Cognitive Interviewing
QUICK REFERENCE

The general strategy of enhanced cognitive interviewing (ECI) is “to guide the witness to those memory records that are richest in relevant information and to facilitate communication when those mental records have been activated.” (Geiselman & Fisher, 2014).

The five phases of an enhanced cognitive interview are as follows:

**Introduction.** Develop rapport with the interviewee and encourage him or her to play an active role by volunteering information; communicate that the interviewee is in control of the interview.

**Reinstate the context.** Encourage the interviewee to mentally reconstruct the physical and personal context that existed at the time of the event.

**Open-ended narration.** Ask the interviewee to report everything he or she can recall even if it is incomplete, seemingly insignificant, out of chronological order, or even inconsistent with something they said earlier. Caution them not to guess or speculate.

**Follow-up questions.** Probe specific scenes and images mentioned in the interviewee’s narrative to elicit further detail and access additional memories.

**Multiple and varied recall.** Ask the interviewee to provide the narrative again, but (1) in reverse order and (2) from a different physical perspective or the perspective of another witness (again, without guessing).


Interviewing Witnesses and Victims

R. Edward Geiselman & Ronald P. Fisher

University of California, Los Angeles & Florida International University


Information is the lifeblood of investigations and it is the ability of investigators to obtain useful and accurate information from victims and witnesses that is most crucial for case solution and effective criminal prosecution. Yet full and accurate memory recall is difficult to achieve during a police interview. We have spent much of the past 25-30 years addressing this issue. In this chapter, we begin by describing a typical police interview, followed by an overview of the origins and underpinnings of the Cognitive Interview (CI), an innovative method to enhance witness recall and the focus of this chapter. We then present a step-by-step description of the main elements of the general CI protocol for the practitioner. Following a description of the empirical tests and some practical applications, we end with a brief overview of one of the more recent extensions of the CI for use with suspects (the CIS).

Past Interview Practices

Two decades ago, Fisher, Geiselman, and Raymond (1987) and George and Clifford (1992) described typical interviewing protocols used by American and British police, respectively. The results were surprisingly uniform and somewhat discouraging. Following a perfunctory effort to establish rapport, interviewers generally began interviews by making an open-ended request to the witness: “Tell me what happened.” After listening to an initial outburst of crime-related facts, often no more than just a few seconds, the interviewer interrupted the witness’s narrative response and asked a series of direct, short-answer questions, on the order of: “How tall was he? How much did he weigh? Did he have a weapon?” These questions, which reflect generically salient crime facts, often were asked in the same order to all witnesses using a standardized checklist. In addition to these neutral questions, the interviewers often asked leading or suggestive questions, such as “He was wearing a red shirt, wasn’t he?” If a witness did not provide the requested details, the interviewer did little or nothing to assist the witness. In practice, police interviewers often dominated the social interaction with the witness by asking many questions, and by asking questions that elicited only brief answers. This relegated the witness to sit passively waiting for interviewers to ask questions (Fisher et al., 1987). To compound the problem, interviewers often discouraged witnesses from taking active roles by interrupting them in the middle of a narrative response.

We note, with a sigh of renewed discouragement, that similar patterns of lesser quality interviewing procedures have been found in more recent interviews conducted by German (Berresheim & Weber, 2003), Canadian (Snook & Keating, 2010), and American police (Schreiber & Fisher, 2005). These practices have the adverse effects of reducing the amount of information witnesses provide and increasing inaccurate responses. This is because these practices entice witnesses to (a) withhold information, (b) not provide any unsolicited information, (c) give abbreviated answers, and (d) volunteer answers they are unsure of. Furthermore, they disrupt the natural process of searching through memory, thereby making memory retrieval inefficient.

We were discouraged by the quality of interviews, in part, because forensic research scientists have known for a while how to conduct interviews effectively. Of these techniques, the most
prominent are the Cognitive Interview, Conversation Management, the Memorandum of Good Practice, and the Step-wise method. Each of these protocols is composed of many specific techniques that have generally been found to (a) increase the amount of information gathered, and/or (b) decrease the likelihood of recalling an event incorrectly. Common to all of these protocols are several core elements, including (a) developing rapport with the witness, (b) asking open-ended questions primarily, (c) asking neutral questions and avoiding leading or suggestive questions, and (d) funneling the interview, beginning with broader questions and narrowing down to more specific questions. We shall focus here on the CI procedure because it is more encompassing than the others and it has been the focus of extensive scientific testing.

The CI is a systematic approach to interviewing witnesses with the goal of increasing the amount of relevant information obtained without compromising the rate of accuracy. The CI is based on scientifically derived principles of memory and communication theory as well as extensive analyses of law-enforcement interviews. Most important, the CI has been found in empirical studies to produce significantly more information than standard question-and-answer type interviews and without decreasing accuracy.

**Origins and Underpinnings of the CI**

In the early 1980s, a constellation of factors contributed to the development and eventual refinement of the CI techniques for investigative interviewing. First, the U.S. Department of Justice set out to fund an effort to produce a protocol for use by law enforcement for purposes of interviewing victims and witnesses of crime (National Institute of Justice - Geiselman & Fisher, 1985). The most notable police procedure at the time was the Reid and Associates confrontational interrogation technique for use with suspects (Inbau, Reid, Buckley, & Jayne, 2001), not for use with witnesses and victims. Second, the one innovative technique employed by law enforcement at the time, forensic hypnosis, was becoming mired in legal issues (e.g., People v Shirley, 1982; Sanders & Simmons, 1983). Third, the RAND Corporation (1975) had just completed a survey of law-enforcement professionals and found that 85% of what police do on a daily basis is talk to citizens, whereas only 2% of the respondents had received any formal training on how to interview people. Fourth, the two of us were eager to steer our respective basic theoretical research programs on memory retrieval in a more applied direction.

The first task in building the CI was to review the literature from cognitive psychology to identify candidate techniques for enhancing memory retrieval. This search led to developing the original version of the CI. Since that time, the basic set of memory retrieval aids was transformed through a progression of research into the enhanced CI, which is described in our “how-to” manual (Fisher & Geiselman, 1992).

The core elements of the CI are organized around three basic psychological processes: memory and cognition, social dynamics, and communication. Some of the memory-enhancing components of the CI protocol attempt to maximize the amount of feature overlap between retrieval strategies and the witness’s memory record (Flexser & Tulving, 1978) or attempt to have the witness explore multiple retrieval routes to the memory record (Tulving, 1974). Other cognitive elements assist witnesses to use their cognitive resources efficiently. The social dynamics include rapport and encouraging active witness participation. The communication elements include promoting extensive, detailed responses and utilizing non-verbal as well as verbal modes of expression.

**Step-by-Step Sequence of the CI**

The CI follows a somewhat flexible order intended to maximize the effectiveness of the individual techniques. The general strategy is to guide the witness to those memory records that
are richest in relevant information and to facilitate communication when these mental records have been activated. The recommended questioning sequence, which is common to the Step-wise method (Yuille, Hunter, Joffe, & Zaparniuk, 1993) and the Memorandum of Good Practice (Home Office, 2002), is to progress from asking open-ended questions to more specific follow-up probing questions. The sequence of the CI protocol provides a general framework for conducting an investigative interview of most persons including civilian eyewitnesses, victims, police, and more recently, suspects. However, it is important to understand that interviewers must be flexible and alter their approaches to meet the needs of each witness rather than to use a rigid template, causing the witness to adapt. In this regard, the CI should be seen as a toolbox of techniques that can be drawn upon as the situation presents itself. The CI is more of a set of general guidelines or a collection of techniques than it is a recipe for conducting an interview. Each witness and situation will call for a slightly different approach. Consequently, it is not surprising that most investigators who have received training on the CI incorporate some, but not all elements of the CI, into any given one of their interviews (Dando, Wilcock, & Milne, 2009; Kebbell, Milne, & Wagstaff, 1999).

The CI protocol is divided into five sections. Although this is the conventional sequence, interviews invariably will deviate from this plan somewhat as unexpected information arises. First, an introduction is made that establishes a relationship between the witness and the interviewer. At this point, the interviewer explains the expected social dynamics for the remainder of the interview, emphasizing a witness-centered approach. The interviewer then gives the witness an opportunity to provide an uninterrupted narration of what s/he experienced. During this time, the interviewer is able to construct a strategy for eliciting additional information. Based on the contents of the uninterrupted narrative, the interviewer guides the witness through several information-rich memory representations (scenes, images). The interviewer then reviews the information generated during the entire interview, followed by the close of the interview in a manner that will extend its functional life.

Introduction

During the introduction phase of the CI, the interviewer will (a) develop rapport with the witness, (b) encourage the witness to play an active role by volunteering information, (c) convey his or her investigative needs for extensive, detailed information, and (d) convey that a thorough search of memory will require concentration. The introduction establishes the appropriate psychological states and interpersonal dynamics to promote efficient memory recall and communication during the remainder of the interview. Victims and witnesses often will be anxious about the interview process because they are uncertain about what is expected of them and how the process will transpire (Sydeman, Cascardi, Poythress, & Ritterbrand, 1997). CI interviewers attempt to reduce this uncertainty by previewing the structure of the interview, and especially by explaining the witness-centered nature of the interview. Furthermore, victims are encouraged to ask questions about the process. Foreshadowing the interview should reduce victims’ anxiety about the process as it reduces uncertainty.

Rapport. Witnesses, and especially victims, often are asked to give detailed descriptions of intimate, personal experiences to police officers, who are complete strangers. They must be psychologically comfortable with the interviewer as a person to go through the mental effort and emotional distress of describing crime-related details. If anything, the police investigator’s official appearance (badge, uniform, gun) may create a psychological barrier between the police officer and the witness. To overcome this natural barrier, CI interviewers will invest time at the outset of the interview to develop meaningful, personal rapport with the witness (Abbe &
Brandon, 2013; Collins, Lincoln & Frank, 2002; Shafer, & Navarro, 2012), a feature often absent in police interviews in the past (Fisher et al., 1987). Interviewers should develop a set of topics that they become comfortable in using to begin the casual conversation. There are popular books written for this purpose (Lowndes, 2003). One strategy is to explain to the witness, “Before we begin, I would like to get to know you a little better – what do you do on a typical day?” The benefits of developing rapport include freeing the witness of some anxiety about being interviewed that might otherwise consume some of the witness’s cognitive resources. Furthermore, the interviewer must interact with the victim not merely as a source of evidence that can be applied toward solving the crime. Rather, the interviewer should express his/her concern about the victim’s plight, as a person who has undergone a potentially life-altering experience (Fisher & Geiselman, 2010).

Transfer Control. The interviewer generally has higher social or expert status, which normally dictates that that the interviewer should control the interview; however, the witness has first-hand knowledge of the crime, which dictates that the witness should control the interview. Resolving this apparent conflict is crucial for a successful interview. Therefore, after developing rapport, the CI interviewer will in effect transfer control of the interview to the witness. The interviewer will openly acknowledge that s/he was not at the scene and that the witness must play an active role in the interview, “I was not there when this happened, so I will be relying on you to do most of the work here.” This clarifies for the witness the role that s/he will be playing during the interview, and that s/he should not wait for the interviewer to ask questions. It is commonly desired that the interviewer contribute only 20% of the talking during an investigative interview (Snook & Keating, 2010), thus preserving the “80-20 rule.”

Detailed Recall. Police interviews require witnesses to describe people, objects and actions in more detail than civilians normally do in casual conversation. Inducing such an extraordinary level of description requires that police convey this goal explicitly. Witnesses often withhold information because they do not know what is relevant for a police investigation. To minimize witnesses’ withholding information, CI interviewers will instruct witnesses to report everything they can recall, whether it is trivial, out of chronological order, or even if it contradicts a statement made earlier. (Note that this is not an invitation to guess, as is sometimes inferred incorrectly, Memon, Wark, Bull, & Kohnken, 1997.) One technique is to explain to the witness that in normal conversation we typically tell others only the highlights about an event, but for purposes of this interview, it is desired that you be as detailed in your recollections as possible. The witness will be instructed not to consider what might have investigative value, but rather to be as complete as possible.

The interviewer will explain further that retrieving memories often is not an easy task, but rather this will require concentration. The interviewer will thank the witness in advance for anticipated effort and co-operation. To promote high accuracy in recall, interviewers should explicitly instruct witnesses not to guess, and to indicate that they “don’t know” or “don’t recall” when that is the case. Interviewers can promote more efficient use of witnesses’ limited mental resources by offering them the option to close their eyes when recalling (Bekerian & Dennett, 1997; Perfect, Wagstaff, Moore, Andrews, Cleveland, Newcombe, Brisbane, & Brown, 2008). Doing so, however, requires that the interviewer has developed good rapport with the witness. Interviewers can further assist witnesses to focus their mental resources more effectively by minimizing physical distractions, such as phone calls during the interview.

Open-ended narration
Following completion of the introduction phase of the CI, in preparation for requesting an open-ended narration of the target event, the interviewer will request that the witness take a few moments to mentally go back to the time and place where the event happened. This process is commonly known as reinstatement of the context. Retrieving information from memory is most efficient when the context of the original event is recreated at the time of recall (Tulving & Thomson, 1973). Interviewers should therefore instruct witnesses to mentally recreate the external factors (weather), emotional factors (mood, fear), and cognitive factors (thoughts) that existed at the time of the original event. Sights, sounds, and smells are relevant as well as the witness’s state of mind leading up to the event. The interviewer will give the witness the time necessary to recreate the period of time leading up to the target event.

Reinstating the context also can help circumvent any additions or contaminations to the witness’s memory record that may have occurred subsequent to the event. By directing the witness’s thoughts back to the encoding of the original event in context, the narrative is less likely to reflect post-event influences, additions during post-event rehearsal, or the last time the witness told the story to someone else. In part for this reason, the CI has been found to reduce misleading question effects in most of the studies where this has been tested (Geiselman, Fisher, Cohen, Holland & Surtes, 1986; Memon, Zaragoza, Clifford & Kidd, 2010; Milne & Bull, 2003 – however, see LaPaglia, Wilford, Rivard, Chan, & Fisher, in press).

Once the witness signals to the interviewer that s/he has the target time frame in mind and has mentally recreated the context, the interviewer should request the open-ended narrative – “Tell me in your own words what happened in detail from beginning to end.” The initial open-ended narration permits the interviewer to infer the witness's overall representation of the event and to develop an efficient strategy for probing the various memory records. The interviewer will note the witness’s "mental images" of the crime (e.g., perpetrator, weapons), and will develop a preliminary plan about which images to probe, in what order, and which questions to ask when probing each image. If the initial narrative is not interrupted, witnesses will convey a large portion of their total recall during this phase (Roberts & Higham, 2002). Not interrupting is one of the more difficult skills for investigators to learn. Interrupting frustrates witnesses by making it difficult for them to narrate their story and to communicate all of their information.

It is preferable that the interviewer record only cursory notes during the witness’s narrative so as to flag segments that the interviewer wants to probe with follow-up questioning. This will allow the interviewer to concentrate on what the witness is saying. It is advisable to use some of the witness’s own words in these notations. By using the witness’s own words in the follow-up questioning, the interviewer will demonstrate to the witness that s/he has been listening and will minimize the potential for miscommunication.

**Follow-up Questions – Probing Scenes and Images**

In the probing stage, the CI interviewer will guide the witness to the richest sources of information (scenes or “mental images”) and thoroughly exhaust these sources of their contents. To accomplish this task, the interviewer will carry out the plan developed while listening to the witness’s open-ended narrative.

**Principle of Detail.** The most promising scene from the narrative should be addressed first (principle of detail – Fisher & Geiselman, 1992). This is because elements in the memory record are associated with the other elements such that recalling one detail can trigger recollection of others, and because both the interviewer and the witness will become less able to fully concentrate as the interview progresses due to fatigue. Each scene as notated by the interviewer will be addressed independently.
There are a number of strategies for asking the probing questions that should improve productivity, as follows. Interviewers can avoid contributing to the cognitive load of witnesses by refraining from asking questions while witnesses are searching through memory (interrupting the thought process) and, in general, by asking fewer, but more open-ended questions. Asking fewer questions and encouraging witnesses to narrate their story in an uninterrupted fashion also makes the interviewer’s task easier—by not having to formulate many questions—and frees the interviewer to listen more effectively to the witness’s narration. Rather than asking many specific questions, interviewers should explicitly instruct witnesses on the importance of describing events in great detail. In general, witness recall is much more accurate when answering open-ended questions (e.g., “Describe the robber’s appearance.”) than closed questions (e.g., “Did the robber have dark or light hair?”). The interviewer should allow the witness the time needed to formulate an answer and to search memory adequately. This can be accomplished by pausing after each of the witness’s answers, perhaps 3-4 seconds, and by using longer pauses strategically. The interviewer should never talk over a witness or appear to want to ask a question while the witness is still answering a previous question. To do so could serve to create an interviewer-centered rather than a witness-centered interview environment.

Interviewers and respondents often exchange ideas using only the verbal medium, but some people are more expressive non-verbally, and some events are better described non-verbally (Leibowitz, Guzy, Peterson, & Blake, 1993). Ideally the response format should be compatible with the witness’s mental record of the event, thereby minimizing the need to transform the mental record into an overt response (Greenwald, 1970). If an event is inherently spatial, (e.g., the location of objects within a room) then witnesses should be allowed to respond spatially, by drawing a sketch of the room or by placing model objects within a (model) room. Encouraging witnesses to sketch out the crime scene could promote more extensive recall (Dando, Wilcock, Milne, & Henry, 2009a, 2009b). Drawing sketches has also been shown to enhance recalling abstract information, for example, how one made earlier decisions (Hirn, Fisher, & Carol, 2012).

Another key to successfully probing a witness’s memory is to ask for the same information repeatedly but with different questions (e.g., visual, auditory; forward, reverse). Following the multi-component conception of memory, multiple searches of memory should lead to more finds, but only if different access routes are explored. Instead of asking, “Tell me more about his appearance” multiple times, ask about the intruder in different ways, e.g., “Did the intruder remind you of anyone you know?” For objects, instead of asking “Tell me more about the objects” multiple times, ask about different properties of the objects, e.g., “How much did it weigh; what kind of material was it made out of?” For speech characteristics, “Were any unusual words used; “did the person sound foreign or native-born; educated or uneducated?” For names of persons or places, suggest that the witness go through the alphabet searching for the first letter of the name as a cue or think about the length of the name or whether it was a typical American name or was it foreign-sounding. For numbers, instruct the witness to think about partial information such as size and orientation (MacKinnon, O'Reilly, & Geiselman, 1990). To learn about additional such tools from the CI toolbox for requesting specific kinds of information, see Fisher and Geiselman (1992).

As a note of caution, interviewers should refrain from applying social pressure on witnesses or otherwise encouraging them to answer questions they are uncertain of. Similarly, interviewers must guard against inducing feelings of inadequacy by formulating questions in a negative tone,
“You don’t recall his name, do you?” Such negative questioning may reinforce the victim’s sense of inadequacy. This form of the question also allows the victim to answer the question easily with a “No” response rather than encouraging a deep search through memory.

**Principle of Momentum.** The CI interviewer should not skip around between scenes (principle of momentum – Fisher & Geiselman, 1992) but rather should ask all questions relevant to one scene at a time, while the witness has that scene in consciousness. Interviewers should try to be aware of the witness’s currently active mental image, so as to time their questions most efficiently. This practice may require interviewers to defer asking certain questions until later in the interview, when the questions are compatible with the witness’s mental image. For instance, if the interviewer needs to learn about the license tag of the getaway car, but the witness is currently thinking about the robber’s face and not about the getaway car, then the interviewer should defer asking the witness to describe the license tag until the witness is thinking about the getaway car. Memory for the perpetrator’s face should be more accessible when the witness is thinking about the perpetrator than when the witness is thinking about the getaway car. In general, event details will be most accessible when they are perceptually related to the witness’s current mental image (Pecher, Zeelenberg, & Barsalou, 2003)

**Multiple and Varied Recall.** As noted above, the more often witnesses search through their memories about an event, the more new details they likely will recall. Interviewers can enhance witness recollection by asking witnesses to describe the event multiple times within the interview, but in ways that promote varied retrieval rather than repeating the same line of questioning. Two techniques from the original CI serve this purpose: (a) Requesting the narrative again but this time in reverse order, and (b) requesting the witness to change physical or conceptual perspectives on the event. Each of these techniques typically is employed near the end of the specific-questions phase of the CI protocol. Each was taken from research on memory for stories (Anderson & Pichert, 1978; Whitten & Leonard, 1981) and is based on the concept that there are multiple access routes to memories.

The reverse-order technique can help the witness generate information that is incidental or atypical to the target event because recalling an event in reverse order is less amenable to thematic-based recall (Geiselman & Callot, 1990). Eliciting these incidental details can be crucial for case solution. Recalling an event in reverse order is more of a frame-by-frame approach compared to recalling the event in forward order where most often there is a clear chronology of events (Bransford & Franks, 1971). The results of a recent study suggest that the reverse-order technique should be employed only after the forward narrative report and the follow-up questioning phase have been completed (Dando, Ormerod, Wilcock, & Milne, 2011). Otherwise, this technique might disrupt the temporal clustering of information stored in the witness’s memory.

One form of the change-perspectives technique (Boon & Noon, 1994) asks the witness to think about the various physical perspectives s/he may have had throughout the event. A second form of the change-perspectives technique asks the witness to consider the perspective of another person at the event (e.g., “What do you think the cashier saw?”). With the latter application of the change-perspectives technique, a caution to the witness against guessing is recommended, especially with children or mentally challenged persons who may have difficulty taking on the perspectives of others (Saywitz, Geiselman, & Bornstein, 1992).

**Review**

Reviewing the information already notated allows the interviewer to check on its accuracy. It also provides the witness with an opportunity to recall additional information. During the review
stage, the interviewer should clarify any uncertainties or discrepancies that occurred earlier in the interview. The interviewer should read his or her notes back to the witness and ask the witness to (a) correct any errors or omissions in the interviewer’s notes, and (b) inform the interviewer of any new recollections. The interviewer should point out in a non-challenging way any ambiguities or contradictions within the witness’s statement and ask the witness to clarify these matters, even if that means to indicate that the witness is not certain about the matter.

Close

When closing the interview, the interviewer will fulfill any official police requirements associated with the investigation, e.g., collecting background information about the witness. Then, the interviewer should thank the witness for the witness’s help and co-operation. The interviewer also should encourage the witness to contact him or her several days later when the witness thinks of new information. We have found that some investigators prefer to explicitly tell witnesses that they will call them in a couple of days to see if they have thought of any additional information. Extending the life of the interview is important, given the likelihood of delayed recollections, especially following incidents that were emotionally arousing for the witness (Fisher, Brewer, & Mitchell, 2009). In one such real-world case described to us recently, a traumatized witness experienced a delayed recollection of a box outside her apartment that had not been there prior to a home invasion. The subsequent latent print analysis of the recovered box revealed the identity of one of the intruders/murderers.

Empirical Tests and Evaluations of the CI

The CI is sometimes presented on popular television crime shows as being magical or mystical. In fact, however, the CI is based on well-founded principles of cognitive psychology. The CI is a good-practice, information-gathering technique that has been tested rigorously in more than 100 laboratory experiments, most of which were conducted in the United States, England, Germany or Australia. In these studies, volunteer witnesses (usually college students, but not always) observed either a live, innocuous event or a videotape of a simulated crime. Shortly thereafter (ranging from a few hours to several days), the witnesses were interviewed by a trained researcher—or in some cases by experienced police officers—who conducted either a CI or a control interview. The control interview was modeled after a typical police interview or after a generally accepted interview protocol such as the UK Memorandum of Good Practice. Across these studies, the CI typically elicited between 25%-50% more correct statements than did the control interview. The effect was extremely reliable: Of the 55 experiments examined in a meta-analysis conducted in 1999 (Koehnken Milne, Memon, & Bull, 1999), 53 experiments found that the CI elicited more information than did the comparison interview (median increase = 34%). A second meta-analysis in 2010, with a larger sample of studies produced similar results (Memon, Meissner, & Fraser, 2010). Equally important, accuracy was just as high in the CI interviews as in the comparison interviews (for reviews, see Bekerian and Dennett, 1993; Fisher and McCauley, 1995 and Fisher & Schreiber, 2007). The CI also typically does not contribute to the creation of false memories (Sharman & Powell, 2013), but instead serves to reduce misleading question effects (Geiselman, Fisher, Cohen, Holland & Surtes, 1986; Memon et al. 2010; Milne & Bull, 2003), in a sense inoculating the witness from inadvertent misleading questions. We attribute these latter findings to the use of context reinstatement, which directs the witness back to the original memory record, and to a greater reliance on open-ended questions.

In addition to the laboratory studies, two field studies that examined interviews with victims and witnesses of real crimes found the advantage of the CI to hold equally well. In Fisher, Geiselman, & Amador (1989), 16 experienced detectives from the Metro-Dade (Florida) Police
Department tape recorded several interviews, mainly from victims or witnesses of purse snatchings or commercial robberies. The detectives then were divided into two equivalent groups based on their supervisors' evaluations and on their objective performance (number of statements elicited on the tape recorded interviews). One of the two groups received training on the CI, whereas the other did not. Although the two groups were comparable before training, the trained group of detectives elicited 63% more information than the untrained group after training. Furthermore, the trained detectives elicited 48% more facts after training than before training. Of the seven trained detectives, six improved dramatically (34% - 115%). Only the one detective who did not change his interviewing style failed to improve.

A parallel field study was conducted in England by George and Clifford (1992, 1996) in which experienced police investigators tape recorded interviews before and after training (or no training for some). The investigators’ questioning styles changed dramatically as a result of CI training. Compared to the untrained group, and also to themselves before training, the CI-trained group (a) asked fewer questions, (b) asked a higher proportion of open-ended questions, (c) asked fewer leading questions, and (d) injected more pauses. These changes in questioning style were also accompanied by an increase in the amount of information elicited. The CI group elicited 55% more information after than before training and 14% more information than did the untrained group.

Since the late 1990’s, several studies have examined whether the CI could be used effectively with children and other “non-standard” witnesses. In some of these studies, the experimental witnesses were young children (from 7-12 years of age: e.g., Milne & Bull, 1996; Larsson, Granhag, & Spjut, 2003; McCauley & Fisher, 1995; Saywitz et al., 1992), in some studies, the witnesses were older people (Mello & Fisher, 1996; McMahon, 2000), and in some studies, the witnesses were young adults with learning disabilities or some other cognitive deficit (Brown & Geiselman, 1990; Geiselman & Padilla, 1988; Milne, Clare, & Bull, 1999). The patterns of results are remarkably similar to those found with “normal” adults, namely, the CI elicits considerably more information than the control interview, and at comparable or slightly higher accuracy rates.

Aside from the positive witness performance data, we were interested in how the CI would appear to an external observer. To address this issue, Fisher, Mello, and McCauley (1999) had research subjects listen to tape recordings (taken from an earlier CI experiment) of witnesses being interviewed with a CI or with a standard police interview. The listeners then rated the credibility of the witnesses on several dimensions (e.g., accuracy of memory, confidence, intelligence, trustworthiness). In two separate experiments, one in which the witnesses were children and the second in which the witnesses were adults, there were no differences in perceived credibility. These non-differences were apparently not due to insensitivity of the data, as other differences were observed: specifically, CI interviewers were perceived to be less manipulative than were conventional interviewers. This was a serendipitous finding, but, with the benefit of hindsight, not all that surprising given that CI interviewers ask fewer questions, ask fewer leading questions, and in general follow a more witness-centered approach than do conventional interviewers. On balance, these results suggest that, if anything, the CI should be more, not less, acceptable as an interview procedure—assuming, of course, that it is desirable for an interviewer not to manipulate the witness’s testimony. Indeed, one study has found police officers in the U.K. to judge complainants who received a CI to be more credible than witnesses who received a standard interview (Westera, Kebbell, & Milne, 2011). There also is reason to believe that the witnesses themselves will view their experience of being interviewed more
favorably (Fisher & Geiselman, 2010). This is important because they also will have a more favorable impression of the policing agency and consequently they will be more likely to help the police in the future. Also, recent research has shown a link between greater victim satisfaction with the police and a reduction in the likelihood of PTSD-type symptoms (Kunst, Rutten, & Knijff, 2013).

In summary, the benefits of the CI have been found repeatedly by different researchers, with different witness populations, and in a variety of settings. The CI elicits considerably more information than the typical police interview while maintaining the same or achieving a higher level of accuracy compared with conventional police interviews. The CI appears to external observers to be less manipulative than standard interview techniques, and witnesses will likely consider their experience of being interviewed more favorably.

For some time now, the CI has been taught and/or implemented by several policing agencies and allied investigative agencies worldwide (Fisher & Geiselman, 1997). Today, those agencies include: FBI, National Transportation Safety Board, Department of Homeland Security, Rural Policing Institute (Federal Law Enforcement Training Center), Defense Intelligence Agency, UK Home Office, Calgary Police Service, Singapore Police Force, ICAC (Hong Kong), as well as several mid-level police departments around the United States. However, unlike in the U.K., a comprehensive, coordinated program for training on the CI has yet to be established in the U.S.

Of course, the CI is not without its limitations. For example, recent research has found that the general CI is not effective for persons diagnosed with Autism Spectrum Disorder (Maras & Bowler, 2010). However, follow-up research suggests that this limitation can be ameliorated in large part by taking these witnesses physically back to the scene of the crime for the CI rather than having them mentally reconstruct the context at a different location as is typically the case (Maras & Bowler, 2012). Work is ongoing for this important group of witnesses, but, to the best of our knowledge, parallel research for persons with different stages of Alzheimer’s disease who must be interviewed as victims or witnesses of crime has yet to begin, although it is sorely needed by police.

Practical Applications

Given the success of the CI in laboratory and field experiments, how does it fare in real-world investigations? Geiselman and Fisher (1997) reported several instances in which the CI was used successfully to solve real-world cases, ranging from a witness to a kidnapping, to a child molestation victim, to a witness of a politically motivated bombing. More recently, an investigator from the U.S. Bureau of Alcohol, Tobacco, and Firearms reported conducting an extended CI with a 38-year old woman who had witnessed a homicide as a 4-year old child. The interview elicited many recollections, most of which were corroborated by police records established at the time of the crime (e.g., location of objects and furniture at the crime scene). A victim who had been raped in her own home in Los Angeles recalled crucial information while using a combination of the reverse-order and change-perspectives techniques during a CI, after previous attempts with standard interview procedures had failed. During the CI, the interviewer asked her to mentally go through her house from different perspectives, ultimately leading to that portion of the memory record with critical investigative value. Similarly, after repeated questioning, a child victim in Los Angeles disclosed verifiable elements of being molested in his room only after the detective suggested that he take on the perspective of a stuffed animal that was on his shelf. We also have received some encouraging feedback from the National Transportation Safety Board, whose investigators received training on the CI and shortly thereafter conducted extremely effective interviews of crew members of the USS Greenville, the
American submarine that collided at sea with a Japanese fishing boat in February, 2001. The CI also has been credited for solving crimes when applied to suspects, although that was not an intended application of the original or enhanced CI. In one case from Texas, the suspect was attempting to avoid implicating his friend by describing a different person as his accomplice, but during the CI he momentarily forgot this deception and fully described his friend. In another case from California, the suspect described elements of the crime during the CI that only the killer could have known, whereas during prior interviews he said that he had seen nothing. This is not surprising in that the CI is designed to elicit large amounts of information from people using a rapport-based information-gathering approach, and sometimes deceptive persons are induced to say too much. Of course, we recognize that these reports are merely anecdotes, which are subject to many biases.

An additional recent application of the CI has been for investigations following use-of-force incidents involving police officers. Interviewing officers in these situations is a sensitive matter given that the officer is a witness, victim, and now in some sense is a suspect until the matter is cleared. State of mind and ongoing threat assessment are important with these law-enforcement witnesses, not just the material facts of the matter (Wilson & Geisleman, 2011). The CI is ideally suited for this circumstance given that it is an information-gathering protocol, rather than a confrontational approach, where thoughts and emotions are addressed. The application of the CI in this arena has been generally welcomed by policing agencies (Force Science News, 2011, #169).

Offsetting these successes, British police reported that the complete CI was sometimes difficult to implement (Kebbell, Milne, & Wagstaff, 1999; Kebbell & Wagstaff, 1996). Officers interviewed by these researchers reported that they found it difficult to communicate to witnesses some of the CI’s mnemonic instructions. Kebbell and Wagstaff (1996) note, however, that it may be possible to overcome these problems with suitable training. Other police officers interviewed about their experiences with the CI also have reported that using the complete CI requires more time than is sometimes available (Kebbell, Milne, & Wagstaff, 1999). This commentary reflects in part a mistaken belief that the CI must be implemented as a whole rather than as a toolbox. Nevertheless, in response to this practical concern, Davis, McMahon, and Greenwood (2005) set out to create a shorter version of the CI that would still capture most of the information gathered by the complete CI. Their laboratory-based research demonstrated that a pared-down version of the CI saved considerable time and yet was almost as effective as the complete CI. Dando et al. (2009a; 2009b) also have created modified versions of the CI for use by patrol officers in time-limited situations. We find this a healthy development in the progression of CI research, trying to make it more efficient and sensitive to real-world conditions.

The two meta-analyses of the available studies of the CI found greater effects of the CI with a shorter delay before the initial interview than following a two-day delay, but there still was a substantial advantage for the CI over control interviews following a delay. Therefore, the general recommendation would be to conduct the CI as soon as possible following an incident but that the CI protocol should be considered regardless of timing. None of the available studies on delaying the CI have included potentially significant extraneous factors such as sleep deprivation or lingering stress, either of which might favor a decision to delay the interview. The CI requires time and requires the witness’s full cooperation and exhaustive participation. It is reasonable to expect that a well-rested witness would produce more complete and more accurate recall than would a less-rested witness, especially following high-stress situations (Geiselman,
Unfortunately, the decision to delay the full investigative interview to allow for rest typically must be made based on surface indicators of witness stress. Therefore, following stressful incidents such as officer-involved shootings, it has been recommended that as a general rule, the involved officers should sleep first and give their statements later (Artwohl, 2002). Delaying the interview would not preclude a limited request to provide enough information to get the investigation started, perhaps using a shortened version of the full CI as noted above.

The CI for Suspects (CIS)

The original and enhanced versions of the CI protocol were developed for use with co-operative victims and eyewitnesses, but as noted in the previous section, the CI also has been used successfully to interview suspects in some real-world cases. We have learned about these cases by way of reports from those practitioners directly involved. These success stories typically have involved either of two scenarios: (a) the subject recalls and inadvertently reports details of the crime that only the perpetrator could have known, or (b) the subject recalls details that contradict details s/he previously reported to the police but that conform to an alternative police theory of the case. The CI is particularly suited for these kinds of revelations because the CI is a witness-centered approach that relies on open-ended questions and narrative responses. That is, the subject generates the information almost exclusively on his/her own rather than responding to leading, close-ended questions from the investigator.

Considerable research has shown that even with some training, distinguishing truthful from deceptive oral narratives based on verbal, vocal, and behavioral indicators alone is not an easy task, with accuracy often just above chance (Bond & Depaulo, 2008; Geiselman, Elmgren, Green, & Rystad, 2011; Geiselman, Musarra, Berezovskaya, Lustig, & Elmgren, 2013). Interactive strategies for detecting deception involving dyadic exchanges with the subject have proven to be more reliable (Dando & Bull, 2011; Granhag, Strömwall, & Hartwig, 2007; Vrij & Granhag, 2012). This is not surprising given that recent meta-analyses have shown that lie detection is more readily improved by increasing behavioral differences between liars and truth tellers than by informing lie-catchers of valid cues to deception (Bond & Hartwig, 2011). In light of these findings, the general CI protocol was merged with techniques from the research literature on detecting deception to produce the cognitive interview for suspects (CIS). The CIS allows for a non-judgmental approach to interviewing subjects in situations where it is not yet clear whether the subject will be honest or deceptive during the interview (Kassin, Goldstein, & Savitsky, 2003). The same principles that drive the memory-enhancement and communications aspects of the CI for co-operative victims and eyewitnesses are at work in the CIS. The subject is encouraged to generate large amounts of information before any challenge is made. This allows for a greater opportunity to observe inconsistent statements as well as any changes in demeanor from baseline (as observed during the rapport stage) should they occur. This approach reduces the likelihood that the investigator will challenge the subject pre-maturely based on any confirmatory bias (Hill, Memon, & McGeorge, 2008; Kassin, Dror, & Kukucka, 2013).

The CIS also employs two techniques from the regular CI for asking for information in ways that are unanticipated by the subject, namely drawing/sketching the story and re-telling the story in reverse chronological order. It is logical that deceptive persons must use some of their limited cognitive resources to maintain the integrity of the false story and to monitor the interviewer’s reactions to the story, whereas truthful persons need only to search memory. The sketch and reverse-recall elements of the CIS provide an opportunity to observe the subject’s performance while carrying out these unexpected tasks with further elevated levels of cognitive load.
(McCormack, Ashkar, Hunt, Chang, Silberkleit, & Geiselman, 2009; Vrij, Leal, Mann, Warmelink, Granhag, & Fisher, 2010). It is important to recognize that performance “signs” of deception are red flags only (sometimes called “hot spots”) and must be monitored throughout the entire interview by the interviewer toward making an accurate overall decision about truth versus deception (Geiselman, 2012). That is, some truthful persons will experience a heavy cognitive load to retrieve true memories and some deceptive persons will experience little cognitive load to confabulate false memories (Vrij, Leal, Granhag, Mann, Fisher, Hillman, & Sperry, 2009).

Step-by-Step Sequence of the CIS

The eight stages of the full CIS are:

Rapport/Introduction. As in the CI for co-operative witnesses, the interviewer first attempts to create rapport with the subject using casual conversation in a non-judgmental manner. In one study, prisoners who confessed during their police interviews reported that their interviewers had developed rapport with them whereas a dominating approach was met with resistance (Holmberg & Christianson, 2002). In addition, the interviewer will observe the subject’s general demeanor during this exchange as a baseline. A change in demeanor later on could indicate deception.

Narrative. Research shows that the vast majority of persons put in a position of being deceptive choose to offer only the highlights of their story or a “bare-bones” account (Colwell, Hiscock-Anisman, Memon, Woods, & Michlik, 2006; Geiselman et al., 2011, 2013). The typical justification for such an abbreviated narrative is that to be more elaborate would appear as an attempt to convince the interviewer, or to “sell” the story (Colwell et al., 2006). In contrast, the elements of the CI leading up to the request for the narrative raise the expectation that the subject will be detailed in the narrative. More information provides a greater opportunity to produce conflicting details and/or details that are incriminating or known to be false, yet interviewers often attempt to control the interview rather than encouraging the suspect to talk extensively (Snook, Luther, Quinlan, & Milne, 2012). To encourage this process, the interviewer should use “extenders” and other prompts to keep the narrative going as long as possible (“Really… tell me more about that.”).

Drawing/Sketch. The interviewer’s request for an illustration of the story following the narrative stage is unexpected by the subject. Research shows that unexpected requests can prove difficult for a subject who is being deceptive, in part because unexpected requests increase cognitive load (Vrij et al., 2009; Vrij et al., 2010). The request for an illustration should be presented to the subject as in the regular CI, as a means to clarify the narrative for greater understanding by the interviewer as well as to give the subject another opportunity to recall additional information. Deceptive persons are more likely than truthful persons to exhibit unusual difficulty in making the drawing/diagram, to leave out significant details from the narrative, produce more inconsistencies, and offer little, if any, additional new details compared with truthful subjects.

Follow-up, Open-Ended Questions. The follow-up, open-ended questions will be presented as information-gathering rather than as confrontational to maintain the momentum toward generating more information from the suspect. Deceivers typically answer these questions with minimal elaboration without offering much that is new in terms of details (Colwell, Hiscock-Anisman, Memon, Taylor, & Prewett, 2008). Truth tellers in contrast are likely to elaborate and provide additional information.

Reverse-Order Technique. When all of the scenes from the narrative have been exhausted with follow-up questioning as in the regular CI, the interviewer will introduce the reverse-order
technique as another means for possibly jogging the subject’s memory for additional details. Research shows that deceptive persons have unusual difficulty telling their fabricated stories backward compared with truthful subjects (McCormack et al., 2009; Vrij, Fisher, Mann, & Leal, 2009). Recalling a story backward increases cognitive load, and the deceptive subject’s cognitive resources already are being strained to maintain the consistency of the story.

**Challenge.** At this point, the subject will be asked about inconsistencies, incriminating statements, and/or external incriminating evidence. None of these elements should be addressed prior to this point in the interview (Granag, Strömwall, & Hartwig, 2007). Consistent with the regular CI, the interviewer should remain soft-spoken, respectful, and use pauses effectively to maintain the focus on the subject. Given that rapport has been developed and the interview has been conducted in an information-gathering style using “teamwork,” the interviewer should ask the subject to, “Help me understand this.” The confrontation process should be conducted in a “drip” (piece by piece) manner rather than in attempt to overwhelm the subject with all of the incriminating information at once (Dando & Bull, 2011). This procedure increases the likelihood for additional inconsistent statements because it does not allow for a single comprehensive explanation from a deceiver at one time.

**Review.** The interview summary should be framed as in the regular CI as an opportunity for the subject to correct any inaccuracies and to recall additional facts. With suspects, the interviewer has the option to intentionally change a non-incriminating element of the subject’s story to see if the subject spontaneously corrects the inaccuracy (Shafer, & Navarro, 2012). Deceptive persons are less likely than truthful persons to correct the inaccuracy, but rather quickly agree with the review, including the changed element of the story, in an attempt to end the interview as soon as possible.

**Close.** If the subject has appeared truthful, the interviewer will thank the subject for having co-operated. If the subject has appeared deceptive, the interviewer will express that s/he feels disappointed and disrespected (exploiting the connection established during the rapport stage). In the latter case, the interviewer will attempt to explain to the subject that it would be better for him/her to tell the truth now rather than later.

**Empirical Test of the CIS**

In the first empirical test of the CIS, 6 trained interviewers conducted interviews with 20 participants who were instructed to describe a recent autobiographical event and also separately to describe a completely fabricated autobiographical event (Geiselman, 2012). The stakes in this preliminary study were relatively low in that the participants could win fifty dollars if they convinced the interviewer that they were telling the truth, and they stood to lose fifty dollars if they failed. The interviewers rated the likelihood of the participants’ truthfulness at each of the first six stages of the CIS protocol. The results showed that the interviewers were only slightly better than chance at assessing deception following the narrative stage (consistent with past research on detecting deception), but increased accuracy systematically throughout the remainder of the CIS to an almost perfect level of discriminability. These findings provide an encouraging initial demonstration of the potential of the CIS protocol for assessing the likelihood of deception during investigative interviews. Further study of the CIS is warranted and welcomed.

**Future Directions**

A myriad of studies over the past two and a half decades has demonstrated clearly that the CI reliably enhances witness recollections under a variety of event and test conditions. While training on the CI thus far has been spotty with no coordinated effort by a national agency, numerous national and local law-enforcement agencies have sponsored training in several U.S.
states and in several countries. We are committed to delivering this training which is sought and needed by investigators who must interview witnesses about important events.

With respect to future research, it is important to identify the limitations as well as the strengths of the CI for different populations and contexts. It is equally important, however, to then offer guidance and to seek solutions for those limitations. We believe the major areas of progress in the future will include (a) developing training programs for investigators to learn the procedures more effectively, (b) streamlining the procedures so that police can use them more efficiently in field situations where resources are often limited, (c) developing new component techniques to expand the CI, (d) refining the extension of the CI for suspect interviews (CIS), (e) exploring the utility of the CI in other (non-criminal) investigations, and (f) modifying and refining the CI for use with persons exhibiting various disabilities. To end this chapter on a positive note, we are happy to report that considerable progress has been made along each of these lines. Powell, Fisher & Wright (2005) have explored various methods to improve training. Davis et al. (2005) and Dando et al. (2009a, 2009b, 2011) have examined ways to streamline the CI for time-constrained situations. Geiselman (2012) has completed the first laboratory test of the CIS, and Morgan et al. (2011) have found a variant of the CI to lead to deception-detection accuracy rates above 80% in ecologically valid settings involving US Army personnel. Chapman and Perry (1995), Brock, Fisher, and Cutler (1999), and Roos (2007) have examined how the CI might be used to investigate auto accidents. Finally, Maras & Bowler (2010; 2012) have pursued workable modifications of the CI for use with persons with certain disabilities. We encourage others to advance the research on the CI protocol for investigative interviewing in these and other directions. We also encourage practitioners to contact us with your case experiences involving the CI.
References


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 (Kohnken, 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; Kohnken et al., 1999; Memon, 2006). Field tests of the CI have also indicated that police officers trained in its techniques gain more information and
more detailed information from eyewitnesses in investigative contexts (Fisher, Geiselman & Amador, 1989; Clifford & George, 1996; Kebbell & Milne, 1998, cf. Fisher & Schriebert, 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,” “sug-
gestibility,” “age,” “crime,” “mock crime,” “memory distortions,” “person de-
scription,” “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.
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 a 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

### Table 1

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

<table>
<thead>
<tr>
<th>Witness population</th>
<th>Original CI</th>
<th>Enhanced CI</th>
<th>Modified CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students/pupils</td>
<td>15 (23%)</td>
<td>8 (12%)</td>
<td>23 (35%)</td>
</tr>
<tr>
<td>Civilians</td>
<td>6 (9%)</td>
<td>6 (9%)</td>
<td>5 (8%)</td>
</tr>
<tr>
<td>Police</td>
<td>0 (0%)</td>
<td>1 (2%)</td>
<td>1 (2%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Witness age</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Children</td>
<td>3 (5%)</td>
<td>2 (3%)</td>
<td>13 (20%)</td>
</tr>
<tr>
<td>Young adults</td>
<td>17 (26%)</td>
<td>12 (18%)</td>
<td>13 (20%)</td>
</tr>
<tr>
<td>Older adults</td>
<td>1 (2%)</td>
<td>1 (2%)</td>
<td>3 (5%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Interviewer background</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Students</td>
<td>1 (2%)</td>
<td>3 (5%)</td>
<td>7 (11%)</td>
</tr>
<tr>
<td>Researchers</td>
<td>16 (25%)</td>
<td>9 (14%)</td>
<td>19 (29%)</td>
</tr>
<tr>
<td>Professionals</td>
<td>2 (3%)</td>
<td>3 (5%)</td>
<td>3 (5%)</td>
</tr>
<tr>
<td>Written script</td>
<td>2 (3%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Control group</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard interview</td>
<td>15 (23%)</td>
<td>8 (12%)</td>
<td>6 (9%)</td>
</tr>
<tr>
<td>Structured interview</td>
<td>4 (6%)</td>
<td>6 (9%)</td>
<td>20 (31%)</td>
</tr>
<tr>
<td>Free recall</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>2 (3%)</td>
</tr>
<tr>
<td>No control group</td>
<td>2 (3%)</td>
<td>1 (2%)</td>
<td>1 (2%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Event duration</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;1 min</td>
<td>4 (6%)</td>
<td>1 (2%)</td>
<td>2 (3%)</td>
</tr>
<tr>
<td>&lt;5 min</td>
<td>11 (17%)</td>
<td>5 (8%)</td>
<td>14 (22%)</td>
</tr>
<tr>
<td>5–10 min</td>
<td>2 (3%)</td>
<td>5 (8%)</td>
<td>9 (14%)</td>
</tr>
<tr>
<td>&gt;10 min</td>
<td>4 (6%)</td>
<td>4 (6%)</td>
<td>4 (6%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Retention interval</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>None or very brief</td>
<td>5 (8%)</td>
<td>5 (8%)</td>
<td>10 (16%)</td>
</tr>
<tr>
<td>24–72 hrs</td>
<td>14 (22%)</td>
<td>4 (6%)</td>
<td>13 (20%)</td>
</tr>
<tr>
<td>1–2 weeks</td>
<td>2 (3%)</td>
<td>5 (8%)</td>
<td>4 (6%)</td>
</tr>
<tr>
<td>2–6 weeks</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>1 (2%)</td>
</tr>
<tr>
<td>1.5–6 months</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>1 (2%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Event medium</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Staged (live)</td>
<td>6 (9%)</td>
<td>5 (8%)</td>
<td>4 (6%)</td>
</tr>
<tr>
<td>Video</td>
<td>14 (21%)</td>
<td>10 (15%)</td>
<td>24 (40%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Event type</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Emotional/arousing event</td>
<td>17 (26%)</td>
<td>11 (17%)</td>
<td>13 (20%)</td>
</tr>
<tr>
<td>Neutral event</td>
<td>4 (6%)</td>
<td>4 (6%)</td>
<td>16 (25%)</td>
</tr>
</tbody>
</table>

*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 a 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; Kebell & Milne, 1998; Kebell, 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

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

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. Ginét 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 Ginét 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 dependent 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.
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 (SDs) 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.

### Table 3

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

<table>
<thead>
<tr>
<th>Recall measure</th>
<th># of articles</th>
<th>k</th>
<th>N</th>
<th>Weighted mean $d$</th>
<th>$p$-value</th>
<th>95% CI</th>
<th>$N_{FS}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct details</td>
<td>46</td>
<td>59</td>
<td>2,887</td>
<td>1.20</td>
<td>$&lt; .001$</td>
<td>(1.12, 1.28)</td>
<td>$&gt; 10,000$</td>
</tr>
<tr>
<td>Incorrect details</td>
<td>43</td>
<td>56</td>
<td>2,645</td>
<td>0.24</td>
<td>$&lt; .001$</td>
<td>(0.16, 0.32)</td>
<td>337</td>
</tr>
<tr>
<td>Confabulated details</td>
<td>29</td>
<td>33</td>
<td>1,940</td>
<td>0.08</td>
<td>$0.10$</td>
<td>$(-0.01, 0.17)$</td>
<td>—</td>
</tr>
</tbody>
</table>
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 ($Z_j$s $\leq 0.83, ns.$). 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.
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.


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


### Appendix

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

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

(Appendix continues)
## Appendix

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

*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