

University of Southern California Law

From the Selected Works of Thomas D. Lyon

Summer July 23, 2021

90. The difficulty of teaching adults to recognize referential ambiguity in children's testimony: The influence of explicit instruction and sample questions.

Breanne E. Wylie, *Brock University*

Jennifer Gongola, *University of Southern California Law School*

Thomas D. Lyon, *University of Southern California Law School*

Angela D. Evans, *Brock University*



Available at: <https://works.bepress.com/thomaslyon/182/>

RESEARCH ARTICLE

The difficulty of teaching adults to recognize referential ambiguity in children's testimony: The influence of explicit instruction and sample questions

Breanne E. Wylie¹  | Jennifer Gongola² | Thomas D. Lyon² | Angela D. Evans¹

¹Department of Psychology, Brock University, St. Catharines, Ontario, Canada

²Gould School of Law, University of Southern California, Los Angeles, California, USA

Correspondence

Breanne E. Wylie, Brock University, St. Catharines, ON, Canada.
Email: bw15gs@brocku.ca

Funding information

National Institute of Child Health and Human Development Grant, Grant/Award Number: HD087685

Abstract

Adults often fail to recognize the ambiguity of children's unelaborated responses to "Do you know/remember (DYK/R) if/whether" questions. Two studies examined whether sample questions and/or an explicit instruction would improve adults' ability to recognize referential ambiguity in children's testimony. In Study 1 ($N = 383$), participants rarely recognized referential ambiguity in the sample questions or in children's testimony, and answering sample questions had no influence on their ability to detect ambiguity in children's testimony. Study 2 ($N = 363$) attempted to clarify the meaning of ambiguity for participants with explicit instructions. Results revealed that although an explicit instruction improved performance on sample questions, this also led to an overcorrection, and instructions and sample questions only modestly improved adults' ability to recognize referential ambiguity in children's testimony. Ultimately, the absence of an effective strategy for alerting adults to the problem of referential ambiguity highlights the dangers of DYK/R if/whether questions.

KEYWORDS

child witness, instruction, jury, referential ambiguity, sample question, training

1 | INTRODUCTION

An often overlooked problem in jurors' assessments of children's reports is whether and how jurors recognize referentially ambiguous responses. One form of question found to lead to ambiguous responding is "Do you know/remember (DYK/R) if/whether" questions (e.g., Do you remember if it was blue?), as these questions explicitly ask if/whether the child knows the information, and at the same time implicitly request that information. Past observational and experimental research (Evans et al., 2014, 2017) has found that children often provide unelaborated "yes" or "no" responses to DYK/R if/whether questions, creating a problem of referential ambiguity, as it is unclear whether the child is answering the explicit or implicit part of the question. For example, an unelaborated "no" response to a DYK/R if/whether question could be answering the explicit question (e.g., No, I do not remember) or the implicit question (e.g., No, it wasn't blue). The danger of ambiguous responding is that

jurors will overlook the ambiguity and make an assumption about which question the child is answering.

Past research by Wylie et al. (2019) examining jury-eligible adults' interpretations of ambiguous responding found that adults rarely recognized the ambiguous nature of unelaborated responses to DYK/R if/whether questions. In that study, adults were presented with child testimony that included referentially ambiguous "yes" and "no" responses to DYK/R if/whether questions. Remarkably, adults identified unelaborated "yes" and "no" responses as unclear less than 10% of the time. The results for "no" responses were particularly problematic, because adults' response patterns suggested they believed they knew whether the child was answering the explicit or the implicit question. Research examining children's "no" responses to DYK/R if/whether questions has not found a systematic pattern in whether children are answering the implicit or explicit question (Evans et al., 2017). "Yes" responses are less problematic, because experimental evidence suggests that when children respond "yes" to DYK/R

if/whether questions, they are answering the implicit question (Ahern et al., 2016), and this is consistent with adults' typical interpretation of children's "yes" responses (Wylie et al., 2019).

Adults' inability to recognize that DYK/R if/whether questions may lead to ambiguous "no" responses is reminiscent of research finding that adults have a limited understanding of the importance of question-type in influencing children's testimony. For example, adults failed to recognize the difference between more leading and less leading questions (Quas et al., 2005; Ruva & Bryant, 2004), between recall and recognition questions (Laimon & Poole, 2008), and between chronological and non-chronological question order (Mugno et al., 2016) in influencing children's reports. This lack of understanding may be explained by how adults remember interviews with children. Research suggests that adults rely on gist memory and as such fail to recall the form of questions used to obtain the information in child interviews (Bruck et al., 1999; Warren & Woodall, 1999). Given their inability to recall question types, adult perceivers may also overlook ambiguity. However, it may be possible to improve adults' performance by educating them about problematic question-types. This was the focus of the current studies.

When examining adults' ability to identify referentially ambiguous responses, Wylie et al. (2019) had participants read transcripts of children testifying in court and then asked participants to recall how children had responded to specific questions. Adults displayed difficulty in recalling that children's "no" responses to DYK/R if/whether questions were ambiguous (often recalling them as answering the direct or indirect question). Adults' difficulty in identifying the ambiguity might have been due to misremembering questions, or they might have failed to detect the ambiguity when the questions were first asked. Thus, the current studies examined whether adults could learn to recognize ambiguity in children's unelaborated "no" responses to DYK/R if/whether questions.

Study 1 provided adults with sample questions in which they were asked to evaluate ambiguous "no" responses to DYK/R if/whether questions in isolation, either before or after reading court transcripts that contained embedded ambiguous "no" responses. This design enabled us to determine if adults could recognize ambiguity when tested immediately after reading ambiguous responses (sample questions), and if answering sample questions improved adults' ability to identify responses as ambiguous embedded in children's testimony. For the sample questions, adults were given examples of DYK/R if/whether questions followed by an unelaborated "no" response and were asked to interpret the response by choosing whether the child's response meant that they answered "no" to the explicit question, the implicit question, or whether their response was unclear. Because adults found the sample questions in Study 1 surprisingly difficult, Study 2 added clarifying language and simplified the sample questions to assess whether this improved performance, while adding an unambiguous response to the sample questions to better identify adult recognition of ambiguity. Additionally, Study 2 tested whether providing explicit instructions about the problem of unelaborated "no" responses to DYK/R if/whether questions resulting in referentially ambiguous responses prior to sample questions would improve performance.

2 | STUDY 1

The first study examined whether adults recognized the problem of referential ambiguity in sample questions, and whether this in turn alerted adults to referentially ambiguous responses in children's testimony. All participants received three sample questions (i.e., examples of ambiguous responses with all possible interpretations) either before or after reading modified court transcripts of child witness testimony. Testimony was manipulated to vary on ambiguity such that for target question-answer pairs, attorneys asked either DYK/R if/whether or Direct questions and children provided either "No" or "IDK" responses, creating four between-subjects conditions (DYK-No, DYK-IDK, Direct-No, Direct-IDK). In the DYK-No condition the child's response was referentially ambiguous (i.e., unclear if the child was answering the explicit "No, I don't remember" or implicit "No, it wasn't blue" question), whereas in the other three conditions the child's response unambiguously answered the question (clearly providing a "No" or "IDK" response). Testing the four conditions enabled us to determine whether giving adults sample questions would help them to identify ambiguous responses without also leading them to false alarm to unambiguous responses.

2.1 | Hypotheses

The sample questions were presented in isolation, rather than embedded in testimony, in order to minimize memory demands. We therefore expected that participants would be able to identify the ambiguity of children's unelaborated "no" responses in the sample questions by more often interpreting these responses as unclear, rather than interpreting them as answering the implicit (i.e., no) or explicit question (i.e., IDK).

Furthermore, we expected that receiving the sample questions first, before evaluating children's testimony, would alert adults to these question-answer pairs and prompt them to consider the possible interpretations (no and IDK), improving adults' ability to recognize referential ambiguity in children's testimony. Specifically, it was expected that participants in the DYK-No condition who received the sample questions before reading testimony (sample first condition), would more often recognize children's unelaborated "no" response to DYK/R if/whether questions as ambiguous (i.e., unclear) in the transcript, compared to participants who did not receive these sample questions (sample last condition).

Finally, it was expected that participants who performed particularly well on the sample questions, recognizing children's unelaborated "no" responses as unclear, would better recognize the ambiguity of "no" responses in the child witness testimony.

3 | METHOD

3.1 | Participants

The final sample included 383 adults recruited from an online participant pool. GPower 3.0.10 was used to calculate the necessary sample

size to detect a moderate effect. Based on the GPower calculation ($N = 424$) and prior research using a similar paradigm using a sample size of approximately 50 participants per cell, a sample size of 383 was deemed appropriate for the present investigation. CloudResearch (formerly Turk Prime) was utilized to obtain a community sample, including only those who were of jury eligible age (18 years or older). Eleven participants were removed from the sample for failing at least one of the two attention check questions (Oppenheimer et al., 2009). Participants were randomly assigned to one of four response-type conditions: DYK-No ($N = 95$), DYK-IDK ($N = 91$), Direct-No ($N = 95$), or Direct-IDK ($N = 100$).

Forty percent of the participants were female, and their ages ranged from 20 to 70 ($M = 37.6$, Median = 35, IQR = 15). The participants were primarily Caucasian (72.8%) as well as 17.8% African American, 3.7% Asian, and 5.7% mixed or another race. One participant did not complete high school, 12.3% held a high school diploma, 22.2% completed some post-secondary education, 41.8% had a university diploma, 10.7% had a trade-school diploma, 12.8% had a graduate or professional degree. Half of the sample (50.9%) was parents. A little over one-third of participants (36.3%) had experience working with children, including teachers, tutors, and other school professionals ($n = 39$), camp and activities counselors ($n = 16$), childcare professionals ($n = 26$), and health care professionals ($n = 3$).

3.2 | Measures

3.2.1 | Sample questions

As a method for alerting participants to referential ambiguity and to assess whether participants could identify referential ambiguity without additional memory demands (i.e., immediately after reading the question-answer pair), participants were asked to interpret children's referentially ambiguous "no" responses to three DYK/R if/whether sample questions. Participants were presented with three question-answer pairs (e.g., Q: "Do you remember if she locked the front door," A: "No") and had the option of selecting one of three interpretations, including (1) a no interpretation (e.g., "No, she didn't lock the front door"), (2) an IDK interpretation (e.g., "No, I don't remember"), or (3) an unclear interpretation (e.g., "it was unclear what the respondent said"). Proportion scores were calculated for the number of "no," "IDK," and "unclear" interpretations by adding the number of responses and dividing it by the total number of trials ($n = 3$).

3.2.2 | Transcripts

The two transcripts in the current study were utilized in a previous study (Wylie et al., 2019). The transcripts had been selected from a larger database, obtained pursuant to the California Public Record Act, which included trial transcripts of child sexual abuse cases that went to trial in Los Angeles County between 1997 and 2001. The two transcripts were selected to include the youngest age group when

referential ambiguity occurs most frequently (Evans et al., 2017), and for being under 20 pages in length (a reasonable length for participants to read). Child 1's transcript included a 7-year-old female witness, the accused was known to the child (friend's father), and the case resulted in an acquittal. Child 2's transcript included a 6-year-old female witness, the accused was known to the child (a neighbor), and the case resulted in a conviction. Only the direct, cross, re-direct, and re-cross examinations of the child's testimony were included. Any competency questions or discussions between the attorneys and the judge were removed from the transcripts.

The two transcripts were manipulated to create four question/response type conditions (DYK-No, DYK-IDK, Direct-No, Direct-IDK), such that approximately every eighth question-answer pair was modified according to the question/response type condition participants were assigned to. As such, a total of 21 question-answer pairs across the two transcripts were manipulated to include either a Do You Know/Remember if/whether question (e.g., Do you know if he was home?) or Direct question (e.g., Was he home?), followed by a "No" or "IDK" response from the child based on the corresponding condition. The transcripts were also manipulated so that participants could not deduce the answer to the target questions from any other question-answer pairs.

3.2.3 | Questionnaire

Participants were told that the questionnaire was assessing their perceptions of the child's report, and they were asked to recall what the child said in response to the attorney's question. Given that children provided unelaborated responses to DYK/R if/whether questions, this allowed us to assess how participants were interpreting children's referentially ambiguous responses. For each target detail, participants had the option of selecting one of five responses in the form of "According to the child..." (1) the event detail did occur (yes interpretation; e.g., "...Joanna was at the swimming pool with her"), (2) the event detail did not occur (no interpretation; e.g., "...Joanna was NOT at the swimming pool with her"), (3) the child did not know if the event occurred (IDK interpretation; e.g., "...she did not know if Joanna was at the swimming pool with her"), (4) it was unclear what the child said (unclear interpretation), and (5) prefer not to answer. Proportion scores were calculated for the number of yes, no, IDK, unclear, and prefer not to answer interpretations by dividing the total number of each type of interpretation by the total number of potentially ambiguous question-answer pairs ($n = 21$, 11 in Child 1's transcript, 10 in Child 2's transcript).

3.3 | Procedure

The procedure was approved by Brock University research ethics board. Prior to commencing participants completed a consent form. Participants completed the Qualtrics survey online via CloudResearch (formerly Turk Prime). Participants were given a maximum of 2 h to

complete the study once the survey was opened ($M = 20.2$ min to complete). Participants were randomly assigned to a sample first (complete the sample questions at the beginning of test) or sample last condition (complete the sample questions after all other tasks). Participants were also randomly assigned to one of the four question/response type conditions, including DYK-No, DYK-IDK, Direct-No, and Direct-IDK. After reading each transcript (the order of presentation of the two transcripts was counterbalanced across participants), participants completed the questionnaire. Upon completion of the study, participants were provided with an online debriefing form outlining the purpose of the study and were compensated \$2.75 for their time.

4 | RESULTS

Preliminary results revealed that there were no differences in the pattern of results for Child 1 and Child 2; therefore, subsequent analyses collapsed across the Child variable. Furthermore, two target question-answer pairs (one question from the Child 1 transcript and one from the Child 2 transcript) were identified as problematic, as participants were consistently answering these questions incorrectly. These errors in responding were likely due to design flaws (e.g., vague and conflicting details in the testimony). Thus, all analyses were performed with the full sample, as well as with dropping the two problematic questions. The pattern of results remained the same when the two problematic questions were dropped. Therefore, all reported analyses below exclude the two questions that were identified to be problematic.

4.1 | Performance on sample questions

Preliminary analyses revealed that there was no difference in performance on the sample questions for those who received these questions before compared to after reading the testimony. Thus, for our examination of participants' performance on the sample questions we collapsed across the sample first and sample last conditions.

A repeated measures ANOVA was conducted with Response Interpretation (unclear, no, IDK) as the within subjects factor, and the proportion of participants' response as the dependent variable. Mauchly's Test of Sphericity indicated that the assumption of sphericity had been violated, $\chi^2(2) = 43.90$, $p < .001$, thus Greenhouse-Geisser correction was used for the repeated measures effects ($\epsilon = 0.90$). There was a significant main effect of Response Interpretation, $F(2, 689.02) = 104.60$, $p < .001$, $\eta_p^2 = 0.215$, where participants most often interpreted children's responses as "IDK" ($M = 0.59$, $SD = 0.39$), followed by "no" ($M = 0.24$, $SD = 0.33$), and lastly "unclear" ($M = 0.17$, $SD = 0.31$; between "IDK" and "no," $p < .001$; "IDK" and "unclear," $p < .001$, "no" and "unclear," $p = .012$). These findings suggest that participants rarely (17% of the time) identified children's unelaborated "no" responses as ambiguous, suggesting a lack of ability to recognize referential ambiguity, even without

memory demands. Furthermore, participants displayed a preference for interpreting children's ambiguous "no" responses as IDK responses (59% of the time). Notably, only 7% of participants were aware that children's unelaborated no responses were unclear across all three sample questions.

4.2 | Influence of sample questions on interpretation of responses

Next, we tested whether receiving the sample questions prior to reading testimony influenced participants' interpretations of referentially ambiguous responses in child witness testimony. A 2 Sample Condition (sample first, sample last) $\times$ 4 Question/Response Type Condition (DYK-No, DYK-IDK, Direct-No, Direct-IDK) $\times$ 4 Response Interpretation (unclear, yes, no, IDK) repeated measures ANOVA was conducted, with Sample condition and Question/Response type condition as the between subjects factors, response interpretation as the within subjects factor, and the proportion of participants' response as the dependent variable. Mauchly's Test of Sphericity indicated that the assumption of sphericity had been violated, $\chi^2(5) = 188.80$, $p < .001$, thus Greenhouse-Geisser correction was used for the repeated measures effects ($\epsilon = 0.77$). There was no significant main effect of Sample Condition, $F(1, 375) = 0.02$, $p = .903$, $\eta_p^2 < 0.001$ or Question/Response Type Condition, $F(3, 375) = 0.56$, $p = .644$, $\eta_p^2 = 0.004$. However, there was a significant main effect of Response Interpretation, $F(2.32, 870.74) = 147.87$, $p < .001$, $\eta_p^2 = 0.283$, which was qualified by a significant Question/Response Type by Response Interpretation interaction, $F(6.96, 870.74) = 90.40$, $p < .001$, $\eta_p^2 = 0.42$. This interaction is interpreted below (see Influence of Question/Response Type on Interpretation of Responses).

None of the other interactions were significant, including the three way interaction, $F(6.97, 870.74) = 0.29$, $p = .957$, $\eta_p^2 = 0.002$. The lack of the three-way interaction is particularly noteworthy because we predicted that the sample questions would uniquely influence responses in the DYK-No condition. Indeed, as shown in Table 1, the pattern of response interpretation was virtually identical for the sample first and sample last conditions. Therefore, the sample questions did not increase participants' identification of "no" responses as unclear.

4.3 | Influence of question/response type on interpretation of responses

Below we interpret the interaction between Question/Response Type and Response Interpretation. See Table 2 for the pattern of response interpretations, recalling that participants ought to interpret children's responses in the DYK-No condition as "unclear", Direct-No as "no," and DYK- and Direct-IDK as "IDK." Table 2 shows that only 10% of participants correctly recognized ambiguous "no" responses as unclear.

TABLE 1 Proportion response interpretation scores (*SD*) by sample condition for the DYK-No condition

	Unclear	Yes	No	IDK	Prefer not to answer
Sample first	0.10 (0.11)	0.26 (0.20)	0.34 (0.17)	0.30 (0.21)	0.00 (0.00)
Sample last	0.10 (0.11)	0.26 (0.20)	0.33 (0.17)	0.31 (0.22)	0.001 (0.01)

TABLE 2 Proportion response interpretation scores (*SD*) by question/response type conditions

	Unclear	Yes	No	IDK	Prefer not to answer
DYK-No	0.10 _{a§} (0.10)	0.26 _{ab‡} (0.20)	0.33 _{b†} (0.20)	0.30 _{b†‡} (0.20)	0.001 (0.01)
DYK-IDK	0.06 _{b¶} (0.10)	0.32 _{a‡} (0.19)	0.12 _{c§} (0.19)	0.50 _{a†} (0.19)	0.001 (0.01)
Direct-No	0.05 _{b§} (0.10)	0.22 _{b‡} (0.19)	0.68 _{a†} (0.19)	0.04 _{c§} (0.19)	0.000 (0.00)
Direct-IDK	0.04 _{b¶} (0.10)	0.26 _{b‡} (0.20)	0.15 _{c§} (0.20)	0.55 _{a†} (0.20)	0.000 (0.00)

Note: Proportion scores were calculated out of the total number of question-answer pairs across the two transcripts ($n = 19$). Significant differences across conditions (within columns) are represented by letter subscripts that do not match, non-significant differences are represented by matching letter subscripts. For comparisons made within conditions (within rows), significant differences are represented by subscript symbols (†, ‡, §, ¶) that do not match, non-significant differences are represented by matching symbols.

4.3.1 | Response interpretations across conditions

First, we examined Question/Response Type conditions for the proportion of unclear, yes, no, and IDK responses separately (see Table 2, within columns). When examining whether participants recognized the ambiguity of “no” responses to DYK/R if/whether questions, participants more often interpreted children's responses as unclear in the condition where children's responses were referentially ambiguous (DYK-No) compared to the conditions where children's responses were not ambiguous (DYK-IDK, Direct-No, Direct-IDK; see Unclear column in Table 2). Furthermore, the proportion of yes interpretations were similar across conditions (see Table 2, Yes column), whereas No interpretations were most common when children provided unambiguous “No” responses (Direct-No; see Table 2 No column) and IDK interpretations were most common when children provided unambiguous “IDK” responses (DYK- and Direct-IDK; see Table 2, IDK column). Therefore, participants were more inclined to identify “no” responses as unclear, but did so only 10% of the time.

4.3.2 | Response interpretations within conditions

Next, response interpretations were examined separately for the DYK-No, DYK-IDK, Direct-No, and Direct-IDK conditions (see Table 2, within rows).

First, we examined participants' response pattern in the referentially ambiguous condition (DYK-No). Although responses were ambiguous, participants rarely interpreted the response as unclear, and instead were equally likely to misinterpret the unelaborated response as no (answering the implicit question) and IDK (answering the explicit question; see Table 2 DYK-No row).

Next, in conditions where children's responses were not ambiguous, participants most often correctly interpreted the child's response (as no in the Direct-No condition, and IDK in the DYK-IDK and Direct-IDK conditions).

We anticipated examining whether the participants who correctly answered the sample questions were better able to identify referentially ambiguous responses in the testimony (i.e., in the DYK-No condition, interpreted responses as unclear). However, as noted above, few participants answered 3/3 sample questions correctly (7% overall, and 5% of subjects in the DYK-No condition). Of this limited sample, participants in the DYK-No condition recognized unambiguous “no” responses as unclear 33% of the time. Participants in the unambiguous conditions rarely identified unambiguous responses as unclear (Direct-No, 6%; DYK-IDK, 12%; Direct-IDK, 2%), suggesting that participants who correctly answered the sample questions were somewhat sensitive to ambiguity in children's testimony and less often tended to overcorrect.

5 | DISCUSSION

The purpose of Study 1 was to examine adults' ability to recognize referential ambiguity when presented with sample questions (i.e., ambiguous “no” responses to DYK/R if/whether questions) and determine whether the opportunity to interpret sample questions would improve their ability to detect this problem in child testimony.

Of concern, participants rarely identified referential ambiguity within the sample questions (in which memory demands were removed). Specifically, referentially ambiguous responses were only identified as unclear 17% of the time. Given participants' poor performance on the sample questions, it is not surprising that completing

sample questions prior to reading the transcripts did not influence participants' interpretations of children's responses. Furthermore, although ambiguous "no" responses in the sample questions were most often interpreted as "IDK" answers, in the transcripts these responses were equally likely to be interpreted as "no" and "IDK." However, we did find a small increase in the "unclear" interpretation in the DYK-No condition compared to other question/response type conditions, unlike Wylie et al. (2019) who found no condition differences, suggesting that participants exhibited some sensitivity to the ambiguity.

One possibility for participants' poor performance may have been a lack of clarity in the response options. In particular, the unclear response option stated that "it was unclear what the child said." Perhaps rather than interpreting this response option as "the child's response was ambiguous" participants interpreted it as not being able to understand the child's statement. Perhaps increasing the clarity of the "unclear" response option would improve participants' performance.

Given that sample questions were not effective in alerting adults to the problem of referential ambiguity, Study 2 sought to examine an additional method of forewarning participants about this problem; providing an explicit instruction. Changes were also made to the wording of the sample questions and the response interpretation options to maximize awareness of referential ambiguity.

6 | STUDY 2

Study 2 examined whether providing an instruction regarding the ambiguous nature of unelaborated "no" responses to DYK/R if/whether questions, with or without accompanying sample questions, would improve adults' ability to detect referential ambiguity in children's testimony.

We also modified the sample questions and the questionnaire to improve participants' ability to identify ambiguity. First, sample questions were modified so that one of the three responses was an unambiguous "IDK" response, thus providing two examples of unelaborated "no" responses (referentially ambiguous) and one unambiguous "IDK" response (clearly answers the question) to DYK/R if/whether questions. We anticipated that this might assist in alerting participants to the differences in the question/response pairs and would ensure that participants did not blindly label all three responses as ambiguous after receiving the explicit instruction describing referential ambiguity. Second, the three sample questions were modified to remove extraneous details (e.g., "Do you remember if she locked the front door" was revised to "Do you remember if the door was locked").

To improve clarity, we also modified the questionnaire. First, the response interpretation options were clarified to alert participants to the distinctions among the three interpretations. Most importantly, the unclear response was modified to highlight referential ambiguity (e.g., "it was unclear what the child said" was revised to "it was unclear whether the witness meant the door wasn't locked or she doesn't remember"). We also modified the other responses to add

clarity. For example, "No, she didn't lock the front door" was revised to "The witness said the door wasn't locked," and "No, I don't remember" was revised to "The witness said she does not remember if the door was locked." Finally, the two problematic questions from Study 1 were dropped.

6.1 | Hypotheses

We predicted that receiving an explicit instruction would increase participants' ability to recognize referential ambiguity on the sample questions. Furthermore, we expected that receiving an explicit instruction and/or sample questions would improve adults' ability to recognize ambiguity in children's testimony in the DYK-No condition, and we anticipated that receiving an instruction and sample questions in combination would be the most effective. We expected the sample questions to improve performance (unlike in Study 1) because of the improved clarity of the response options. Again, in line with Study 1, the Direct- and DYK-IDK conditions were included as control conditions, to assess whether the manipulations of the explicit instruction and sample questions might lead participants to erroneously label unambiguous responses as ambiguous.

7 | METHOD

7.1 | Participants

The final sample included 363 adults recruited from CloudResearch. GPower 3.0.10 was used to calculate the necessary sample size to detect a moderate effect. Based on the GPower calculation ($N = 424$) and prior research using a similar paradigm using a sample size of approximately 50 participants per cell, a sample size of 363 was deemed appropriate for the present investigation. Eighteen participants were removed from the sample for failing at least one of the two attention check questions (Oppenheimer et al., 2009). In addition, 15 participants were removed because they completed the survey in less than 5 min (two standard deviations below the mean completion time of 20 minutes). Participants were randomly assigned to one of four response-type conditions: DYK-No ($N = 93$), DYK-IDK ($N = 91$), Direct-No ($N = 90$), or Direct-IDK ($N = 89$).

Participants' ages ranged from 21 to 81 years ($M = 36.9$, Median = 33, IQR = 17), and 64% of the sample was female. Participants were primarily Caucasian (78.5%) as well as 7.4% African American, 4.7% Hispanic, 3.3% Asian, and 5% mixed or another race, and 1.1% did not report their race. Thirty-nine participants (10.7%) had a high school diploma, 24% completed some post-secondary education, 43.5% had a university diploma, 9.1% had a trade-school diploma, 12.4% had a graduate or professional degree, and one participant did not report their level of education. Approximately half of the sample (48.8%) were parents, and 28.9% had experience working with children, including as teachers, tutors, and other school professionals ($n = 51$), camp and activities counselors ($n = 13$), child-care professionals ($n = 31$), and health care professionals ($n = 10$).

7.2 | Materials

7.2.1 | Explicit instruction

Participants were given the following instruction:

Children's answers to questions in court can sometimes be hard to interpret. One type of question that can lead to such problems is questions that begin with Do You Remember or Do You Know and includes the words "if" or "whether." For example, "Do you remember if it was dark?" These questions can lead to problems in interpreting children's answers because they are really asking two questions: "Do you remember" and "Was it dark?" If a child says "no," then it is unclear what question they are answering. They might be saying "No, I don't remember," or "No, it wasn't dark."

7.2.2 | Sample questionnaire

Participants received three sample questions, revised from Study 1, to include two DYK/R if/whether questions with ambiguous "No" responses and one DYK/R if/whether question with an unambiguous "IDK" response (Q: Do you know if it was raining? A: I don't know.). Additionally, the language used for the response interpretation options was revised for clarity. For example, participants were presented with a question-answer pair (e.g., Q: Do you remember if the door was locked? A: No.) and had the option of selecting one of three interpretations, including (1) a no interpretation (e.g., "The witness said the door wasn't locked"), (2) an IDK interpretation (e.g., "The witness said she doesn't remember if the door was locked"), and (3) an unclear interpretation (e.g., "It was unclear whether the witness meant the door wasn't locked or she does not remember").

Proportion scores were calculated for the number of no, IDK, and unclear interpretations by adding the number of responses, separately for the "no" and "IDK" question-answer pairs. For the ambiguous no response trials, the summed responses were then divided by the number of trials ($n = 2$).

7.2.3 | Transcripts

The transcripts used in Study 2 were the same as those used in Study 1.

7.2.4 | Questionnaire

The questionnaire remained consistent with the first study, except two questions were dropped (one for Child 1, and Child 2), and the wording for the unclear response interpretation was modified to

suggest it was unclear what the child meant (e.g., "It is unclear whether Celina meant Joanna wasn't at the swimming pool or she didn't know"). Proportion scores were calculated for the number of yes, no, IDK, unclear, and prefer not to answer interpretations by dividing the total number of each type of interpretation by the total number of questions ($n = 19$, 10 in Child 1's transcript, 9 in Child 2's transcript).

7.3 | Procedure

As in Study 1, participants were randomly assigned to receive the sample questions before or after evaluating the testimony (sample first or sample last). Additionally, half of the participants were randomly assigned to receive the explicit instruction at the beginning of the study (instruction or no-instruction). The 2 sample questions (sample first, sample last) $\times$ 2 instructions (instruction, no-instruction) design resulted in four conditions: instruction/sample first, instruction/sample last, no-instruction/sample first, no-instruction/sample last. Upon completion of the study, participants were provided with an online debriefing form outlining the purpose of the study and were compensated \$2.75 for their time. This procedure was approved by the Brock University research ethics board.

8 | RESULTS

Preliminary results revealed no differences in the pattern of results across Child 1 and Child 2, thus all further analyses collapsed across the Child variable.

8.1 | Performance on sample questions

Notably, only 19% of all participants correctly interpreted children's responses across all three sample questions (see the pattern of responses in Table 3). As can be seen from the percentages (see Table 3, Unclear column), the explicit instruction increased the proportion of unclear responses, but it appeared to do so indiscriminately, that is, both in response to the truly ambiguous and the unambiguous sample question.

First, we examined whether receiving an explicit instruction influenced performance on the two *ambiguous* sample questions. A 4 Instruction/Sample condition (instruction/sample first, instruction/sample last, no-instruction/sample first, no-instruction/sample last) $\times$ 3 Response Interpretation (unclear, no, IDK) repeated measures ANOVA was conducted, with Instruction/Sample condition as the between subjects factor, response interpretation as the within subjects factor, and the proportion of participants' response as the dependent variable. Mauchly's Test of Sphericity indicated that the assumption of sphericity had been violated, $\chi^2(2) = 56.54$, $p < .001$, thus Greenhouse–Geisser correction was used for the repeated measures effects ($\epsilon = 0.87$). There was no significant main effect of

TABLE 3 Proportion response interpretation scores (SD) by Study 2 conditions for ambiguous and unambiguous sample questions

	Unclear	No	IDK
Ambiguous No Trials			
Instruction/sample first	0.64a* (0.41)	0.10* (0.30)	0.26* (0.41)
Instruction/sample last	0.44b (0.42)	0.15* (0.30)	0.42 (0.41)
No-instruction/sample first	0.23c (0.41)	0.13 (0.30)	0.64* (0.41)
No-instruction/sample last	0.32bc* (0.42)	0.18* (0.30)	0.50* (0.41)
Unambiguous IDK Trial			
Instruction/sample first	0.35a* (0.35)	0.03* (0.23)	0.62* (0.41)
Instruction/sample last	0.19b* (0.35)	0.03* (0.23)	0.78* (0.41)
No-instruction/sample first	0.09bc (0.35)	0.05 (0.23)	0.86* (0.40)
No-instruction/sample last	0.05c (0.36)	0.10 (0.24)	0.85* (0.41)

Note: Proportion scores were calculated out of the total number of sample questions (ambiguous, $n = 2$; unambiguous, $n = 1$). Comparisons made within conditions (within rows), significant differences are represented by * (ambiguous trials, $p < .03$; unambiguous trials, $p < .003$). For comparisons across instruction conditions ('Unclear' column), significant differences are represented by letter subscripts that do not match, non-significant differences are represented by matching letter subscripts, separately for ambiguous and unambiguous trials.

Instruction/Sample condition, $F(3, 359) = 0.00$, $p = 1.00$, $\eta_p^2 = 0.00$. There was a significant main effect of Response Interpretation, $F(1.75, 626.47) = 48.88$, $p < .001$, $\eta_p^2 = 0.12$, qualified by a significant Instruction/Sample condition by Response Interpretation interaction, $F(5.24, 626.47) = 12.00$, $p < .001$, $\eta_p^2 = 0.09$ (see Table 3, Ambiguous No Trials). Participants who received an explicit instruction immediately preceding the sample questions (instruction/sample first condition) were significantly more likely to identify the ambiguity of "No" responses compared to all other conditions ($ps \leq .001$), followed by participants who received the instruction with a delay (instruction/sample last condition), and participants who received no explicit instruction (between "instruction/sample last" and "no instruction/sample first," $p = .001$; "instruction/sample last" and "no-instruction/sample last," $p = .062$; both no-instruction conditions, $p = .145$, see Table 3, within "Unclear" column). The explicit instruction increased performance by reducing participants' tendency to classify "No" responses as "IDK" (with immediate explicit instruction, more likely to interpret as "unclear"; with delayed explicit instruction, equally likely to interpret as "unclear" and "IDK"; without explicit instruction, more likely to interpret as "IDK"; see Table 3, within rows). Therefore, the instruction increased participant's recognition of the ambiguous "no" responses as such in the sample questions.

Next, we examined performance on the unambiguous sample question to assess whether participants were over applying the

"unclear" option. A 4 Instruction/Sample condition (instruction/sample first, instruction/sample last, no-instruction/sample first, no-instruction/sample last) $\times$ 3 Response Interpretation (unclear, no, IDK) repeated measures ANOVA was conducted, with Instruction/Sample condition as the between subjects factor, response interpretation as the within subjects factor, and the proportion of participants' response as the dependent variable. Mauchly's Test of Sphericity indicated that the assumption of sphericity had been violated, $\chi^2(2) = 136.77$, $p < .001$, thus Greenhouse-Geisser correction was used for the repeated measures effects ($\epsilon = 0.76$). There was no significant main effect of Explicit Instruction, $F(3, 359) = 0.00$, $p = 1.00$, $\eta_p^2 = 0.00$. There was a main effect of Response Interpretation, $F(1.52, 544.96) = 316.58$, $p < .001$, $\eta_p^2 = 0.47$, qualified by a significant Instruction/Sample condition by Response Interpretation interaction, $F(4.55, 544.96) = 7.94$, $p < .001$, $\eta_p^2 = 0.06$ (see Table 3, Unambiguous IDK Trials). Participants who received an explicit instruction immediately preceding the sample questions (instruction/sample first) were significantly more likely to interpret unambiguous responses as unclear compared to all other conditions ($ps \leq .003$), followed by participants who received an explicit instruction with a delay (instruction/sample last), and participants who did not receive an explicit instruction (between "instruction/sample last" and "no-instruction/sample first," $p = .062$; "instruction/sample last" and "no-instruction/sample last," $p = .010$; both no-instruction conditions, $p = .484$, see Table 3, within "Unclear" column). Hence, the instruction led to an overcorrection of identifying the unambiguous sample question as ambiguous.

8.2 | Influence of instructions and/or sample questions on interpretation of responses in children's testimony

Next, we tested whether receiving an explicit instruction and/or sample questions influenced participants' interpretations of referentially ambiguous responses in child testimony. A 4 Instruction/Sample Condition (instruction/sample first, instruction/sample last, no-instruction/sample first, no-instruction/sample last) $\times$ 4 Question/Response Type Condition (DYK-No, DYK-IDK, Direct-No, Direct-IDK) $\times$ 4 Response Interpretation (unclear, yes, no, IDK) repeated measures ANOVA was conducted, with Instruction/Sample condition and Question/Response type condition as the between subjects factors, response interpretation as the within subjects factor, and the proportion of participants' response as the dependent variable. Mauchly's Test of Sphericity indicated that the assumption of sphericity had been violated, $\chi^2(5) = 18.80$, $p = .002$, thus Greenhouse-Geisser correction was used for the repeated measures effects ($\epsilon = 0.96$). There were no significant main effects of Instruction/Sample Condition, $F(3, 347) = 1.41$, $p = .238$, $\eta_p^2 = 0.01$ or Question/Response Type Condition, $F(3, 347) = 1.52$, $p = .209$, $\eta_p^2 = 0.01$. There was a significant main effect of Response Interpretation, $F(2.89, 1003.20) = 70.52$, $p < .001$, $\eta_p^2 = 0.17$, that was qualified by a Question/Response Type Condition by Response Interpretation interaction, $F(8.67,$

TABLE 4 Proportion response interpretation scores (*SD*) by question/response type conditions and instruction/sample conditions

	Unclear	Yes	No	IDK	Prefer not to answer	<i>n</i>
DYK-No						
Instruction/sample first	0.27 (0.18)	0.23 (0.17)	0.28 (0.14)	0.22 (0.21)	0.001 (0.02)	20
Instruction/sample last	0.26 (0.28)	0.17 (0.21)	0.25 (0.20)	0.32 (0.27)	0.002 (0.01)	22
No-instruction/sample first	0.23 (0.25)	0.18 (0.15)	0.32 (0.19)	0.27 (0.21)	0.000 (0.00)	26
Average	0.25 (0.24)*	0.19 (0.17)	0.28 (0.18)*	0.27 (0.23)	0.002 (0.01)	68
No-instruction/sample last	0.14 (0.20)*	0.21 (0.20)	0.38 (0.22)*	0.25 (0.23)	0.008 (0.04)	25
DYK-IDK						
Instruction/sample first	0.21 (0.26)	0.24 (0.20)	0.20 (0.16)	0.33 (0.21)	0.02 (0.05)	21
Instruction/sample last	0.13 (0.22)	0.18 (0.16)	0.21 (0.18)	0.45 (0.28)	0.02 (0.07)	22
No-instruction/sample first	0.09 (0.10)	0.23 (0.20)	0.19 (0.17)	0.48 (0.25)	0.006 (0.02)	24
Average	0.14 (0.20)	0.22 (0.19)	0.20 (0.16)	0.42 (0.25)	0.02 (0.05)	67
No-instruction/sample last	0.20 (0.21)	0.23 (0.22)	0.16 (0.14)	0.40 (0.25)	0.01 (0.04)	24
Direct-No						
Instruction/sample first	0.10 (0.12)	0.20 (0.21)	0.64 (0.24)	0.04 (0.08)	0.02 (0.05)	24
Instruction/sample last	0.05 (0.08)	0.10 (0.12)	0.81 (0.16)	0.02 (0.05)	0.01 (0.03)	23
No-instruction/sample first	0.08 (0.11)	0.15 (0.17)	0.74 (0.21)	0.02 (0.03)	0.01 (0.04)	19
Average	0.08 (0.11)	0.15 (0.17)	0.73 (0.21)	0.03 (0.06)	0.01 (0.04)	66
No-instruction/sample last	0.07 (0.06)	0.14 (0.13)	0.74 (0.19)	0.05 (0.12)	0.000 (0.00)	24
Direct-IDK						
Instruction/sample first	0.17 (0.15)	0.18 (0.16)	0.20 (0.13)	0.41 (0.23)	0.03 (0.11)	22
Instruction/sample last	0.14 (0.11)	0.18 (0.12)	0.25 (0.17)	0.42 (0.18)	0.000 (0.00)	19
No-instruction/sample first	0.14 (0.15)	0.21 (0.19)	0.16 (0.11)	0.47 (0.26)	0.01 (0.39)	23
Average	0.15 (0.14)	0.19 (0.16)	0.20 (0.14)	0.44 (0.22)	0.01 (0.07)	64
No-instruction/sample last	0.10 (0.10)	0.27 (0.17)	0.20 (0.10)	0.43 (0.18)	0.004 (0.02)	25

Note: * $p < .05$ significant differences presented in the manuscript. *n* represents the number of participants in each condition. Average rows were calculated by collapsing across the instruction/sample first, instruction/sample last, and no-instruction/sample first conditions.

1003.20) = 69.78, $p < .001$, $\eta_p^2 = 0.38$, where the pattern of results replicated what was found in Study 1 (see Table 2 for Study 1, and Table 4 in the “None” rows for Study 2). None of the other interactions were significant.

Planned comparisons examined the influence of the instruction and/or sample first conditions on participants' ability to correctly identify referential ambiguity in the DYK-No condition, compared to the unambiguous conditions (DYK-IDK, Direct-No, Direct-IDK). First, we examined differences between receiving an instruction and/or sample questions (i.e., collapsing across instruction/sample first, instruction/sample last, and no-instruction/sample first conditions), compared to receiving no information before reading children's testimony (no-instruction/sample last condition; see Table 4, “Average” and “None” rows). As expected, when responses were ambiguous (DYK-No), participants that received an instruction and/or sample questions first ($M = 0.25$, $SD = 0.24$) were significantly more likely to identify ambiguity in children's testimony as unclear, compared to participants that received no instruction or sample questions prior to reading children's testimony (no-instruction/sample last, $M = 0.14$, $SD = 0.20$), $t(91) = 2.02$, $p = .047$. Receiving an instruction and/or sample questions first reduced participants' tendency to classify ambiguous “No”

responses as “No,” $t(91) = 2.14$, $p = .035$ (see Table 4 “No” column for DYK-No). For the unambiguous conditions (DYK-IDK, Direct-No, Direct-IDK), there were no significant differences in participants' interpretations of unambiguous responses as unclear when they received an instruction and/or sample questions first, compared to no information (no-instruction/sample last condition; see Table 4, Unclear column for unambiguous conditions), $ps > .08$. Therefore, participants who received the instruction and/or sample questions first were more likely to identify ambiguous questions as such, and did not interpret unambiguous questions as ambiguous (i.e., they did not overcorrect).

Next, we examined differences within the three conditions in which participants received an instruction or answered sample questions before reading the transcripts (instruction/sample first, instruction/sample last, and no-instruction/sample first). Contrary to our prediction that participants who received both an explicit instruction and sample questions before reading children's testimony (instruction/sample first) would perform the best, the results revealed no significant difference, $F(2, 65) = 0.167$, $p = .846$, in participants' identification of referentially ambiguous responses as unclear when receiving both an instruction and sample questions (instruction/sample first, $M = 0.27$, $SD = 0.18$), only an instruction (instruction/sample last, $M = 0.26$,

$SD = 0.28$), or only sample questions (no-instruction/sample first, $M = 0.23$, $SD = 0.25$) before reading children's testimony, $ps > .60$. There were no differences in participants' interpretation of unambiguous responses as unclear, $ps > .07$. Therefore, either sample questions or an instruction-improved participants' ability to identify ambiguous questions as such, with no evidence of overcorrection.

8.3 | A comparison of study 1 and 2: Condition and response interpretation

Given the findings in Study 2 (and lack thereof in Study 1), we explored whether the manipulation of revising the unclear response option in Study 2 improved participants' performance compared to Study 1, by clarifying what the unclear option was intended to reflect. A 2 Study (Study 1, Study 2) $\times$ 4 Question/Response Type Condition (DYK-No, DYK-IDK, Direct-No, Direct-IDK) $\times$ 4 Response Interpretation (unclear, yes, no, IDK) repeated measures ANOVA was conducted, with study and Question/Response type condition as the between subjects factors, response interpretation as the within subjects factor, and the proportion of participants' response as the dependent variable. Mauchly's Test of Sphericity indicated that the assumption of sphericity had been violated, $\chi^2(5) = 109.42$, $p < .001$, thus Greenhouse-Geisser correction was used for the repeated measures effects ($\epsilon = 0.91$).

There was a significant main effect of Study, $F(1, 738) = 4.76$, $p = .029$, $\eta_p^2 = 0.01$, and Response Interpretation, $F(2.72, 2005.24) = 185.41$, $p < .001$, $\eta_p^2 = 0.20$, qualified by a significant Study by Response Interpretation, $F(2.72, 2005.24) = 34.11$, $p < .001$, $\eta_p^2 = 0.04$, and Condition by Response Interpretation interaction, $F(8.15, 2005.24) = 156.29$, $p < .001$, $\eta_p^2 = 0.39$, and a significant three-way Study by Condition by Response Interpretation interaction, $F(8.15, 2005.24) = 2.04$, $p = .038$, $\eta_p^2 = 0.01$. There was no significant main effect of Condition, $F(3, 738) = 0.65$, $p = .583$, $\eta_p^2 = 0.003$, or Study by Condition interaction, $F(3, 738) = 1.95$, $p = .121$, $\eta_p^2 = 0.01$. Pairwise comparisons were performed to examine Study differences, separately for each Condition and Response Interpretation.

For the DYK-No condition, participants were significantly more likely to interpret children's ambiguous "No" responses as unclear in Study 2, compared to Study 1, $p < .001$, suggesting the revised language for the unclear option may have improved participants' ability to detect ambiguity. Performance improved by reducing participants' tendency to misinterpret ambiguous "no" responses as "yes" in Study 2, compared to Study 1, $p < .02$.

However, participants in the DYK-IDK and Direct-IDK conditions were also significantly more likely to interpret children's unambiguous "IDK" response as unclear in Study 2, compared to Study 1, $ps < .001$, suggesting the revised unclear option may have led to over interpretation of unclear when responses were not ambiguous. Participants' performance decreased, whereby they were also more likely to misinterpret children's IDK response as "no," $p \leq .03$, and less likely to correctly interpret children's responses as IDK in Study 2, compared to Study 1, $ps < .01$.

Participants in the Direct-No condition did not demonstrate an increased tendency to interpret unambiguous "no" responses as unclear, when comparing Study 1 and 2, $p = .318$.

In sum, the revised unclear option increased participants' recognition of ambiguous responses as such, but also increased overcorrection, at least with respect to participants' interpretation of do not know responses.

9 | DISCUSSION

The purpose of the current set of studies was to examine whether highlighting referential ambiguity (via sample questions or an explicit instruction) could forewarn adults about the problem, and in turn influence their interpretations of children's ambiguous "no" responses as unclear within child witness testimony. Study 1 found that adults had difficulty identifying referential ambiguity, even when asked sample questions designed to highlight ambiguity. In turn, sample questions did not help when identifying ambiguity in children's testimony. Study 2 found that giving an explicit instruction, and simplifying and clarifying response options, led adults to identify ambiguity in the sample questions, but led them to do so indiscriminately. Instructions and/or sample questions improved adults' ability to identify ambiguity in children's testimony, though the rate of improvement was modest.

As seen in the results from Study 1, participants most often failed to recognize referential ambiguity (10%) in the sample questions. Therefore, in Study 2, we examined the influence of an explicit instruction on participants' performance on the sample questions. Results revealed that participants who received the instruction were more likely to interpret referentially ambiguous "no" responses as unclear in the sample questions, compared to participants who did not receive the instruction. However, while this instruction improved performance on the sample questions, ambiguous responses were only identified as such 54% of the time, suggesting that while this instruction may be beneficial, misinterpretations remain pervasive. Of concern, the explicit instruction also resulted in participants over-using the "unclear" label, identifying unambiguous IDK responses as unclear. In line with past research, these findings suggest that providing an educational instruction can lead to overcorrection (e.g., Wegener & Petty, 1997; Wilson et al., 2002).

Given that receiving the explicit instruction influenced participants' performance on the sample questions, we were also interested in whether receiving this instruction influenced participants' interpretation of children's responses within witness testimony. Promisingly, the instruction and sample questions improved adults' ability to identify referential ambiguity in children's testimony without leading to overcorrection, though there were no differences in receiving the instruction, sample questions, or both. The overapplication of the explicit instruction in the sample questions but not witness testimony may suggest that the effect of the instruction was short-lived, and that respondents were unthinkingly identifying responses as ambiguous following the instruction. Furthermore, whereas the sample questions were not effective in highlighting the problem of referential ambiguity in Study 1, they were beneficial in Study 2. This may be due to the improved clarity of the sample questions, as the response options were simplified and clarified. However, as seen when comparing results from Study 1 and 2, the revised language of the unclear response may have led to an overcorrection of identifying

unambiguous responses as unclear, as seen in the unambiguous IDK conditions. These findings suggest that it may be possible to alert adults to the problem of ambiguity, though they may overcorrect when interpreting unambiguous responses.

It is important to note that the rate of improved performance when assessing child testimony is still concerning. That is, participants receiving an instruction and/or sample questions prior to reading the transcripts only recognized ambiguity in children's responses 24% of the time (compared to 14% with no-instruction/sample last), suggesting adults are most often failing to recognize the problem of referential ambiguity. Instead, adults were most often forming their own interpretations of the intentions of children's responses, again replicating the general pattern of results as found in Study 1 and by Wylie et al. (2019).

10 | CONCLUSION

Research suggests that given jurors' lack of sensitivity to problematic question and response types, it may be necessary to help jurors in cases where errors and miscommunications are more likely to occur (McAuliff & Kovera, 2007; Quas et al., 2005). However, the results of the studies reported here suggest that highlighting the problem of referential ambiguity in children's responses to DYK/R if/whether questions may not be sufficient to enable adults to identify ambiguities in children's testimony. Despite instruction and sample questions illustrating ambiguity, adults continued to misinterpret most ambiguous responses. In fact, our findings suggest that some strategies may lead to overcorrection such that adults interpret children's unambiguous responses as unclear. Future studies might modify the strategies for alerting adults to this problem, such as by providing still more explicit instructions distinguishing between ambiguous and unambiguous responses, and by giving participants feedback on sample questions. At some point, however, it is unrealistic to expect that this type of education could take place in courtrooms in which laypeople assess children's credibility. Given the current lack of an effective strategy to highlight the problem of referential ambiguity, DYK/R if/whether questions should be avoided whenever possible.

ACKNOWLEDGMENT

Preparation of this article was supported in part by National Institute of Child Health and Human Development Grant HD087685 awarded to Thomas D. Lyon.

CONFLICT OF INTEREST

The authors declare there is no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

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

REFERENCES

- Ahern, E. C., Stolzenberg, S. N., McWilliams, K., & Lyon, T. D. (2016). The effects of secret instructions and yes/no questions on maltreated and non-maltreated children's reports of a minor transgression. *Behavioral Sciences & the Law*, 34, 784–802. <https://doi.org/10.1002/bsl.2277>
- Bruck, M., Ceci, S. J., & Francoeur, E. (1999). The accuracy of mothers' memories of conversations with their preschool children. *Journal of Experimental Psychology: Applied*, 5, 89–106. <https://doi.org/10.1037/1076-898X.5.1.89>
- Evans, A. D., Stolzenberg, S. N., Lee, K., & Lyon, T. D. (2014). Young children's difficulty with indirect speech acts: Implications for questioning child witnesses. *Behavioral Sciences & the Law*, 32, 775–788.
- Evans, A. D., Stolzenberg, S. N., & Lyon, T. D. (2017). Pragmatic failure and referential ambiguity when attorneys ask child witnesses "Do you know/remember" questions. *Psychology, Public Policy, and Law*, 23, 191–199.
- Laimon, R. L., & Poole, D. A. (2008). Adults usually believe young children: The influence of eliciting questions and suggestibility presentations on perceptions of children's disclosures. *Law and Human Behavior*, 32, 489–501. <https://doi.org/10.1007/s10979-008-9127-y>
- McAuliff, B. D., & Kovera, M. B. (2007). Estimating the effects of misleading information on witness accuracy: Can experts tell jurors something they don't already know? *Applied Cognitive Psychology*, 21, 849–870.
- Mugno, A. P., Klemfuss, J. Z., & Lyon, T. D. (2016). Attorney questions predict jury-eligible adult assessments of attorneys, child witnesses, and defendant guilt. *Behavioral Sciences & the Law*, 34, 178–199.
- Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, 45, 867–872.
- Quas, J. A., Thompson, W. C., & Clarke-Stewart, K. A. (2005). Do jurors "know" what isn't so about child witnesses? *Law and Human Behavior*, 29(4), 425–456. <https://doi.org/10.1007/s10979-005-5523-8>
- Ruva, C. L., & Bryant, J. B. (2004). The impact of age, speech style, and question form on perceptions of witness credibility and trial outcome. *Journal of Applied Social Psychology*, 34, 1919–1944. <https://doi.org/10.1111/j.1559-1816.2004.tb02593.x>
- Warren, A. R., & Woodall, C. E. (1999). The reliability of hearsay testimony: How well do interviewers recall their interviews with children? *Psychology, Public Policy, and Law*, 5, 355–371. <https://doi.org/10.1037/1076-8971.5.2.355>
- Wegener, D. T., & Petty, R. E. (1997). The flexible correction model: The role of naive theories of bias in bias correction. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 29, pp. 141–208). Academic Press.
- Wilson, T. D., Centerbar, D. B., & Brekke, N. (2002). Mental contamination and the debiasing problem. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 185–200). Cambridge University Press. <https://doi.org/10.1017/CBO9780511808098.012>
- Wylie, B. E., Lyon, T. D., O'Connor, A. M., Aidy, C. L., & Evans, A. D. (2019). Adults' perceptions of children's referentially ambiguous responses. *Psychology, Crime & Law*, 25, 729–738. <https://doi.org/10.1080/1068316X.2018.1552757>

How to cite this article: Wylie, B. E., Gongola, J., Lyon, T. D., & Evans, A. D. (2021). The difficulty of teaching adults to recognize referential ambiguity in children's testimony: The influence of explicit instruction and sample questions. *Applied Cognitive Psychology*, 35(5), 1297–1307. <https://doi.org/10.1002/acp.3863>