2012

The Effects of Repeated Questioning and Negative Feedback on Eyewitness Consistency after a Cognitive Interview (CI)

Lauren Wysman
University of Windsor

The Effects of Repeated Questioning and Negative Feedback on Eyewitness Consistency after a Cognitive Interview (CI)

Lauren M. Wysman

A Thesis
Submitted to the Faculty of Graduate Studies
through the Department of Psychology
in Partial Fulfillment of the Requirements for
the Degree of Master of the Arts at the
University of Windsor

Windsor, Ontario, Canada

2012

The Effects of Repeated Questioning and Negative Feedback on Eyewitness Consistency after a Cognitive Interview (CI)

by

Lauren M. Wysman

APPROVED BY:

_____

Dr. M. El-Masri
Department of Nursing

_____

Dr. D. Jackson
Department of Psychology

_____

Dr. A. Scoboria, Advisor
Department of Psychology

_____

Dr. S. Hibbard, Chair of Defense
Department of Psychology

September 6, 2012

DECLARATION OF ORIGINALITY

I hereby certify that I am the sole author of this thesis and that no part of this thesis has been published or submitted for publication.

I certify that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis and have included copies of such copyright clearances to my appendix.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

ABSTRACT

The Cognitive Interview (CI), an interview technique used with cooperative eyewitnesses of crime, has been shown to lead to the receipt of more correct information than control interviews, with stable errors and accuracy. The present study was conducted to determine if the CI conveys benefits protecting against the effects of problematic interview techniques such as repeated questioning and/or negative feedback. Undergraduates (n = 98) watched one of two crime videos and were interviewed with either a CI or a Free Recall. One week later, a second interviewer asked a set of questions. Half of the participants received negative feedback about their performance in questioning and all participants were then questioned a second time. Findings indicated that the CI was protective against inconsistencies due to repeated questioning but only in the absence of negative feedback. Relevance of the findings to investigative interviewing is discussed.

ACKNOWLEDGEMENTS

I would like to extend much gratitude to my supervisor, Dr. Alan Scoboria, whose knowledge, guidance, and patience have been invaluable for me to achieve my academic goals. I would also like to thank our collaborators, Dr. Amina Memon and Dr. Julie Gawrylowicz, my committee members, Dr. Jackson and Dr. El-Masri, and my lab members for their valuable feedback on this project. In addition, I would like to acknowledge the team of research assistants who have helped in collecting, transcribing, and coding this project's data.

Finally, I would like to thank those peers, friends, and family members who have been a source of reassurance and encouragement throughout this process. I must acknowledge Mia and Michelle for providing me with some much needed camaraderie and moral support as we each toiled on our respective projects. I must also extend the utmost thanks to my brother, Colin, for providing me with a healthy dose of wit and reminders of balance, and to my parents, Tom and Jane, for their wisdom, sacrifice, and relentless support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Overview

Imagine you have just witnessed a home invasion. You are interviewed by a police officer immediately afterward. You report many details about the invader; but, in a follow up interview a few days later, you are asked more specific questions about the crime. Since the police officer is not exactly sure about what you did or did not see, he/she asks some questions that you could not possibly answer. You may even start giving information that you are unsure about because you feel pressured to; the fact that you are being interviewed and asked some of the same questions for a second time may be indicative that you were not trying hard enough in the first interview. The police officer may tell you that the information you had given her/him was not sufficient to apprehend the criminal, and that everyone would really appreciate it if you could work even harder to remember everything this time around.

This interview scenario is not unlikely for eyewitnesses of real-life crimes. As reviewed below, research shows that questioning individuals in such a way may lead to changes in testimony that undermine the accuracy of information recalled and the credibility of witnesses in the courtroom. Repeated interviewing or questioning may or may not be recommended, depending on the interviewing techniques used. Some of these more problematic techniques also include asking pointed, specific questions and applying social pressure or negative feedback to encourage witnesses to "work harder" or to give more correct information.

Certain crimes may leave a trail of important evidence, such as videotapes of the

crime in progress or DNA. In many cases, however, crimes lack physical evidence and can only be solved with detailed information from eyewitnesses. Valuable eyewitness information can be lost in the face of poorly-conducted investigative interviews. Unfortunately, eyewitness memory, and memory processes in general, do not operate to record fixed and stable representations of events. As such, one cannot expect that memory functions in such a way that a video recording operates (Bartlett, 1932). Interviewees may only spontaneously report some of the details that they witnessed and they are rarely able to report every single detail of a previously witnessed event. Certain interview techniques may lead to the retrieval of more of this valuable eyewitness information; other less-than-ideal techniques may result in distorted or fabricated information.

The Cognitive Interview (CI) guideline systematically outlines techniques that enable witnesses to remember as much as they can, and deters interviewers from using potentially counterproductive techniques. The CI is considered to be one of the best practice investigative interviews for cooperative adult eyewitnesses (Fisher, 2010). Developed by Geiselman et al. (1984), it blends principles of cognition, social psychology, and knowledge of interviewing, and was designed as a systematic approach for police officers to interview cooperative witnesses. A recent study-space analysis (Memon, Meissner, and Fraser, 2010) found that the use of the CI results in a significant increase in correct information garnered from witnesses and a much smaller, non-significant increase in errors compared to various control interviews. Furthermore, the research indicates that the CI does not lead to increases in the amount of confabulated information provided by interviewees. In other words, the CI typically obtains more

information without changing the ratio of correct to erroneous information. As will be discussed below, the CI also provides extended protective benefits for memory in subsequent interviews. The purpose of the current study is to broaden the investigation of such extended benefits, by examining whether the CI provides protection against subsequent repeated questioning and the application explicit negative social pressure to change responses.

Some research does not appear to support the use of multiple interviews or repeated questioning within or across interviews. For example, when individuals repetitively recall information, it can lead to inflated confidence (Odinot, Wolters, & Lavender, 2009; Shaw, 1996). This can be problematic because witness confidence significantly affects not only whether mock jurors believe that a witness has made an appropriate identification of a culprit, but also the verdict that the mock jurors give for a case (Brewer & Burke, 2002; Cutler, Penrod, & Stuve, 1988). Repetitive recall can also exacerbate misinformation effects, whereby individuals incorporate post-event information into their later reports (Chan, Thomas, & Bulevich, 2009; Davis & Loftus, 2007). Memory for aspects of witnessed events that are not questioned can be inhibited in future recall attempts (Wright, Loftus, & Hall, 2001). Multiple recalls can also lead to "retrieval induced forgetting," whereby witnesses who are questioned multiple times about certain details may forget or fail to report other related details, even if there was no misinformation present during questioning (Shaw, Bjork, & Handal, 1995). Repeated questioning can also lead to changes in responses. For example, in one study of children who were repeatedly questioned, approximately one quarter of responses changed, leading to a decrease in witness accuracy (Krähenbühl, Blades, & Eiser, 2009).

Furthermore, adults who felt uncertain were more likely to speculate when repetitively questioned (Poole & White, 1991), especially after time had passed (Poole & White, 1993).

Therefore, a number of problems arise from repeating questioning. One is that the content of an answer may change (i.e., "shift") from one time point to another. Shifts can be desirable or undesirable (e.g., Howie, Sheehan, Mojarrad, Wrzesinska, 2004); an answer may change from an error to a correct response, or from a correct response to an error, etc. Despite the fact that a shift can objectively be "good" or "bad," it is also known that a witness who makes shifts (i.e., is not consistent) might be perceived as lacking credibility (Gilbert & Fisher, 2006). Whether shifts are good or bad is also rarely possible to determine in practice.

Gudjonsson, in his development of his interrogative Suggestibility Scale (GSS; 1984), theorized regarding two components of suggestibility: yielding when asked suggestive questions, and shifting responses after negative feedback. The latter is one of the foci of the present study. Gudjonsson noted that when pressure is placed on a witness via negative feedback, true responses that are undesirable to the interviewer (e.g., a witness saying that he or she does not know the answer to a question) may shift to untrue or speculative responses. This negative feedback can affect a witness' mood and behaviour (Gudjonsson, 2003). Using a negative feedback manipulation, Gudjonsson (1986) found that giving this feedback to participants led to increases in acquiescence; that is, changing answers in an affirmative direction that is consistent with interviewers' expectations. He argued that was due to an increased sense of uncertainty and decreased self-esteem. Moreover, the pressure to alter responses could be felt implicitly, simply by

repeating questions. Gudjonsson further argued that repeated questioning may be a type of implicit negative feedback (2003). Witnesses who are repetitively questioned about the same topic may begin to understand the implicit message that the interviewer thinks they have more information that they are simply not providing to the interviewer. This notion of repeated questioning as a form of negative feedback is supported by Register and Kihlstrom (1988) and replicated by Linton and Sheehan (1994) who found that asking witnesses a set of questions twice led to shifting of responses. Simply repeating questions may have encouraged participants to think that they had previously given erroneous or undesirable information, or that the questioner was seeking other information. The effects of both explicit pressure and implicit encouragement to shift responses have also been documented with children (e.g., Moston, 1987; Warren, Hulse-Trotter, & Tubbs, 1991).

The aforementioned research demonstrates that allowing witnesses multiple opportunities to recall events might be perceived as problematic. However, there are also benefits of repeated interviews. Reminiscence effects (i.e., remembering something at a later time that was not recalled before) are consistently found in tests about word-lists or pictures (Erdelyi & Becker, 1974; Roediger & Payne, 1982; Roediger, Payne, Gillespie, & Lean, 1982; Shapiro & Erdelyi, 1974). Gilbert and Fisher (2006) found that, no matter what condition participants were assigned to in their study, 98% of their overall sample made reminiscent statements. They also found that the number of consistent statements made by participants was not necessarily related to accuracy. But, the average accuracy of consistent information was significantly higher than both forgotten and contradictory information. La Rooy, Pipe, and Murray (2005) found that up to 39% of information given is new information when a high quality second interview follows shortly after a

first high quality interview. Therefore, not only is reminiscence a normal phenomenon, but conducting high quality second interviews is also advisable. Even with long lengths of time between the first and second interviews, the same amount of information may be recalled in a secondary interview (e.g., Ackil, Van Abbema, & Bauer, 2003), or the amount of information may even increase (e.g., Fivush, McDermott Sales, Goldberg, Bahrick, & Parker, 2004). As noted, though, by La Rooy, Lamb, and Pipe (2009), reminiscent information tends to be less accurate than information that is consistently recalled depending on lengths of delay. Therefore, researchers ought to examine consistency and reminiscence effects together rather than simply sheer quantity of information output at each session. While there are other reasons for encouraging repeated interviews (e.g., a victim is distressed at the time of the initial interview, the interviewee has a short attention span, the interviewee is a child), the mere fact that important information is reminisced in secondary interviews may be reason enough to encourage the use of multiple interviews.

Despite some debate, experts argue for repeated interviews if they are done well (i.e., that use methods from best practice interview techniques such as the CI, such as avoiding closed-questioning and allowing witnesses to go through a free recall) in order to elicit the most complete accounts possible (La Rooy, Katz, Malloy, & Lamb, 2010). If repetition leads to consistency, this is only useful if the information being repeated was accurate to begin with (Gilbert & Fisher, 2006). The costs that may come with repeated questioning arise from ineffective interviewing practices, not the act of conducting multiple interviews (La Rooy et al., 2010).

The effects of an initial CI followed by a second CI have also been examined.

Only three published studies have examined the effectiveness of repeated interviews with the CI with adults (Brock, Fisher, & Cutler, 1999; McCauley & Fisher, 1995; Memon, Wark, Bull, & Köhnken, 1997). In these studies, the delays between the initial viewing of the event in question and the actual initial interview varied from five minutes to two days; the second interview took place from between 10 to 14 days following the first. The studies cannot be directly compared because of differences in the types of CIs used (i.e., the techniques selected to use). However, overall, the studies seem to be indicative that the CI leads to more correct details when compared to control conditions at the time of the first interview. However, these studies do not conclusively indicate any added advantages to having a second CI. A question still remains of what occurs if a well-conducted interview is followed up by subsequent repeated questioning or makes use of social pressure or negative feedback in an attempt to get the desired information from the witness, as in the scenario outlined earlier. The CI may be a tool for addressing some of these concerns, since it is plausible that an initial, well-conducted interview may be followed up with poor techniques.

The quality of interviews can also be assessed by the types of questions used (e.g., appropriate vs. inappropriate, productive vs. unproductive) and the times at which those questions are used (Griffiths, 2012). An example of a potentially inappropriate question type is that of closed questioning (e.g., "Was the robber wearing a red shirt?"), which has been noted to be a poor questioning technique for eliciting information relative to open ended questions (e.g., "Please describe the robber"). Further to this, Fisher, Falkner, Trevisan, and McCauley (2000) note the benefits of open questioning. They found that techniques such as those in the CI led to the receipt of more and more precise

information. In this study, one important difference between the CI and a standard control interview is that that those interviewing with the CI asked more open ended questions. This led to the apparent superiority of the CI. They note that open ended questions permit a more elaborate response to the questions asked and create an implicit expectation for witnesses to provide elaborate responses and detail in the interview. Overall, closed questioning appears to lead to fewer and sometimes inaccurate answers (Fisher, Geiselman, Raymond, & Jurkevich, 1987). The CI's instruction to interviewers to use open ended questions and to avoid closed ended questions therefore usually leads to the receipt of more correct information from witnesses. However, despite these recommendations, open ended questions are at best inconsistently used in practice (Clarke & Milne, 2001; Lamb et al., 2002; Schreiber Compo, Gregory, & Fisher, 2012).

Furthermore, when individuals underreport, answer questions by saying "I don't know," or give less information than the interviewer is seeking, they may be encouraged to say more. Koriat and Goldsmith (1996) proposed a "quantity-accuracy trade-off model" that asserts that as motivation to respond increases, the amount of information will typically increase as well, but this increase often occurs at the expense of the accuracy of the information. Encouraging output leads to the provision of lower confidence responses and guessing. As individuals engage in a search of memory, they weigh candidate responses and eventually select a best candidate response. With this best candidate in mind, they then weigh the quality of this response against a response criterion which is affected by the costs associated with responding and not responding. Allowing witnesses to freely report their recollections allows witnesses to regulate their responding, since it is up to the witnesses to regulate the information they output; as

questions asked become more focused, interviewees may feel increased pressure to answer.

Another influence on responding is acquiescence (La Rooy et al., 2009), whereby individuals tend to agree with other people. When there is social pressure from interviewers, individuals may respond in ways that they otherwise would not. For example, Garven, Wood, and Malpass (2000) found that simple positive reinforcement of children's answers led them to make a significant number of false allegations. Children are more inclined to incorrectly acquiesce to misleading questions when they are facing social pressure (e.g., Greenstock & Pipe, 1997; Pipe & Wilson, 1994). This effect is exacerbated by differences in social power, which may explain why younger children are disproportionately more susceptible to make these errors than older children (Ceci & Bruck, 1993). Effects are still found for older participants; as noted earlier, studies have found that merely asking adults questions twice also produces shifting of responses (Register & Kihlstrom, 1988), possibly because of the implicit social pressure that is felt by witnesses and the inherent differences in social status between interviewer and interviewee.

Hence, repeating questions and applying explicit pressure frequently lead to the receipt of more information. However, such techniques also lead to lower overall quality of memory reports. Repeating questions, pressuring, or forcing a witness to respond leads to reduced accuracy in responding and unwarranted increases confidence in answers (Ackil & Zaragoza, 1995; Bruck, Ceci & Hembrooke, 2002; Memon & Vartoukian, 1996; Poole & Lindsay, 2001; Poole & White, 1991; Register & Kihlstrom, 1988). It is worth noting that in less than ideal interviews, interviewer confirmatory biases may drive

their questioning agendas. A goal and benefit of the CI is that it minimizes the ability for an interviewer's agenda to interfere with the interview.

Clearly, there are issues for eyewitness interviews conducted with techniques such as repetitive questioning, social pressure, and negative feedback. Because of the high quality report that results from an initial CI, it may offer protective effects for those witnesses whose interviews are followed up with poor interviewing techniques. While the efficacy of the CI has been demonstrated in the empirical literature, the effects of the CI in preserving memory in the face of subsequent poorly conducted repetitive interviews are not as well understood. This issue is relevant because, as noted, witnesses are typically interviewed multiple times by different interviewers (Tuckey & Brewer, 2003).

As detailed above, witnesses may feel pressured to respond by either being asked specific questions, or by facing subtle social pressure or direct negative feedback from interviewers. The CI has offered protective effects in other situations that place pressure on witnesses. Memon, Zaragoza, Clifford and Kidd (2010) found that the CI protected against the negative effects of forced confabulation, a procedure in which participants were told to fabricate details about certain components of a witnessed event. They noted that this procedure can lead to the creation of false memories about witnessed events. Older adults previously interviewed with a CI had been found to be less susceptible to misinformation effects (Holliday et al., 2011). If the CI offered protective effects in these examples, it may also be useful for buffering against the effects of repeated questioning and negative feedback. A preliminary CI may lead to positive carryover effects for witnesses who are questioned repeatedly and when witnesses are pressured to respond. The CI can lead to a more thorough retrieval process in memory and one's ability to

convey the contents of memory may be enhanced with the CI. Therefore, it was thought that it may offer protective effects for the problematic questioning scenarios of interest in this study.

The Present Study

The goals of the present study were to examine whether the CI offered protective effects to individuals who were repeatedly questioned with answerable and unanswerable questions at a later date. Specifically, the study aimed to examine if an initial CI would lead individuals to remain more consistent in the face of repeated questioning and explicit negative feedback when questioned. Past research indicates that these manipulations can lead to higher suggestibility, shifting of responses, and decreased witness credibility (e.g., Gilbert & Fisher, 2006; Goodman & Quas, 2008). Thus, every participant was questioned twice; however, only half of participants received negative feedback. Answerable and unanswerable questions were examined separately. Answerable questions are those for which a witness is able to give an answer, such that the information being asked of them was actually present to be witnessed. An example is, if in the video there is a clear image of the culprit of the crime, asking what colour shirt the culprit was wearing. An unanswerable question is a question for which a response such as "it was not in the video" was a correct answer; for example, a video clip commenced with a robber already in a house, committing a burglary. Asking how the robber entered the home is an unanswerable question; it forced the witness to guess the answer, reject the question, or indicate that he or she does not know. Question types are delineated as such and have been examined separately in this study because the process of recollecting each type of information differs. The Koriat and Goldsmith model (1996), as discussed earlier, can be

used to conceptualize some of the cognitive processes involved in searching memory for knowledge that one can remember (i.e., for answerable questions). Individuals search memory, weigh candidate responses, select their best choice and weigh the quality of this choice against a changeable response criterion. In contrast, Mazzoni and Kirsch (2002) argue that in those situations where individuals cannot retrieve a memory or belief from autobiographical memory (i.e., for unanswerable questions), they must assess whether this lack of memory is diagnostic. They may attempt to determine if they merely forgot the information. Or, if the event is non-distinctive or common, they may maintain that it might have occurred, but that they simply have forgotten it, since it is much more common to forget non-distinct events. Thus, if the lack of memory about an event is not diagnostic as to whether the event did or did not occur, the individuals in question may use inferential processes to determine the answer to a question or determine the likelihood of an event. It is for this reason that answerable and unanswerable questions are considered separately; the cognitive processes involved for both differ.

Young adult participants were interviewed at two time points about a video clip. One half of the participants were interviewed with the CI at Time One. The other half of participants underwent a Free Recall where they were asked to tell the interviewer what they could remember from the video. One week later, participants returned and were questioned by a different interviewer. A set of answerable and unanswerable questions were asked twice. Half of the participants were told "I'm going to ask you the same questions again." In contrast, those in the negative feedback condition were questioned once, and then told, "You have made a number of errors. It is therefore necessary to go through the questions once more, and this time try to be more accurate," as per the

instructions in Singh and Gudjonsson (1984).

To examine the effects of repeated interviewing, the number of responses that shifted between the first and second questioning were examined. Group differences (e.g., CI vs. Free Recall, No Negative Feedback vs. Negative Feedback, etc.) were also examined. The nature of shifts (whether they are from correct to error, error to correct, etc.) was examined, as well as correct information, errors, accuracy, and output (i.e., the number of substantive responses an individual gives) in interviews and the first question set to examine whether the CI offered protective effects for later questioning.

Hypotheses

See Table 1 and 2 for hypotheses. Hypotheses were delineated as "main" or "supporting" hypotheses. Supporting hypotheses, while interesting, were expected based on prior findings in the literature and are relevant for placing the main hypotheses into context. Main hypotheses, in contrast, were those that reflected the repeated questioning and negative feedback components of the study.

Main hypotheses. The dependent variable for the three primary main hypotheses was the consistency of responses between the two questioning periods at Time Two. Note that both consistency and shifting of responses are discussed; one is the inverse of the other, in that the more consistent the participant, the fewer shifts he or she made.

The first main hypothesis was that the CI would lead to higher rates of consistency of responses in general, whether or not negative feedback was provided. This was because the CI would lead to more thorough retrieval and better consolidation of the event in memory. This effect (i.e., lower rates of shifting/higher consistency after a CI) was predicted to be most evident for unanswerable questions, whereby less shifting

Table 1

Main Hypotheses

Hypothesis	IVs	DVs	Statistical Test
Those initially interviewed with CI will have significantly fewer shifts overall than those interviewed with a Free Recall for unanswerable questions and answerable questions (with larger effect sizes for unanswerable).	Interview manipulation (CI/FR)	Consistency	t-test
Those who receive negative feedback will shift more than those who do not (examined separately for answerable and unanswerable questions; expect larger effect sizes for unanswerable questions).	Feedback manipulation (No negative feedback/Negative feedback)	Consistency	t-test
Effects of negative feedback will be stronger in the Free Recall condition than for the CI (for answerable and unanswerable questions, with larger effect sizes for unanswerable questions).	Interview manipulation (CI/FR) Feedback manipulation (No negative feedback/Negative feedback)	Consistency	Interaction in 2x2 ANOVA
Consistent responses will have higher confidence ratings than inconsistent responses.	Consistency (Consistent/Inconsistent)	Confidence	t-test

Table 2

Supporting Hypotheses

Hypothesis	IVs	DVs	Statistical Test
CI will lead to more correct information at Time One than Free Recall.	CI/FR	Amount of correct information	t-test
CI at Time One will lead to higher accuracy in Question Set 1 (Q1) for answerable questions.	CI/FR	Accuracy of answerable questions	t-test
CI at Time One will lead to more correct responses in Q1 than Free Recall (examined separately for answerable and unanswerable questions).	CI/FR	Correct Responses	t-test
CI at Time One will lead to higher confidence in Q1 than Free Recall (examined separately for answerable and unanswerable questions).	CI/FR	Confidence	t-test
CI at Time One will lead to more DK responses for unanswerable questions in Q1.	CI/FR	DK responses to unanswerable questions	t-test
CI at Time One will lead to fewer DK responses for answerable questions in Q1.	CI/FR	DK responses to answerable questions	t-test
CI at Time One will lead to higher accuracy for answerable and unanswerable questions in Q1.	CI/FR	Accuracy	t-test

would occur for those who originally had a CI. Shifting in general was predicted to be lower for answerable questions; however, the CI was still predicted to possibly help to enhance the material that one holds in memory. Therefore, this effect for answerable questions was predicted to potentially be significant, but with smaller effect sizes than for unanswerable questions. With respect to unanswerable questions, it was hypothesized that the CI may aid individuals to identify that that lack of memory is diagnostic. This would lead to more initial rejections, more DK responses, and/or fewer errors to unanswerable questions and importantly, less shifting.

The second main hypothesis was that negative feedback would lead to higher rates of shifting of responses. Two outcomes were predicted to potentially be observed. First, while it has been noted that mere repetition of questions is a form of implicit social pressure, negative feedback is explicit and ought to lead to an increased sense of uncertainty and therefore increase proneness to shifting. Thus, one can predict that there would be a significantly larger amount of shifting for unanswerable questions and answerable questions for those who are provided with negative feedback, with larger effect sizes for unanswerable questions. A second potential outcome is that when questions were repeated, participants may have assumed that it was because they needed to change answers. Thus, it is possible that there might not be significant differences between those who received negative feedback and those who did not.

Third, it was anticipated that the CI and Negative Feedback conditions would interact. While there were anticipated main effects for interview type and negative feedback manipulation, it was hypothesized that that the effects of negative feedback would be stronger for those in the Free Recall condition, rather than those in the CI

condition. This is because those in the Free Recall condition would not have gone through the enhanced recollective experience of the CI; and, those who were given negative feedback would feel more uncertain than those who were not. See Figures 1 and 2 for a depiction of the predicted patterns in average consistency of responses. As per the previous hypotheses, effect sizes were expected to be larger for unanswerable questions.

Further to these hypotheses focused on main effects and an interaction effect for consistency, t-tests were planned to examine whether consistent responses received higher initial confidence ratings than inconsistent responses. If those responses that were consistent had significantly higher initial confidence ratings, this might serve to explain why some responses shifted while others did not.

Supporting hypotheses. The first supporting hypothesis was that the CI would lead to the provision of more correct information at Time One than those who did not receive the CI. Accuracy was predicted to remain stable in comparing the CI to the Free Recall and it was predicted that there would not be significant differences in the number of erroneous statements made. This hypothesis would serve to confirm that the CI was indeed effective in leading to superior performance of participants, as frequently demonstrated in prior research.

The effects of the CI on responding to the first question set were examined. Based on prior work, it was anticipated that the CI would result in higher accuracy for answerable questions, more correct answers, and higher confidence than those who had engaged in a Free Recall (second, third, and fourth supporting hypotheses). Those who were interviewed with the CI were predicted to have a lower number of responses to

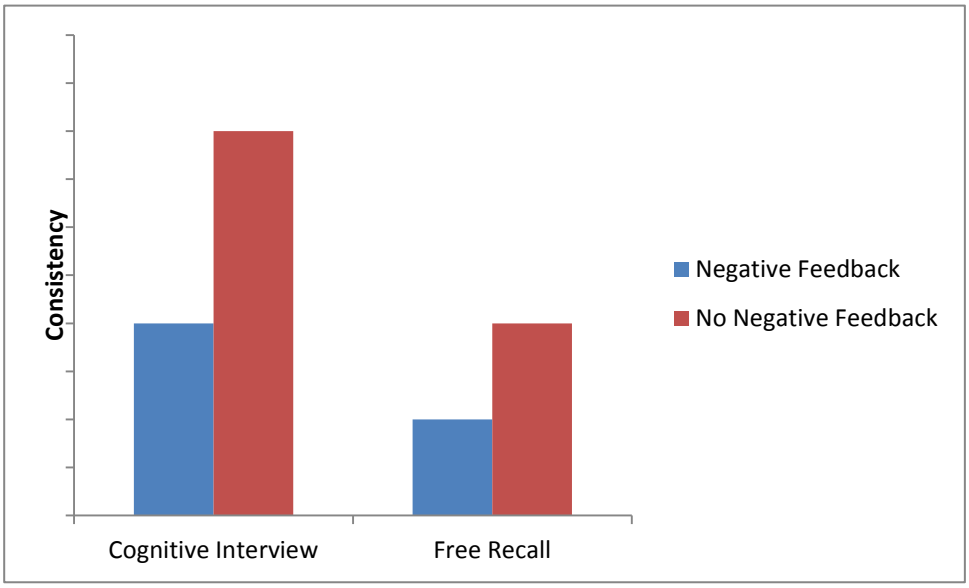

Figure 1. Anticipated patterns of consistency of responses for unanswerable questions.

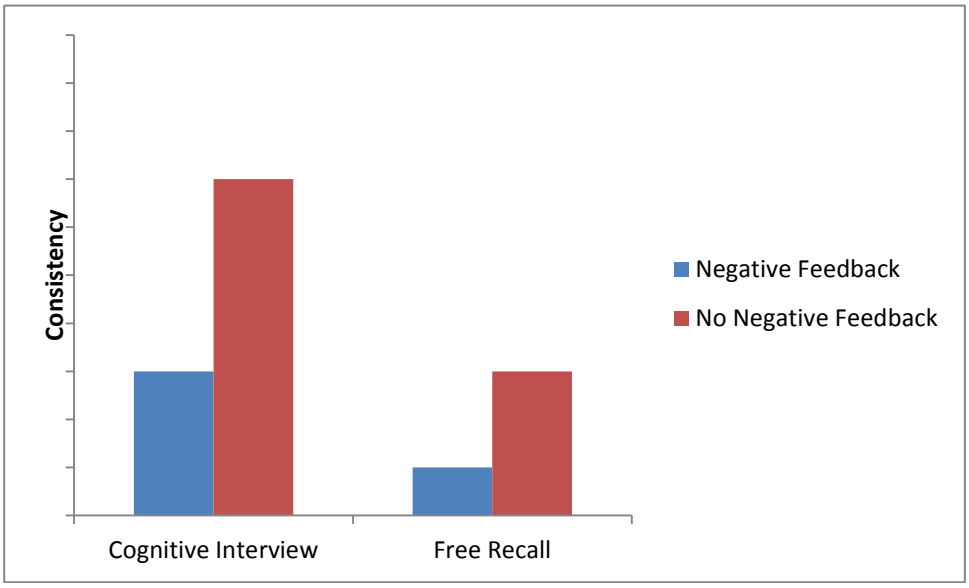

Figure 2. Anticipated patterns of consistency of responses for answerable questions.

unanswerable questions, seen as a higher number of "don't know" (DK) responses. It was also thought that they might have a lower number of DK responses to answerable questions (supporting hypotheses five and six). If the CI contributed to not only memory but also had metacognitive benefits (i.e., benefits to how one thinks about one's own thinking), more broadly, one should see higher accuracy for both answerable and unanswerable questions than the Free Recall group (supporting hypothesis seven), and higher output for answerable questions, as delineated above. Overall, if the CI aids participants in being more consistent and therefore resist the pressures that come with repeated questioning, the value of the CI is reinforced.

CHAPTER II

METHOD

Participants

Memon, Meissner, and Fraser (2010) outlined average effect sizes for the CI compared to control interviews in terms of differences in correct details, *weighted Cohen's d* = 1.21, *95% CI* =1.12, 1.28. It was decided to use the lower bound of the confidence interval to be more conservative. The lower bound of this confidence interval, $d = 1.12$, was converted to an *f hat* statistic (*f hat* = 0.56) using stat-help.com spreadsheets adapted from Cohen (1988). Using this *f hat*, an ideal sample was understood to be 87 participants, calculated with G*Power (Faul, Erdfelder, Lang, & Buchner, 2007). The effect size and therefore power necessary for the negative feedback manipulation, and the examination of response shifting is not readily apparent because many of the studies examined were within-subjects designs with small samples. Therefore, to be conservative, it was planned to collect data from approximately 100 participants in order to have sufficient statistical power. Participants were recruited through the University of Windsor's Psychology Department's Participant Pool and received academic credit. One-hundred and five participants completed session one. Seven participants did not attend the second session. The final sample consisted of 98 participants (79.6% female, 19.4% male, 1 missing data; age range: 17-56 years, M = 23.15; 76.5% Caucasian).

Design

The study is a two by two between-subjects ANOVA. The two independent variables were interview method (i.e., Cognitive Interview vs. Free Recall) and feedback

condition (i.e., No Negative Feedback vs. Negative Feedback). T-tests and ANOVA procedures were both used as appropriate.

Participants were randomly assigned (i.e., randomized to which condition they would receive) to one of two conditions: Cognitive Interview (CI) or Free Recall (FR). They were also randomly assigned to whether they received or did not receive negative feedback following the first questioning in the second session. The use of Video One or Video Two was randomly counterbalanced among participants (see Figure 3 for assignment of participants and design of study).

Materials and Measures

Video. The videos were developed and were in use at Royal Holloway University of London, United Kingdom, for studies on the CI. Participants were randomly counterbalanced to watch one of two videos. The first video depicted a man tampering with a young woman's drink in a bar, ending with a suggested date-rape. The second video depicted an elderly man being robbed while he was in another room in his home. The films were not graphic, but were credible; participants therefore were alerted to the fact that they were permitted to withdraw their participation at any point in time throughout the study. The films were mild in their content, for example, relative to popular television crime dramas.

Distractor task. Participants completed word scramble tasks (see Appendix B) for 30 minutes in between the video and the interview (CI or Free Recall). This was to prevent participants from actively rehearsing the contents of the video clip in the time leading up to the interview.

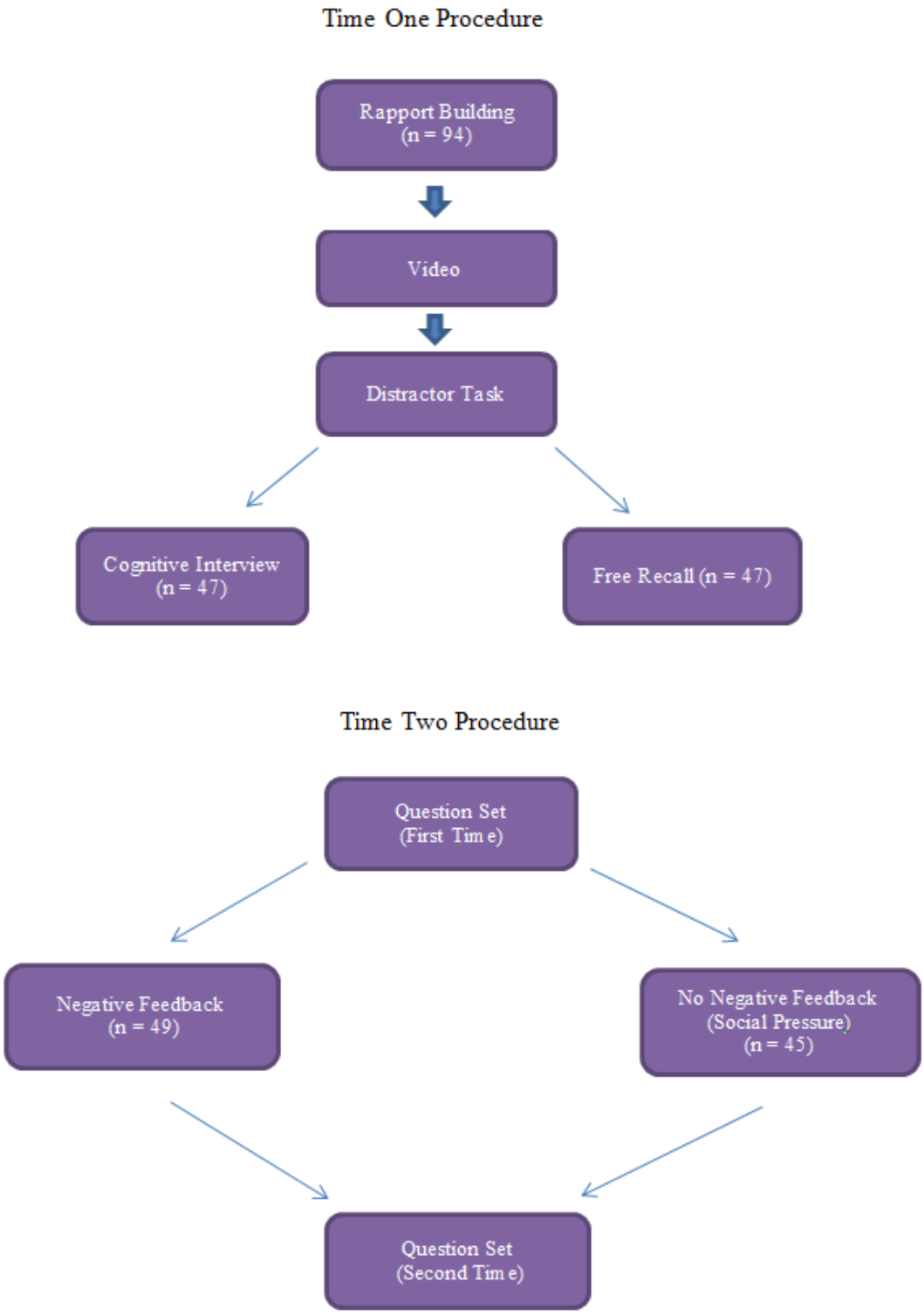

Figure 3. Experimental design of study.

The Cognitive Interview (CI). The CI has different iterations: the original CI, initially developed by Geiselman et al. (1984), the Enhanced Cognitive Interview (ECI; Fisher & Geiselman, 1992), and the Modified Cognitive Interview, developed out of the ECI. The MCI is modified as a researcher/interviewer sees fit and therefore, which CI techniques are used in a study is inconsistent between many research studies. The present study maintained use of certain parts of the CI, making it a Modified CI.

All sessions began with rapport building in an attempt to make the participant feel comfortable. Then, for those assigned to the CI, the purpose of the interview was explained. This was followed by a mental focus instruction in which the participant was asked to focus and concentrate on retrieving information. The interviewer told the interviewee to report everything, even if it seemed unimportant, and conveyed that he or she (i.e., the participant) was in control of the interview. The interviewer also instructed the witness not to guess. She also instructed the witness to say "I don't know" or to ask for clarification if the participant was ever confused. Then, the process of mental context reinstatement commenced, whereby the interviewer encouraged the witness to bring himself or herself back to the time at which he or she encoded the original crime video, and then had the witness engage in a free recall in which he or she reported everything he or she could think of. The interviewer then proceeded with specific questioning about certain pre-selected, broad topics (e.g., the people, prominent objects, and location of the scene in the video clip). Throughout this process, the interviewer avoided interrupting the witness or changing the course of the interview by probing with questions that were not directly related to the topic that the witness was focused on; when a witness has a particular image in mind, details relating to that image alone are most readily available

(Fisher & Schreiber, 2007). Therefore, the interviewers knew that they must not hastily change the direction of the conversation. See Appendix C for instructions used.

Free recall. The free recall commenced with the same rapport building used with those interviewed with the CI. Participants were then asked to "please tell me what you can remember about the video clip." They were given as long as they needed to complete this task, usually in the same amount of or less time than it takes to be interviewed with a CI.

Question set. At the second session, participants were asked a standard set of questions. There were 24 questions for each video: 14 were answerable and 10 were unanswerable for each, in order to remain consistent with question sets used at Royal Holloway University in London. See Appendix D for questions. Unanswerable questions in the present study were not explicitly misleading, in that they did not suggest specific answers within the question. They were suggestive only in the sense that, in some cases, participants were asked about information that was not actually in the video clip. Therefore, these questions mirror many real-life interview settings, in which interviewers typically do not know what the witness really knows or witnessed.

Dependent measures. The dependent measures in the first session came from the recording and transcription of the session. A master list of relevant facts was put together by researchers at Royal Holloway University for the videos. The number of correct and incorrect details were enumerated, and accuracy was calculated. Coding was conducted by trained independent raters.

At Time Two, the dependent measures came from the answers to questions in both question sets one and two. Correct, error, or don't know (DK) responses were coded

for answerable questions. Correct rejections, errors, or DK responses were coded for unanswerable questions. These dependent measures (e.g., correct responses, errors, etc.) were examined separately by type of question (i.e., answerable and unanswerable). Further to this, the number of questions answered, accuracy rates, and confidence were examined separately for each type of question for those who had a CI compared to those who had a Free Recall. All coding at Time Two was conducted by trained independent raters.

The amount of shifting (i.e., consistency) was measured between the first and second times the questions were asked. Response change from Time Two Question Set One to Time Two Question Set Two was calculated. Several types of shifts were made: correct to error; correct to DK; error to correct; error to DK; DK to correct; and, DK to error. The examination of the type of shift was exploratory.

Confidence ratings. At Time Two, participants rated their confidence for each response they made on a scale of 0 to 100. Ratings were made for each question set.

Social desirability scale. At the end of Time Two, participants filled out a social desirability scale. With this, it could be assessed if individuals who aimed to appear in a socially desirable way made more shifts between answers. Stöber's (2001) Social Desirability Scale-17 was selected. It correlates highly with other social desirability measures, impression management, agreeableness, and conscientiousness. It also correlates highly with the Marlowe-Crowne Scale, with the exception of Stöber's oldest participants, but has significantly smaller age effects than the Marlowe-Crowne Scale.

Procedure

All procedures were reviewed by the University of Windsor Research Ethics

Board. Participants signed up for two sessions, separated by approximately one week. The Time One session lasted approximately one hour; the Time Two session lasted approximately 30 minutes. In the first session participants viewed the video and were interviewed with the CI or Free Recall procedure. The interview was audio recorded for the purposes of coding the information output by participants. Participants were randomly assigned to watch one of the two designated videos. The intent of using both videos was to ensure that effects found in the study were not due to use of one video. The videos were shown on a 26 inch high quality video monitor. Before watching the videos, the interviewer engaged in a brief period of rapport building, usually consisting of discussions about school or plans for the weekend. After this, participants' attention was directed to the video monitor. They were told to pay close attention, and that they would be asked questions about the video after. Then the interviewer left the room for the participant to watch the video alone. After the video ended, they engaged in a 30 minute filler task (i.e., a word unscrambling task). After this, half of the participants were randomly assigned to be interviewed using the CI and the other half engaged in the Free Recall procedure. Participants were then thanked and reminded of their next session in the following week.

At Time Two, participants were asked the set of answerable and unanswerable questions by a different interviewer in the same room. They were given instructions adapted from previous work (Fisico & Scoboria, unpublished manuscript): "I'll be asking you some questions about the video-clip that you watched last session. This might help you to remember more of the details of the video. Even though you might have already given us the information, please answer every question to the best of your ability. Please

also indicate your level of confidence for each question that you answer: 0% means not confident at all and 100% means you are very confident in your answer." They were asked the question set and gave their confidence ratings for each question. Then, half of the participants were randomly assigned to be asked the exact same set of questions without receiving any feedback, simply being told, "I'm going to ask you the same questions again." The other half of the participants were asked the same set of questions but first were told, "You have made a number of errors. It is therefore necessary to go through the questions once more, and this time try to be more accurate." After being asked the set of questions and receiving confidence ratings for the second time, for purposes unrelated to the current study, all participants were asked to return to every question for which they said that they did not know the answer, and were asked to guess or confabulate an answer and then rate their confidence. After this, participants filled in the social desirability scale, were given a letter of information (Appendix E), and were thanked for participating.

Interviewers and Interviews

Interviewers. The interviewers conducting the first session were fully trained in the use of the CI. The interviewers for the second session were trained to administer the questions and to transcribe answers verbatim. Different interviewers were used for Time One and Time Two. Other studies (e.g., Bjorklund et al., 2000) found that those who were interviewed by a different interviewer at their second session had more incorrect recognition than those interviewed by the same interviewer. Issues of familiarity with the interviewer have not been examined in great depth for the CI. Odinot, Memon, La Rooy, and Millen (unpublished manuscript) found no differences in the number of correct

details conveyed, based on the use of same or different interviewers; both interviewers were similar in age, sex, and appearance, but had different accents (i.e., one Dutch, one Scottish). Furthermore, the current procedure is an analogue to a typical real-world scenario in which a good interview is followed by a poor interview at a later time by another person. Based on this and the fact that the second interview made use of poor interviewing tactics, different interviewers were used for Session One and Session Two of the study.

Scoring of Interviews. Interviews at Time One were audio recorded and transcribed verbatim. Certain pieces of text were ignored in the transcripts: un-measurable subjective statements (e.g., "He was ugly") and utterances (e.g., "Uhhh, I think," "like"). The information was separated by unit of information per standard procedures for coding free recall narratives (e.g., Gilbert & Fisher, 2006). For example, "The man had brown hair and was wearing a blue hat" would be divided into "The man" "had brown hair" "and was wearing a blue hat." This information was compared to a master list of facts for each video about characteristics of people in the video, actions, objects, and locations throughout the video. Items were coded as correct information, errors, confabulations, or suppositions, and accuracy was calculated from that (correct information divided by total information provided). Inter-rater reliability was calculated for a subset of transcripts by dividing the number of coding agreements by the number of coding disagreement per transcript (Tinsley & Weiss, 2000).

CHAPTER III

RESULTS

Coding

Cognitive Interview and Free Recall. Twenty transcripts (i.e., 20.4% of sample) were double coded, with a strong intraclass correlation coefficient among the number of correct details coded for these transcripts (r = .88).

Time Two Questioning. Data for forty participants (i.e., 40.8% of sample) at Time Two were double coded. Consistency rates were calculated as the proportion of exact agreement between the raters (Question Set 1: 0.94; Question Set 2: 0.93) and deemed to be reasonable. From inspection of the coding, one question from the second video was recoded for every participant due to differences in rater interpretation of the coding manual.

Data Cleaning and Preparation

The assumptions of Analysis of Variance (ANOVA) were tested with two-tailed tests and several iterations of data cleaning were engaged in. The assumption of independence of observations was met in the experimental design of the study, which included random assignment to all conditions, and by the participant recruitment procedures (i.e., opening up the study to nearly all participants of the Psychology Department's Participant Pool). Before the removal of outliers, the normality and homogeneity of variance assumptions of ANOVA were analyzed. Because of the very large number of dependent variables and their interdependence, examination was focused primarily on the dependent variables for the main hypotheses, and the key dependent variables for the CI effect. Thus the dependent variables examined were the percentage of

consistent answerable questions, the percentage of consistent unanswerable questions, confidence in answerable questions that remained the same, confidence in unanswerable questions that remained the same, confidence in answerable questions that shifted, confidence in unanswerable questions that shifted, accuracy at Time One, total correct information output at Time One, and total errors output at Time One.

When all participants were included, all skewness and kurtosis values for these variables were in appropriate ranges (i.e., 2 to -2 skewness, 3 to -3 kurtosis), with the exception of one group on one variable (i.e., the Free Recall, Negative Feedback group for percentage of consistent answerable questions). Of the variables, 10 of 72 Shapiro-Wilk scores (calculated for 4 conditions by 18 dependent variables), were found to be significant, indicating potential violations of normality for these cells. Regarding homogeneity of variance, Levene's test was significant for one variable (i.e., percentage of consistent answerable questions), indicating that group variances were not sufficiently similar. For all variables, when comparing groups the largest variance did not exceed the smallest above a ratio of 4:1, so the analyses are likely robust to violations. Group sizes were also roughly equal (i.e., Condition 1, Free Recall/Negative Feedback: 26 participants, Condition 2, Free Recall/No Negative Feedback: 23 participants, Condition 3, Cognitive Interview/Negative Feedback: 25 participants, Condition 4, Cognitive Interview/No Negative Feedback: 24 participants), aiding robustness.

Outliers were then evaluated. Several attempts at outlier removal were made. Assumptions were reassessed and results calculated for each iteration of data removal and compared to the assumption findings and results of the study with all cases included. One attempt included removing all cases with z-scores exceeding a ±2.5 cut off (Kirk, 1995)

for any dependent variable, regardless of whether it was for a main hypothesis, supporting hypothesis, or exploratory analysis. This led to finding 9 outliers. Being 9.1% of the sample, it was decided that this was too considerable a portion of the sample. A second attempt looked at five outliers: one univariate and four multivariate outliers on those specific dependent variables that were analyzed for normality and homogeneity of variance. The final attempt involved an examination of only those four multivariate outliers. After reviewing the assumptions and the results with and without these four multivariate outliers and the other iterations of outlier removal, it was decided to remove only these four multivariate outliers (i.e., 4.1% of the sample). Their removal led to a reduction in the number of significant Shapiro-Wilk scores (i.e., from 10 to 8), and improved the kurtosis value for the variable demonstrating the percentage of consistent responding to answerable questions. The Levene's value for this same variable remained significant. This limitation will be considered further in the discussion. With these outliers removed, the group sizes were: Free Recall/Negative Feedback: 24, Free Recall/No Negative Feedback: 23, Cognitive Interview/Negative Feedback: 25, and Cognitive Interview/No Negative Feedback: 22.

The assumptions of Analysis of Covariance (ANCOVA) were also assessed to determine whether social desirability was suitable as a covariate in the analyses. While social desirability had a significant relationship with one of the independent variables (i.e., interview manipulation; $p = .033$), it correlated significantly with just one of the dependent variables (i.e., total correct information at Time One). This lack of correlation between social desirability and the dependent variables indicates that it is not relevant as a covariate in the analysis. Further examination indicated that it also did not meet the

assumption of homogeneity of regression slopes, demonstrating that the relationship between the dependent variable of interest and the potential covariate is not consistent across the different experimental levels, further reinforcing that ANCOVA is not the desired method for the data and, therefore, that social desirability is not a relevant covariate in this analysis. One can assume that the pattern of responding in this data set is not significantly influenced by social desirability.

Video was examined as a potential factor in the analyses. Inclusion of video as a factor did not reveal any notable effects. Thus, video was not included as a factor in the results reported below.

Main Hypotheses

Cohen's d was calculated as the primary effect size in this study, due to the suitable standard deviations (i.e., no group's standard deviation exceeding a comparison group at a ratio of 4:1), the nature of most comparisons (i.e., t-tests), and the fact that many of these comparisons were outlined before the study was executed (Fritz, Morris, & Richler, 2012). For interactions in the study, partial omega squared (*partial* ω^2) was calculated so as to not overestimate effect sizes compared to other effect size measures, such as eta squared.

Hypotheses one through three involved main effects and the interaction between the independent variables (i.e., interview condition, feedback condition) when predicting consistency of responding across the two questionings. See Table 3 below for means and standard deviations of the dependent variables pertaining to consistency. The first hypothesis, that those interviewed with a CI would be more consistent than those interviewed with a Free Recall was not significant for either question type. This indicated

Table 3

Descriptive Statistics for Consistency Findings

Dependent Variable	*Question Type*	*Free Recall*		*Cognitive Interview*	
		No Feedback *n = 23* *M (SD)*	*Neg. Feedback* *n = 24* *M (SD)*	*No Feedback* *n = 22* *M (SD)*	*Neg. Feedback* *n = 25* *M (SD)*
Percentage of Consistent Responses between Q1 and Q2*	Answerable	.88 (.12)	.86 (.13)	.93 (.07)	.74 (.17)
	Unanswerable	.83 (.17)	.79 (.17)	.86 (.13)	.79 (.16)
Average Q1 Confidence in Consistent Responses	Answerable	65.60 (13.54)	65.45 (14.62)	64.06 (11.78)	66.54 (12.60)
	Unanswerable	76.76 (11.51)	74.25 (15.20)	75.92 (15.30)	78.29 (16.96)
Average Q1 Confidence in Shifted Responses	Answerable	59.43 (27.78)	37.23 (21.43)	55.56 (29.10)	55.05 (23.24)
	Unanswerable	63.29 (27.20)	60.01 (28.47)	62.61 (15.79)	73.94 (21.65)
Average Overall Confidence	Answerable	64.50 (14.33)	62.61 (15.79)	63.56 (11.30)	63.51 (13.41)
	Unanswerable	74.80 (12.85)	71.80 (17.00)	75.40 (13.36)	73.72 (17.29)

Note. *Mean percentage represented as decimal (e.g., .90 = 90%). Question set was composed of fourteen answerable and ten unanswerable questions.

no significant group differences in consistency when looking at whether participants were interviewed with a CI or a Free Recall. In contrast, for the second hypothesis there were group differences in consistency for answerable questions pertaining to whether one received negative feedback; specifically, those who did not receive negative feedback shifted fewer responses than those who did (*Mean consistency* = .80, *SD* = .16 vs. .91, *SD* = .10), $F(1, 92) = 13.67$, $p < .001$, $d = .77$. No significant effect was found for the same test with unanswerable questions. For the third main hypothesis, a significant interaction was found between interview type and feedback for consistency for answerable questions, $F(1, 90) = 9.19$, $p = .003$, *partial* $\omega^2 = .08$. Inspection of the interaction (see Figure 4) indicates that the Free Recall groups did not differ in terms of whether they had received negative feedback (*contrast estimate for the Free Recall group comparison* = -.022, $p = .563$). In contrast, those who were interviewed with a CI and who did not receive negative feedback showed the highest consistency of the four groups, whereas the CI group that received negative feedback showed the lowest consistency (*contrast estimate for the CI group comparison* = -.183, $p = .001$, $d = 1.46$). Thus, the hypothesis was only partially supported; the CI group outperformed the Free Recall groups in terms of consistency when no negative feedback was present (*contrast estimate of CI/NF vs. FR/NF and FR/No NF* = .058, $p = .046$, $d = .57$). However, what was not consistent with the original prediction was that the CI group provided with negative feedback was the most inconsistent of the four groups. Pertaining to the final main hypothesis, significant differences for Time Two, Question Set One's confidence ratings emerged between responses that were consistent versus inconsistent across the two questionings. Individuals who did not make any shifts were excluded from this analysis.

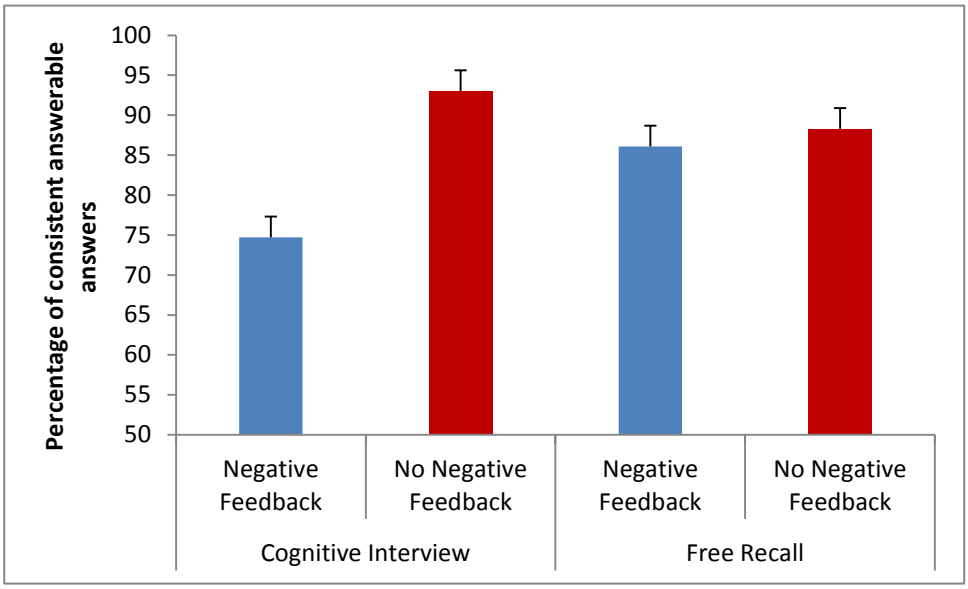

Figure 4. Patterns of consistency of responses for answerable questions between Question Set One and Question Set Two.

Note. Standard error as error bars.

Confidence was higher for consistent responses, for both question types: answerable, $t(55) = 4.25$, $p < .001$, $d = .65$; unanswerable questions, $t(58) = 3.63$, $p = .001$, $d = .57$. As predicted, average confidence was higher for consistent responses for both types of questions.

Supporting Hypotheses

See Table 4 for descriptive statistics pertaining to Time One interviews and Table 5 for the first questioning at Time Two. Of the supporting hypotheses, two yielded interesting results. Importantly, the results pertaining to the first supporting hypothesis were found to be significant. Those who were initially interviewed with a CI produced more correct information ($M = 63.21$, $SD = 15.64$ vs. $M = 45.51$, $SD = 13.55$), $t(92) = 5.87$, $p < .001$, $d = 1.21$, with a stable number of errors ($M = 5.15$, $SD = 3.03$ vs. $M = 4.06$, $SD = 3.25$), $t(92) = -1.68$, $p = .097$, and stable accuracy ($M = .86$, $SD = .08$ vs. $M = .85$, $SD = .10$), $t(92) = .67$, $p = .50$. Finding this effect is important to show that those who were interviewed with a CI actually output information in the way that is typical for those who have been interviewed with a CI.

Second, it was found that those participants initially interviewed with a CI had more correct responses ($M = 4.43$, $SD = 1.70$ vs. $M = 3.57$, $SD = 1.78$) to answerable questions in the first question set at Time Two than those interviewed with a Free Recall (i.e., supporting hypothesis three, $t(92) = 2.37$, $p = .02$, $d = .49$), indicating some benefits for the CI before observing the effects of negative feedback. All other supporting hypotheses (i.e., regarding interview group differences in the Time Two Questioning before negative feedback and repeated questioning) had non-significant group differences. In addition, a significant group difference was found between CI and Free

Table 4

Descriptive Statistics for Time One, Interview Performance for Cognitive Interview versus Free Recall

Dependent Variable	*Free Recall*		*Cognitive Interview*	
	No Feedback n = 23 M (SD)	*Neg. Feedback n = 24 M (SD)*	*No Feedback n = 22 M (SD)*	*Neg. Feedback n = 25 M (SD)*
Total Output	53.78 (11.95)	54.08 (18.21)	78.55 (18.82)	69.40 (16.93)
Total Correct	45.00 (11.63)	46.00 (15.40)	66.95 (16.47)	59.92 (14.41)
Total Error	4.17 (8.51)	3.96 (3.59)	5.50 (3.13)	4.84 (2.97)
Time One Accuracy	.84 (.10)	.86 (.10)	.85 (.08)	.87 (.08)

Note. Question set was composed of fourteen answerable and ten unanswerable questions.

Table 5

Descriptive Statistics for Time Two, Question Set One

Dependent Variable	*Question Type*	*Free Recall*		*Cognitive Interview*	
		No Feedback *n =23* *M (SD)*	*Neg. Feedback* *n = 24* *M (SD)*	*No Feedback* *n = 22* *M (SD)*	*Neg. Feedback* *n = 25* *M (SD)*
Correct Responses	Answerable	3.70 (1.46)	2.46 (2.06)	4.55 (1.50)	4.32 (1.87)
	Unanswerable	1.74 (1.51)	1.29 (1.63)	2.05 (1.84)	1.68 (1.91)
Don't Know Responses	Answerable	4.65 (2.67)	4.13 (2.56)	3.59 (2.34)	3.40 (2.51)
	Unanswerable	3.57 (2.00)	3.67 (2.28)	4.14 (1.96)	4.00 (2.42)
Erroneous Responses	Answerable	5.65 (2.52)	6.42 (2.70)	5.82 (2.52)	6.28 (3.01)
	Unanswerable	4.70 (1.49)	5.04 (1.99)	3.81 (1.65)	4.32 (2.21)
Accuracy	Answerable	.41 (.16)	.35 (.20)	.46 (.17)	.42 (.20)
	Unanswerable	.24 (.20)	.18 (.21)	.32 (.25)	.25 (.24)

Note. These results are one week following the interview (CI or Free Recall), but preceding the administration of negative feedback. Question set was composed of fourteen answerable and ten unanswerable questions.

Recall groups in the number of errors made to unanswerable questions, whereby the participants interviewed with a CI made fewer of these errors (M = 4.09, SD = 1.97 vs. M = 4.87, SD = 1.75), $t(92) = 2.05$, $p = .043$, $d = .42$.

Exploratory Analyses

See Table 6 for descriptive statistics for exploratory analyses of the number of shifts from one answer type to another answer type. When analyzing the types of shifts, group differences were found only for shifts for answerable questions from correct to erroneous responses. Both main effects were significant: CI vs. Free Recall, $F(1, 90) = 4.42$, $p = .038$, d = .44, Negative Feedback vs. No Negative Feedback, $F(1, 90) = 8.98$, $p = .004$, d = .61. A statistically significant interaction qualified both main effects, $F(1, 90) = 5.78$, $p = .018$, *partial* $\omega^2 = .05$. The CI/Negative Feedback group showed more shifts of this type than the other three groups, which did not differ. This coincides with the prior finding that the CI/Negative Feedback group was the least consistent of the four conditions for answerable questions. It is also of interest to note that overall there were more shifts involving "don't know" responses (i.e., 181 shifts) compared to any other shift that did not include a "don't know" response (i.e., 79 shifts).

The planned analyses found that those interviewed with a CI and who received negative feedback were the most inconsistent. This led to the consideration of whether this shifting was problematic. As will be discussed later, shifting of any type can be troublesome in courtrooms. A witness who changes his or her answers, regardless of the answers' correctness, may be viewed as inconsistent and therefore less credible. Despite this, the current findings led to the question of whether or not there were group differences in "good" or "bad" shifts when considered altogether, as opposed to looking

Table 6

Descriptive Statistics for Exploratory Analyses

Dependent Variable	*Question Type*	*Free Recall*		*Cognitive Interview*	
		No Feedback *n = 23* *M (SD)*	*Neg. Feedback* *n = 24* *M (SD)*	*No Feedback* *n = 22* *M (SD)*	*Neg. Feedback* *n = 25* *M (SD)*
Correct to Error Shifts	Answerable	.13 (.34)	.21 (.59)	.09 (.29)	.80 (1.00)
	Unanswerable	.04 (.21)	.21 (.51)	.23 (.53)	.20 (.50)
Don't Know to Error Shifts	Answerable	.57 (1.31)	.38 (.58)	.32 (.65)	.48 (.92)
	Unanswerable	.35 (.71)	.38 (.65)	.32 (.65)	.36 (.70)
Error to Don't Know Shifts	Answerable	.26 (.54)	.29 (.69)	.09 (.29)	.48 (1.19)
	Unanswerable	.30 (.56)	.17 (.48)	.09 (.29)	.28 (.68)
Error to Correct Shifts	Answerable	.17 (.39)	.13 (.34)	.09 (.29)	.40 (.65)
	Unanswerable	.04 (.21)	.13 (.34)	.14 (.47)	.28 (.61)
Correct to Don't Know Shifts	Answerable	.09 (.29)	.08 (.28)	.05 (.21)	.12 (.33)
	Unanswerable	.22 (.52)	.29 (.69)	.09 (.29)	.28 (.61)
Don't Know to Correct Shifts	Answerable	.09 (.29)	.25 (.53)	.09 (.29)	.16 (.47)
	Unanswerable	.17 (.49)	.13 (.34)	.23 (.53)	.20 (.50)
Consistent Correct Responses	Answerable	3.43 (1.38)	3.00 (1.91)	4.32 (1.35)	3.36 (1.98)
	Unanswerable	1.35 (1.30)	.67 (1.09)	1.64 (1.79)	1.16 (1.43)
Consistent Erroneous Responses	Answerable	4.83 (2.76)	5.54 (2.67)	5.41 (2.37)	4.32 (2.54)
	Unanswerable	3.91 (1.78)	3.96 (1.76)	3.36 (1.62)	3.24 (1.59)
Consistent Don't Know Responses	Answerable	4.00 (2.66)	3.50 (2.32)	3.18 (2.30)	2.76 (2.37)
	Unanswerable	3.00 (2.07)	3.13 (2.11)	3.59 (2.06)	3.40 (2.48)
Correct to Correct Shifts	Answerable	.04 (.21)	.17 (.48)	.09 (.29)	.04 (.20)
	Unanswerable	.13 (.34)	.04 (.20)	.09 (.29)	.04 (.20)
Error to Error Shifts	Answerable	.30 (.70)	.46 (.72)	.18 (.39)	1.08 (1.26)
	Unanswerable	.39 (.72)	.71 (1.04)	.23 (.53)	.48 (1.05)

Note. Question set composed of fourteen answerable and ten unanswerable questions.

at each individual type of shift, as was done above. Thus, three approaches were taken to explore this. First, shifts away from accuracy and towards accuracy were examined by grouping shifts as "good" and "bad." A second approach was looking at shifts *to* errors, corrects, or don't know responses from any other response. A third approach taken involved examining groups of responses that shifted *away* from correct, from errors, or from don't know responses.

The first approach was to examine "good" and "bad" shifts by looking at groups of desirable and undesirable shifting. Here good shifts were defined as the sum of shifts from errors to either correct responses or don't know responses considered together, and bad shifts as the sum of shifts from either correct responses or don't know responses to errors considered together. Significant Interview Type by Feedback Condition interactions were found for both good shifts, $F(1, 90) = 4.14$, $p = .045$, *partial* $\omega^2 = .03$ and bad shifts, $F(1, 90) = 4.76$, $p = .032$, *partial* $\omega^2 = .04$ for answerable questions. Visual inspection of the interactions indicated that the CI/Negative Feedback group made more shifts than the other three groups. Thus further approaches to examine types of shift were undertaken.

The second approach was used to determine whether inconsistency was due to shifts to being more correct, to making more errors, or to don't know responses between groups. Thus, the rates of shifting from anything to errors, anything to correct responses, and anything to don't know responses were examined. In this set of tests, the only significant finding was an interaction for the number of shifts from any response to an erroneous response for answerable questions, $F(1, 90) = 4.76$, $p = .032$, *partial* $\omega^2 = .04$. A significant contrast, $F(1, 90) = 7.62$, $p < .05$ indicated that this interaction was driven

by the CI/Negative Feedback group which made more of these shifts (*M* =1.28, *SD* =1.40) compared to the other groups combined (*combined M* = .57, *combined SD* = .95).

Last, using the third approach, an analysis of group differences in initial answers that were more prone to shifting was also examined; more plainly, this set of analyses focused on correct responses shifting to any other response, errors shifting to any other response, or don't know responses shifting to any other response. A significant interaction indicated that correct answers were more apt to shift to anything else for answerable questions, $F(1,90) = 6.75$, $p = .011$, *partial* $\omega^2 = .06$. Again, a significant difference, $F(1, 90) = 21.54$, $p < .05$ appeared between the CI/Negative Feedback group (*M* =.92 , *SD* = 1.00) compared to the other groups (*combined M* = .22, *combined SD* = .48), indicating that the CI /Negative Feedback group shifted its initial correct answers to any other response (i.e., to errors or don't know responses) for answerable questions more than the other groups. There was also a significant interaction for responses to answerable questions that shifted from errors to any other type of response, $F(1, 90) = 4.14$, $p = .045$, *partial* $\omega^2 = .03$. Again, the group that drove this interaction was the CI/Negative Feedback group (*M* = .88, *SD* = 1.20) compared to the other three groups (*combined M* = .35, *combined SD* = .68), with the CI/Negative Feedback group making more of these types of shifts when contrasted with the other three groups, $F(1, 90) = 6.99$, $p < .05$.

Based on these additional exploratory analyses, it appears that the CI/Negative Feedback group was more apt to make both problematic (i.e., any answer shifting to an error, correct answers shifting to any other answer) and some less problematic shifts (i.e., errors to any other answer) for answerable questions.

Also of note, exploratory comparisons were made between answerable and unanswerable questions to assess why most of the effects in the study were found for only answerable questions. It was determined that there were significantly more correct answers output, $t(92) = 10.65$, $p < .001$, and significantly higher accuracy, $t(92) = 5.58$, $p < .001$, for answerable questions when compared to unanswerable counterparts at the first point of questioning at Time Two.

CHAPTER IV

DISCUSSION

The CI is well-established as a tool for eliciting large amounts of correct information in investigative interviews. The current study corroborated this by finding an anticipated CI effect (i.e., more correct information with stable error rates and accuracy), and went further to explore its effects on subsequent repeated questioning. Furthermore, before any repeated questioning occurred (i.e., the first time participants were questioned about the video clip), those interviewed with a CI output more correct answers at Time Two for answerable questions and avoided errors to unanswerable questions. Thus, one can see further benefits of using the CI as an interviewing tool; not only did it lead to the output of more correct information in the initial interview, but it also led to the higher quality responding to questions one week later.

As noted earlier, in real-world interview settings, interviewers might be required to question a witness multiple times. Furthermore, they might engage in potentially problematic interviewing techniques such as giving negative feedback in questioning scenarios. As anticipated, those responses that remained consistent, whether in the face of negative feedback or not, had higher confidence ratings compared to the responses that shifted. This indicated that the higher the confidence, the less apt the respondent was to change the answer, even when these problematic interviewing techniques are used.

Further to the other predicted results, a significant interaction was found for answerable questions that remained consistent when examining interview condition and feedback condition together. However, the pattern of results for this interaction was not as initially predicted. Specifically, of interest, was that the CI group that did not receive

negative feedback performed the best of the four groups, in that these participants were the most consistent in responding to repeated answerable questions. Hence, it appears that the thorough memory processing and retrieval tasks involved in the CI led to benefits not only for initial questioning, but also for those times when interviewees were repeatedly questioned about the same topics. In contrast, the CI group that received negative feedback was the group that showed the lowest consistency in responding across the two questionings. This type of difference (i.e., between feedback groups) was not observed in the groups interviewed with a Free Recall. Thus, negative feedback appeared to have a stronger effect on consistency of responding for individuals who had been interviewed with a CI. A possible explanation for this finding is that those individuals who were initially interviewed with the CI may have understood that they engaged in a very thorough memory search and reporting process, moreso than those who engaged in a Free Recall. Thus, upon being told that they made errors, these participants may have felt cognitive dissonance (i.e., they had thought they performed well), leading them to change more responses than any of the other groups.

A question may then be asked about whether these shifts were problematic, in terms of changes in content. As noted earlier, it is possible that shifts can be "good" (e.g., an error changing to a correct response) or "bad" (e.g., a correct response changing to an error). Exploratory analyses indicated certain types of shifts in both of categories occurred more frequently in the CI/Negative Feedback group. Specifically, this group had significantly more "bad" shifting in terms of any answer shifting to an error, and correct answers shifting to any other answer. They also had significantly more "good" shifting in terms of errors shifting to any other answer. Therefore, not all the shifts made by those in

the CI/Negative Feedback group were entirely problematic in terms of content; there were changes in the quality of both towards and away from being accurate, a type of distinction noted by Howie, Kurukulasuriya, Nash, and Marsh (2009).

It is also worth noting that there were many more responses that either began as or shifted to a "don't know" (DK) response than those which shifted to and from more substantive responses (i.e., correct responses or errors). Koriat and Goldsmith's (1996) model would suggest that those responses that are initially of a quality below the participant's response criterion would be withheld, with the participant outputting DK responses instead. Perhaps the repeated questioning or the negative feedback provided to participants served to lower the response criterion, making participants apt to shift these responses to substantive answers. Or, repeated questioning and/or negative feedback might have led to changes in monitoring of memory by prompting a further memory search and evaluation, which could lead to changes in confidence in potential responses. Thus, an explanation for the number of shifts to and from DK could be based on a change in response criterion or in the way memory is monitored due to the implicit pressures of repeated questioning or the explicit negative feedback manipulation in the study.

In the context of the Koriat and Goldsmith (1996) model of metacognitive monitoring, if a response criterion shifted due to negative feedback, one would expect to see higher output in general, and hence more of both correct and erroneous responses. In the current study, output did not change across the repeated questionings and shifts were more likely to be toward errors in the CI/Negative Feedback group. Such shifting to erroneous information might be better explained by a change in the quality of monitoring of the contents of memory. After negative feedback, the monitoring processes of

participants in the CI/Negative Feedback condition might have altered to accept "noise" (i.e., errors) in lieu of appropriate responses. In other terms, this group appears to have been less able to discriminate signal (correct responses) from noise (any other information). This further reinforces why the provision of negative feedback in interviews is unwise.

A set of predicted findings that were not observed in the data were those pertaining to unanswerable questions; no significant group differences were found for unanswerable questions barring two effects (i.e., higher confidence for consistent unanswerable responses; fewer errors made to unanswerable questions at Time Two, First Questioning for those interviewed with a CI). Thus, there were not many group differences due to repeated questioning or negative feedback for unanswerable questions. Perhaps the unanswerable questions in the study were challenging for every participant, no matter the type of initial interview; note that there were significantly more correct responses output to, and significantly higher accuracy for answerable than unanswerable questions. Further, if repeated questioning and negative feedback do not impact output, then one would not expect to see many effects for unanswerable questions. As noted earlier, a distinction exists between memory for occurrence and memory for non-occurrence, involving differing memory processes (Mazzoni & Kirsch, 2002). In the present study, it appeared that memory for occurrence was the variable for which the group differences in the processing and output of information became pronounced; the effects of repeated questioning and negative feedback were observed more prominently for answerable questions. This is not to make the argument that there is no room for improvement for participants in properly answering unanswerable questions; one saw, for

example, that the Free Recall groups made more errors to unanswerable questions, and that unanswerable responses held in higher confidence did not shift as much as lower confidence responses. Perhaps if one asked more obviously unanswerable questions, significant group differences would be observed with respect to consistency when repeatedly questioned or given negative feedback.

Another finding worth considering is the lack of difference between the Free Recall groups, in that the provision of negative feedback did not lead to significantly more shifting than simply repeating questions did. This lack of difference does not necessarily indicate that the negative feedback did not have an effect. Rather, one could consider that in these groups, merely repeating questioning without a substantial explanation as to why might have indicated to participants that they did, in fact, make many errors, and that the expectation of the interviewer was that they would make some changes to improve their answers. In contrast, those who were provided with negative feedback after a Free Recall might not have felt the effects of this feedback as strongly since they did not experience themselves as having provided thorough information during the initial interview with a supportive interviewer. Regardless of the cognitive processes underlying this finding, the results of the present study are not intended to advocate for the use of less thorough interview techniques (i.e., a free recall) in order to avoid the risks associated with the provision of negative feedback. Rather, the use of empirically validated interview techniques and the avoidance of negative feedback in interviews is encouraged.

The findings of this study have applications in real-world settings. The results indicated memory benefits for individuals interviewed with the CI when later questioned

repeatedly. However, these benefits are partially undercut by the provision of negative feedback. This effect is counter to the original prediction that the CI would be protective in the face of negative feedback due to the thorough initial recollective process that the interviewee underwent. This finding reinforces the problematic nature of an interviewer explicitly passing judgment on the quality of an eyewitness' report. If urging a witness to be more accurate leads to shifting, this can lead to a perception of that witness lacking credibility. Considering the value that the current legal system places on consistency, a good interview technique (i.e., the CI) can look bad simply because of the shifts a person might make after being told that he or she must work harder to give more correct information. Thus, even if a witness output more correct information than someone interviewed with a less thorough technique, and continued to output more correct information in repeated questioning, the mere fact that he or she shifted responses can lead to the appearance of lacking reliability and credibility. As observed in this study, the provision of negative feedback was the key variable that contributed to this shifting. Thus, while negative feedback is problematic in interviews, perhaps the current legal system's interest in consistency is equally problematic. As noted earlier, it is not uncommon for someone to output information at a later time that he or she failed to either remember or state earlier. Perhaps the legal system's focus should move from witness or victim consistency to a focus on the quality of interview and appropriate questioning techniques the interviewer used in his or her attempt to retrieve accurate information from a witness or victim.

While any inconsistency could be problematic for appearances of witness credibility, one can contextualize the types of shifting that took place in this study.

Specifically, as noted above, those in the CI/Negative Feedback group had significantly more specific kinds of both "good" and "bad" shifts compared to the remaining three groups. They changed more responses of all types, and did not solely generate more errors. Thus, one cannot say that all shifting is bad shifting; however, one must consider that police interviewers do not have as much knowledge about a crime as the coders did for the crime videos employed in this study. Therefore, if a real-life interviewee shifts his/her answer from an erroneous one to a correct response, this "wavering" on his/her part may be perceived as problematic, regardless of the content of his/her answer. While this insistence on consistency may not be the most defensible position, police interviewers should consider that their provision of negative feedback may lead to otherwise credible witnesses or victims appearing inconsistent, even if they had been initially interviewed with an empirically-supported technique such as the CI.

Limitations

Limitations of the study include challenges with homogeneity of variance for one of the key variables (i.e., percentage of consistent answerable responses). Furthermore, results of the exploratory analyses should be considered tentatively until they are replicated to assess their stability. Also, the present study made use of a university-based sample. Further replication with different demographic groups is advised. A further limitation is that the research was conducted in a lab-based setting; the participants were likely not as emotionally aroused by the crime videos as they would have been as real victims or witnesses of crime. They also were interviewed in a university-based laboratory setting, not in a video-taped police interview room. They were not given an explanation regarding legal procedures the way one would in a police interview. No

crime was committed against them; thus, they would not have had the same vested interest in providing information as they might have had it been their homes robbed, for example. All participants also engaged in a 30 minute distractor task between the time they watched the crime and the time they were interviewed with a CI or Free Recall. While this was important to mimic the fact that witnesses are rarely interviewed immediately after a crime, victims or witnesses may wait more than 30 minutes to engage in a thorough interview. However, what the study lacks in external validity, it makes up for with the internal validity of being a well-controlled experimental study with random assignment to condition. Results of this study must be considered with caution and replication is encouraged.

Future Directions

Replication in contexts that are more similar to real-life crime contexts (e.g., emotionally-arousing events) is recommended. The present study was also conducted with a sample of young adults attending university. Examination of the efficacy of the CI for later questioning and the effects of negative feedback in questioning with different populations (e.g., the elderly, individuals with learning challenges) is recommended since these groups are thought to potentially encode and retrieve memories in a different way.

Further studies might also refine understanding about negative feedback by examining gradations of social pressure or negative feedback. The present study included explicit, verbal negative feedback which predictably led to shifting of responses. Not all interviewers in the real world use such direct types of feedback. An examination of more nuanced types of feedback (e.g., changes in facial expressions, in the tone of conversation, etc.) may be warranted in the future. Considering that the CI/Negative

Feedback group was the least consistent of the groups pertaining to consistency for answerable questions, it might be of particular interest to note if this susceptibility to shifting persists with less direct or obvious feedback. It might also be of interest to attempt to separate which component(s) of the CI contribute to this susceptibility to shifting.

Another future step that could help in clarifying the results of the present study would be a post-interview assessment of the interviewees' perception of their experiences in the interview and in the primary and secondary questioning. In doing so, one could assess if individuals interviewed with a CI actually perceived their experience to be more thorough than those interviewed with a Free Recall; if they actually believed that they made as many errors as might have been suggested; and, if using more nuanced forms of social pressure or negative feedback, whether or not they felt the pressure from the interviewer to shift their responses. In future, which interviewer asks the questions a second time can be manipulated; as noted by Howie et al. (2004), when a different interviewer repeats a question, it has the potential to reduce the perception that the original questioner merely was unhappy with the first answer.

Furthermore, in this study, the questions asked, while not "closed," were specific. Lamb and Fauchier (2001) noted that contradictions did not arise when broad questions were asked in real-life interviews; rather, directive open questions, such as ones asked in the present study, did elicit inconsistencies. Future examination of different types of repeated questioning after a CI or a Free Recall may be warranted. Another factor that contributes to quality of memory reports is time delay between encoding and interview/questioning. Thus, a manipulation of time delay would be of interest as well.

Finally, as noted above, there was a lack of difference in consistency between the two Free Recall groups. This may have been because repeating questions for this group without negative feedback might have been perceived by participants as an indication of having made mistakes, which is comparable to the direct, explicit negative feedback given to the other group of participants. Of interest would be a future study examining if group differences appear between three groups: a No Negative Feedback and Negative Feedback group, as per the present study, and a Positive Feedback group. Instructions to this group might indicate that the participants had many correct responses, but that the interviewer is still required to ask the questions a second time. Perhaps it will take actually telling participants that they are performing well for them to avoid making shifts in repeated questioning.

Conclusion

The present study exists as a first step in filling the gap in the literature regarding following up best-practice interviews with later repeated questioning and negative feedback. It demonstrated that the CI can serve to protect interviewees from being inconsistent when asked questions repeatedly. However, upon the receipt of negative feedback, individuals interviewed with a CI made more shifts than any other experimental group in this study. These shifts were both towards and away from accuracy, thus indicating that the negative feedback did not clearly lead to only problematic or only good shifts. However, in the context of a legal system that values consistency, this propensity to shift responses may lead individuals to be perceived as lacking credibility in the courtroom.

APPENDICES

APPENDIX A

Consent Forms

CONSENT TO PARTICIPATE IN RESEARCH

Title of Study: Eyewitness memory for crime

You are asked to participate in a research study conducted by Lauren Wysman (Master's Candidate) under the supervision of Dr. Alan Scoboria from the Psychology Department at the University of Windsor as part of the principal researcher's Master's thesis.

If you have any questions or concerns about the research, please feel to contact Lauren Wysman (*email address removed*) or Dr. Alan Scoboria (*email address and phone number removed*).

PURPOSE OF THE STUDY

The purpose of the study is to examine how individuals remember eye-witnessed events.

PROCEDURES

If you volunteer to participate in this study, you will be asked to watch a short video and answer questions about it. You will be asked to participate at two time points. The first session will take one hour, and the second session will take thirty minutes. The sessions will be one week apart.

POTENTIAL RISKS AND DISCOMFORTS

There are no known risks associated with this research.

POTENTIAL BENEFITS TO SUBJECTS AND/OR TO SOCIETY

You may benefit from increased knowledge about research in psychology. This research will expand knowledge about the function of memory and interviewing eyewitnesses, and may lead to improved knowledge in the best ways to gather information.

COMPENSATION FOR PARTICIPATION

Participants will receive 1 bonus point for their participation in session one, and 0.5 bonus points for their participation in session two; for a total of 1.5 bonus points; if enrolled in the psychology participant pool and a course that offers bonus points. You must attend both sessions to receive the full amount of credit.

CONFIDENTIALITY

Any information that is obtained in connection with this study and that can be identified with you will remain confidential and will be disclosed only with your permission. Upon being credited on the Participant Pool and once recorded portions of the procedures are transcribed, the information that you provide will no longer be associated with your identity and will not be linked to you in any manner. After transcription audio recordings

will be deleted. Data is stored securely and can only be accessed by the investigators and members of their research teams. Audiotapes of interviews will not be associated with your name, and will be stored on a computer without access to internet. Data will be retained indefinitely for research purposes.

PARTICIPATION AND WITHDRAWAL

You can choose whether to be in this study or not. If you volunteer to be in this study, you may withdraw at any time without consequences of any kind. You may also refuse to answer any questions you don't want to answer and still remain in the study. The investigator may withdraw you from this research if circumstances arise which warrant doing so. Once the study is completed your identity is not associated with the data and cannot be withdrawn.

FEEDBACK OF THE RESULTS OF THIS STUDY TO THE SUBJECTS

Results will be available in approximately December 31 2012 at www.uwindsor.ca/reb

SUBSEQUENT USE OF DATA

This data will be used in subsequent studies.

RIGHTS OF RESEARCH SUBJECTS

If you have questions regarding your rights as a research subject, contact: Research Ethics Coordinator, University of Windsor, Windsor, Ontario, N9B 3P4; Telephone: 519-253-3000, ext. 3948; e-mail: ethics@uwindsor.ca

SIGNATURE OF RESEARCH SUBJECT/LEGAL REPRESENTATIVE

I understand the information provided for the study **Eyewitness Memory for Crime** as described herein. My questions have been answered to my satisfaction, and I agree to participate in this study. I have been given a copy of this form.

_____
Name of Subject

_____ _____
Signature of Subject Date

SIGNATURE OF INVESTIGATOR

These are the terms under which I will conduct research.

_____ _____
Signature of Investigator Date

CONSENT FOR AUDIO TAPING

Research Participant's Name: _____

Title of the Project: Eyewitness memory for crime

I consent to the audio-taping of interviews.

I understand these are voluntary procedures and that I am free to withdraw at any time by requesting that the taping be stopped. I also understand that my name will not be revealed to anyone and that taping will be kept confidential. Tapes are filed by number only and stored on a computer or external hard-drives that are not connected to the internet.

I understand that confidentiality will be respected and that the audio tape will be for professional use only.

_____ _____

(Research Participant) (Date)

APPENDIX B

Distractor Tasks

Please unscramble the following words:

1. cliog
2. itmenoo
3. hibevuaor
4. delhi
5. rodw
6. lomesoacrs
7. pelse
8. nbira
9. diemniec
10. lutda
11. mahnu
12. ryteho
13. iktnh
14. macidea
15. oecurs
16. golysyhpoe
17. kobo
18. ranel
19. rxtpmeiene
20. arppe
21. lklis
22. dtecatnane
23. drega
24. aadcelrn
25. atleamo
26. ralndeca
27. aerpporkw
28. fiticetaretc
29. cckkaabp
30. ecofef
31. peonetlhe
32. krnap
33. ecrsbalm
34. bnaana
35. pipetperm
36. tchoewrtaw
37. leenif
38. cmgeuaaol
39. sdsoipah
40. atrhaonm

APPENDIX C

CI and Free Recall Instructions

CI	FR
Explain purpose of the Interview ***I will now start asking you some questions about the video clip you saw.***	Explain purpose of the Interview ***I will now start asking you some questions about the video clip you saw.***
Report everything Instruction ***First, what I want you to do is to please tell me everything that comes to your mind in as much detail as possible, even things you think might be unimportant and even if you cannot remember something completely. Don't leave anything out. Please tell me everything that you remember.***	
Transfer of Control ***Also, please keep in mind that I didn't see the video clip. So I am relying on you to provide as much detail as possible and tell me everything so I can know exactly what happened and what you could see in the video.***	
Also, please do not guess or make something up, just tell me everything that you can actually remember.	
Mental Context Reinstatement ***I will now give you further instructions I would like you to just listen and follow my instructions closely. If you feel comfortable, close your eyes. Sometimes it helps. Now please picture yourself back when you arrived here at the room with me. Think about how you were feeling when you arrived here. Also think about our conversation when you arrived here. Now picture yourself back in front of the monitor and think about what your first impressions were when you saw the***	

video clip. Now play back the video clip in your mind. Once you have a really clear picture in your mind about what you could see in the video clip, please tell me everything you can remember about it in as much detail as possible. But for now, make sure you play the video clip through in your mind.	
(after participant is done) ***Is there anything else you can remember about it?***	
	Initiate Free Recall ***Please tell me what you can remember about the video clip.***
	(after participant is done) ***Is there anything else you can remember about it?***
Questioning phase	Questioning phase
I am now going to ask you some more questions about the video clip. If you do not know the answer to any of these please say so and if you do not understand a question please say so as well. a) You mentioned a couple of people; could you please tell me more about them? What did they look like, what were they wearing, what were they doing? b) Please tell me more about the location. What it looked like, what you could see, any objects?	***I am now going to ask you some more questions about the video clip. If you do not know the answer to any of these please say so and if you do not understand a question please say so as well.*** a) You mentioned a couple of people; could you please tell me more about them? What did they look like, what were they wearing, what were they doing? b) Please tell me more about the location. What it looked like, what you could see, any objects?
Closure ***We are now finished with the interview. Thanks you very much for answering all my questions. That was very helpful.***	Closure ***We are now finished with the interview. Thanks you very much for answering all my questions. That was very helpful.***

APPENDIX D

Question Sets for Video One and Two

Bar Video

1) What was the name of the bar?
2) What was the relationship between the young girl and boy who entered the bar in the middle of the video clip?
3) What colour was the hat of the girl who entered the bar?
4) Where did the boy and the girl sit after they entered the bar?
5) What colour was the cigarette pack of the smoker?
6) What drink did the smoker order at the bar?
7) What did the barman say to the smoker?
8) What was being shown on the television in the bar?
9) What was on the sign on the door behind the woman in the green dress?
10) Who was the woman in the green dress waiting for?
11) Where did the woman in the green dress keep her cellphone?
12) What was the barman doing when the stranger entered the bar?
13) What did the stranger say to the barman?
14) How much was the wine the stranger ordered?
15) What colour was the stranger's shirt?
16) What did the woman in the green dress have in her hand when she returned to the table?
17) What was the name of the woman in the green dress?
18) What was directly behind where the stranger was sitting at the table?
19) What was on the picture above the table in the bar?
20) What colour was the coat of the woman in the green dress?
21) What kind of shoes was the woman in the green dress wearing?
22) How often did the woman in the green dress look at her wristwatch?
23) What did the smoker do with his cigarette after the stranger and the woman in the green dress have entered the toilet?
24) Where was the purse of the woman in the green dress after she has been dragged into the toilet?

Burglary Video

1) How many rings was the homeowner wearing?
2) What company does the robber work for?
3) How many locks does the homeowner open?
4) What is showing on the television?
5) What pet does the homeowner have?
6) What type of footwear is the robber wearing?
7) How many sugars does the robber have in his tea?
8) Which newspaper is in the fireplace?
9) What house number does the homeowner live at?
10) Where is the sofa located?
11) How many cookies are on the plate?
12) How many times does the doorbell ring?
13) What time does the robber leave?
14) What did the robber take from the man's office?
15) Who else lives in the home?
16) What colour is the coffee machine in the kitchen?
17) How did the robber get to the house?
18) What colour is the cushion the homeowner is sitting on?
19) Where does the homeowner put his glasses?
20) What colour is the carpet in the study?
21) What receipt does the robber ask to see?
22) What object does the robber break?
23) What book is the homeowner reading?
24) What brand of tea does the homeowner drink?

APPENDIX E

Letter of Information

The study that you have participated in was examining several variables. First, some participants at Time One were interviewed using a type of interview that has been shown to be effective with eyewitness memory. This interview is called the Cognitive Interview. Other participants were asked to tell the researcher everything that he or she could remember about the video clip. This is a control interview that does not include the procedures that are key to the effectiveness of the Cognitive Interview.

Everyone returned for questioning one week later. Everyone was asked a set of questions about the video twice. Some questions had clear answers. Others might have led the participant to guess the answer to the question because it would be impossible to know the correct response, based on the content of the video. Between question sets, some participants were told that they had to go through the set of questions again because they made a number of errors the first time. Others were given no feedback.

This project is being conducted to examine the efficacy of particular types of interviews in helping interviewees remember more correct information and be less susceptible to pressure or feedback from interviewers.

If you have any further questions about the study, feel free to contact Lauren, the principal researcher for the study, whose email address is on the consent form that you were given at the beginning of the study.

Thanks again for participating in the study. You will be credited on the Participant Pool shortly for your participation in the study.

APPENDIX F

Demographic Form

Demographic Form

Age:

Gender:

Ethnicity (please select)
__ Black/African/Caribbean
__ Chinese
__ Filipino
__ First Nations
__ Japanese
__ Latin American
__ Mixed
__ South Asian (e.g., East Indian, Pakistani, Sri Lankan, etc.)
__ Southeast Asian (e.g., Vietnamese, Cambodian, Malaysian, etc.)
__ White
__ Other: _____

REFERENCES

Ackil, J. K., Van Abbema, D. L., & Bauer, P. J. (2003). After the storm: Enduring differences in mother-child recollections of traumatic and nontraumatic events. *Journal of Experimental Child Psychology, 84*(4), 286-309. doi:10.1016/S0022-0965(03)00027-4

Ackil, J.K. & Zaragoza, M.S. (1995). Developmental differences in eyewitness suggestibility and memory for source. *Journal of Experimental Child Psychology. Special Issue: Early Memory, 60*(1), 57-83. doi:10.1006/jecp.1995.1031

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. New York, NY, US: Cambridge University Press.

Brewer, N., & Burke, A. (2002). Effects of testimonial inconsistencies and eyewitness confidence on mock-juror judgments. *Law and Human Behavior, 26*(3), 353-364. doi:10.1023/A:1015380522722

Brock, P., Fisher, R. P., & Cutler, B. L. (1999). Examining the cognitive interview in a double-test paradigm. *Psychology, Crime & Law, 5*, 29-45. doi:10.1080/10683169908414992

Bruck, M., Ceci, S. J., & Hembrooke, H. (2002). The nature of children's true and false narratives. *Developmental Review. Special Issue on Forensic Developmental Psychology, 22*(3), 520-554. doi:10.1016/S0273-2297(02)00006-0

Bjorklund, D. F., Cassel, W. S., Bjorklund, B. R., Brown, R. D., Park, C. L., Ernst, K., & Owen, F. A. (2000). Social demand characteristics in children's and adults' eyewitness memory and suggestibility: The effect of different interviewers on free recall and recognition. *Applied Cognitive Psychology, 14*(5), 421-433.

doi:10.1002/1099-0720(200009)14:5<421::AID-ACP659>3.0.CO;2-4

Ceci, S. J., & Bruck, M. (1993). Suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin, 113*(3), 403-439. doi:10.1037/0033-2909.113.3.403

Chan, J. C. K., Thomas, A. K., & Bulevich, J. B. (2009). Recalling a witnessed event increases eyewitness suggestibility: The reversed testing effect. *Psychological Science, 20*(1), 66-73. doi:10.1111/j.1467-9280.2008.02245.x

Clarke, C., & Milne, R. (2001). *National evaluation of the PEACE investigative interviewing course*. London: Home Office.

Cohen J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.), Hillsdale, NJ: Erlbaum. pp. 281, 284, 285

Cutler, B. L., Penrod, S. D., & Stuve, T. E. (1988). Juror decision making in eyewitness identification cases. *Law and Human Behavior, 12*(1), 41-55. doi:10.1007/BF01064273

Davis, D., & Loftus, E. F. (2007). Internal and external sources of misinformation in adult witness memory. In M. P. Toglia, J. D. Read, D. F. Ross & R. C. L. Lindsay (Eds.), (pp. 195-237). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Erdelyi, M. H., & Becker, J. (1974). Hypermnesia for pictures: Incremental memory for pictures but not words in multiple recall trials. *Cognitive Psychology, 6*(1), 159-171. doi:10.1016/0010-0285(74)90008-5

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical

sciences. *Behavior Research Methods, 39*, 175-191. doi:10.3758/BF03193146

Fisher, R. P. (2010). Interviewing cooperative witnesses. *Legal and Criminological Psychology, 15*(1), 25-38. doi:10.1348/135532509X441891

Fisher, R. P., Falkner, K. L., Trevisan, M., & McCauley, M. R. (2000). Adapting the cognitive interview to enhance long-term (35 years) recall of physical activities. *Journal of Applied Psychology, 85*(2), 180-189. doi:10.1037/0021-9010.85.2.180

Fisher, R. P. & Geiselman, R. E. (1992). *Memory-enhancing techniques in investigative interviewing: The cognitive interview.* Springfield, IL: C. C. Thomas.

Fisher, R. P., Geiselman, R. E., Raymond, D. S., & Jurkevich, L. M. (1987). Enhancing enhanced eyewitness memory: Refining the cognitive interview. *Journal of Police Science & Administration, 15*(4), 291-297.

Fisher, R. P., & Schreiber, N. (2007). Interview protocols for improving eyewitness memory. In M. P. Toglia, J. D. Read, D. F. Ross & R. C. L. Lindsay (Eds.), (pp. 53-80). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Fisico, S., & Scoboria, A. (2011) *Encouragement and clarification of 'don't know' responses enhances interview response quality.* Manuscript submitted for publication.

Fritz, C.O., Morris, P.E., & Richler, J.J. (2012). Effect size estimates: Current use, calculations, and interpretation. *Journal of Experimental Psychology: General, (141)*1, 2-18. doi: 10.1037/a0024338

Fivush, R., McDemott Sales, J. M., Goldberg, A., Bahrick, L., & Parker, J. (2004). Weathering the storm: Children's long-term recall of hurricane Andrew. *Memory, 12*(1), 104-118. doi:10.1080/09658210244000397

Garven, S., Wood, J. M., & Malpass, R. S. (2000). Allegations of wrongdoing: The effects of reinforcement on children's mundane and fantastic claims. *Journal of Applied Psychology, 85*(1), 38-49. doi:10.1037/0021-9010.85.1.38

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

Gilbert, J. A. E., & Fisher, R. P. (2006). The effects of varied retrieval cues on reminiscence in eyewitness memory. *Applied Cognitive Psychology, 20*(6), 723-739. doi:10.1002/acp.1232

Goodman, G. S., & Quas, J. A. (2008). Repeated interviews and children's memory: It's more than just how many. *Current Directions in Psychological Science, 17*(6), 386-390. doi:10.1111/j.1467-8721.2008.00611.x

Greenstock, J. & Pipe M-E. (1997). Are two heads better that one? Peer support and children's eyewitness reports. *Applied Cognitive Psychology, 11*, 461-483. doi:10.1002/(SICI)1099-0720(199712)11:6<461::AID-ACP473>3.0.CO;2-T

Griffiths, A. (2012, May) *The GQM: A training and assessment tool.* Paper presented at the Annual Masterclass of the International Investigative Interviewing Research Group, Toronto, Canada.

Gudjonsson, G. H. (1984). A new scale of interrogative suggestibility. *Personality and Individual Differences, 5*(3), 303-303-314. doi:10.1016/0191-8869(84)90069-2

Gudjonsson, G. H. (1986). The relationship between interrogative suggestibility and acquiescence: Empirical findings and theoretical implications. *Personality and*

Individual Differences, 7, 195-199. doi:10.1016/0191-8869(86)90055-3

Gudjonsson, G.H. (2003). *The psychology of interrogations and confessions: A handbook.* West Sussex, England: John Wiley & Sons Ltd.

Holliday, R. E., Humphries, J. E., Milne, R., Memon, A., Houlder, L., Lyons, A., & Bull, R. (2011). Reducing misinformation effects in older adults with cognitive interview mnemonics. *Psychology and Aging.* doi:10.1037/a0022031

Howie, P., Kurukulasuriya, N. K., Nash, L., & Marsh, A. (2009). Inconsistencies in children's recall of witnessed events: The role of age, question format and perceived reason for question repetition. *Legal and Criminological Psychology, 14*(2), 311-329. doi: 10.1348/135532508X383879

Howie, P., Sheehan, M., Mojarrad, T., & Wrzesinska, M. (2004). 'Undesirable' and 'desirable' shifts in children's responses to repeated questions: Age differences in the effect of providing a rationale for repetition. *Applied Cognitive Psychology, 18*(9), 1161-1180. doi: 10.1002/acp.1049

Kirk, R. E. (1995). *Experimental design* (3rd ed.). Pacific Grove, CA: Brooks/Cole.

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review, 103*, 490-517.1996-01780 -00410.1037//0033 -295X.103.3.490

Krähenbühl, S., Blades, M., & Eiser, C. (2009). The effect of repeated questioning on children's accuracy and consistency in eyewitness testimony. *Legal and Criminological Psychology, 14*(2), 263-278. doi:10.1348/135532508X398549

La Rooy, D., Katz, C., Malloy, L. C., & Lamb, M. E. (2010). Do we need to rethink guidance on repeated interviews? *Psychology, Public Policy, and Law, 16*(4),

373-392. doi:10.1037/a0019909

La Rooy, D., Lamb, M. E., & Pipe, M. (2009). Repeated interviewing: A critical evaluation of the risks and potential benefits. In K. Kuehnle, & M. Connell (Eds.), (pp. 327-361). Hoboken, NJ, US: John Wiley & Sons Inc.

La Rooy, D., Pipe, M., & Murray, J. E. (2005). Reminiscence and hypermnesia in children's eyewitness memory. *Journal of Experimental Child Psychology, 90*(3), 235-254. doi:10.1016/j.jecp.2004.11.002

Lamb, M. E., & Fauchier, A. (2001). The effects of question type on self-contradictions by children in the course of forensic interviews. *Applied Cognitive Psychology, 15*(5), 483-491. doi: 10.1002/acp.726

Lamb, M.E., Sternberg, K., Orbach, Y., Hershkowitz, I., Horowitz, D., & Esplin, P.W. (2002). The effects of intensive training and ongoing supervision on the quality of investigative interviews with alleged sex abuse victims. *Applied Developmental Science, 6(3)*, 114-125. doi:10.1207/S1532480XADS0603_2

Linton, C. P., & Sheehan, P. W. (1994). The relationship between interrogative suggestibility and susceptibility to hypnosis. *Australian Journal of Clinical & Experimental Hypnosis, 22*(1), 53-64.

Mazzoni, G., & Kirsch, I. (2002). Autobiographical memories and beliefs: A preliminary metacognitive model. In T. J. Perfect, & B. L. Schwartz (Eds.), *Applied metacognition.* (pp. 121-145). New York, NY, US: Cambridge University Press, New York, NY. doi: 10.1017/CBO9780511489976.007

McCauley, M. R., & Fisher, R. P. (1995). Facilitating children's eyewitness recall with the revised cognitive interview. *Journal of Applied Psychology, 80*(4), 510-516.

doi:10.1037/0021-9010.80.4.510

Memon, A., Meissner, C. A., & Fraser, J. (2010). The cognitive interview: A meta-analytic review and study space analysis of the past 25 years. *Psychology, Public Policy, and Law, 16*(4), 340-372. doi:10.1037/a0020518

Memon, A., & Vartoukian, R. (1996). The effects of repeated questioning on young children's eyewitness testimony. *British Journal of Psychology, 87*(00071269), 403-403. doi:10.1111/j.2044-8295.1996.tb02598.x

Memon, A., Wark, L., Bull, R., & Köhnken, G. (1997). Isolating the effects of the cognitive interview techniques. *British Journal of Psychology, 88*(2), 179-197. doi:10.1111/j.2044-8295.1997.tb02629.x

Memon, A., Zaragoza, M., Clifford, B. R., & Kidd, L. (2010). Inoculation or antidote? The effects of cognitive interview timing on false memory for forcibly fabricated events. *Law and Human Behavior, 34*(2), 105-117. doi:10.1007/s10979-008-9172-6

Moston, S. (1987). The suggestibility of children in interview studies. *First Language, 7*(19), 67-78. doi: 10.1177/014272378700701904

Odinot, G., Memon, A. La Rooy D., & Millen, A. *Are two interviews better than one? Eyewitness memory across repeated cognitive interviews.* Unpublished manuscript.

Odinot, G., Wolters, G., & Lavender, T. (2009). Repeated partial eyewitness questioning causes confidence inflation but not retrieval-induced forgetting. *Applied Cognitive Psychology, 23*(1), 90-97. doi:10.1002/acp.1443

Pipe, M.-E., & Wilson, J. C. (1994). Cues and secrets: Influences on children's event

reports. *Developmental Psychology, 30*, 515-525. doi:10.1037/0012-1649.30.4.515

Poole, D. A., & Lindsay, D. S. (2001). Children's eyewitness reports after exposure to misinformation from parents. *Journal of Experimental Psychology: Applied, 7*(1), 27-50. doi:10.1037/1076-898X.7.1.27

Poole, D. A., & White, L. T. (1991). Effects of question repetition on the eyewitness testimony of children and adults. *Developmental Psychology, 27*(6), 975-986. doi:10.1037/0012-1649.27.6.975

Poole, D. A., & White, L. T. (1993). Two years later: Effect of question repetition and retention interval on the eyewitness testimony of children and adults. *Developmental Psychology, 29*(5), 844-853. doi:10.1037/0012-1649.29.5.844

Register, P. A., & Kihlstrom, J. F. (1988). Hypnosis and interrogative suggestibility. *Personality and Individual Differences, 9*(3), 549-558. doi:10.1016/0191-8869(88)90152-3

Roediger, H. L., & Payne, D. G. (1982). Hypermnesia: The role of repeated testing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*(1), 66-72. doi:10.1037/0278-7393.8.1.66

Roediger, H. L., Payne, D. G., Gillespie, G. L., & Lean, D. S. (1982). Hypermnesia as determined by level of recall. *Journal of Verbal Learning & Verbal Behavior, 21*(6), 635-655. doi:10.1016/S0022-5371(82)90810-6

Schreiber Compo, N., Gregory, A.H., & Fisher, R.P. (2012). Interviewing behaviors in police investigators: A field study of a current U.S. sample. *Psychology, Crime and Law, 18*(4), 359-375. doi: 10.1080/1068316X.2010.494604

Singh, K. K., & Gudjonsson, G. H. (1984). Interrogative suggestibility, delayed memory and self-concept. *Personality and Individual Differences, 5*, 203-209. doi:10.1016/0191-8869(84)90052-7

Shapiro, S. R., & Erdelyi, M. H. (1974). Hypermnesia for pictures but not words. Journal of Experimental Psychology, 103, 1218-1219 .44575961975-06674 -001

Shaw, J. S. (1996). Increases in eyewitness confidence resulting from postevent questioning. *Journal of Experimental Psychology: Applied, 2*(2), 126-146. doi:10.1037/1076-898X.2.2.126

Shaw, J. S., Bjork, R. A., & Handal, A. (1995). Retrieval-induced forgetting in an eyewitness-memory paradigm. *Psychonomic Bulletin & Review, 2*(2), 249-253.

Stöber, J. (2001). The social desirability scale-17 (SDS-17): Convergent validity, discriminant validity, and relationship with age. *European Journal of Psychological Assessment, 17*(3), 222-232. doi:10.1027//1015-5759.17.3.222

Tinsley, H. E. A., & Weiss, D. J. (2000). *Interrater reliability and agreement*. San Diego, CA, US: Academic Press. doi:10.1016/B978-012691360-6/50005-7

Tuckey, M. & Brewer, N. (2003). How schemas affect eyewitness memory over repeated retrieval attempts. *Applied Cognitive Psychology, 17,* 785–800. doi: 10.1002/acp.906

Warren, A., Hulse-Trotter, K., & Tubbs, E. C. (1991). Inducing resistance to suggestibility in children. *Law and Human Behavior, 15*(3), 273-285. doi: 10.1007/BF01061713

Wright, D. B., Loftus, E. F., & Hall, M. (2001). Now you see it; now you don't: Inhibiting recall and recognition of scenes. *Applied Cognitive Psychology, 15*(5),

471-482. doi:10.1002/acp.719

VITA AUCTORIS

NAME:	Lauren M. Wysman
PLACE OF BIRTH	Windsor, Ontario
YEAR OF BIRTH	1988
EDUCATION	F.J. Brennan High School, Windsor, Ontario 2001-2005 University of Windsor, Ontario 2005-2009 B.A. (Hons.) Psychology University of Windsor, Ontario 2010-present M.A. Adult Clinical Psychology (Candidate)