color name or a number, coupled with their ability to generate color names and numbers (Wagner et al., 2013), leads to high rates of guessing in response to color/number questions.

At the same time, the rates of error in response to questions about false details were still very high, which suggests that children were guessing in response to the other kinds of wh- questions at high rates as well. Most of the wh- detail questions in Study 1 were about objects, and to control for content across number/color and wh- detail question, all of the wh- detail questions in Study 2 asked about the identity ("what") and location ("where") of various objects. We suspect that questions asking for descriptions may elicit more guessing than questions asking about actions. There is evidence that children provide more productive responses when asked wh- questions about actions than when asked for descriptions (Ahern et al., 2018; Andrews et al., 2016), and that children remember actions better than descriptions (Peterson et al., 1999). Even when questions about actions appear to presuppose information (e.g., "What did he do?" and "What did you do?" presuppose that the actors did something), the availability of a "nothing" response may lower rates of guessing and thus lower the risks.

In practical terms, actions are often the most important type of detail in investigations of physical and sexual abuse against young children, because of the frequency with which they involve familiar suspects and familiar locations. Hence, interviewers may be able to avoid extensive questioning about descriptions. When color and number information is essential, it may be possible for interviewers to elicit color and number information through exhaustive use of broader questions, including invitations that call for free recall (e.g., "Tell me everything that happened the last time [abuse occurred]") or cued recall (e.g., "You said [child generated detail]; tell me more about that"). Furthermore, other types of recall questions may also prove more productive and less susceptible to guessing, such as "Tell me how everything looked" (Poole & Lindsay, 1995, 2001) or "Tell me everything you saw."

Future work can identify other topics with high response availability and thus greater risks of guessing. For example, researchers studying children's language development have identified some types of temporal terms as a lexical domain; young children will readily, but inaccurately, answer questions about duration (i.e., "How long does it take . . .?"; [Shatz et al., 2010](#)). It also seems likely that there are individual differences in response availability; children with greater knowledge about a topic have the wherewithal, and may also have the inclination, to guess more often.

Interview Instructions Encouraging IDK Responses

The results highlight the dangers in encouraging IDK responses. If children are emboldened to answer "IDK" but are not able (or inclined) to monitor their certainty closely, they may answer indiscriminately, and both accurate and inaccurate responses will be reduced. A reduction in both accurate responses and errors has been noted in some research ([Gee et al., 1999](#) [Study 1]; [Nesbitt & Markham, 1999](#); [Saywitz & Moan-Hardie, 1994](#) [Study 1]). A novel problem, identified here, is that when children recognize that the interviewer's question presupposes false information, IDK instructions will encourage them to substitute their corrective responses with IDK responses.

To counteract the negative effects of IDK instructions, researchers have encouraged children to provide informative responses when they know the answers ([Gee et al., 1999](#) [Study 2]; [Saywitz & Moan-Hardie, 1994](#) [Study 2]). Here, we also encouraged both types of answers and still encountered some negative effects. A number of factors are likely responsible for the different outcomes.

An obvious difference is the age and status of the children studied. In [Saywitz and Moan-Hardie \(1994](#) [Study 2]), the participants were second graders, and averaged 7½ years of age, and in [Gee et al. \(1999](#) [Study 2]) the participants were 9 to 11 years of age. Our samples were 4 to 7 years of age, with an average age of 5½ years of age. Often, researchers have found IDK instructions have limited efficacy with younger children ([Ellis et al., 2003](#); [Geddie et al., 2001](#); [Peterson & Grant, 2001](#)).

Neither [Saywitz and Moan-Hardie \(1994\)](#) nor [Gee et al. \(1999\)](#) recruited maltreated children. We tested young maltreated children, who are an obviously relevant population to study, but one missing from prior experimental work on interview instructions. Observational work has examined maltreated children's IDK responding to instructions ([Earhart et al., 2014](#)), but that research is unable to assess accuracy, and examined instructions without practice or feedback. Prior work with the maltreated population from which our children were drawn has identified deficits in executive functioning ([Williams et al., 2020](#)) and receptive vocabulary ([Lyon & Saywitz, 1999](#)), which likely limit their metacognitive ability to recognize when they know the answers to questions, their working memory capacity to keep the instruction in mind, and their executive capacity to apply the instruction in formulating their response ([Brubacher et al., 2015](#)). Future researchers can enroll maltreated and comparable nonmaltreated children with a larger age range and directly test capacities hypothesized to underlie performance.

A second difference between the current studies and studies avoiding negative effects is that the prior work included substantially more practice. As [Gee et al. \(1999\)](#) noted, [Saywitz and](#)

[Moan-Hardie \(1994\)](#) intervention was quite extensive, a "multi-component package consisting of an illustrative story, discussion, instructions about response strategies, a practice session with questions about a video, visual reminders about the acceptable responses, feedback and a review session" (p. 112). Although the researchers did not note the total amount of time, the review session alone lasted 15 min ([Saywitz & Moan-Hardie, 1994](#)). [Saywitz and Moan-Hardie \(1994\)](#) Study 2, which reduced inaccuracies without decreasing accuracies, had "more opportunity for and reinforcement for 'telling the answer when you know the answer'" (p. 419). [Gee et al. \(1999\)](#) described their own intervention as "brief," though they asked children multiple practice questions, and to minimize negative effects in Study 2, "the training package was extended to reinforce and encourage correct answers" ([Gee et al., 1999](#), p. 121). Specifically, they added four answerable questions to the training package.

In the current studies, we asked children one unanswerable question ("What is my dog's name?") and one answerable question ("Do you have a dog?"), and then periodically reminded children of the utility of both IDK and substantive answers. Moreover, in the correct me conditions, we gave children practice correcting an incorrect assertion by the interviewer, and we periodically reminded children of the value of correcting the interviewer.

It might be the case that had we given children more extensive practice their performance would have improved. However, from a practical perspective, our intervention was more akin to the kind of practice that children will receive in the field. We modeled our intervention after the Ten-Step interview ([Lyon, 2014](#)) and the forensic interview guidelines of the [American Professional Society on the Abuse of Children Taskforce \(2012\)](#), which both recommend that interviewers practice with one unanswerable and one answerable question as well as the correct me instruction. We are not aware of any interview protocol that recommends practice as extensive as that used by [Gee et al. \(1999\)](#), let alone [Saywitz and Moan-Hardie \(1994\)](#). Other protocols limit practice to encouraging IDK responses through one question ([Lamb et al., 2018](#) [NICHD protocol]), or additional questions if the child attempts to guess ([Dickinson et al., 2015](#) [Michigan protocol]), without countering IDK instructions with instructions to answer when children know the answer. Some protocols fail to recommend any practice at all (see protocols discussed in [Dickinson et al., 2015](#)). In the future, researchers can explore a happy medium between instructions with no practice, which are most often ineffective, and instructions with extensive practice, which may be impractical.

An important limitation is that we assessed children's responses to questions about videotaped interactions immediately after viewing the videos, which of course is quite different than questioning children about experienced events sometime after those events occur, and thus unlike abuse investigations. How this may have affected the efficacy of instructions is unclear, in part because the short retention span would have benefited children's memory but the videotaped presentation would have impaired it (compared to a live event; [Roebbers et al., 2004](#)), and in part because the relation between memory strength and the effect of IDK instructions is unclear ([Brubacher et al., 2015](#)). It is also significant that children had little reason to be reluctant about disclosing the contents of the videos. A well-understood phenomenon among survey researchers is that reluctant respondents often resort to IDK responding ([Kaminska et al., 2010](#)). Although experimental work assumes

children are honest when they claim that they IDK, researchers conducting observational research on forensic interviewing have identified IDK responses as a type of reluctance (Blasbalg et al., 2019). A future direction for experimental work is to examine whether IDK instructions might be associated with children's non-responsiveness about transgressions and other sensitive topics.

Implications for Practice

Researchers routinely warn forensic interviewers that recognition (yes–no and forced-choice) questions are risky, in part because children only very infrequently give IDK answers to such questions. These results show that some wh- questions, despite tapping recall memory, share similar risks, and that children are particularly likely to guess when asked questions about color and number. Interviewers should be mindful of these risks and attempt to elicit color and number information through more open-ended questions. When they feel compelled to ask questions about color and number, they should treat children's responses with caution, and follow-up brief responses with requests for elaboration to assess the likelihood that the child's response was based on memory rather than a guess.

The results counsel caution with respect to the use of IDK instructions because they may reduce inaccurate responses at the cost of reducing accurate responses, including true corrective responses in which children correct interviewer misconceptions. However, we would hesitate to make strong recommendations regarding interviewers' use of instructions, given research supporting their use with older and nonmaltreated children. Independently of their effect on children's responding, instructions enjoy a number of benefits. When used at the beginning of an interview, they allow for a structured introduction to the interview during which the interviewer can do most of the talking, allowing both the interviewer and the child to settle in before the interviewer turns to questions designed to elicit longer narrative responses from the child. Children's answers to instruction practice questions may have diagnostic value; their answers to "do not understand" instructions in forensic interviews predict how they exhibit incomprehension during substantive questioning (Henderson & Lyon, 2021). When children demonstrate undue eagerness to guess in response to instruction practice questions, this may provide a useful warning to the interviewer (and others who observe the interview) regarding the interviewer's use of questions with high response availability. Instructions thus serve many purposes, and researchers should consider their rapport-building and diagnostic qualities in addition to further exploring their effects on children's acknowledgment of ignorance.

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>
- American Professional Society on the Abuse of Children Taskforce. (2012). *Practice guidelines: Forensic interviewing in cases of suspected child abuse*. <https://www.apsac.org/guidelines>
- 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>
- Andrews, S. J., Ahern, E. C., & Lamb, M. E. (2017). Children's uncertain responses when testifying about alleged sexual abuse in Scottish courts. *Behavioral Sciences & the Law*, 35(3), 204–224. <https://doi.org/10.1002/bsl.2286>
- Blasbalg, U., Hershkowitz, I., Lamb, M. E., Karni-Visel, Y., & Ahern, E. C. (2019). Is interviewer support associated with the reduced reluctance and enhanced informativeness of alleged child abuse victims? *Law and Human Behavior*, 43(2), 156–165. <https://doi.org/10.1037/lhb0000303>
- Bloom, P., & Wynn, K. (1997). Linguistic cues in the acquisition of number words. *Journal of Child Language*, 24(3), 511–533. <https://doi.org/10.1017/S0305000997003188>
- Brown, D. A., Lewis, C. N., Lamb, M. E., Gwynne, J., Kitto, O., & Stairmand, M. (2019). Developmental differences in children's learning and use of forensic ground rules during an interview about an experienced event. *Developmental Psychology*, 55(8), 1626–1639. <https://doi.org/10.1037/dev0000756>
- Brubacher, S. P., Malloy, L. C., Lamb, M. E., & Roberts, K. P. (2013). How do interviewers and children discuss individual occurrences of alleged repeated abuse in forensic interviews? *Applied Cognitive Psychology*, 27(4), 443–450. <https://doi.org/10.1002/acp.2920>
- Brubacher, S. P., Poole, D. A., & Dickinson, J. J. (2015). The use of ground rules in investigative interviews with children: A synthesis and call for research. *Developmental Review*, 36, 15–33. <https://doi.org/10.1016/j.dr.2015.01.001>
- Danby, M. C., Brubacher, S. P., Sharman, S. J., & Powell, M. B. (2015). The effects of practice on children's ability to apply ground rules in a narrative interview. *Behavioral Sciences & the Law*, 33(4), 446–458. <https://doi.org/10.1002/bsl.2194>
- Dickinson, J. J., Brubacher, S. P., & Poole, D. A. (2015). Children's performance on ground rules questions: Implications for forensic interviewing. *Law and Human Behavior*, 39(1), 87–97. <https://doi.org/10.1037/lhb0000119>
- Durkin, K., Shire, B., Riem, R., Crowther, R. D., & Rutter, D. R. (1986). The social and linguistic context of early number use. *British Journal of Developmental Psychology*, 4(3), 269–288. <https://doi.org/10.1111/j.2044-835X.1986.tb01018.x>
- Earhart, B., La Rooy, D. J., Brubacher, S. P., & Lamb, M. E. (2014). An examination of "don't know" responses in forensic interviews with children. *Behavioral Sciences & the Law*, 32(6), 746–761. <https://doi.org/10.1002/bsl.2141>
- Ellis, L. M., Powell, M. B., Thomson, D. M., & Jones, C. (2003). Do simple "ground rules" reduce preschoolers' suggestibility about experienced and nonexperienced events? *Psychiatry, Psychology and Law*, 10(2), 334–345. <https://doi.org/10.1375/pplt.2003.10.2.334>
- Endres, J., Poggenpohl, C., & Erben, C. (1999). Repetitions, warnings and video: Cognitive and motivational components in preschool children's suggestibility. *Legal and Criminological Psychology*, 4(1), 129–146. <https://doi.org/10.1348/135532599167725>
- Evans, A. D., O'Connor, A. M., & Lee, K. (2018). Verbalizing a commitment reduces cheating in young children. *Social Development*, 27(1), 87–94. <https://doi.org/10.1111/sode.12248>
- Fessinger, M. B., McWilliams, K., Bakth, F. N., & Lyon, T. D. (2021). Setting the ground rules: Use and practice of ground rules in child forensic interviews. *Child Maltreatment*, 26(1), 126–132. <https://doi.org/10.1177/1077559520910783>
- Geddie, L. F., Beer, J., Bartosik, S., & Wuensch, K. L. (2001). The relationship between interview characteristics and accuracy of recall in young children: Do individual differences matter? *Child Maltreatment*, 6(1), 59–68. <https://doi.org/10.1177/1077559501006001006>
- Gee, S., Gregory, M., & Pipe, M.-E. (1999). What colour is your pet dinosaur? The impact of pre-interview training and question type on children's answers. *Legal and Criminological Psychology*, 4(1), 111–128. <https://doi.org/10.1348/135532599167716>
- Goodman, G. S., Quas, J. A., Bulkley, J., & Shapiro, C. (1999). Innovations for child witnesses: A national survey. *Psychology, Public Policy, and Law*, 5(2), 255–271. <https://doi.org/10.1037/1076-8971.5.2.255>

- 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>
- Heyman, G. D., Fu, G., Lin, J., Qian, M. K., & Lee, K. (2015). Eliciting promises from children reduces cheating. *Journal of Experimental Child Psychology*, 139, 242–248. <https://doi.org/10.1016/j.jecp.2015.04.013>
- Horgan, D. (1978). How to answer questions when you've got nothing to say. *Journal of Child Language*, 5(1), 159–165. <https://doi.org/10.1017/S0305000900002014>
- Kaminska, O., McCutcheon, A. L., & Billiet, J. (2010). Satisficing among reluctant respondents in a cross-national context. *Public Opinion Quarterly*, 74(5), 956–984. <https://doi.org/10.1093/poq/nfq062>
- Kloo, D., Rohwer, M., & Perner, J. (2017). Direct and indirect admission of ignorance by children. *Journal of Experimental Child Psychology*, 159, 279–295. <https://doi.org/10.1016/j.jecp.2017.02.014>
- Lamb, M. E., Brown, D. A., Hershkovitz, I., Orbach, Y., & Esplin, P. W. (2018). *Tell me what happened: Questioning children about abuse* (2nd ed.). Wiley-Blackwell.
- Lamb, M. E., & Brown, D. A. (2006). Conversational apprentices: Helping children become competent informants about their own experiences. *British Journal of Developmental Psychology*, 24(1), 215–234. <https://doi.org/10.1348/026151005X57657>
- Lyon, T. D. (2014). Interviewing children. *Annual Review of Law and Social Science*, 10(1), 73–89. <https://doi.org/10.1146/annurev-lawsocsci-110413-030913>
- Lyon, T. D., & Saywitz, K. J. (1999). Young maltreated children's competence to take the oath. *Applied Developmental Science*, 3(1), 16–27. https://doi.org/10.1207/s1532480xads0301_3
- Marquis, K. H., Marshall, J., & Oskamp, S. (1972). Testimonial validity as a function of question form, atmosphere, and item difficulty. *Journal of Applied Social Psychology*, 2(2), 167–186. <https://doi.org/10.1111/j.1559-1816.1972.tb01270.x>
- McWilliams, K., Narr, R., Goodman, G. S., Ruiz, S., & Mendoza, M. (2013). Children's memory for their mother's murder: Accuracy, suggestibility, and resistance to suggestion. *Memory*, 21(5), 591–598. <https://doi.org/10.1080/09658211.2013.763983>
- Ministry of Justice. (2011). *Achieving best evidence in criminal proceedings*.
- Mulder, M. R., & Vrij, A. (1994). Explaining conversation rules to children: An intervention study to facilitate children's accurate responses. *Child Abuse & Neglect*, 20(7), 623–631. [https://doi.org/10.1016/0145-2134\(96\)00050-6](https://doi.org/10.1016/0145-2134(96)00050-6)
- Nesbitt, M., & Markham, R. (1999). Improving young children's accuracy of recall for an eyewitness event. *Journal of Applied Developmental Psychology*, 20(3), 449–459. [https://doi.org/10.1016/S0193-3973\(99\)00027-1](https://doi.org/10.1016/S0193-3973(99)00027-1)
- Peterson, C., Dowden, C., & Tobin, J. (1999). Interviewing preschoolers: Comparisons of yes-no and wh- questions. *Law and Human Behavior*, 23(5), 539–555. <https://doi.org/10.1023/A:1022396112719>
- Peterson, C., & Grant, M. (2001). Forced-choice: Are forensic interviewers asking the right questions? *Canadian Journal of Behavioural Science/Revue Canadienne Des Sciences du Comportement*, 33(2), 118–127. <https://doi.org/10.1037/h0087134>
- Poole, D. A., & Lindsay, D. S. (1995). Interviewing preschoolers: Effects of nonsuggestive techniques, parental coaching, and leading questions on reports of nonexperienced events. *Journal of Experimental Child Psychology*, 60(1), 129–154. <https://doi.org/10.1006/jecp.1995.1035>
- Poole, D. A., & Lindsay, D. S. (2001). Children's eyewitness reports after exposure to misinformation from parents. *Journal of Experimental Psychology: Applied*, 7(1), 27–50. <https://doi.org/10.1037/1076-898X.7.1.27>
- Roebers, C. M., Gelhaar, T., & Schneider, W. (2004). It's magic!" The effects of presentation modality on children's event memory, suggestibility, and confidence judgments. *Journal of Experimental Child Psychology*, 87(4), 320–335. <https://doi.org/10.1016/j.jecp.2004.01.004>
- Rohwer, M., Kloo, D., & Perner, J. (2012). Escape from metaignorance: How children develop an understanding of their own lack of knowledge. *Child Development*, 83(6), 1869–1883. <https://doi.org/10.1111/j.1467-8624.2012.01830.x>
- Sandhofer, C. M., & Smith, L. B. (1999). Learning color words involves learning a system of mappings. *Developmental Psychology*, 35(3), 668–679. <https://doi.org/10.1037/0012-1649.35.3.668>
- Saywitz, K. J., & Moan-Hardie, S. (1994). Reducing the potential for distortion of childhood memories. *Consciousness and Cognition*, 3(3–4), 408–425. <https://doi.org/10.1006/ccog.1994.1023>
- Scoboria, A., Mazzoni, G., & Kirsch, I. (2008). Don't know" responding to answerable and unanswerable questions during misleading and hypnotic interviews. *Journal of Experimental Psychology: Applied*, 14(3), 255–265. <https://doi.org/10.1037/1076-898X.14.3.255>
- Scoboria, A., & Fisco, S. (2013). Encouraging and clarifying "don't know" responses enhances interview quality. *Journal of Experimental Psychology: Applied*, 19(1), 72–82. <https://doi.org/10.1037/a0032067>
- Shatz, M., Tare, M., Nguyen, S. P., & Young, T. (2010). Acquiring non-object terms: The case for time words. *Journal of Cognition and Development*, 11(1), 16–36. <https://doi.org/10.1080/15248370903453568>
- Sumner, E., DeAngelis, E., Hyatt, M., Goodman, N., & Kidd, C. (2019). Cake or broccoli? Recency biases children's verbal responses. *PLoS ONE*, 14(6), E0217207. <https://doi.org/10.1371/journal.pone.0217207>
- Tare, M., Shatz, M., & Gilbertson, L. (2008). Maternal uses of non-object terms in child-directed speech: Color, number, and time. *First Language*, 28(1), 87–100. <https://doi.org/10.1177/0142723707085316>
- Wagner, K., Dobkins, K., & Barner, D. (2013). Slow mapping: Color word learning as a gradual inductive process. *Cognition*, 127(3), 307–317. <https://doi.org/10.1016/j.cognition.2013.01.01>
- Waterman, A. H., Blades, M., & Spencer, C. P. (2000). Do children try to answer nonsensical questions? *British Journal of Developmental Psychology*, 18(2), 211–225. <https://doi.org/10.1348/026151000165652>
- Waterman, A. H., Blades, M., & Spencer, C. P. (2001). Interviewing children and adults: The effect of question format on the tendency to speculate. *Applied Cognitive Psychology*, 15(5), 521–531. <https://doi.org/10.1002/acp.741>
- Waterman, A. H., Blades, M., & Spencer, C. P. (2004). Indicating when you IDK the answer: The effect of question format and interviewer knowledge on children's "don't know" responses. *British Journal of Developmental Psychology*, 22(3), 335–348. <https://doi.org/10.1348/0261510041552710>
- Williams, S., McWilliams, K., & Lyon, T. (2020). Children's concealment of a minor transgression: The role of age, maltreatment, and executive functioning. *Journal of Experimental Child Psychology*, 191, Article 104664. <https://doi.org/10.1016/j.jecp.2019.104664>
- Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology*, 24(2), 220–251. [https://doi.org/10.1016/0010-0285\(92\)90008-P](https://doi.org/10.1016/0010-0285(92)90008-P)

Received May 11, 2020

Revision received January 29, 2021

Accepted February 6, 2021 ■