2010

The Cognitive Interview: A meta-analytic review and study space analysis of the past 25 years

Amina Memon, *Royal Holloway College, University of London*
Christian A Meissner, *University of Texas at El Paso*
Joanne Fraser, *University of Portsmouth*

Psychology, Public Policy, and Law
2010, Vol. 16, No. 4, 340–372

1076-8971/10/$12.00 DOI: 10.1037/a0020518

THE COGNITIVE INTERVIEW: A Meta-Analytic Review and Study Space Analysis of the Past 25 Years

Amina Memon
Royal Holloway College, University of London

Christian A. Meissner
University of Texas at El Paso

Joanne Fraser
University of Portsmouth

The Cognitive Interview (CI) is a well-established protocol for interviewing witnesses. The current article presents a study space analysis of laboratory studies of the CI together with an empirical meta-analysis summarizing the past 25 years of research. The study space comprises 57 published articles (65 experiments) on the CI, providing an assessment of the boundary conditions underlying the analysis and application of this interview protocol. The current meta-analysis includes 46 published articles, including 20 articles published since the last meta-analysis conducted a decade earlier (Köhnken, Milne, Memon, & Bull, 1999). Reassuringly for practitioners, the findings of the original meta-analysis were replicated with a large and significant increase in correct details and a small increase in errors. In addition we found that there were no differences in the rate at which details are confabulated. Importantly, the effect sizes were unaffected by the inclusion of recent studies using modified versions of the CI. The CI appeared to benefit older adult witnesses even more than younger adults. We highlight trends and gaps in research and discuss how our findings can inform policy and training decisions.

Keywords: Cognitive interview, review, meta-analysis, eyewitness memory

The Cognitive Interview (or CI) is perhaps one of the most successful developments in psychology and law research in the last 25 years. It is a method that comprises a series of memory retrieval and communication techniques designed to increase the amount of information that can be obtained from an interviewee. The CI was initially developed 25 years ago by psychologists Ed Geiselman and Ron Fisher as a response to the many requests they received from police officers and legal professionals for a method of improving witness interviews. It is based upon established psychological principles of remembering and retrieval of information from memory, and empirical laboratory research on the CI has documented its ability to dramatically improve the number of correct details while only slightly increasing the number of incorrect details (Schrieber & Fisher, 2006; Köhnken et al., 1999; Memon, 2006). Field tests of the CI have also indicated that police officers trained in its techniques gain more information and

Amina Memon, Department of Psychology, Royal Holloway College, University of London; Christian A. Meissner, Department of Psychology, University of Texas at El Paso; Joanne Fraser, Department of Psychology, University of Portsmouth.

This research was funded by a grant from the European Union awarded to the first author.

Correspondence concerning this article should be addressed to Amina Memon, Department of Psychology, Royal Holloway College, University of London, Egham, Surrey TW20 0EX, England. E-mail: Amina.Memon@rhul.ac.uk

more detailed information from eyewitnesses in investigative contexts (Fisher, Geiselman & Amador, 1989; Clifford & George, 1996; Kebbell & Milne, 1998, cf. Fisher & Schrieber, 2007). The CI is also useful in other contexts where accurate information gathering is the goal, such as in the investigation of accidents and near-miss events in organizations (see Flin, O'Connor, & Crichton, 2008). Furthermore, Fisher, Falkner, Trevisan, and McCauley (2000) used an adapted CI to elicit accurate information from survey respondents about their physical activities 35 years earlier. Köhnken (1995), in a review of the information processing approach to interviewing, highlights other uses of the CI ranging from interviewing children and adults during competency assessments and custody disputes, to obtaining information about present moods, attitudes, and opinions of respondents. The potential of the CI has not as yet been fully explored in all these domains. Finally, there is the benefit that interviewers' memories will be enhanced too with the CI (Köhnken, Thurer, & Zorberbier, 1994).

Over the past 25 years some 65 studies have been published on the CI. The current article provides an up-to-date review of the literature using both meta-analysis and study space analysis, the latter being an approach advocated by Malpass and colleagues (2008) intended to supplement a meta-analytic approach. A study space analysis allows us to identify the breadth and adequacy of an empirical literature base and to assess trends and gaps in the area that individual researchers might otherwise not see. We will use the study space to provide an in-depth review of the study attributes or variables that have been investigated in the published literature on the CI. We will also highlight areas that warrant further investigation. A mechanism that can evaluate the adequacy of the research and its scope can provide policy makers with information about the suitability and applicability of the research. Currently, one of the problems with applying research on the CI is that practitioners are reluctant to use the techniques either because they are insufficiently trained in its use or because they have concerns about the efficacy of some of the techniques (see Wells, Memon, & Penrod, 2006, for a recent review). A study space analysis may provide those responsible for the provision of interview training with useful information about the research base supporting the CI, as well as provide justification for devoting resources to such training.

What Is the Cognitive Interview?

In line with Tulving's (1983) notion that memory is a joint product of stored memory traces and cues that are available at retrieval, the CI engages the witness in a detailed retrieval of the original event. The original CI (Geiselman et al., 1984) was comprised of four techniques designed to enhance participants' recall of a prior event. The first technique involves *context reinstatement*, in which the interviewee is encouraged to mentally reconstruct the physical and personal context that existed at the time of the event. The second technique is to ask participants to *report everything* they can recall even if it is partial or incomplete. The third method is based on the premise that different retrieval cues may access different aspects of a complex event (Anderson & Pichert, 1978). Witnesses are instructed to recall from a *variety of perspectives*—from their own perspective and to adopt the perspective of others. Finally, witnesses engage in further retrieval

attempts in a different *temporal order*—from the start, from the end working backwards in time, the middle or any other point in time that may be salient to the individual (see Memon, 2006).

In 1992, Fisher and Geiselman published an *enhanced* version of the CI (ECI) that included a framework for building rapport and communicating effectively with the witness. Throughout the interview process, the interviewer is discouraged from interrupting the witness, and is instructed to allow the witness to control the flow of information and to listen actively to what the witness has to say. This witness-centered interview procedure is a major characteristic of the enhanced CI. The interviewer facilitates this process by use of open-ended questions about neutral topics. The next phase of the interview involves context reinstatement followed by the interviewee's free narrative account of the incident. The interviewer reminds them at this point of the importance of providing a detailed account (report everything) and requests that they do not guess or fabricate, but simply tell the interviewer if they don't know. Research had shown that witnesses are more likely to maintain high accuracy if they are reminded not to guess (Koriat & Goldsmith, 1996). Following the free narrative, the interviewer questions the witness about details provided, facilitated by the use of *focused* memory techniques, which involve instructing the witness to concentrate on mental images of the various parts of the event such as the suspect's face and using these images to guide recall. An important principle of the ECI is that event details will be most accessible when they are perceptually related to the witness's image and thus interviewers should time their questions accordingly (Fisher & Schreiber, 2007).

Köhnken et al. (1999) Meta-Analysis

A meta-analysis of research on the CI and ECI was published a decade ago (Köhnken et al., 1999). It included 42 studies (29 of which were published) and 55 individual comparisons of the CI to a control interview. The meta-analysis examined different methodological variables across studies such as the control interview used (standard untrained vs. structured interview), medium of event presentation (staged vs. video), age of the interviewee (adults vs. children), and witness involvement (passive viewing vs. active participation in the event). Köhnken et al. reported a large overall effect size for the increase in correctly recalled details generated by the CI ($d = 0.87$). The overall effect size for the increase in incorrect details, although considerably smaller, was also significant with more incorrect details reported in the CI ($d = 0.28$). In terms of the various methodological variables and moderators, the authors found that effect sizes were larger for live events (as compared to video) and if the interviewees actively participated in the event. No differences in effect size were observed as a function of the age of the participant. The authors did find a decrease in effect size for correct details as the delay between the event and the interview increased; however, there were few studies that actually manipulated delay and the average delay was just 2 days. There were no significant differences in effect sizes for the OCI and ECI. Also contrary to expectation the effect of the CI did not decrease when trained vs. untrained interviewers were used for the control condition. Finally, with respect to incorrect recall, there was a larger increase in incorrect

recall for adults as compared to children, and a larger effect of the ECI on incorrect details as compared to the original version.

Another Decade of Research on the CI

Since the 1999 meta-analysis another version of the CI has become increasingly popular—namely, the *modified* CI (MCI) which is an adaption of the ECI. For example, Holliday (2003a, 2003b) has modified the CI so that it is suitable for use with young children (4 to 9 year olds). In her version, which follows the ECI procedure of building rapport establishing ground rules and transferring control, the change perspective instruction is removed. Davis, McMahon, and Greenwood (2005) present another version of the MCI involving a shortened variation of the interview in which the change order and change perspective techniques are omitted and *replaced* with an additional prompt to go through the event once more in chronological order. The fact that some researchers have come up with their own versions of the MCI is potentially problematic. The previous meta-analysis was based on the original and enhanced versions only—thus, we do not know if we can continue to make recommendations for policy and practice based on the 1999 meta-analysis when the procedure that is typically tested in laboratory studies has changed.

There has also been a corresponding change in the control or comparison group. Whereas early studies tended to compare the CI with an untrained control group, an increasing number of recent studies have used what is typically referred to as a "Structured Interview," which in some studies is either based on, or closely resembles, nationally agreed guides to interviewing such as the British Achieving Best Evidence, 2001 (formerly the Home Office Memorandum). Whether a Cognitive Interview has any benefits over a structured interview based on a nationally approved protocol is an important question for policy makers. In coding the variables included in the study space (and meta-analysis) we paid close attention to both the type of Cognitive Interview and the control group. Finally, the fact that an increasing number of studies have used children and older adults as interviewees is noteworthy. Once again, this issue is an important one for policy makers in the U.K. who in recent years have made efforts to ensure vulnerable witnesses (which include children and adolescents) can give evidence in criminal proceedings (for example, the Youth Justice & Criminal Evidence Act, 1999). These developments motivate a re-examination of the effect sizes including the new studies.

Overview of the Current Analytic Approach

In the current article we combine the use of meta-analytic and study space methods to provide researchers and practitioners with data on the efficacy and robustness of the CI. The study space analysis provides a more exhaustive review that includes studies which failed to meet the strict inclusion criteria for our meta-analysis, while the meta-analysis provides a much needed update on the statistical effectiveness of the CI, including an expanded analysis of possible moderator variables. Together, these analytic approaches will assess the efficacy and robustness of the CI based on the contemporary literature, and seek to identify gaps in our knowledge of the conditions under which the CI is most effective. We

will ask whether researchers have sufficiently explored variables that are critical to determining when and where generalizations to field settings are warranted. The consequences for policy and practice will also be discussed. We present the study space and meta-analysis in succession with a combined general discussion.

Study Space Analysis

According to Malpass and colleagues (2008), the study space concept relies on the identification of elements and sub-elements of studies that assess a particular topic. These in turn are defined by the intersections of the levels of study attributes, namely the independent and dependent variables as well as any methodological and procedural strategies used across the studies. Malpass et al. maintain that "examining the study space using the variables, methods and procedures present in an existing literature can assist in identifying regions of concentration and inattention, alerting investigators to territories that have been well worked over and to others where new contributions can be made" (p. 794). Where researchers have made recommendations for training and public policy on the basis of empirical research on eyewitness testimony, it is essential not to exaggerate knowledge (Turtle, Read, Lindsay, & Brimacombe, 2008). As pointed out by Malpass et al. (2008) the size or consistency of the literature becomes redundant if important dimensions of the study space are unrepresented in published work.

Methods

Studies. The studies that were considered for inclusion in the study space analysis were primarily obtained via searches of on-line databases. The cognitive interview study space and meta-analysis was part of a larger meta-analysis on eyewitness descriptions, so both general searches for papers on eyewitness recall and a specific search for CI articles was undertaken. The two main databases used were *PsycARTICLES and PsycINFO*. The databases were searched using the key words: "cognitive interview," "interview," "eyewitness," "testimony," "memory," "memory event," "recall," "cued recall," "episodic memory," "accuracy," "suggestibility," "age," "crime," "mock crime," "memory distortions," "person description," "emotion," "emotional," "race," and "alcohol." For specific articles on the CI, searches were made using the names of authors who had previously published articles on the CI. Review articles on the CI were also assessed for additional references. In addition to on-line searches, researchers in the field were contacted via obtaining email lists from professional bodies (e.g., the Society for Applied Research in Memory and Cognition, the American Psychology-Law Society, etc.). A request was made for in press and published papers on the variables influencing eyewitness memory, with a focus on studies that contain measures of eyewitness recall. Both senior and junior authors of the published papers available to us were also contacted to further request any papers that may have been missed in the search. A total of 57 published articles (65 experiments) that empirically assessed the effectiveness of the CI were located based upon this search process.

Inclusion/exclusion criteria. To be included in the study space, studies must have conducted an experimental analysis of the cognitive interview in

comparison to a control or other interview protocol. In addition, the research had to be published or accepted for publication in a peer-reviewed journal. Legal standards for proffered scientific testimony in the United States and other countries stress the importance of conducting a review of the literature based on well conceived, well executed and retrievable studies (see *Daubert vs. Merrell Dow Pharmaceutical,* 1993). Moreover, one of the problems with alternative sources (e.g., conference papers, unpublished manuscripts available online, etc.) is that they frequently do not provide the data needed to conduct an appropriate study space or meta-analysis. A full set of study descriptors, including independent and dependent variables, as well as a host of methodological characteristics of each study was compiled for each study and these can be obtained from the following website: http://www.pc.rhul.ac.uk/sites/rheg/.

Coding of studies. For each study the independent and dependent variables were identified and were listed into an individual matrix. Taking the Akehurst, Milne, and Köhnken (2003) article as an example, the independent variables included type of interview (enhanced CI vs. structured interview), retention interval (four hours vs. six weeks), and age of witness (8–9 vs. 11–12 year olds). The dependent variables included total correct recall, total incorrect recall (e.g., describing a red coat as blue), and total confabulations (saying there was a coat when there was no coat), and each of these were split into type of detail (i.e., action, person, and object details). In addition, we identified the cross-study or methodological variables (i.e., those variables that are controlled and held constant in a given study, but may vary across studies such as population, type of target event, interviewing condition, etc). Each of these variables was added to the individual matrix together with the corresponding independent and dependent variables for that study. As such, an individual matrix was created for every study included in the study space, with the matrices subsequently merged to reflect the total sample of studies. This cross-study matrix included all the independent, dependent, and cross-study variables noted in each individual study, and involved a frequency count of the number of studies falling into each category intersection. The cross-study matrix can be downloaded from the following website: http://www.pc.rhul.ac.uk/sites/rheg/.

Results

Table 1 provides an overview of data gathered in the study space. It shows the number of studies (and percentages) classified by type of interview (OCI, ECI, or MCI). The first two sections display the number of studies as a function of the background of the witnesses and interviewers, followed by studies distinguished by the age of the witness. The frequency with which different control groups have been used in the OCI, ECI, and MCI studies is also shown. Finally some study variables are included such as retention interval, event duration and event type and mode of presentation (live versus video). We have set the cells in bold that contain fewer observations than might be expected by an even distribution of the study space to denote areas that have been understudied by researchers. We turn now to a discussion of these areas and their implications for generalizability and application of the CI.

Table 1
Frequency of CI Studies (and Percentages) by Study Characteristics and Test Variables

	Original CI	Enhanced CI	Modified CI
Witness population			
Students/pupils	15 (23%)	8 (12%)	23 (35%)
Civilians	**6 (9%)**	**6 (9%)**	**5 (8%)**
Police	**0 (0%)**	**1 (2%)**	**1 (2%)**
Witness age			
Children	**3 (5%)**	**2 (3%)**	13 (20%)
Young adults	17 (26%)	12 (18%)	13 (20%)
Older adults	**1 (2%)**	**1 (2%)**	**3 (5%)**
Interviewer background			
Students	**1 (2%)**	**3 (5%)**	7 (11%)
Researchers	16 (25%)	9 (14%)	19 (29%)
Professionals	**2 (3%)**	**3 (5%)**	**3 (5%)**
Written script	**2 (3%)**	**0 (0%)**	**0 (0%)**
Control group			
Standard interview	15 (23%)	8 (12%)	6 (9%)
Structured interview	**4 (6%)**	6 (9%)	20 (31%)
Free recall	**0 (0%)**	**0 (0%)**	**2 (3%)**
No control group	2 (3%)	1 (2%)	1 (2%)
Event duration			
<1 min	**4 (6%)**	**1 (2%)**	**2 (3%)**
<5 min	11 (17%)	5 (8%)	14 (22%)
5–10 min	**2 (3%)**	5 (8%)	9 (14%)
>10 min	**4 (6%)**	**4 (6%)**	**4 (6%)**
Retention interval			
None or very brief	5 (8%)	5 (8%)	10 (16%)
24–72 hrs	14 (22%)	4 (6%)	13 (20%)
1–2 weeks	**2 (3%)**	5 (8%)	4 (6%)
2–6 weeks	**0 (0%)**	**0 (0%)**	**1 (2%)**
1.5–6 months	**0 (0%)**	**0 (0%)**	**1 (2%)**
Event medium			
Staged (live)	6 (9%)	**5 (8%)**	**4 (6%)**
Video	14 (21%)	10 (15%)	24 (40%)
Event type			
Emotional/arousing event	17 (26%)	11 (17%)	13 (20%)
Neutral event	**4 (6%)**	**4 (6%)**	16 (25%)

Note. Cells that are set in bold represent percentages below that expected if studies were evenly distributed across regions of the study space.

Who are the witnesses and interviewers? One of the first questions that a policy maker may ask is whether research on the CI is based on representative sample of witnesses, and whether the effects can be generalized beyond the typical participant (young adult, college educated) in laboratory studies. As indicated by the "bold" areas of Table 1, young adults drawn from college populations are overly represented in the CI studies with the exception that children are well represented in more recent studies of the MCI. Of the 65 published experiments included in the current sample, 42 studies (or 64%) used young adult witnesses. The sample of studies also included 28% that involved younger (pre-school) or older children, while 8% used older adults and 6% used

special populations (learning disabled). The interviews were conducted by researchers (68%) or students (17%), the latter were overly represented in recent studies on the MCI. Only 12% of studies included professional law enforcement as interviewers. The restricted use of professionally relevant samples as interviewers is clearly a limitation when it comes to the question of generalization to field settings. Several important lessons have been learnt from the few studies that have included police samples and from surveys of police officers. We know that training police officers to change the techniques they normally use is far more challenging than training researchers to adopt new ones (Memon, Bull, & Smith, 1995; Memon, Milne, Holley, Bull, & Köhnken, 1994, see also Fisher, 2010). When questioned about their use of the CI in surveys British police officers state they use some of the individual CI components such as the "report everything" instruction while some techniques (such as the "change perspective" and "recall in reverse order") are seldom used (e.g., Dando, Wilcock, & Milne, 2008; Dando, Wilcock, & Milne 2009c; Kebbell & Milne, 1998; Kebbell, Milne, & Wagstaff, 1999; Wright & Holliday, 2005). It is only in recent years that efforts have been directed towards developing an adapted version of the CI that addresses the basic training needs of police officers (see Dando, Wilcock, Milne, & Henry, 2009a). Researchers have also developed a tool which police officers can give witnesses so they can self-administer the cognitive interview (Gabbert, Hope, & Fisher, 2009). Field tests of the SAI are currently underway (Gabbert, Hope, & Jamieson, 2010). Thus while the research on the CI has largely relied on student interviewers till now, researchers are now working more closely with the police to ensure that the CI is implemented. We will elaborate on this further and consider the implications for policy and practice in the general discussion.

The type of Cognitive Interview. Analysis of the type of CI used suggests that there have been substantial deviations from the original and enhanced interviews in recent years. Early studies (through the '80s and early '90s) examining the CI consistently used the original version (OCI)—however, only 32% of the studies in the current analysis used the OCI. Following the publication of the Fisher and Geiselman (1992) text on the CI, there was a move to the ECI (23% of studies), but in the last 10 years various modified and shortened versions of the CI have emerged in the literature (45% of studies).

There have been good reasons for modifying the CI. First, one of the purposes of modification is to adapt the CI to meet the individual needs of the witness, with vulnerable witnesses (e.g., children, elderly, or mentally disabled) providing a good example. Consistent with this hypothesis, 45% of the MCI studies have used child witnesses and 10% older adult witnesses. Four studies have directly compared the ECI and MCI, reporting similar increases in correct details with each version of the CI (Wright & Holliday, 2007; Dando et al., 2009a; Davis, McMahon, & Greenwood, 2005; Mello & Fisher, 1996). Here we briefly discuss a sample of these studies to illustrate how different versions of the MCI have evolved over the last 15 years.

One of the earliest studies to use a modified version of the CI was Saywitz, Geiselman, and Bornstein (1992) who adapted the CI so it could be used with children (ages 7–12 years). Saywitz et al. (1992) modified the CI to ensure that the children were aware they could use the "I don't know" response. They also modified the wording of the change perspective instruction so that the children

could understand the instruction by using the phrasing "*Put yourself in the body of _____ and tell me what that person saw*." In another study using an MCI adapted for child witnesses (ages 8–9 years), Memon, Holley, Wark, Bull, and Köhnken (1996) omitted both change perspective and order techniques. In each study, these MCI versions were found to be effective with children. Several other studies have followed suit (Ginet & Verkampt, 2007, with young adults; Allwood, Ask, & Granhag, 2005, college students; Searcy, Bartlett, Memon, & Swanson, 2001, older adults). Table 2 shows the frequency of subcomponents of the CI that have been retained or removed across the various MCI variations based on the 29 studies included in this study space analysis. We present what appear to be the three main versions of the MCI based on descriptions of the interview procedure in research articles plus an "other" category crossed with population (child versus adult). It is clear from this analysis that there are numerous variations of the MCI.

The control or comparison group. An inspection of the study space shows that whereas early studies of the CI tended to use a standard interview as a comparison (21%), more recent studies (69%) refer to a structured interview control. Interestingly, early OCI studies (e.g., Geiselman, Fisher, MacKinnon, & Holland, 1985; Geiselman et al., 1984) appeared to lack a commonly agreed definition of a standard interview except that such a condition denoted the absence of any training of interviewers. For instance, Chapman and Perry (1995) simply

Table 2
Number of CI Studies Using Different Versions of the Modified Cognitive Interview and the Percentage Use of Each of the Techniques Mentioned in the Published Papers

	Version 1	Version 2	Version 3	Other versions
# of studies employed	10	5	3	11
Establish rapport	+	+	+	72%
Establish ground rules*	+	+		45%
Transfer control	+		+	50%
Concentrate	+			27%
Report everything/detail	+	+	+	100%
Context reinstatement	+	+	+	100%
Sketch/draw				18%
Free recall	+	+	+	72%
Prompt ("Can you tell me more")		+		27%
Remind not to guess		+		9%
Open questions (based on free recall)	+	+	+	45%
Open questions (predetermined)				50%
Generate and probe images		+		27%
Cued recall	+	+		64%
Additional retrieval attempts	+			36%
Change temporal order	+	+		27%
Change perspectives				27%

Note. These techniques are approximately in the order in which they are used but there is considerable variability across the studies in the order and some studies do not provide sufficient information.

* Ground rules vary in the extent to which they are used. Some researchers imply that they tell the witness they must not guess, some remind them of this prior to the questioning. Others also encourage witnesses to say "I don't know" and "I don't understand."

indicated that a standard interview involved the interviewer asking the witness to give an account in their own words followed by specific questions intended to obtain more information. Similarly, Geiselman and colleagues (1985) (using a law enforcement sample) instructed the standard interview group to use the questioning procedures they would normally use. Other studies have used professional interviewers where the standard interview was described as one in which interviewers were told to use "questioning procedures during the interviews they would normally use with children" (e.g., Saywitz, Geiselman, & Bornstein, 1992, p. 746).

The structured interview typically follows an identical format to the CI in that the interview begins with open-ended questions and a free narrative, and only after free recall has been exhausted are specific questions asked. It also generally includes all of the techniques that have to do with building rapport and communicating effectively with the witness. It was Köhnken and colleagues (e.g., Köhnken, Schimmossek, Aschermann, & Höfer, 1995) who first introduced the structured interview as a control and this coincides with the bold areas in Table 1 and accounts for why the structured interview is represented mostly in the ECI and MCI studies. The increase in the number of studies using a structured interview control suggests that researchers are aware of the importance of comparing the CI with a procedure that holds constant both interview structure and communication strategy. It can also provide an important test of the current policy by showing that any additional gains in information with the CI (over and above any nationally approved protocol) would improve upon current practice. Unfortunately only a few studies make reference to any national guidance in their description of the Structured Interview. Many of the British researchers (e.g., Akehurst et al., 2003; Memon, Wark, Bull, & Köhnken, 1997a; Wright & Holliday, 2007; Holliday, 2003a, 2003b) referred to national guidance (e.g., Achieving Best Evidence, Home Office & Department of Health, 2001) when describing the structured interview. However researchers from other European countries (e.g., Mantwill, Köhnken, & Aschermann, 1995; Larsson, Granhag, & Spjut, 2003) and the United States (e.g., Mello & Fisher, 1996) tended not to specify whether or not their version of the structured interview was based on any nationally approved guidance. A couple of exceptions are an early study reported by Brock, Fisher, and Cutler (1999) where they refer to a "standard" interview based on the protocol used by the National Transportation Safety Board. It is interesting to note that the structured interview resembles the Step-Wise Interview (Yuille, Hunter, Joffre, & Zaparniuk, 1993) and National Institute of Child Health and Development (NICHD) protocol (e.g., Sternberg, Lamb, Esplin, Orbach, & Hershkowitz, 2002) although the two have never been directly compared.

The type of event: witness involvement and arousal. Following the Köhnken meta-analysis, we coded for witness involvement namely whether the witness participated in a staged event or passively viewed a video-taped scenario. We excluded two studies from the percentages presented in Table 1 on the basis that they had used a slide and narrative as the to-be remembered event. The majority (74%) of studies have tended to use video to present events with 83% of the more recent MCI studies falling into this category. As indicated by the bold areas within Table 1, research on the potential of the MCI when the witness is recalling events in which they are involved and live events is lacking.

We also coded for event type based upon the description of the event scenario provided by the researchers. Any incidents which were made up of a crime or accident scenario was coded as "emotional" or "arousing" and any other scenarios were coded as "neutral." The main purpose of this classification was to see whether the CI generalizes across different types of events. A meta-analysis of the effects of heightened stress indicates that it negatively impacts eyewitness recall of details of a crime, as well as identification of a perpetrator or target person (Deffenbacher, Bornstein, Penrod, & McGorty, 2004). The majority of CI studies (64%) have attempted to use emotionally arousing scenarios but again only 45% of the more recent MCI studies fell into this category. Moreover, only one study has systematically manipulated the degree of arousal experienced by the witness. Ginet and Verkampt (2007) showed college students a video of an accident. Some participants were led to believe they would receive electric shocks during the video (high arousal); while others were told electrodes were being attached to measure physiological signs (low arousal). The authors compared an MCI with a Structured Interview and found that the MCI elicited more correct central and peripheral details regardless of the level of arousal. A manipulation check indicated that participants did report feeling more aroused and threatened in the electric shock condition. The Ginet and Verkampt study was conducted in France and presumably the authors met ethical standards for conducting research. Ethical issues, however, may dissuade other authors from pursuing such studies. A possible solution is to consider the use of the CI to test memory for situations in which interviewees will experience arousal during the normal course of their work (see Morgan et al., 2004, for an example) or during training sessions involving threatening encounters (Hulse & Memon, 2006). We would urge researchers to conduct more studies to examine the potential of the CI as an information gathering tool in situations where the witnesses or interviewees have been subjected to stressful situations (see Valentine & Mesout, 2010, for one example).

In addition to the practical importance of demonstrating the efficacy of the CI under situations of high stress or arousal, it would be theoretically important to see what effect use of the CI might have on the retrieval of emotional memories. Two theoretical perspectives on the effects of emotion on memory predict that emotion will improve the recall of central details, but at a cost to peripheral details. One perspective involves the Easterbrook cue-utilization hypothesis (Easterbrook, 1959), which suggests that high arousal narrows the focus of attention so that central details that are attended to are recalled (see Christianson, 1992). Similarly, arousal is thought to elicit consciously-controlled processing regarding the cause of the arousal (e.g., a weapon). Moreover, Christianson (1992) speculated that effortful elaboration after an arousing experience (or post-stimulus elaboration) results in a focus on the central actions and events in a scenario. It has been suggested that post-stimulus elaboration elicited by the sight of weapons or gruesome injuries might consist of reliving and evaluating the critical moment (Christianson & Lindholm, 1998; Hulse, Allan, Memon, & Read, 2007), which can in turn enhance recall. This raises an interesting question as to whether the CI can improve recall even further and the potential of the CI as way of eliciting detailed recall of traumatic events has not as yet been systematically explored.

The effectiveness of the CI across long delays and multiple interviews. The effects of multiple interviews and long delays on accuracy are an important practical question for policy makers because witnesses (and in particular children) are often interviewed repeatedly during the course of a criminal investigation (Goodman & Quas, 2008; La Rooy, Lamb, & Pipe, 2009) and long delays between the interviews are not unusual. The previous CI meta-analysis by Köhnken et al. (1999) noted that relatively few studies at the time had employed delays extending beyond 48 hours. This situation has changed little—the current sample of studies suggests that only 17% of studies have employed a delay of 1–2 weeks, with only 3% utilizing a delay of over 2 weeks. The majority (48%) of studies incorporate a delay period between 24–72 hours, and 31% have either no delay or a very brief (minutes) delay as illustrated by the "grey" areas within Table 1. This is unfortunate because the moderating effects of delay have not been investigated systematically within the eyewitness literature (see Dysart & Lindsay, 2007) despite the fact that it is a key variable (see Deffenbacher, Bornstein, McGorty, & Penrod, 2008).

While it is known that interviews occurring relatively soon after an event can serve as a buffer against forgetting (Goodman et al., 1992; La Rooy, Pipe, & Murray, 2005), delay can cause errors to increase over time particularly when biased questioning procedures are used (Goodman & Quas, 2008). Given the importance of the two variables (delay and repeated testing) in real world contexts (see Fisher, Brewer, & Mitchell, 2009), it is surprising that these variables have been almost entirely overlooked by researchers studying the CI. Only three studies in the current sample have examined the effects of repeated recall using the CI in the initial and subsequent interview (Brock, Fisher, & Cutler, 1999; McCauley & Fisher, 1995; Memon, Wark, Bull, & Köhnken, 1997a). All three of the repeated retrieval studies involved delays ranging from 5 min to 2 days for the initial interview, and 10 to 14 days for the follow-up interview. The type of interviews conducted limit certain comparisons that might be made across the studies—for example, McCauley and Fisher used an ECI and standard interview control, while Memon et al. used an MCI and structured interview control The overall findings in each study was an advantage in terms of correct details at first interview when a CI was used as compared to a control but no apparent benefit of having two Cognitive Interviews (Time 1 and Time 2) and no carry over effects. A lack of research on the effects of repeated testing with a CI and carryover effects of an early CI on subsequent recall weakens the generalizability of the CI and more research is needed on this issue.

Conclusions regarding the study space analysis. In sum, the present study space analysis identified several shortcomings in the CI literature, gaps that have both theoretical and practical significance. Most of the studies to date have used college populations and researchers as interviewers. While studies have employed both crime and neutral scenarios, an increasing number of studies now rely on videotaped scenarios. Moreover, only one study has directly manipulated the presence of arousal and no studies have as yet compared the CI with a control procedure in obtaining details from people who experience stress and arousal in real life contexts. Furthermore, studies have typically included short delays ranging from 48 hours to a week, and only three studies have examined the effect of a repeated CI. This raises some concerns about generalizing the current body

of research to relevant forensic field settings. In addition to concerns about ecological validity, several theoretical questions about the CI remain unanswered: When the witness has been subject to multiple interviews does the CI elicit new details not originally recalled by the witness? Does the CI rely on a strong memory trace for the original event to be effective? Does emotional arousal reduce the effectiveness of the CI? Do older adults benefit from the CI as much as younger adults and children? Finally, in terms of impact on policy and practice, researchers (Dando et al., 2009b; Dando, Wilcock, & Milne, 2009c; Dando, Wilcock, Behnkle, & Milne, in press) are currently exploring the potential of a CI that has been modified so it can be efficiently administered by frontline police investigators. We return to this issue in the general discussion.

Meta-Analysis

The aim of the current meta-analysis was to update our understanding of the statistical effect of the CI on eyewitness recall, including the analysis of potential moderating variables across samples. Since the first meta-analysis of the CI was published in 1999, 20 additional studies that meet our inclusion criteria have been published. Many of these studies use a version of the MCI, and a significant number of new studies have utilized non-student samples. In addition, more studies have begun to examine the impact of the CI on confabulated recall, a dependant measure that we include in the present analysis.

Methods

Studies. A total of 59 independent effect sizes described in 46 research articles were included in the meta-analysis, representing the responses of 2,887 subjects. These research articles were primarily obtained via searches of on-line databases. The articles included in the meta-analysis were a subset of those published papers used in the study space involving the selection procedure described earlier.

Inclusion and exclusion criteria. Criteria for including studies in the final sample were that: (i) studies must have been published in a peer-reviewed journal; (ii) studies must have required participants to provide verbal recall of an event or a verbal description of a person; (iii) a cognitive interview (either original, enhanced, or modified version) was conducted; (iv) the control interview was either a standard interview, a structured-interview, or a free recall task; and (v) dependent measures of recall (correct, incorrect, and/or confabulated) were provided in a manner that permitted the computation of an appropriate effect size comparing the CI and control interviews.

For the reasons specified previously, we only included published articles in the meta-analysis (see Appendix A for a listing of the articles). Details of articles that were excluded from the meta-analysis were recorded and set aside for the study space or the general discussion. Examples of studies which were excluded because they did not fit our criteria included those failing to include a control group (e.g., Fisher, Geiselman, Raymond, Jurkevich, & Warhaftig, 1987b; Memon, Holley, Wark, Bull, & Köhnken, 1996) and those that provided insufficient technical or statistical information (e.g., Holliday, 2003a; Searcy et al., 2001). Studies that only examined one CI component in isolation (e.g., context

reinstatement, report in detail) were also excluded (e.g., Dietze & Thomson, 1993; Dando, Wilcock, & Milne, 2009b; Memon, Cronin, Eaves, & Bull, 1996; Milne & Bull, 2002). Field research studies were excluded due to the difficulty of determining ground truth in measures of accuracy (e.g., Clifford & George, 1996; Fisher et al., 1989); nevertheless, the importance of obtaining data from the field is considered in the general discussion.

Estimate of effect size and meta-analytic approach. Our primary measure of effect size was Cohen's *d*, consistent with the previous meta-analysis (Köhnken et al., 1999). Cohen's *d* effect size was computed from *M*s and *SD*s, *F*-tests ($df = 1$), or *t*-tests reported in each published article. In some cases, the effect size computed based upon a study's reported results failed to match that reported by Köhnken et al. (1999)—all computations, however, were double-checked and authors were contacted for detailed statistical information when necessary. Finally, some authors simply reported that a specific effect was "not significant." If a directional effect could be determined in such instances (based upon the observed means) a positive or negative $d = .01$ was assigned as appropriate, whereas when no directional effect could be determined a $d = .00$ was assigned. Such instances constituted only 4% of effect sizes in the current sample. Appendix A provides the estimates of effect size calculated for each experimental comparison across studies for each measures of recall.

Our meta-analysis involved estimating the mean weighted effect size for the sample of studies, followed by prediction of effect size based upon moderating variables (Hedges & Olkin, 1985; see Johnson, Mullen, & Salas, 1995, for a discussion of various approaches). We examined the impact of the CI across recall measures of correct details (total number of correctly recalled details), incorrect details (errors in reporting detail; e.g., describing a coat as black when it was red), and confabulated details (a commission error; e.g., describing a coat when there was no coat). Moderator variables, discussed below, were coded and used to predict the variance in effect size across samples via a weighted least squares regression analysis.

Coding of study characteristics as moderator variables. Articles were assessed by two independent coders, including the senior author who is an experienced researcher in the CI field. An initial screening was performed on all of the articles selected according to the search and inclusion criteria noted above, and to assess whether each article provided the relevant statistical information. Moreover, the studies were coded on several variables of interest, including: age of the participants (children vs. young adults vs. older adults), medium of event presentation (live vs. video/slide/narrative), type of event (crime/accident scenario vs. neutral scenario), retention interval between the presentation of the event and the interview (hours), type of cognitive interview that was employed (standard vs. enhanced vs. modified), and type of control interview that was employed (standard vs. structured). Any discrepancies in the coding of these variables were resolved by the third author.

Results and Discussion

Effect size analysis. Table 3 provides the mean weighted effect sizes calculated for measures of correct recall, incorrect recall, and confabulated recall.

Table 3
Mean Weighted Effect Sizes Calculated for Measures of Correct Recall, Incorrect Recall, and Confabulated Recall

Recall measure	# of articles	*k*	*N*	Weighted mean *d*	*p*-value	95% CI	N_{FS}
Correct details	46	59	2,887	1.20	<.001	(1.12, 1.28)	>10,000
Incorrect details	43	56	2,645	0.24	<.001	(0.16, 0.32)	337
Confabulated details	29	33	1,940	0.08	.10	(−0.01, 0.17)	—

Across the sample of studies, the CI produced a large and significant increase in *correct details* ($d = 1.20$) when compared with a control interview. The size of this effect is somewhat larger than that reported previously by Köhnken et al. (1999). The fail safe N associated with this effect size was substantial ($N_{FS} > 10{,}000$) suggesting a very robust effect of the CI in improving correct recall. In fact, only one of the 59 effect sizes proved negative in value (indicating a benefit of the control condition over the CI). Consistent with the previous meta-analysis, we also found a small, but significant, effect of the CI on *incorrect details* ($d = 0.24$), suggesting that the CI increased the frequency of incorrect details reported by participants when compared with a control condition. The fail safe *N* associated with this effect ($N_{FS} = 337$) indicated that it was rather robust and unlikely to be altered by future studies. Finally, the analysis of *confabulated details* produced a non-significant effect across studies ($d = 0.08$), indicating that the CI did not significantly differ from the control condition on this measure. We note here that too few studies (e.g., Gabbert et al., 2009) provided a statistical analysis of *recall accuracy* (i.e., % of correct recall as a function of total recall), thereby precluding a formal effect size analysis of this measure. Only 19 studies reported mean accuracy rates for the relevant conditions, often to the exclusion of other statistical information (*SD*s) that might enable computation of an effect size. An informal analysis of these mean estimates across studies showed that average accuracy for the CI and control conditions differed by < 1% ($M_{DIFF} = 0.32\%$), with 95% confidence intervals suggesting no-significant difference between the two interview conditions (−0.74%, 1.38%).

Moderator analyses. Results for all effect size analyses were heterogeneous: correct details: $Q(58) = 254.08$, $p < .001$; incorrect details: $Q(55) = 117.73$, $p < .001$; and confabulated details: $Q(32) = 91.02$, $p < .001$. As a result, weighted least squares regression models were conducted to predict the variance in effect sizes across studies based upon the following moderators: age of the participants (young adults vs. children or older adults), medium of event presentation (live vs. video/slide/narrative), type of event (crime/accident scenario vs. neutral scenario), retention interval between the presentation of the event and the interview (log of hours), type of cognitive interview that was employed (standard vs. enhanced or modified), and type of control interview that was employed (standard vs. structured). Regression models were run for each dependent variable, including all predictor variables and with studies weighted by sample size. Results of the models are described below as a function of each moderator.

First, *age of the sample* proved significant when considering correct and incorrect details, but was non-significant in the analysis of confabulated details. Specifically, children produced significantly smaller effect sizes for both correct details ($d = 0.91$, $p < .001$, $k = 19$; $Z_j = 1.96$, $p = .05$) and incorrect details ($d = 0.07$, *ns*., $k = 19$; $Z_j = 3.34$, $p < .001$) when compared with young adults (ds = 1.21 and 0.29, ps $< .001$, ks = 35 and 32, respectively), but showed no differences with regard to confabulated details ($Z_j = 0.65$, *ns*.). In contrast, older adults showed a significantly larger effect size for correct details ($d = 1.99$, $p < .001$, $k = 5$, $Z_j = 2.25$, $p < .05$) when compared with young adults, but showed no differences on the measure of incorrect details ($Z_j = 1.17$, *ns*.). Only two studies involving older adults included a measure of confabulated details, so this condition was excluded from the analysis. Overall, it appears that the CI produces greater correct recall for adults and the elderly, while eliciting fewer incorrect details (compared with a control condition) for children.

The *medium of event presentation* (live vs. video/slide/narrative) failed to significantly predict effect size across studies for either correct details ($Z_j = 1.37$, *ns*.) or incorrect details ($Z_j = 1.16$, *ns*.). Only two studies employing live events provided a measure of confabulated details, so this variable was excluded from that analysis.

With regard to the *type of event*, results indicated a significant difference for correct details ($Z_j = 2.97$, $p < .01$), but no differences with regard to incorrect details ($Z_j = 1.09$, *ns*.) or confabulated details ($Z_j = 1.27$, *ns*.). Events that likely evoked greater arousal via a crime or accident scenario ($d = 1.06$, $p < .001$, $k = 34$) produced smaller effect sizes for correct details when compared with events involving more neutral conditions ($d = 1.43$, $p < .001$, $k = 25$) – though the benefit of the CI remained substantial regardless of event type (i.e., $d > 1.00$), supporting its potential for improving recall in a forensic context.

The *retention interval* between viewing the event and recalling information proved significant for measures of correct details ($Z_j = 3.32$, $p < .001$) and confabulated details ($Z_j = 4.65$, $p < .001$), but was non-significant for incorrect details ($Z_j = 1.16$, *ns*., $r = .02$, $k = 56$). Specifically, effect sizes decreased as delay increased for correct details ($r = -0.29$, $k = 59$), while effects increased commensurate with delay for confabulated details ($r = .50$, $k = 33$). As displayed in Figure 1, it appears that the benefit of the CI does decrease as delay increases, though a rather substantial advantage for the CI remains in terms of correct details ($d > 1.0$) following the most extreme delay.

The *type of CI* employed across studies showed effects only for the MCI on estimates of incorrect details ($Z_j = 2.53$, $p < .01$), such that the MCI produced significantly greater effect sizes for incorrect details ($d = 0.30$, $p < .001$, $k = 25$) when compared with the original CI ($d = 0.12$, *ns*., $k = 15$). The MCI showed no significant differences with respect to correct ($Z_j = 0.22$, *ns*.) or confabulated details ($Z_j = 0.43$, *ns*.), and the ECI showed no significant differences from that of the standard CI for any dependent measure (Z_js ≤ 0.83, *ns*.). These results suggest that modifications to the CI (MCI) produced greater incorrect details when compared with the effects of the original CI.

Finally, the *control interview* used across studies significantly predicted effect sizes on measures of correct details ($Z_j = 2.76$, $p < .01$) and confabulated details ($Z_j = 2.40$, $p < .01$), but proved non-significant for estimates of incorrect details

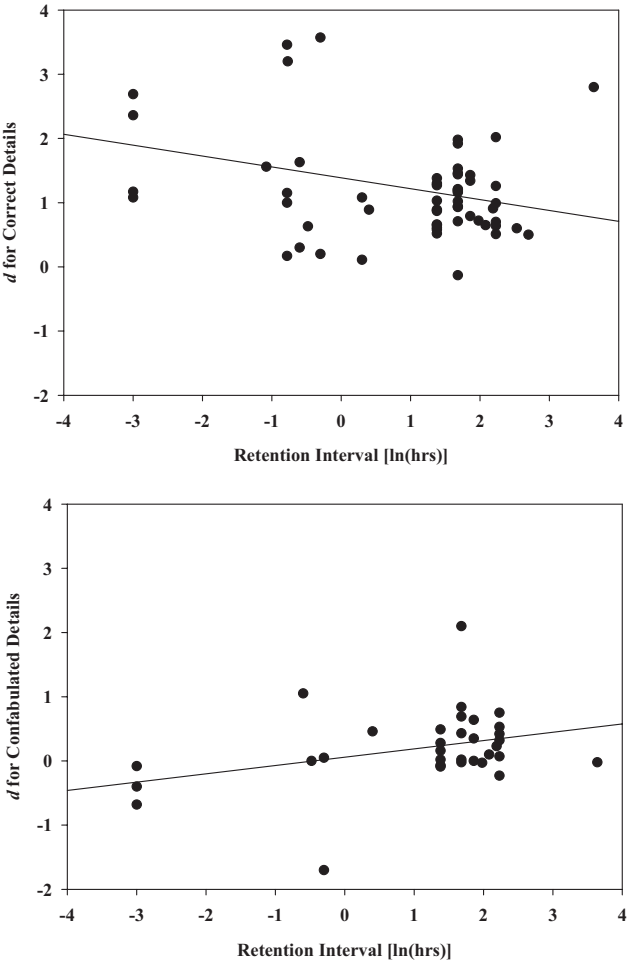

Figure 1. Scatterplot of the relationship between retention interval (involving the natural log of hours between encoding and retrieval episodes) and effect size (d) for correct details (top) and confabulated details (bottom).

($Z_j = 1.07$, *ns.*). Studies employing a standard interview ($d = 1.38$, $p < .001$, $k = 34$) produced larger effect sizes for correct details when compared with those employing a structured interview ($d = 1.09$, $p < .001$, $k = 25$). In addition, standard interview studies ($d = 0.32$, $p < .001$, $k = 13$) produced larger effect sizes for confabulated details when compared with structured interview studies ($d = -0.06$, *ns.*, $k = 20$).

General Discussion

An extensive body of empirical literature on the CI has emerged over the past 25 years. While a previous meta-analysis, conducted 10 years ago, has examined this literature for the statistical (and moderator) effects of the CI, there has been no systematic review of the research literature with an eye towards the sufficiency

of the literature and policy implications therein. The study space and meta-analysis indicate a rather robust literature with a substantial number of studies using a modified version of the CI (44% of the current sample of studies) and an increase in the number of studies sampling from vulnerable populations (children and older adults) since the publication of the Köhnken et al. (1999) meta-analysis. However, several key areas require further study. Only a few studies have included police officers and civilians as witnesses. Events which elicit emotional reactions and are arousing have not been included and the efficacy of the CI over long delays has not been fully explored. Only a small number of studies have used police or professional interviewers to conduct the cognitive interviews in their studies. The results of the meta-analysis indicated a large and significant increase in *correct details*, a smaller but significant increase in incorrect *details* and no differences in confabulated details with the CI when compared with a control interview. The moderator analysis indicated that the CI produces greater correct recall for adults and the elderly. The benefit of the CI remained substantial regardless of event type (emotional versus neutral) and medium (live versus video). There was an effect of retention interval with the benefit of the CI decreasing as retention interval increased (but see Figure 1). The modified version of the CI (see Table 2) produced greater incorrect details when compared with the original CI. Below we consider these issues in greater depth and discuss the implications of our systematic review for policy development.

Substantial Increases in Correct Recall

The current (and previous) meta-analysis suggests rather substantial increases in correct recall with the CI as compared with a structured interview. The effect holds even when lessened a bit by variables known to decrease memory recall (such as retention interval and emotional/arousing events). This is an important finding in view of the fact that development of internationally recognized protocols have been influenced by research on the CI (see Lamb, Hershkowitz, Orbach, & Esplin, 2008, for a review). From a policy standpoint, this is significant and argues in favour of adopting a CI approach for everyday investigative interviewing (a point reinforced by Fisher, 2010). At the same time, with some qualifications, more recent studies do suggest that simpler versions of the CI can be quite effective and this is a finding that will hold much appeal for practitioners.

Small Increases in Incorrect Recall—A Cause for Concern?

In both the original and current meta-analysis we noted a small increase in the recall of erroneous details. The 1999 meta-analysis indicated an increase in output was not accompanied by a drop in accuracy. The current meta-analysis which included a subset of studies that reported accuracy rates concludes the same. Whether or not quantity comes with a drop in accuracy is likely to depend on monitoring processes which evaluate the quality of the contents of memory and control processes which regulate memory output (Koriat, Goldsmith, & Halamish, 2008). Monitoring effectiveness relies on mnemonic cues derived from the on-line process of remembering, one's own beliefs about factors that can affect memory performance as well as the motivation to be accurate (Koriat et al., 2008). By improving monitoring effectiveness it is possible to increase quantity and accu-

racy of reported information (Koriat & Goldsmith, 1996). Instructions to witnesses that they should not guess or make up details but tell the interviewer if they do not know (or do not remember) ought to improve monitoring and control of output promoting high accuracy (Koriat & Goldsmith, 1996; Fisher, 2010). We are only just beginning to understand how this is achieved in a CI. In one recent study, an early CI conducted before a witness was asked misleading questions led to more accurate source monitoring at the item level in a delayed recognition test (Memon, Zaragoza, Clifford, & Kidd, 2010). Our main focus in that study was on whether the retrieval of information (prior to suggestion) enhanced the salience of that information and reinforced associated retrieval cues such that it increased resistance to suggested details when they were presented. An alternative or additional account of the Memon et al. (2010) findings is that they are a result of improved monitoring of correct versus incorrect items at retrieval. The extent to which the CI improved monitoring and control of output was not examined in the Memon et al. (in press) study because all participants were *forced* to answer every question in the misinformation phase, in other words there was no "don't know" option. There are various ways which researchers could proceed in the future. A way forward would be to examine how the CI might improve monitoring and control processes increasing the recollection of source specifying details. Recently, Scoboria, Trang, Shapero, and Frey (2009) found brief training in which interviewees were encouraged to thoroughly search their memory and to weight confidence in potential responses increased sensitivity to unanswerable questions. Clarification of the meaning of "I don't know" responses also led to an increased accuracy in response to unanswerable questions (Scoboria, Mazzoni, & Kirsch, 2008).

Our meta-analysis suggests that the use of the MCI was associated with a greater number of incorrect and confabulated details. The errors may be related to the modifications of the CI which have resulted in interviewers overlooking some of the key components of the ECI such as the instruction not to guess (see Table 2). Regardless of which version of the CI is being used, we strongly advocate that interviewers remind witnesses throughout the interview to use the "I don't know" and "do not guess" instructions. Future research should explore the strategic control of memory reporting with the CI and in particular how the use of specific instructions and additional measures such as confidence can improve accuracy without reducing output.

Does the CI Generalize to Children and Older Adults?

Since the last meta-analysis was conducted in 1999, more researchers have begun to explore the potential of the CI to improve the recall of children and a few studies have been conducted with older adults as witnesses. Taking the child studies first, contrary to the Köhnken et al. meta-analysis, we found that children produced significantly smaller effect sizes for both correct details and incorrect details when compared with young adults. In the past one of the concerns about using the CI with young children (6 years and under) was that the techniques that form the original CI were difficult for the children to use (Memon, Cronin, Eaves, & Bull, 1996). Subsequently, researchers (e.g., Holliday, 2003a, 2003b) modified the CI so it could be suitable for younger children (including 4 year-olds).

However, as indicated by Table 2, these modified versions vary from study to study which makes it difficult to identify whether the addition or omission of a particular component of the CI or some characteristic of the children sampled accounts for the smaller effect size for correct details. In terms of whether or not we advocate the use of CI in investigative interviews with children, we suggest interviewers are trained in modified versions of the CI which include ground rules (the use "I don't know" and "I don't understand") and appropriate use of the "do not fabricate" instruction for the reasons discussed earlier.

The older adult sample is an important addition to the literature in view of the fact that the population is ageing and senior citizens are active in society for longer. There is evidence that older adults are more likely to come into contact with law enforcement and fear crime even though they are less at risk of being victims of crime (Lachs et al., 2005). Based on a small set of studies, the current meta-analysis suggests that older adults benefit even more from the CI than younger adults in generating correct details (with no observed differences for incorrect details). The gains in correct details seen in older adult witnesses are consistent with the *environmental support hypothesis* which predicts that older adults rely more on and can make more effective use of, external support at the time of remembering due to a depletion of cognitive resources that are needed to initiate their own retrieval strategies (Craik, 1994; Craik, Byrd, & Swanson, 1987). Hence older adults will benefit more from any additional cues provided by the context reinstatement instruction of the CI. Research is currently underway to develop a modified version of the CI for use with older witnesses (Holliday, Ferguson, Milne, Bull, & Memon, 2009). As the number of studies continues to increase there is an opportunity to learn more about the how the CI can compensate for age related deficits in recall.

Does the CI Generalize to a Real-World Context?

In contrast to the robust findings on the effectiveness of the CI based on the published literature and the meta-analyses reported here, our study space has made it clear that only small subsets of studies have examined the efficacy of interview procedures under conditions that closely approximate those of real life witnesses. For example, while a number of studies have used crime relevant and emotionally arousing scenarios, these tend to be presented via video with memory being tested after relatively short delays. Our meta-analysis showed that events involving a crime or accident scenario did produce smaller effect sizes for correct details when compared with events involving more neutral conditions, however, in both cases there were rather substantial increases in correct details with a CI relative to a control group. Importantly, there were no differences as a function of event medium (live versus videotape). Moreover, the substantial benefits of CI were retained at the long delays sampled in the studies (see Figure 1). From a theoretical perspective it is important to note that we might expect these factors (event type and delay, for example) to reduce the effectiveness of the CI as access to memories degrade over time and key retrieval cues are lost. That the CI can be shown to, despite these memory failures—increase correct recall and produce a large effect therein is noteworthy. In terms of the effectiveness of using a CI over long delays, given that context reinstatement is a key component we could make

some predictions. The use of this technique has a strong theoretical basis drawing upon the notion that reinstatement of the original encoding context increases the accessibility of stored information (Tulving and Thomson's Encoding Specificity Hypothesis, 1973). The literature on context dependent memory would lead us to expect that MCR would aid recall at long retention intervals. For example the "outshining hypothesis" predicts that when memory cues are impoverished (such as after a long delay) context reinstatement is more likely to aid retrieval (Smith, 1988). In support of this, there is evidence that children's recall of an event experienced 6 months earlier benefits from the provision of context cues (returning to the location of the original event) 24 hours before an interview (Priestley, Roberts, & Pipe, 1999). Context reinstatement was particularly beneficial for the youngest children in the study (5–6 year olds). Thus, the relevant literature suggests a CI after a lengthy delay is likely to aid recall.

Despite the potential of the CI, we currently lack sufficient data on witnesses who are interviewed following lengthy delays and neither do we know enough about the effects of repeated interviews. It is in addressing these real world issues that the empirical literature falls short. It could be argued using Neisser's (1978) characterization that the CI literature has been dominated by "high road" research conducted in highly controlled laboratory contexts as compared to "low road" studies conducted in more ecologically valid settings. Perhaps what is needed now is a "middle road" approach that bridges basic and applied research (Lane & Meissner, 2008). A good example of the middle road approach is the current research of Dando and colleagues. The study space analysis revealed that few studies have examined the performance of police officers trained in the use of CI in the field and the laboratory. However, this has been offset somewhat by the recent efforts on the part of Dando and colleagues to develop protocols based on their experience of working with the police and to test them under controlled conditions using police officers as interviewers whenever possible (Dando et al., in press, 2009a, 2009b, 2009c).

Which of the CI Components are Necessary to Yield a Significant Benefit?

Do we need to use all components of the CI? Our meta-analyses suggests that it is possible to see gains in correct information with a much simpler shorter version of the CI, namely the MCI and this supports the findings of studies that have examined the individual cognitive components of the CI. Milne and Bull (2002) for example, examined the relative effectiveness of each of the four original CI mnemonics in a study where the participants were adults and children (aged 8–9 years and 5–6 years). For all age groups, they found a combination of Mental Reinstatement of Context (MCR) and Report Everything occasioned more recall compared to the individual use of the other techniques. Importantly, there was no significant difference when MCR was used on its own confirming the determinant role of context reinstatement in the CI. There is evidence from field research indicating that a structured interview with MCR can be an effective procedure. Hershkowitz, Orbach, Lamb, Sternberg, and Horowitz (2001) interviewed alleged victims (aged 4 to 13 years) of abuse using the NICHD protocol with or without MCR. They found the MCR resulted in proportionally more details when it was followed by an open-ended invitation to elaborate.

Researchers have recently suggested a modification of the modified cognitive interview to encourage the police to use it in investigative interviews. Dando et al. (2009b) in a mock eyewitness study found that asking witnesses to draw a detailed sketch of what they saw and talk while doing so was as effective as an instruction to mentally reinstate context. The authors also found fewer confabulations when sketch was used which they attribute to a witness's generating their own cues to help them remember rather than relying on the interviewer to direct them towards relevant cues.

It appears that we do not always need to use the full procedure to see the benefits of the CI although we would urge caution in reaching this conclusion for three reasons. First, our study space indicated there was a great deal of variability in how the MCI is operationalized and put into practice (see Table 2). We also noted that some researchers left out critical components such as the transfer of control and communication of ground rules, the consequences (association with increased error rates) were discussed earlier. Secondly, there was a lack of detail in the published papers about how interviewers were trained and how familiar they were with basic principles of communication such as establishing rapport with the witness. Our advice would be to ensure that interviewers are familiar with the basic principles of a structured approach to interviewing. Thirdly, the MCI appears to have been used primarily with vulnerable populations such as children and as we argue later in this discussion a higher level of skill may be required from interviewers to effectively adapt the CI to meet the needs of vulnerable individuals.

Policy Implications of This Systematic Review

One of the challenges that researchers continually face involves convincing policy makers and practitioners to adopt empirically-derived methods and thereby alter their everyday practice. In his 1996 review, Fisher made recommendations to improve the quality of interview training that were based at the time on what Malpass et al. (2008) refer to as a *Best Practice* model (i.e., the best evidence currently available). This model contains no criteria for assessing the strength of the empirical base, though it can be used as a guide for policy development with the assumption that developments in research may result in changes in policy recommendations. The alternative model is what Malpass et al referred to as *Well Established Knowledge*, which assumes that (a) the studies forming the literature base are of a high scientific standard, (b) the question to be evaluated has been extensively studied, and (c) the findings are well established. Our evaluation based upon the current systematic review is that current research on the CI comes close to meeting the standards of the WEK model. The research on the CI meets the "adequacy criteria" set out by Malpass et al. in that our conclusions are based on peer reviewed publications and the literature is extensive in terms of the volume of studies. Moreover, the positive effects of the CI have been well replicated and are robust. Furthermore, there is general agreement in the scientific community as to its effectiveness (see Wells, Memon, & Penrod, 2006) and reference was made to the Cognitive Interview in the U.S. Department of Justice *Eyewitness Evidence: A Guide for Law Enforcement* (1999) as well as in the British Home Office, *Achieving Best Evidence* (2001). The literature is diverse in

terms of the manner in which CI has been implemented and the design of future studies should give attention to real world application.

Moreover, to date there have only been two published field tests of the CI (Clifford & George, 1996; Fisher, Geiselman, & Amador, 1989). A more recent (unpublished) comparison of 9 police officers pre- and post-CI training did not show any benefits of CI training (Schreiber & Fisher, 2006) and it has been police and other investigators have not made best use of the scientific advances in the field of investigative interviewing (Fisher, 2010). Not only is more field data needed, but it is critical to identify ways of improving training to increase uptake of the CI among police interviewers.

The U.K. Model as a "Way Forward" to Implementation and Training

A way forward has been suggested by the developments in training in the U.K. One of the factors that has contributed to changes in police training in CI in the U.K. has been the evaluation of effectiveness of police training. There are two issues of interest here firstly, are they changing their behavior and secondly, is this change in behavior making the evidence they collect more accurate/diagnostic? As indicated earlier, we have limited data on the second question. With respect to behavior change, Clark and Milne (2005) found no evidence of the CI procedure having been used at all in the vast majority (83%) of British interviews they examined. Subsequent research highlighted that the CI procedure, as it has been taught to novice police officers, is either too complex or is administered too early in their police career to provide a foundation for their investigative work (Dando et al., 2009c). Police officers also felt the interviews they were conducting related mostly to less serious crimes where the additional time and resources involved were not warranted. It was also apparent that more emphasis was being placed in training on suspect interviewing and less on witness interviews (see Dando et al., 2009a).

The structure for CI training under the U.K. Home Office investigative interviewing framework, referred to as the *PEACE* (Planning and Preparation, Engage and Explain, Account, Closure and Evaluation) model, underwent major revision to address these and related issues. Since 2009, PEACE training has become part of a new *Professionalizing Investigation Programme* (PIP). The most basic PIP level is the standard expected for police interviewing victims and witnesses in volume crime (e.g. Robbery), the second level is the standard required for witnesses in serious and complex investigations, with further markers for those carrying out specialist interviews and those managing and coordinating interviews for major investigations. Competency at PIP Level 1 is a prerequisite for the development of specialist interview skills (National Investigative Interviewing Strategy, 2009).

The developments in training in the U.K. are significant because they deal with two critical issues with respect to the confidence and ability of a police officer when it comes to using the CI. Firstly, it is clear that police officers find the cognitive interview to be a demanding interview protocol. Not only does the CI take longer to administer, but it involves instructing witnesses in the use of several sophisticated techniques (e.g., context reinstatement). Questioning does not comprise a set of pre-determined questions, but instead involves active

listening to the free narrative and basing questions on what the witness has provided. Moreover, interviewers need both basic social skills in communicating effectively with a witness as well as a higher level of skills to gauge the needs of particular types of witnesses (for example, very young children, and victims of sexual offences). For example, rapport building is not only a technique that could be used at the start of an interview, but also in a later phase when a witness may become too distressed to speak. Similarly, the context reinstatement technique can be used on more than one occasion to focus retrieval. As pointed out by Dando et al. (2009c), it is essential that basic interview skills are confidently mastered and regularly applied as this will provide a foundation upon which to build some of the more complex CI components (Dando & Milne, 2009). It will also encourage police officers to use the Cognitive Interview techniques flexibly as Fisher and Geiselman (1992) intended (See also Fisher, 1995).

In sum, the U.K. model is one that could be adopted by police departments around the world. The revisions to the U.K.'s national strategy and the introduction of PIP levels may go far in improving the quality of training and increasing use of the CI among practitioners. It remains to be seen if these developments in training are accompanied by increases in the use of CI by police officers and indeed if this influences the accuracy and diagnostic of the evidence obtained with a CI.

Conclusions

Twenty-five years of empirical research has shown the CI to be an effective method of interviewing witnesses. While some gaps in the literature remain to be filled, the current literature provides a strong basis from which policymakers and law enforcement should seriously consider altering their everyday practices to allow for introduction of the CI. We strongly encourage U.S. researchers and policy makers to take advantage of the foundational model offered by the U.K. and further develop their collaborations with law enforcement and intelligence personnel. It is critical that research continue to evaluate both training and implementation of the CI in the field as we monitor its successful application to investigative interviews.

References

* Denotes that study was used in both the Study Space and Meta-Analysis.
** Denotes that the study was used in the Study Space Analysis only.

Achieving Best Evidence. (2001). Retrieved October 1, 2009, from http://www.homeoffice.gov.uk/documents/achieving-best-evidence/

*Akehurst, L., Milne, R., & Köhnken, G. (2003). The effects of children's age and delay on recall in a cognitive or structured interview. *Psychology, Crime, & Law, 9,* 97–107. doi:10.1080/1068316021000057686

*Allwood, C. M., Ask, K., & Granhag, P. A. (2005). The cognitive interview: Effects on the realism in witnesses' confidence in their free recall. *Psychology, Crime, & Law, 11,* 183–198. doi:10.1080/10683160512331329943

Anderson, R. C., & Pichert, J. W. (1978). Recall of previously unrecalled information following a shift in perspective. *Journal of Verbal Learning & Verbal Behavior, 17,* 1–12. doi:10.1016/S0022-5371(78)90485-1

*Aschermann, E., Mantwill, M., & Köhnken, G. (1991). An independent replication of the effectiveness of the cognitive interview. *Applied Cognitive Psychology, 5,* 489–495. doi:10.1002/acp.2350050604

*Brock, P., Fisher, R. P., & Cutler, B. L. (1999). Examining the cognitive interview in a double-test paradigm. *Psychology, Crime, & Law, 5,* 29–45. doi10.1080/10683169908414992

*Brown, C., & Geiselman, R. E. (1992). Eyewitness testimony of mentally retarded: Effect of cognitive interview. *Journal of Police & Criminological Psychology, 6,* 14–21. doi:10.1007/BF02806593

*Campos, L., & Alonso-Quecuty, M. (1999). The cognitive interview: Much more than simply "try again". *Psychology, Crime, & Law, 5,* 47–59. doi:10.1080/10683169908414993

*Campos, L., & Alonso-Quecuty, M. L. (1998). Knowledge of the crime context: Improving the understanding of why the cognitive interview works. *Memory, 6,* 103–112. doi:10.1080/741941602

*Centofanti, A. T., & Reece, J. (2006). The cognitive interview and its effect on misleading post event information. *Psychology, Crime, & Law, 12,* 683–699. doi:10.1080/10683160600558394

*Chapman, A. J., & Perry, D. J. (1995). Applying the cognitive interview procedure to child and adult eyewitnesses of road accidents. *Applied Psychology: An International Review, 44,* 283–294. doi:10.1111/j.1464-0597.1995.tb01081.x

Christianson, S.-Å. (1992). Emotional stress and eyewitness memory: A critical review. *Psychological Bulletin, 112,* 284–309. doi:10.1037/0033-2909.112.2.284

Christianson, S.-Å., & Lindholm, T. (1998). The fate of traumatic memories in childhood and adulthood. *Development & Psychopathology, 10,* 761–780. doi:10.1017/S0954579498001850

Clarke, C., & Milne, R. (2005). *National evaluation of the PEACE investigative interviewing course.* London: Home Office.

Clifford, B. R., & George, R. (1996). A field evaluation of training in three methods of witness and victim investigative interviewing, *Psychology, Crime, & Law, 2,* 231–248. doi:10.1080/10683169608409780

Craik, F. I. M. (1994). Memory changes in normal aging. *Current Directions in Psychological Science, 5,* 155–158. doi:10.1111/j.1467-8721.1994.tb00166.x

Craik, F. I. M., Byrd, M., & Swanson, J. M. (1987). Patterns of memory loss in three elderly samples. *Psychology & Aging, 2,* 79–86. doi:10.1037/0882-7974.2.1.79

Dando, C., Wilcock, R., Behnkle, C., & Milne, R. (in press). Modifying the Cognitive Interview: Countenancing forensic application by enhancing practicability. *Psychology, Crime, & Law*.

Dando, C., Wilcock, R., & Milne, R. (2008). Victims and witnesses of crime: Police officers' perceptions of interviewing practices. *Legal & Criminological Psychology, 13,* 59–70. doi:10.1348/135532506X162498

*Dando, C., Wilcock, R., Milne, R., & Henry, L. (2009a). An adapted cognitive interview procedure for frontline police investigators. *Applied Cognitive Psychology, 23,* 698–716. doi:10.1002/acp.1501

**Dando, C., Wilcock, R., & Milne, R. (2009b). The cognitive interview: The efficacy of a modified mental reinstatement of context procedure for frontline police investigators. *Applied Cognitive Psychology, 23,* 138–147. doi:10.1002/acp.1451

**Dando, C., Wilcock, R., & Milne, R. (2009c). Novice police officers application of the Cognitive Interview procedure. *Psychology, Crime, & Law, 15,* 679–696. doi:10.1080/10683160802203963

Dando, C. J., & Milne, R. (2009). The Cognitive Interview. In R. N. Kocsis (Ed.), *Applied*

criminal psychology: A guide to forensic behavioral sciences. (pp. 147–168). Springfield: Charles C Thomas.

Daubert v. Merrell Dow Pharmaceuticals. (1993). 509 US 579.

*Davis, M. R., McMahon, M., & Greenwood, K. M. (2005). The efficacy of mnemonic components of the cognitive interview: Towards a shortened variant for time-critical investigations. *Applied Cognitive Psychology, 19,* 75–93. doi10.1002/acp.1048

Deffenbacher, K. A., Bornstein, B. H., McGorty, K., & Penrod, S. D. (2008). Forgetting the once-seen face: Estimating the strength of an eyewitness's memory representation. *Journal of Experimental Psychology: Applied, 14,* 139–150. doi:10.1037/1076-898X.14.2.139

Deffenbacher, K. A., Bornstein, B. H., Penrod. S. D., & McGorty, E. K. (2004). A meta-analytic review of the effects of high stress on eyewitness memory. *Law & Human Behavior, 28,* 687–706. doi:10.1007/s10979-004-0565-x

Dietze, P. M., & Thomson, D. M. (1993). Mental reinstatement of context: A technique for interviewing child witnesses. *Applied Cognitive Psychology, 7,* 97–108. doi: 10.1002/acp.2350070203

*Dornburg, C. C., & McDaniel, M. A. (2006). The cognitive interview enhances long-term free recall of older adults. *Psychology & Aging, 21,* 196–200. doi:10.1037/0882-7974.21.1.196

Dysart, J. E., & Lindsay, R. C. L. (2007). The effects of delay on eyewitness identification accuracy: Should we be concerned? In R. C. L. Lindsay et al. (Eds.), *The handbook of eyewitness psychology, Vol II: Memory for people* (pp. 361–376). Mahwah, NJ: Lawrence Erlbaum.

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review, 66,* 183–201. doi:10.1037/h0047707

Eyewitness Evidence: A guide for Law Enforcement. (1999). U.S. Department of Justice Office of Justice Programs. Retrieved February 1, 2010, from http://www.ncjrs.gov/pdffiles1/nij/178240.pdf

*Finger, K., & Pedzek, K. (1999). The effect of the cognitive interview on face identification accuracy: Release from verbal overshadowing. *Journal of Applied Psychology, 84,* 340–348. doi:10.1037/0021-9010.84.3.340

Fisher, R., Brewer, N., & Mitchell, G. (2009). The relation between consistency and accuracy of witness testimony: Legal versus cognitive explanations. In T. Williamson, R. Bull, & T. Valentine (Eds.), *Handbook of psychology of investigative interviewing: Current developments and future directions* (pp. 121–136). Chichester, UK: John Wiley.

Fisher, R. P. (1995). Interviewing victims and witnesses of crime. *Psychology, Public Policy, & Law, 1,* 732–764. doi:10.1037/1076-8971.1.4.732

Fisher, R. P. (2010). Interviewing cooperative witnesses. *Legal and Criminological Psychology, 15,* 25–38. doi:10.1348/135532509X441891

Fisher, R. P., Falkner, K. L., Trevisan, M., & McCauley, M. R. (2000). Adapting the Cognitive Interview to enhance long-term (35-years) recall of psychical activities. *Journal of Applied Psychology, 85,* 180–189. doi:10.1037//0021–9010.85.2.180

Fisher, R. P., & Geiselman, R. E. (1992). *Memory enhancing techniques for investigative interviewing: The Cognitive Interview.* Springfield, IL: Charles C. Thomas.

Fisher, R. P., Geiselman, R. E., & Amador, M. (1989). Field test of the cognitive interview: Enhancing the recollection of actual victims and interviewees of crime. *Journal of Applied Psychology, 74,* 722–727. doi:10.1037/0021-9010.74.5.722

**Fisher, R. P., Geiselman, R. E., Raymond, D. S., Jurkevich, L. M., & Warhaftig, M. L. (1987b). Enhancing enhanced eyewitness memory: Refining the cognitive interview. *Journal of Police Science & Administration, 15,* 291–297.

*Fisher, R. P., & Quigley, K. L. (1992). Applying cognitive theory in public health

investigations: Enhancing food recall with the cognitive interview. In J. M. Tanur (Ed.), *Questions about questions: Inquiries into the cognitive bases of surveys* (pp. 154–169). New York: Russell Sage Foundation.

Fisher, R. P., & Schreiber, N. (2007). Interviewing protocols to improve eyewitness memory. In M. Toglia et al. (Eds.), *The handbook for eyewitness psychology, Vol. I: Memory for events* (pp. 53–80). Mahwah, NJ: Lawrence Erlbaum.

Flin, R., O'Connor, P., & Crichton, M. (2008). *Safety at the sharp end: A guide to non-technical skills.* Aldershot: Ashgate.

*Gabbert, F., Hope, L., & Fisher, R. P. (2009). Protecting eyewitness evidence: Examining the efficacy of a self-administered interview tool. *Law & Human Behavior, 33,* 298–307. doi:10.1007/s10979-008-9146-8

Gabbert, F., Hope, L., & Jamieson, K. (2010). In the field: Trials of self-administered interview recall tool with witnesses to serious crime. Paper presented at the American Psychology Law Society Meeting, Vancouver.

*Geiselman, R. E., Fisher, R. P., Cohen, G., Holland, H., & Surtes, L. (1986). Eyewitness responses to leading and misleading questions under the cognitive interview. *Journal of Police Science & Administration, 14,* 31–39.

*Geiselman, R. E., Fisher, R. P., Firstenberg, I., Hutton, L. A., Sullivan, S. J., Avetissian, I. V., & Prosk, A. L. (1984). Enhancement of eyewitness memory: An empirical evaluation of the cognitive interview. *Journal of Police Science & Administration, 12,* 74–80.

*Geiselman, R. E., Fisher, R. P., MacKinnon, D. P., & Holland, H. L. (1985). Eyewitness memory enhancement in the police interview: Cognitive retrieval mnemonics versus hypnosis. *Journal of Applied Psychology, 70,* 401–412. doi:10.1037/0021-9010.70.2.401

*Geiselman, R. E., Fisher, R. P., MacKinnon, D. P., & Holland, H. L. (1986). Enhancement of eyewitness memory with the cognitive interview. *American Journal of Psychology, 99,* 385–401. doi:10.1080/09658210601092670

*Geiselman, R. E., & Padilla, J. (1988). Cognitive interviewing with child witnesses. *Journal of Police Science & Administration, 16,* 236–242.

*Geiselman, R. E., Taras, L., Schaap, R., & Woodruf, N. (1994). Recall of constituent elements from multiple episodes of the Cognitive Interview, *Journal of Police & Criminal Psychology, 10,* 1–7.

*Ginet, M., & Verkampt, F. (2007). The cognitive interview: Is its benefit affected by the level of witness emotion? *Memory, 15,* 450–464. doi:10.1080/09658210601092670

Goodman, G. S., & Quas, J. (2008). Repeated interviews and children's memory: It's more than just how many. *Current Directions in Psychological Science, 17,* 386–390. doi:10.1111/j.1467-8721.2008.00611.x

Goodman, G. S., Taub, E. P., Jones, D. P. H., England, P., Port, L. K., Rudy, L., & Prado, L. (1992). Testifying in criminal court: Emotional effects on child sexual assault victims. *Monographs of the Society for Research in Child Development, 57,* 5, Serial No. 229. doi:10.1111/j.1540-5834.1992.tb00324.x

*Granhag, P. A., Jonsson, A., & Allwood, M. (2004). The cognitive interview and its effect on witness' confidence. *Psychology, Crime, & Law, 10,* 37–52. doi:10.1080/1068316021000030577

*Gwyer, P., & Clifford, B. R. (1997). The effects of the cognitive interview on recall, identification, confidence and the confidence/accuracy relationship. *Applied Cognitive Psychology, 11,* 121–145. doi:10.1002/(SICI)1099-0720(199704)11:2<121::AID-ACP443>3.0.CO;2-L

*Hayes, B. K., & Delamothe, K. (1997). Cognitive interviewing procedures and suggestibility. *Journal of Applied Psychology, 82,* 562–577. doi:10.1037/0021-9010.82.4.562

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Orlando, FL:Academic Press.

*Hernandez-Fernaud, E., & Alonso-Quecuty, M. (1997). The cognitive interview and lie detection: A new magnifying glass for Sherlock Holmes? *Applied Cognitive Psychology, 11,* 55–68. doi:10.1002/(SICI)1099-0720(199702)11:1<55::AID-ACP423>3.0.CO;2-G

Hershkowitz, I., Orbach, Y., Lamb, M. E., Sternberg, K. J., & Horowitz, D. (2001). The effects of mental context reinstatement on children's accounts of sexual abuse. *Applied Cognitive Psychology, 15,* 235–248. doi:10.1002/acp.699

**Holliday, R. (2003a). Reducing misinformation effects in children with cognitive interviews: Dissociating recollection and familiarity. *Child Development, 74,* 728–751. doi:10.1111/1467-8624.00565

*Holliday, R. (2003b). The effect of a prior cognitive interview on children's acceptance of misinformation. *Applied Cognitive Psychology, 17,* 443–457. doi:10.1002/acp.879

**Holliday, R., & Albon, A. J. (2004). Minimizing misinformation effects in young children with cognitive interview mnemonics. *Applied Cognitive Psychology, 18,* 263–281. doi:10.1002/acp.973/

Holliday, R., Ferguson, J., Milne, R., Bull, R. H. C., & Memon, A. (March, 2009). *Isolating the effect of the Cognitive Interview instructions on witness testimony.* Poster presentation at the American Psychology Law Society Conference, San Antonio, TX.

Home Office & Department of Health. (1992). *Memorandum of good practice on video recorded interviews with child witnesses for criminal proceedings.* London: HMSO.

Hulse, L., & Memon, A. (2006). Fatal impact? The effects of emotional arousal and weapon presence on police officers' memories for a simulated crime. *Legal & Criminological Psychology, 11,* 313–325. doi:10.1348/135532505X58062

Hulse, L. M., Allan, K., Memon, A., & Read, J. D. (2007). Emotional arousal and memory: A test of the post stimulus processing hypothesis. *American Journal of Psychology, 120,* 73–90.

Johnson, B., Mullen, B., & Salas, E. (1995). A comparison of the three major meta-analytic approaches. *Journal of Applied Psychology, 80,* 94–106. doi:10.1037/0021-9010.80.1.94

Köhnken, G., Milne, R., Memon, A., & Bull, R. (1999). A meta-analysis on the effects of the Cognitive Interview. Special Issue of *Psychology, Crime, & Law, 5,* 3–27. doi:10.1080/10683169908414991

*Köhnken, G., Schimossek, E., Aschermann, E., & Hofer, E. (1995). The cognitive interview and the assessment of the credibility of adults' statements. *Journal of Applied Psychology, 80,* 671–684. doi10.1037/0021-9010.80.6.671

*Köhnken, G., Thuerer, C., & Zoberbier, D. (1994). The cognitive interview: Are the interviewers' memories enhanced, too? *Applied Cognitive Psychology, 8,* 13–24. doi:10.1002/acp.2350080103

Köhnken, G. (1995). Interviewing adults. In R. Bull and D. Carson (Eds.), *Handbook of psychology in legal contexts.* (pp. 215–233). Chichester: Wiley.

Kebbell, M., & Milne, R. (1998). Police officers' perception of eyewitness factors in forensic investigations. *Journal of Social Psychology, 138,* 323–330. doi:10.1080/00224549809600384

Kebbell, M., Milne, R., & Wagstaff, G. (1999). The Cognitive Interview: A survey of its forensic effectiveness. *Psychology, Crime, & Law, 5,* 101–115. doi:10.1080/10683169908414996

*Kebbell, M. R., & Wagstaff, G. F. (1997). An investigation into the influence of hypnosis on the confidence and accuracy of eyewitness recall. *Contemporary Hypnosis, 14,* 157–166. doi:10.1002/ch.97

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

Koriat, A., Goldsmith, M., & Halamish, V. (2008). Control processes in voluntary remembering. In H. L. Roediger, III (Ed.), *Cognitive psychology of memory. Learning and memory: A comprehensive reference* (J. Byrne, Editor; pp. 307–324). Oxford: Elsevier.

Lachs, M., Bachman, R., Williams, C., Kossac, A., Bove, B., & O'Leary, J. (2005). Older adults as crime victims, perpetrators, witnesses, and complainants: A population-based study of police interactions. *Journal of Elder Abuse & Neglect, 16,* 25–40. doi:10.1300/J084v16n04_02

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

Lane, S. M., & Meissner, C. A., (2008). A "middle road" approach to bridging the basic-applied divide in eyewitness identification research. *Applied Cognitive Psychology, 22,* 779–787. doi:10.1002/acp.1482

La Rooy, D., Lamb, M. E., & Pipe, M. E. (2009). Repeated Interviewing: A critical evaluation of the risks and potential benefits. In K. Kuehnle & M. Connell (Eds.), *The evaluation of child sexual abuse allegations: A comprehensive guide to assessment and testimony* (pp. 327–361). Chichester: Wiley.

La Rooy, D., Pipe, M. E., & Murray, J. E. (2005). Reminiscence and hypermnesia in children's eyewitness memory. *Journal of Experimental Child Psychology, 90,* 235–254. doi:10.1016/j.jecp.2004.11.002

*Larsson, A. S., Granhag, P. A., & Spjut, E. (2003). Children's recall and the cognitive interview: Do the positive effects hold over time? *Applied Cognitive Psychology, 17,* 203–214. doi:10.1002/acp.863

Malpass, R. S., Tredoux, C. G., Schreiber Compo, N., McQuiston-Surrett, D. E., MacLin, O. H., Zimmerman, L. A., & Topp, L. D. (2008). Study space analysis for policy development. *Applied Cognitive Psychology, 22,* 789–801. doi:10.1002/acp.1483

*Mantwill, M., Köhnken, G., & Ascherman, E. (1995). Effects of the cognitive interview on the recall of familiar and unfamiliar events. *Journal of Applied Psychology, 80,* 68–78. doi:10.1037/0021-9010.80.1.68

*McCauley, M. R., & Fisher, R. P. (1995). Facilitating children's eyewitness recall with the revised cognitive interview. *Journal of Applied Psychology, 80,* 510–516. doi:10.1037/0021-9010.80.4.510

*McMahon, M. (2000). The effect of the enhanced cognitive interview on recall and confidence in elderly adults. *Psychiatry, Psychology, & Law, 7,* 9–32. doi:10.1080/13218710009524968

*Mello, E. W., & Fisher, R. P. (1996). Enhancing older adult eyewitness memory with the cognitive interview. *Applied Cognitive Psychology, 10,* 403–417. doi:10.1002/(SICI)1099-0720(199610)10:5<403::AID-ACP395>3.0.CO;2-X

Memon, A. (2006). The Cognitive Interview. In O. Hargie (Ed.), *A handbook of communication skills* (3rd ed.). London: Routledge.

*Memon, A., Bull, R., & Smith, M. (1995). Improving the quality of the police interview: Can training in the use of the cognitive techniques help? *Policing & Society, 5,* 53–68.

Memon, A., Cronin, O., Eaves, R., & Bull, R. (1996). An empirical test of the mnemonic components of the Cognitive Interview. In G. M. Davies et al. (Eds), *Psychology and law: Advances in research* (pp. 135–145). Berlin: De Gruyter.

Memon, A., Milne, R., Holley, A., Bull, R., & Köhnken, G. (1994). Towards understand-

ing the effects of interviewer training in evaluating the cognitive interview. *Applied Cognitive Psychology, 8,* 641–659. doi:10.1002/acp.2350080704

**Memon, A., Holley, A., Wark, L., Bull, R., & Köhnken, G. (1996). Reducing suggestibility in child witness interviews. *Applied Cognitive Psychology, 10,* 503–518. doi:10.1002/(SICI)1099-0720(199612)10:6<503::AID-ACP416>3.0.CO;2-R

*Memon, A., Wark, L., Bull, R., & Köhnken, G. (1997a). Isolating the effects of the cognitive interview techniques. *British Journal of Psychology, 88,* 179–198.

*Memon, A., Wark, L., Holley, A., Bull, R., & Köhnken, G. (1997b). Eyewitness performance in cognitive and structured interviews. *Memory, 5,* 639–656. doi:10.1080/741941481

*Memon, A., & Yarmey, D. (1999). Earwitness recall and identification: Comparison of the Cognitive Interview and Structured Interview. *Perceptual & Motor Skills, 88,* 797–807. doi:10.2466/PMS.88.3.797-807

*Memon, A., Zaragoza, M., Clifford, B., & Kidd, L. (2010). Inoculation or antidote? The effects of cognitive interview timing on false memory for forcibly fabricated events. *Law & Human Behavior, 34,* 105–117. doi:10.1007/s10979-008-9172-6

**Milne, R., & Bull, R. (1996). Interviewing children with mild learning disability with a cognitive interview. In N. K. Clark & G. M. Stephenson (Eds.), *Investigative and forensic decision making: issues in criminological psychology* (No. 26). Leicester: British Psychological Society.

**Milne, R., & Bull, R. (2002). Back to basics: A componential analysis of the original cognitive interview mnemonics with three age groups. *Applied Cognitive Psychology, 16,* 743–753. doi:10.1002/acp.825

**Milne, R., & Bull, R. (2003). Does the cognitive interview help children to resist the effects of suggestive questioning? *Legal & Criminological Psychology, 8,* 21–38. doi:10.1002/acp.825

*Milne, R., Bull, R., Köhnken, G., & Memon, A. (1995). The cognitive interview and suggestibility. In N. Clark & G. M. Stephenson (Eds.), *Criminal behavior: Perceptions, attributions and rationality, issues in criminological and legal psychology* (No. 22, pp. 21–27). Leicester: British Psychological Society.

*Milne, R., Clare, I. C. H., & Bull, R. (1999). Using the cognitive interview with adults with mild learning disabilities. *Psychology, Crime, & Law, 5,* 81–99. doi:10.1080/10683169908414995

Morgan, C. A., Hazlett, G., Doran, A., Garrett, S., Hoyt, G., Thomas, P., Baranoski, M., & Southwick, S. M. (2004). Accuracy of eyewitness memory for persons encountered during exposure to highly intense stress. *International Journal of Law & Psychiatry, 27,* 265–279. doi:10.1016/j.ijlp.2004.03.004

National Investigative Interviewing Strategy. (2009). Retrieved October 1, 2009, from http://www.npia.police.uk/en/docs/National_Investigative_Interviewing_Strategy_09.pdf

Neisser, U. (1978) What are the important questions? In M. Grunerberg & R. N. Sykes. (Eds.), *Practical Aspects of Memory* (pp. 3–24). London: Academic Press.

Priestley, G., Roberts, S., & Pipe, M. E. (1999). Returning to the scene: Reminders and context reinstatement enhances children's recall. *Developmental Psychology, 35,* 1006–1019. doi:10.1037/0012-1649.35.4.1006

*Py, J., Magali, G., Desperies, C., & Cathey, C. (1997). Cognitive encoding and cognitive interviewing in eyewitness testimony. *Swiss Journal of Psychology, 56,* 33–41.

**Roberts, W. T., & Higham, P. A. (2002). Selecting accurate statements from the cognitive interview using confidence ratings. *Journal of Experimental Psychology: Applied, 8,* 33–43. doi:10.1037//1076-898X.8.1.33

*Saywitz, K. J., Geiselman, R. E., & Bornstein, G. K. (1992). Effects of the cognitive

interview and practice on children's recall performance. *Journal of Applied Psychology, 77,* 744–756. doi:10.1037/0021-9010.77.5.744

Schreiber, N., & Fisher, R. P. (2006, March). *Evaluating police training in Cognitive Interview techniques: Possibilities and difficulties.* Paper presented at the American Psychology and Law Society Conference, St. Petersburg, FL.

**Searcy, J. H., Bartlett, J. C., Memon, A., & Swanson, K. (2001). Aging and lineup performance at long retention intervals: Effects of metamemory and context reinstatement. *Journal of Applied Psychology, 86,* 207–214. doi:10.1037/0021-9010.86.2.207

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

Scoboria, A., Trang, H., Shapero, D., & Frey, M. (2009). *The effects of metacognitive monitoring and clarifying 'don't know' statements upon memory accuracy.* Manuscript under review.

Smith, S. M. (1988). Environmental context-dependent memory. In G. M. Davies and D. M. Thomson (Eds.), *Memory in context: Context in memory* (pp. 13–34). New York: Wiley.

*Stein, L. M., & Memon, A. (2006). Testing the efficacy of the cognitive interview in a developing country. *Applied Cognitive Psychology, 20,* 597–605. doi:10.1002/acp.1211

Sternberg, K. J., Lamb, M. E., Esplin, P. W., Orbach, Y., & Hershkowitz, I. (2002). Using a structured protocol to improve the quality of investigative interviews. In M. L. Eisen et al. (Eds.), *Memory and suggestibility in the forensic interview* (pp. 409–436). Mahwah, NJ: Erlbaum.

Tulving, E. (1983). *Elements of episodic memory.* New York: Oxford University Press.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80,* 352–373. doi:10.1037/h0020071

Turtle, J., Read, J. D., Lindsay, D. S., & Brimacombe, C. A. E. (2008). Towards a more informative psychological science of eyewitness evidence. *Applied Cognitive Psychology, 22,* 769–778. doi:10.1002/acp.1481

Valentine, T., & Mesout, J. (2009). Eyewitness identification under stress in the London Dungeon. *Applied Cognitive Psychology, 23,* 151–161. doi:10.1002/acp.1463

Wells, G. L., Memon, A., & Penrod, S. D. (2006). Eyewitness evidence: Improving its probative value. *Psychological Science in the Public Interest, 7,* 45–75. doi:10.1111/j.1529-1006.2006.00027.x

Wright, A. M., & Holliday, R. E. (2005). Police officers' perceptions of older eyewitnesses. *Legal & Criminological Psychology, 10,* 211–223. doi:10.1348/135532505X37001

*Wright, A. M., & Holliday, R. E. (2007). Enhancing the recall of young, young-old and old-old adults with cognitive interviews. *Applied Cognitive Psychology, 21,* 19–43. doi:10.1002/acp.1260

Yuille, J. C., Hunter, R., Joffe, R., & Zaparniuk, J. (1993). Interviewing children in sexual abuse cases. In G. S. Goodman & B. L. Bottoms (Eds.), *Child victims, child witnesses: Understanding and improving children's testimony* (pp. 95–115). New York: Guilford Press.

Appendix

Listing of Studies and Computed Effect Sizes (*d*) Included in the Cognitive Interview Meta-Analysis

Study	Exp./Cond.	N	Correct details	Incorrect details	Confabulated details
Akehurst, Milne, & Köhnken (2003)		64	0.72	0.04	−0.03
Allwood, Ask, & Granhag (2005)		56	1.00	0.36	—
Aschermann, Mantwill, & Köhnken (1991)		29	1.15	0.54	—
Brock, Fisher, & Cutler (1999)		145	1.56	—	—
Brown & Geiselman (1992)		22	1.53	0.01*	2.09
Campos & Alonso-Quecuty (1998)		69	1.17	1.31	−0.08
Campos & Alonso-Quecuty (1999)		170	1.08	−0.30	−0.68
Centofanti & Reece (2006)		40	2.69	−0.41	−0.40
Chapman & Perry (1995)	Exp. 1	48	1.30	−0.49	—
Chapman & Perry (1995)	Exp. 2 (4–5 Y.O.)	16	1.27	−0.88	—
Chapman & Perry (1995)	Exp. 2 (9–10 Y.O.)	16	1.38	−0.38	—
Chapman & Perry (1995)	Exp. 2 (14–15 Y.O.)	16	1.03	−0.44	—
Dando, Wilcock, Milne, & Henry (2009a)		40	1.20	−0.10	0.69
Davis, McMahon, & Greenwood (2005)		45	0.89	0.31	0.46
Dornburg & McDaniel (2006)		40	0.50	0.41	—
Finger & Pedzek (1999)	Exp. 1	75	3.45	0.62	—
Finger & Pedzek (1999)	Exp. 2	69	4.32	0.66	—
Fisher & Quigley (1992)		26	2.02	−0.16	0.07
Gabbert et al. (2009)		35	2.36	0.58	—
Geiselman, Fisher, et al. (1984)		16	1.44	0.32	—
Geiselman, Fisher, et al. (1985)		59	1.21	0.36	0.01*
Geiselman, Fisher, et al. (1986)		51	1.16	−0.15	−0.01*
Geiselman & Padilla (1988)		15	1.34	−0.51	0.00*
Geiselman, Taras, Schaap, & Woodruf (1994)		60	0.52	—	—
Ginet & Verkampt (2007)	Low arousal condition	35	0.70	0.93	0.53
Ginet & Verkampt (2007)	High arousal condition	35	0.51	0.85	−0.23
Granhag, Jonsson, & Allwood (2004)		26	0.60	0.52	—
Gwyer & Clifford (1997)		70	1.46	0.43	—
Hayes & Delamothe (1997)	6 Y.O. sample	64	0.79	0.15	0.64
Hayes & Delamothe (1997)	10 Y.O. sample	64	1.43	0.22	0.35
Hernandez-Fernaud & Alonso-Quecuty (1997)		73	1.63	0.08	1.05
Holliday (2003a)		64	0.87	0.01*	0.01*

(*Appendix continues*)

Appendix (*continued*)

Study	Exp./Cond.	N	Correct details	Incorrect details	Confabulated details
Kebbell & Wagstaff (1997)		38	0.17	—	—
Köhnken, Schimossek, et al. (1995)		28	0.93	0.56	0.84
Köhnken, Thurer, & Zoberbier (1994)		30	0.91	0.63	0.23
Larsson et al. (2003)	7-day retention condition	24	0.99	−0.62	0.75
Larsson, Granhag, Spjut (2003)	6-month retention condition	25	2.80	−0.08	−0.02
Mantwill et al. (1995)		90	0.64	0.58	0.42
McCauley & Fisher (1995)		28	1.08	1.06	—
McMahon (2000)		38	0.20	0.08	0.05
Mello & Fisher (1996)		50	3.20	1.68	—
Memon & Yarmey (1999)		77	0.30	0.07	—
Memon et al. (1995)		38	0.11	0.43	—
Memon, Wark, Bull, & Köhnken (1997)		54	0.64	0.51	−0.07
Memon, Wark, Holley, et al. (1997)		45	0.65	0.53	0.10
Memon, Zaragoza, et al. (in press)		80	0.63	0.12	0.00
Milne & Bull (1996)		82	0.61	0.43	−0.09
Milne et al. (1995)		84	0.58	0.40	0.28
Milne et al. (1999)	Mild L.D. sample	47	0.66	0.40	0.49
Milne et al. (1999)	Normal sample	38	0.89	0.23	0.16
Py et al. (1997)		71	1.18	0.39	0.43
Saywitz et al. (1992)	Exp. 1 (7–8 Y.O. witnesses)	10	1.98	−0.84	—
Saywitz et al. (1992)	Exp. 1 (7–8 Y.O. participants)	10	0.94	−0.25	—
Saywitz et al. (1992)	Exp. 1 (10–11 Y.O. witnesses)	10	1.92	−0.28	—
Saywitz et al. (1992)	Exp. 1 (10–11 Y.O. participants)	10	−0.13	−0.29	—
Saywitz et al. (1992)	Exp. 2 (8–9 Y.O.)	23	0.71	−0.45	—
Saywitz et al. (1992)	Exp. 2 (11–12 Y.O.)	38	1.02	−0.26	—
Stein & Memon (2006)		64	1.26	0.08	0.32
Wright & Holliday (2007)		102	3.57	−0.24	−1.70

Note. Asterisk indicates that study was assigned a non-significant effect size.

Received October 14, 2009
Revision received April 27, 2010
Accepted May 4, 2010 ■