Improving Responding to Questioning Using a Brief Retrieval Training

Alan Scoboria [1]

Amina Memon [2]

Hoa Trang [1]

Marc Frey [1]

1 University of Windsor

2 Royal Holloway University of London

Abstract

This research examined the effect of a brief training procedure for enhancing responding to questions about witnessed events. The training was based in research on metacognition and memory, and emphasized: attending to questions, searching for multiple responses, and weighing confidence in and considering the source of responses. In the main study, adult participants viewed a video of a burglary and after a 25 min. delay half received the training. All participants were then asked answerable and unanswerable questions about the video. The training resulted in fewer errors and more rejections to unanswerable questions. Analysis of response diagnosticity indicated that responses made by the trained group were more likely to be correct responses to answerable questions. A second study showed that these findings were not due to awareness of the presence of unanswerable questions. The procedure has potential as a supplement when questioning is pursued.

Keywords: memory, interview, questioning, unanswerable, don't know

Improving Responding to Questioning Using a Brief Retrieval Training

Considerable interest is currently directed at promoting best practices in forensic interviewing. A large body of research documents the effects of poor interviews on the quality of the information obtained. Studies document negative effects of poor rapport, social pressure, suggestive interviewing, and closed ended questioning (see Fisher, Milne, & Bull, 2011; Milne, Poyser, Williamson, & Savage, 2010). In light of these risks, researchers have directed their attention to developing effective interviewing techniques. Extensive research documents that, following adequate training, best practices can result in substantially higher quality of information gathered (see Memon, Meissner, & Fraser, 2011, for a review of the Cognitive Interview [CI]; and Lamb, Orbach, Hershkowitz, Esplin, & Horowitz, 2007, for a review of the National Institute of Childhood Health and Human Development [NICHD] protocol).

While the development of research-based guidance is noteworthy, evidence to date suggests that guidelines and training do not readily translate to better interviews in the field. Trained interviewers continue to use suggestive questions and tend to fall back on techniques they used prior to training (Cederborg, Orbach, Sternberg, & Lamb, 2000; Cyr & Lamb, 2009; Davies, Wilson, Mitchell, & Milsom, 1995; La Rooy, Lamb, & Memon, 2010; Schreiber Compo, Gregory, & Fisher, 2011). This appears to be partly due to the fact that best practice models are somewhat complex and may be challenging for interviewers to learn and deploy (Dando, Wilcock, Milne, & Henry, 2009). Hence despite empirically supported best-practice recommendations, both directed and specific questioning appear likely to occur, and new tools are needed to help bridge the gap between research and practice. Recognizing this need to identify ways of improving interviewing in practice, Fisher and colleagues (2011) recently called for research advances that promote better interviews.

The current work examined a complementary procedure to enhance witnesses attention to the availability and quality of potential responses. This approach was taken in part due to the problem of questions being asked about non-witnessed information. Interviewers rarely know what occurred during the events about which they ask questions, hence they cannot know which of the questions they ask can be answered based on what occurred during the event. Yet it is well documented that witnesses often provide substantive responses to such questions (see below). When questions are asked about information that was not present in the event (questions about non-occurring information), any substantive response is very likely to be erroneous. Effective responding to such questions requires that individuals either reject the question (state that the question cannot be answered), or choose not to respond.

Hence while questioning can result in gaining valuable information, all types of interviewers, forensic or otherwise, face a common challenge: they do not know what occurred, and hence, may ask questions about information that was never in the witnessed event. Such queries about information that cannot be provided because it was not present to witness are termed *unanswerable questions* in the interviewing literature (Roebers & Fernandez, 2002; Waterman & Blades, 2011). When interviewees respond to such questions, the reliability and utility of the information gathered is reduced. The literature shows that child and adult respondents often provide substantive responses to the unanswerable questions that they are asked, if the questions are reasonably plausible (Mulder & Vrij, 1996; Poole & White, 1991; Scoboria, Mazzoni, & Kirsch, 2008; Waterman & Blades & Spenser, 2004). These questions would ideally be rejected but this work shows that this tends not to be the case. Interview procedures are needed to address such concerns. For example, Scoboria and Fisico (2013) found that explicit encouragement of the use of "don't know" (DK) responses enhanced avoidance of

responding to unanswerable questions, and that such encouragement accompanied with clarifying the meanings of DK responses increased the number of unanswerable questions that were correctly rejected.

The more questions that are asked, and the more specific the questions, the greater the likelihood that unanswerable questions will be asked. The likelihood of unanswerable questions also increases when interviews are agenda driven. For example, interviewers may fall prey to confirmatory bias because they form theories about events, acquire information from other sources, or query typical aspects of crime scenarios (Hill, Memon, & McGeorge, 2008; Meissner & Kassin, 2002; Nickerson, 1998). Simply because a piece of information is acquired from another source does not guarantee that the detail was in fact present, or that the witness was in a position to observe the detail if it was there. Interviewers may assume knowledge and insist that witnesses answer questions about items that were not present, in which case the well documented detrimental effects of forced confabulation can undermine memory (Chrobak & Zaragoza, 2008).

Perhaps interviewers assume interviewees will give accurate reports and that they will say when they are unable to answer a question. However, there are reasons to think that interviewees are unlikely to do this with regularity. To explain this, we turn to a discussion of the role of don't know responses in interviews.

Don't Know Responding

Don't know (DK) responses are widely recognized as a crucial component of effective responding and form part of many best practice interviewing techniques (Milne & Bull, 1999). The DK response option permits individuals to regulate when they answer questions. This is demonstrated, for example, by Koriat and Goldsmith's (1996) work on memory regulation, which has been extended to interviews about witnessed scenes (Pansky, Goldsmith, & Koriat,

2009; Roebers & Fernandez, 2002). When people are forced to respond, output (the total number of questions answered) is artificially maximized, but at the cost of accuracy because people must provide low confidence responses or guess. This body of work supports the assertion that free report (having the option to choose to answer or not answer any given question) is an essential component of best practice interviewing.

When free report to questions is allowed, quantity tends to decrease while accuracy tends to increase (assuming that metacognition is generally effective); this is termed the quantity/accuracy trade-off (Koriat & Goldsmith, 1996). One reason that people choose to not respond (e.g., say DK) is to avoid guessing or reporting low confidence information. This is often beneficial to the quality of responding. But people also say DK to avoid responding to unanswerable questions, and they do this more frequently than they reject questions by stating that information was not present (Scoboria & Fisico, 2013). Hence, under free report adults underreport their knowledge that details queried were not present. This may be because DK responses provide a socially acceptable alternative to appearing unhelpful or uninformative rather than saying 'I did not see that' or 'that was not there' too frequently (Ackerman & Goldsmith, 2008). Ideally, individuals would report their knowledge that something was not present as this information has greater utility.

Developing a Novel Training Procedure

In light of the issues surrounding unanswerable questions, we developed a brief training protocol to enhance responding to such questions. The training procedure was designed to achieve a number of goals. The primary goal was to encourage careful evaluation of questions, potential responses to said questions, and selection of responses. The procedure (reviewed next) drew on established research findings regarding memory and metacognition, to provide

interviewees with general tools as to how to evaluate their responding with care. Associated

goals, drawn from best practice principles (Fisher & Geiselman 1993), were to transfer

responsibility for choosing what to report to interviewees, thereby reducing demands to respond

because interviewers inject fewer statements into the interview. This approach should also

encourage interviewees to use their own idiosyncratic memory search and monitoring strategies,

rather than imposing any specific search strategy. We selected four principles that have the

potential to enhance responding:

Principle 1. Interrupting automatic responding. Sometimes people respond to questions

based on inference that they must have seen something, or because information comes to mind

rapidly (Whittlesea, 1993; Zaragoza & Lane, 1994). The purpose here was to encourage further

evaluation of initial 'quick' responses prior to outputting them as an answer, and more generally,

to encourage careful thought about any response.

Principle 2. Attending to the features of the question. People sometimes respond to

questions that they do not understand, simply because they know that a question has been asked.

An extreme example is that children and adults sometimes provide responses to entirely

nonsensical questions (Pratt, 1990; Waterman, Blades, & Spenser, 2000). This may be related to

interviewees wanting to be helpful or due to compliance with the interviewer (Roper & Shewan,

2002). Attention to the question should encourage individuals to examine whether they

understand what is being asked. This may also aid communication, in that the interviewee may

choose to seek clarification from the interviewer. Directing attention to the question also conveys

that the question may not be well structured, or may query details that were not witnessed.

Principle 3. Encouraging thoughtful search, response monitoring, and response

selection. From Koriat and Goldsmith (1996) we borrowed the concepts of searching memory for

multiple candidate responses, weighing the quality of candidates, and selecting the best response. These are key steps posited as contributing to the selection of correct responses. By referring to the possibility that more than one response may come to mind, individuals are oriented to the chance that some candidates must be erroneous. Encouraging selection of the best response and weighing if the response is "good enough" also raises the possibility that no response that comes to mind may be of sufficient quality. This further reinforces that choosing to not respond is acceptable, and may also heighten awareness that a question has no answer.

Principle 4. Attending to source. From work on source monitoring (Johnson, Hashtroudi, & Lindsay, 1993) we borrowed the idea of attending to the source of potential responses. By drawing attention to the fact that candidates may come from sources other than the event in question, interviewees are invited to decide if a potential response originates from the target event. This cue was included in part due to concerns about the intrusion of post-event misinformation (Zaragoza & Lane, 1994) or event schematic but non-presented information in reports (García-Bajos, Migueles, & Aizpurua, 2012).

This research tested a brief training designed to supplement existing best practice, to promote higher quality responding to questions. We anticipated that trained participants would more carefully weigh potential responses prior to answering, resulting in improved responding. In particular, we anticipated more effective responding to unanswerable questions in the form of more spontaneous statements that the information was not present (preferably), and or greater avoidance of responding to unanswerable questions via DK responses.

Study 1 Method

Participants

Fifty-four undergraduate students (84% female; average age 21.4) participated in

exchange for course credit.

Materials

     Video. The 5 min. video depicted a non-violent burglary and police chase and was the same as used in prior work (Zaragoza & Mitchell, 1996; Scoboria et al., 2008; Scoboria & Fisico, 2013).**Interview questions.** We asked participants 10 open-ended answerable (querying details in the video) and 10 unanswerable (querying details not in the video) questions. The questions were the same as in Scoboria and Fisico (2013). The answerable questions queried information that was present in the video (e.g., "What did the robber's vest say on the back?"). The unanswerable questions asked about plausible information that was not in the video (potentially observable information that was not present; e.g. "Where did the robbers say they were going after the robbery?"), and did not query unobservable information (e.g., internal states). All of the questions were designed to be open-ended and non-leading; very few could be answered with a yes/no response (and these were typically answered with more information). Potential responses were not presented within the questions. The questions varied somewhat in breadth. Some questions queried specific details (e.g., description of a patch on a jacket), while others were broader (e.g., where did the robbers say that they were going). We note that no single question was responsible for the pattern of results; there was reasonable variability in responses across all of the questions.

     Response instructions. The training procedure was comprised of a set of prompts to encourage thoughtful memory search, monitoring, and response selection. We labeled the prompts *Review, Retrieve, References, Reflect,* and *Reply* (see Appendix A for exact wordings for each). The first cue, *Review,* was included to interrupt automatic responding. The second cue, *Retrieve*, encouraged thorough memory search and review of candidate responses. The third,

References, prompted attending to the source of candidates. The fourth, *Reflect*, encouraged

weighing of confidence in candidates. The final cue, *Reply*, asked for selection of the best

answer. The interviewer explained and demonstrated the prompts with a sample question, and

answered any queries. Participants then practiced with two simple answerable questions, during

which the interviewer encouraged participants to verbally articulate their use of the cues. The

total time for explanation and practice ranged from two to four minutes.

Procedure

A research assistant administered the individual sessions, and the procedures were

administered verbally (to mirror face-to-face interviews). After consent, participants were told to

watch a video that they would be questioned about later. They watched the video on a 21" color

monitor from a distance of 5 feet, completed a 5 min free recall as rehearsal, and then completed

a distractor (word scramble) for 25 min.

Participants were randomly assigned to the groups. For participants in the instruction

group, the prompts were introduced and explained, and a printed list of the prompts was

provided for use during questioning. Participants were encouraged to use the list for the first two

questions, after which the prompts were no longer mentioned. Control participants continued

with the distractor for 5 min to equalize the time.

We next questioned all participants in the same manner. The interviewer told participants

that they could respond that they did not know. The interviewer asked the questions aloud and

recorded responses verbatim. No feedback was given during questioning.

Data Coding

Responses were coded as per Scoboria et al. (2008) using a standard manual; two

independent raters showed high agreement ($r > .95$) for all variables. Responses to answerable

questions were coded as correct, error, or DK. Responses to unanswerable questions were coded as correct rejection, error, or DK. Output was calculated as the ratio of questions answered to questions asked; and accuracy was calculated as the ratio of correct responses to questions answered. DK responses were treated as a choice to not answer the question, as DK responses are known to be heterogeneous in their meaning (see Scoboria et al., 2008).

Study 1 Results

Our main interest was whether individuals in the instructed group provided higher quality responses relative to controls. Instructed participants responded more effectively to unanswerable questions (see Table 1 for descriptives and effect sizes). Specifically, trained participants made fewer errors and rejected more unanswerable questions, resulting in higher accuracy. This difference was not due to differences in output, which was equal between the groups. In statistical terms, univariate analyses of variance (ANOVAs) revealed significant main effects for correct rejections, $F(1,52) = 4.65$, $p = .036$, Cohen's $d = .59$; errors, $F(1,52) = 4.93$, $p = .031$, $d = .60$, and accuracy, $F(1,52) = 6.76$, $p = .012$, $d = .83$; but no effect for output , $F(1,52) = .01$, $p = .98$, , $d = .02$. No between-group differences for answerable questions were found (all $p > .10$).

While direct comparison of answerable and unanswerable questions is problematic due to differences in the meaning of each question type (for example, correct responses to answerable questions consist of substantive responses, whereas correct to unanswerable questions involve rejections), the questions can be combined to evaluate the diagnosticity of responses. Diagnosticity addresses the degree to which substantive answers indicate that a correct answer has been made to an answerable question. As in Scoboria and Fisico (2013), diagnosticity was calculated as the hit rate for answerable questions over the false alarm rate for unanswerable

questions (see Wells & Lindsay, 1980; Weber & Perfect, 2012), which was log-transformed

(Agresti, 2002). Diagnosticity was higher in the trained group ($M = .89$, $SD = .77$ vs. $M = .41$, SD

$= .79$); this difference was statistically significant, per one-way ANOVA; $F (1,52) = 4.97$, $p =$

$.030$, $d = .61$ (95% CI .42,.83).

Study 1 Discussion

The training resulted in enhanced responding to unanswerable questions. The training led

to more unanswerable questions being overtly rejected, and improved the likelihood that

substantive responses were correct responses to answerable questions. These findings are also

notable in that the gains occurred for accuracy without a loss in output (questions attempted).

Before discussing the findings further, we address one potential confound that could

qualify the findings of Study 1. Perhaps the training procedure made respondents aware that

unanswerable questions were present. This awareness alone might be sufficient to produce the

pattern of findings. Hence we conducted a second study that examined whether a warning about

the possibility of unanswerable questions produced similar results.

Study 2 Method

Participants

Fifty students (72% female; mean age 19.8) participated in exchange for course credit.

Study 2 Method

The methods were identical to Study 1, with the exception that the trained group was

replaced with a group that was warned about the potential presence of unanswerable questions.

Immediately prior to questioning, those in the warning group were told, "It is possible that some

of the questions that I ask may not have an answer." This phrasing was selected intentionally to

reflect that interviewers typically do not know if and when they ask unanswerable questions.

Study 2 Results

The warned group responded to fewer unanswerable questions (Table 2), $F(1,48) =$ 59.26, $p < 001$, $d = .62$ (95% CI .27,1.10), and trended toward answering fewer answerable questions, $F(1,48) = 2.17$, $p = .147$, $d = .43$ (95% CI .01,.85). The remaining variables (i.e., correct, error, accuracy) for unanswerable questions showed no statistically significant group differences (all $p > .10$). Hence the warned group simply answered fewer questions, with no benefit to response quality. The diagnosticity index confirmed this finding; diagnosticity did not differ between the warned and control groups ($M = .47$, $SD = .26$ vs. $M = .53$, $SD = .25$), and numerically favored the control group, $F(1, 48) = 0.51$, $p = .497$, $d = -.14$ (95% CI -.34,.01).

General Discussion

A brief training that emphasized attention to metacognitive features associated with memory enhanced the quality of responding to unanswerable interview questions. This occurred at no cost to output (the number of questions attempted). Trained individuals were more likely to correctly reject and were less likely to make errors to unanswerable questions. The instructions did not otherwise affect reports for information that was present, hence the gains for responding to unanswerable questions occurred at no cost to the quality of responses made to answerable questions. The training was also associated with an increase in the likelihood that substantive responses were correct responses to answerable questions.

The fact that the instructions affected reporting about non-occurring items but had little impact on memory for details present in the video indicates that the training procedure did not notably impact the processes that contribute to retrieving information that is potentially in memory. This research provides further evidence that the processes required when responding to unanswerable questions are at least partially distinct from those involved in responding to

answerable questions. In their discussion of memory for what they term "non-occurrence",

Strack and Bless (1994) proposed that when a memory search fails to produce a candidate

response, individuals must decide if the lack of memory is diagnostic that the information was

never available. Decision making processes are used to make inferences about the lack of

memory, based on features such as distinctiveness (Dodson & Schacter, 2002), or plausibility

(Ghetti, 2003). Mazzoni and Kirsch (2001) extended Koriat & Goldsmith's (1996) strategic

regulation model to accommodate instances in which a memory search results in no retrieved

information. They proposed that individuals engage inferential reasoning processes in order to

decide the basis for the lack of memory, and can come to any number of conclusions which

determine whether they continue or terminate the memory search. Examples include: (1) the

information needed to answer the question has been forgotten; (2) the information is in memory

but is temporarily unavailable (as in tip-of-the-tongue states); (3) the information is believed to

have been present despite the lack of episodic memory; and, (4) the information cannot be

retrieved because it was never present to be encoded. The difference between the final example

and the others is the conclusion that the information was never available. The training described

in this research enhanced the likelihood that respondents would come to this conclusion

effectively.

This is further supported by the difference in how the groups in Study 1 arrived at their

accuracy rates for unanswerable questions. The trained group traded errors for correct rejections

for unanswerable questions, relative to the control group. It is interesting to speculate why this

may have occurred. These studies permit ruling out several competing explanations such as a

lack of awareness of the option to not respond (all individuals were informed that "don't know"

was an acceptable answer), a lack of awareness of 'not present' as a response option (all

participants made one or more spontaneous rejections during free report, even in the control group), or awareness that unanswerable questions might be asked (as shown in Study 2). Hence, the training itself uniquely enhanced the effectiveness of metacognitive processing and/or communication for unanswerable questions, leading to overt rejections of questions that would have otherwise resulted in errors.

We intentionally included a number of facets in the training that are thought contribute to metacognitive monitoring and control. It is not possible to say here whether the entire training is needed, or if a smaller number of prompts would be sufficient. The training may have impacted a single process which is not isolated here, or may have impacted multiple processes. Because errors to unanswerable questions (about information that might have plausibly been in the video) might arise via the acceptance of gist-related information, interrupting rapid responding may have disrupted the automatic acceptance of gist information. Prompting people to locate multiple candidate responses may have drawn attention to the fact that candidates were not always available. Considering the source of information may have encouraged weighing whether candidate responses did come from the video. This is consistent with prior work (with children) which has shown that training in discriminating source can lead to enhanced source monitoring (Poole & Lindsay, 2002; Thierry, Lamb, Pipe, & Spence, 2010).

Practical Application

These results contribute to the literature that suggests that models of interviewing should include procedures to handle the possibility of unanswerable questions. If best practice protocols are followed and interviews begin with free recall and follow-up questions based on the contents of free recall, unanswerable questions are less likely occur. Once questioning deviates from the free recall material, the degree to which questions can and cannot be answered is unknown.

Existing interviewing practices such as *Achieving Best Evidence in Criminal Proceedings* advise that interviewers make clear so-called ground rules such as the option to say "I don't know" and so forth in introductory and questioning phases of interviews. A recent survey of police officers in Scotland suggests that communicating ground rules is considered by most of those sampled to be an important phase of forensic interviews with children. However, only a quarter of respondents stated they always communicate ground rules suggesting interviewers are not aware of their role in reducing demand characteristics (La Rooy et al., 2010).

Existing best practice could be augmented if interviewers and interviewees are aware of the potential implications when they ask questions that the witness cannot possibly answer. One option is for interviewers to effectively transfer control using the procedures described here, which makes it clear that the interviewer does not know what happened and which alerts interviewees to the possibility the questions are based on limited knowledge of the event. Translating these findings into practice will require research in naturalistic settings where practitioners are more likely to be pressed for time and resources when conducting interviews. The brevity of the training may therefore prove to be useful in this regard. The current training procedure presented should prove the most useful in instances in which interviewers find it necessary to ask specific and detailed questions to obtain useful responses from adult witnesses.

It is also important to emphasize that forensic interviewers are not the only group of interviewers who face challenges in using best practice interviewing techniques. For example, the field of medicine has long grappled with similar issues. Diagnostic interviewing and the gathering of medical histories tend to follow a style similar to poor forensic interviews. Research documents frequent use of close-ended questions, problem-focused interviewing, frequent interruptions, and redirections by medical practitioners (Marvel, Epstein, Flowers, & Beckman,

1999). This can result in patients feeling excluded from decision making about their health, and failures to collect relevant information. When medical interviews are conducted empathically and use open-ended prompts, a greater amount of key information is obtained, and diagnostic errors are reduced (Roter & Hall, 1987). A call for 'patient centered' interviews in the medical literature echoes calls for the 'witness centered' interviews. It seems that when people interview other people, they tend to default to asking detailed questions. The degree to which any interviewer asks questions from his or her own frame (e.g., theory of a forensic case, or symptoms that are likely present in a medical case), the more likely it is that unanswerable questions will be asked.

In summary, providing brief training resulted in enhanced rejection of unanswerable questions, and increased the likelihood that substantive responses were in fact correct responses to answerable questions. This likely occurred via increased sensitivity to the availability of information in memory, reduced social pressure, enhanced attention to the intent of questions, and by altering interviewee expectations about responding. This work has potential to augment current best practice interviewing methods by incorporating procedures that are designed to protect against responding to questions than cannot be answered based on what was witnessed.

References

Ackerman, R., & Goldsmith, M. (2008). Control over grain size in memory reporting - with and

without satisficing knowledge. *Journal of Experimental Psychology: Learning, Memory,

and Cognition , 34,* 1224-1245. doi:10.1037/a0012938

Agresti, A. (2002). *Categorical data analysis* (2nd ed.). New York: Wiley.

doi:10.1002/0471249688

Cederborg, A., Orbach, Y., Sternberg, K. J., & Lamb, M. E. (2000). Investigative interviews of

child witnesses in Sweden. *Child Abuse & Neglect, 24*, 1355-1361. doi:10.1016/S0145-

2134(00)00183-6

Chrobak, Q. M., & Zaragoza, M. S. (2008). Inventing stories: Forcing witnesses to fabricate

entire fictitious events leads to freely reported false memories. *Psychonomic Bulletin and

Review, 15,* 1190-1195. doi:10.3758/PBR.15.6.1190

Cyr, M., & Lamb, M. E. (2009). Assessing the effectiveness of the NICHD investigative

interviewing protocol when interviewing French-speaking alleged victims of child sexual

abuse in Quebec. *Child Abuse & Neglect, 33,* 257-268. doi: 10.1016/j.chiabu.2008.04.002

Dando, C. J., Wilcock, R., Milne, R., & Henry, L. (2009). An adapted Cognitive Interview

procedure for front line police investigators. *Applied Cognitive Psychology, 23,* 698-716.

Davies, G., Wilson, C., Mitchell, R., & Milsom, J. (1995). *Videotaping children's evidence: An

evaluation.* London: Home Office.

Dodson, C. S., & Schacter, D. L. (2002). When false recognition meets metacognition: The

distinctiveness heuristic. *Journal of Memory and Language, 46,* 782-803.

doi:10.1006/jmla.2001.2822

Fisher, R. P., & Geiselman, R. E. (1992). *Memory enhancing techniques for investigative*

interviewing: The cognitive interview. Springfield, IL: Charles C. Thomas.

Fisher, R. P., Milne, B. A., & Bull, R. (2011). Interviewing cooperative witnesses. *Current Directions in Psychological Science, 20,* 16-19. doi:10.1177/ 0963721410396826

García-Bajos, E., Migueles, M., & Aizpurua, A. (2012). Bias of script-driven processing on eyewitness memory in young and older adults. *Applied Cognitive Psychology, 26,* 737-745. doi:10.1002/acp.2854

Ghetti, S. (2003). Memory for nonoccurrence: The role of metacognition. *Journal of Memory and Language, 48,* 722–739. doi:10.1016/S0749-596X(03)00005-6

Johnson, M. K., Hashtroudi, S., & Lindsay, S. D. (1993). Source monitoring. *Psychological Bulletin, 114*, 3-28. doi:10.1037/0033-2909.114.1.3

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review, 103,* 490-517. doi:10.1037/0033-295X.103.3.490

Lamb, M. E., Orbach, Y., Hershkowitz, I., Esplin, P., & Horowitz, D. (2007). A structured forensic interview protocol improves the quality and informativeness of investigative interviews with children: A review of research using the NICHD Investigative Interview Protocol. *Child Abuse & Neglect, 31,* 1201-1231. doi:10.1016/j.chiabu.2007.03.021

La Rooy, D., Lamb, M. E., & Memon, A. (2010) Forensic interviews with children in Scotland: A survey of interview practice among police. *Journal of Police and Criminal Psychology, 26,* 26-34. doi: 10.1007/s11896-010-9072-9

Marvel, M. K., Epstein, R. M., Flowers, K., & Beckman, H. B. (1999). Soliciting the patient's agenda: Have we improved? *Journal of the American Medical Association, 20,* 283-297. doi:10.1001/jama.281.3.283

Mazzoni, G., & Kirsch, I. (2001). Autobiographical memories and beliefs: A preliminary

metacognitive model. In T. Perfect & B. Schwartz (Eds.), *Applied metacognition* (pp.

121–145). Cambridge, UK: Cambridge University Press.

Meissner, C. A., & Kassin, S. M. (2002). "He's guilty!": Investigator bias in judgments of truth

and deception. *Law and Human Behavior, 26,* 469-480. doi:10.1023/A:1020278620751

Memon, A., Meissner, C. A., & Fraser, J. (2011). The cognitive interview: A meta-analytic

review and study space analysis of the past 25 years. *Psychology, Public Policy, and*

Law, 16, 340-372. doi:10.1037/a0020518

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

Wiley.

Milne, R., Poyser, S., Williamson, T., & Savage, S. (2010). Miscarriages of justice: What can we

learn? Invited chapter to appear in J. Adler & J. Gray (Eds.), *Forensic psychology:*

Concepts, debates and practice. Cullompton: Willan Publishing.

Mulder, M. R., & Vrij, A. (1996). Explaining conversation rules to children: An intervention to

facilitate children's accurate responses. *Child Abuse & Neglect, 20,* 623–631.

doi:10.1016/0145-2134(96)00050-6

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review*

of General Psychology, 2, 175-220. doi:10.1037/1089-2680.2.2.175

Pansky, A., Goldsmith, M., & Koriat, A. (2009). Memory accuracy in old age: Cognitive,

metacognitive, and neurocognitive determinants. *European Journal of Cognitive*

Psychology, 21, 303-329. doi:10.1080/09541440802281183

Poole, D. A., & Lindsay, D. S. (2002). Reducing child witnesses' false reports of misinformation

from their parents. *Journal of Experimental Child Psychology, 18,* 117-140.

doi:10.1006/jecp.2001.2648

Poole, D. A., & White, L. T. (1991). Effects of question repetition on the eyewitness testimony

of children and adults. *Developmental Psychology, 27,* 975–986. doi:10.1037/0012-

1649.27.6.975

Pratt, C. (1990). On asking children – and adults – bizarre questions. *First Language, 10,* 167–

175.

Roebers, C. M., & Fernandez, O. (2002). The effects of accuracy motivation on children's and

adults' event recall, suggestibility, and their answers to unanswerable questions. *Journal*

of Cognition and Development, 3, 415-443. doi:10.1080/15248372.2002.9669676

Roper, R., & Shewan, D. (2002). Compliance and eyewitness testimony: Do eyewitnesses

comply with misleading 'expert pressure' during investigative interviewing? *Legal and*

Criminological Psychology, 7, 155-163.

Roter, D. K., & Hall, J. A. (1987). Physicians' interviewing styles and medical information

obtained from patients. *Journal of General Internal Medicine, 2,* 325-329.

doi:10.1007/BF02596168

Schreiber Compo, N., Gregory, A.H., & Fisher, R.P. (2012). Interviewing behaviors in police

investigators: A field study of a current U.S. sample. *Psychology, Crime and Law, 18,*

359-375.

Scoboria, A., & Fisico, S. (2013). Encouraging and clarifying 'don't know' responses enhances

interview quality. *Journal of Experimental Psychology: Applied, 19,* 72-82.

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, 255-265. doi:10.1037/1076-898X.14.3.255

Strack, F., & Bless, H. (1994). Memory for non-occurrence: Metacognitive and presuppositional strategies. *Journal of Memory and Language, 33,* 203-217. doi:10.1006/jmla.1994.1010

Thierry, K. L., Lamb, M. E., Pipe, M., & Spence, M. J. (2010). The flexibility of source-monitoring training: Reducing young children's source confusions. *Applied Cognitive Psychology, 24,* 626-644. doi:10.1002/acp.1574

Waterman, A. H., Blades, M., & Spenser, C. P. (2000). Do children try to answer nonsensical questions? *British Journal of Developmental Psychology, 18,* 211-226.

Waterman, A. H., Blades, M., & Spencer, C. (2004). Indicating when you do not know the answer: The effect of question format and interviewer knowledge on children's 'don't know' responses. *British Journal of Developmental Psychology, 22,* 335–348. doi:10.1348/0261510041552710

Waterman, A. H., & Blades, M. (2011). Helping children correctly say 'I don't know' to unanswerable questions. *Journal of Experimental Psychology: Applied, 17,* 396-405. doi:10.1037/a0026150

Weber, N., & Perfect, T. J. (2012). Improving eyewitness identification accuracy by screening out those who say they don't know. *Law and Human Behavior, 36,* 28–36. doi:10.1037/h0093976

Wells, G. L., & Lindsay, R. C. (1980). On estimating the diagnosticity of eyewitness nonidentifications. *Psychological Bulletin, 88,* 776–784. doi:10.1037/0033-2909.88.3.776

Whittlesea, Bruce W. A. (1993). Illusions of Familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 1235–1253.

Zaragoza, M. S. & Lane, S. M. (1994). Sources of misattribution and suggestibility of eye witness testimony. *Journal of Experimental Psychology: Learning, Memory, and*

Cognition, 20, 934–945.

Zaragoza, M. S., & Mitchell, J. (1996). Repeated exposure to suggestion and the creation of

false memories. *Psychological Science*, *7,* 294-300.

Table 1

Study 1, Average Responses by Group and Question type

Question type	Response	Trained		Control		Cohens d [95% CI]
		Mean	*SD*	*Mean*	*SD*	
Answerable	Correct	4.43	1.86	4.60	1.32	.11 [-.32,.53]
	DK	2.84	1.80	2.14	1.30	.44 [-.03,.87]
	Error	2.72	1.74	3.26	1.74	.32 [-.14,.77]
	Accuracy	0.62	0.20	0.60	0.18	.10 [.06,.16]
Unanswerable	Correct rejection	4.86	1.64	3.92	1.55	.60 [.18,1.02]
	DK	3.12	1.83	3.08	1.92	.02 [-.46,.51]
	Error	2.02	1.50	3.00	1.76	.61 [.19,1.04]
	Accuracy	0.72	0.18	0.58	0.21	.73 [.68,.78]

Table 2

Study 2, Average Responses by Group and Question type

Question type	Response	Control		Warned		Cohens d [95% CI]
		Mean	*SD*	*Mean*	*SD*	
Answerable	Correct	5.25	1.57	4.79	1.94	-.21 [-.74,.21]
	DK	2.04	1.36	2.69	1.74	.43 [.01,.85]
	Error	2.71	1.68	2.52	1.81	.11 [-.38,.59]
	Accuracy	0.67	0.19	0.67	0.23	.00 [-.06,.06]
Unanswerable	Correct rejection	5.04	1.54	4.33	1.52	-.47 [-.89,.06]
	DK	1.65	1.60	2.67	1.43	.68 [.27,1.10]
	Error	3.31	1.64	3.00	1.14	.22 [-.16,.61]
	Accuracy	0.61	0.17	0.59	0.15	-.13 [-.17,.08]

.

Appendix A

Cues Provided During the Training Procedure

Review: "Can I respond completely to the question based upon information available to me?"

Retrieve: "What pieces of information are available to me to answer the question?"

References: "From what source(s) am I getting this information?"

Reflect: "How certain am I that each piece of information is correct?"

Reply: "What is the best answer that I can make?

Participants in the trained group received a brief training. The cues were introduced and explained, and participants practiced using the cues to answer two simple, answerable questions. The interviewer answered any questions, following which the main questioning commenced. The interviewer provided the cues on a card for trained participants to use during questioning.