

## The Enhanced Cognitive Interview: Testing Appropriateness Perception, Memory Capacity and Error Estimate Relation with Report Quality

RUI M. PAULO<sup>1\*</sup>, PEDRO B. ALBUQUERQUE<sup>1</sup>, MAGDA SARAIVA<sup>1</sup> and RAY BULL<sup>2</sup>

<sup>1</sup>*School of Psychology, University of Minho, Campus de Gualtar, Braga, Portugal*

<sup>2</sup>*Department of Psychology, University of Portsmouth, Portsmouth, United Kingdom*

*Summary: The Enhanced Cognitive Interview (ECI) has been widely studied. However, research has overlooked witnesses' attitudes toward the interview and how error estimate and memory capacity relate to report quality. Participants watched a mock robbery video and were interviewed 48 hours later with either the Portuguese version of the ECI or a Structured Interview (SI). Participants interviewed with the ECI provided more information without compromising accuracy, particularly in free recall. Report accuracy was stable across interview phases and information categories. A higher perception of interview appropriateness (how witnesses evaluate the appropriateness of the interview procedure used) was linked with more detailed reports and more interest in being an interviewee. Participants over-estimated their error rate, and their memory capacity was not related to witnesses' recall. It is essential to take into account their perception of interview appropriateness and use alternative methods to evaluate report quality. Major implications for real-life investigations are discussed. Copyright © 2015 John Wiley & Sons, Ltd.*

### INTRODUCTION

Interviewing witnesses is a key procedure that frequently determines the success of a police investigation (Prescott, Milne, & Clark, 2011). However, what witnesses report rarely fully corresponds with what they remember (Bower, 1967). The difference between what happened and what is remembered can be even greater when inadequate interviewing techniques, such as leading questions, are used. These techniques produce a reduced amount of accurate information, as well as incorrect information. To address this issue, Geiselman et al. (1984) developed the Cognitive Interview.

The Cognitive Interview originally included four cognitive mnemonics: report everything, mental reinstatement of context, change order and change perspective. The report everything mnemonic consists of instructing witnesses to report everything they can remember, whether it seems trivial or not (Fisher & Geiselman, 2010). This procedure is very important because otherwise the witness might withhold valuable information that she considers to be irrelevant. Moreover, our memories for any given event may overlap, and 'irrelevant' recall might activate 'relevant' recall (Tulving, 1991). The mental reinstatement of context consists of asking witnesses to mentally recreate the to-be-recalled event and their physiological, cognitive and emotional states at the time of the crime. This mnemonic was derived from the premise that memory retrieval is more effective when the context of the original event is recreated during recall (Tulving & Thomson, 1973). Lastly, because memory may be accessed by using several different memory cues and paths (Tulving, 1991), the change order (asking the witness to recall the event in a different chronological order) and change perspective mnemonics (to recall the event from a different perspective) can be used to try to obtain new information.

This interview was further developed some years later by Fisher and Geiselman (1992) as the Enhanced Cognitive Interview (ECI). Social and communicative components crucial for conducting good investigative interviews, such as rapport building, were added to the original procedure. For more information about these components, see Paulo, Albuquerque, and Bull (2013). Several studies have repeatedly demonstrated that this interview technique is able to increase the amount of correct information recalled by witnesses, while maintaining report accuracy, that is, the amount of correct items of information proportionate to all recalled items of information (Aschermann, Mantwill, & Köhnken, 1991; Rivard, Fisher, Robertson, & Mueller, 2014). Therefore, the ECI has been widely acknowledged as one of the most successful procedures for enhancing witness recollection (Paulo et al., 2013). The ECI has been found to be effective in different countries—for example, USA, UK, Australia and Brazil (Stein & Memon, 2006); with different types of witness—for example, children, adults and elderly (Verkampft & Ginet, 2009; Wright & Holliday, 2006); with different intervals between the crime and the interview—minutes to weeks (Larsson, Granhag, & Spjut, 2002); with different type of event—for example, crime, traffic accident and phone call (Campos & Alonso-Quecuty, 2008) both in laboratory (Colomb & Ginet, 2012) and field studies (Colomb, Ginet, Wright, Demarchi, & Sadler, 2013) and is now being widely used by police forces in a variety of locations (e.g., UK and Australia).

Most authors (Aschermann et al., 1991; Rivard et al., 2014) have focused on how to increase the amount of produced information without decreasing report accuracy. However, Fisher and Geiselman (2010) recently suggested that using the ECI is more than using cognitive techniques to enhance recall, recognizing the need for future research on witnesses' attitudes toward the interview process and the interviewer. Other authors recently acknowledged that witnesses' perceptions toward the interview process might determine how rapport and working alliance, which are two similar constructs, are established and maintained throughout the interview (Vanderhallen & Vervaeke, 2014;

\*Correspondence to: Rui M. Paulo, School of Psychology – University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal.  
E-mail: id4299@alunos.uminho.pt

Vanderhallen, Vervaeke, & Holmberg, 2011). Establishing and maintaining rapport is also crucial during investigative interviews and was associated with better recall (Read, Powell, Kebbell, & Milne, 2009; Walsh & Bull, 2012). Ballardin, Stein, and Milne (2013) postulated that witnesses' consider variables such as the interviewers' capacity to actively listen, welcome the witness at the interview setting and provide a comfortable environment to be very important for the witness during the course of the interview. Such findings are consistent with literature from other fields of Psychology. For instance, Ackerman and Hilsenroth (2003) reviewed a set of studies that suggested that the therapist's personal attributes, such as being flexible, honest, respectful, trustworthy, confident, warm and interested, are important for establishing a positive therapeutic alliance, this is, a positive relation between the therapist and the client. Accordingly, several studies suggested that a positive therapeutic alliance is one of the most important factors for a positive outcome of the therapy (Martin, Garske, & Davis, 2000). However, in the forensic field, researchers have not yet addressed how witnesses' perceptions toward the interviewer and the interview process can directly influence witnesses' report (Fisher & Geiselman, 2010). Therefore, in this study, we focused on how witnesses' perception of interview appropriateness might influence the amount of produced information, as well as their interest in being an interviewee. We measured interview appropriateness perception by asking participants to evaluate, in a post-interview questionnaire, how appropriate they considered the interview procedure to which they had previously been submitted (ECI or Structured Interview (SI)) to be. This research topic is important because even the most theory-driven procedure could be harmful for the success of the investigation, if perceived as inappropriate by the witness.

In a very recent study, Paulo, Albuquerque, and Bull (2015) emphasized that although increasing ECI recall is important for police investigations, increasing and/or evaluating the accuracy of different aspects of each witness report, for instance, through metacognitive techniques, is a topic that ECI literature has largely disregarded. These authors found that witnesses are able to spontaneously and validly differentiate, in an interview setting, between information that they are sure about ('certainties') and information that they are unsure about ('uncertainties'). That is, participants successfully used spontaneous expressions of uncertainty (e.g., I think, Maybe, I believe, etc.) to identify less accurate information. Furthermore, they did this spontaneously while recalling the event (i.e., they were not instructed to do so), successfully performing real-time memory monitoring. Paulo *et al.* (2015) found that correct recall proportion for recalled 'uncertainties' (amount of correct 'uncertainties' over all produced 'uncertainties'—correct, incorrect and confabulated) is significantly lower (.65) than the accuracy proportion for 'certainties' only (amount of correct 'certainties' over all produced 'certainties'), which has an impressive value of .90. The authors conclude that differentiating 'uncertainties' from 'certainties' is a straightforward and time-saving process to increase and evaluate ECI report accuracy. Such results are consistent with the metacognitive/metamemory literature, which generally finds that metacognitive techniques can be

used to improve witnesses' accuracy (Higham, Luna, & Bloomfield, 2010; Roberts & Higham, 2002). Several studies (Allwood, Ask, & Granhag, 2005; Brewer, Weber, Wootton, & Lindsay, 2012; Lindsay *et al.*, 2013; Luna & Martín-Luengo, 2012) show that in different situations (e.g., selections from lineups, cued recall or free recall), when using adequate measures (e.g., calibration approach), a positive relation between confidence and accuracy can be found. Higher accuracy for a given response can be expected when witnesses provide a higher confidence judgment. Other authors (Evans & Fisher, 2010; Koriat & Goldsmith, 1996) suggest that witnesses can improve their accuracy by using metacognitive control techniques, such as exercising 'report option' or adjusting 'report precision'. Nonetheless, only two studies have focused on how this procedure can be used to increase ECI report accuracy (Allwood *et al.*, 2005; Roberts & Higham, 2002). These last authors interviewed witnesses with an ECI and asked them to provide confidence judgments, using a Likert scale, for a small portion of their statements. Using this procedure, participants were also able to distinguish between more and less reliable information. However, to use such procedure, a considerable amount of the interviewer's time is required, for instance, for applying these scales and selecting the information that will be evaluated by the interviewee. Therefore, it would be difficult to use such procedure, in a holistic manner, at a real police interview setting.

Because witnesses seem to be capable of estimating their accuracy using either spontaneous expressions of uncertainty, or post-interview numerical scales, we tested if they are also able to estimate the proportion of errors they committed for each interview phase (e.g., free recall or questioning phase) and for the whole interview. Similar research on frequency judgments (participants' estimates of how many items of information are correct for a given part of their statement) found that, usually, when confidence judgments result in overconfidence (subjective confidence ratings are higher than 'real' accuracy), frequency judgments are reasonably accurate, because these are distinct types of procedures (Lieberman, 2004). Therefore, when confidence judgments are accurate, frequency judgments usually result in underconfidence. This has been found in several studies (Gigerenzer, Hoffrage, & Kleinbölting, 1991; Sniezek & Buckley, 1991). Sniezek and Buckley (1991) proposed a dual-process account stating that confidence judgments and frequency judgments have very different natures. Confidence judgments are based on item-specific considerations, such as evaluation of information about the item content, whereas frequency judgments are influenced by other variables such as one's perception about himself/herself (e.g., perception of expertise). To our knowledge, no previous study has evaluated witnesses' ability to perform accurate frequency judgments on an interview, or ECI, condition. If witnesses are able to make such assessment, it can have important consequences in court (e.g., if a witness is extremely confident that she committed very few errors on her report, should a judge/juror evaluate such report as highly reliable?). Furthermore, providing a frequency judgment, or an error frequency judgment, for a given part of the statement is less time demanding than providing

confidence judgments for all the information units recalled in that portion of the statement.

Lastly, several studies support that witnesses need to access different types of memory when recalling a crime, (Bower, 1967). For instance, when using ECI mnemonics, such as Change Order or Change Perspective, witnesses need to manipulate memory information while recalling, processes that are supported by working memory (Baddeley, 2002; Vrij et al., 2008). Therefore, one could expect that witnesses with a higher working memory capacity might have a better report, because of their ability to successfully use these mnemonics. However, to our knowledge, such research has never been conducted. Furthermore, when describing the face of the criminal, witnesses need to access their memory for faces. Morgan et al. (2007) found that witnesses who achieved higher scores at the Faces Recognition Test (Wechsler, 1997) were also more accurate on an eyewitness task: selecting a target person that they had previously met under stressful conditions, at a sequential photo presentation. These authors believe that trait ability to remember human faces, measured with the Faces Recognition Test, is related to witnesses' ability to recall faces under stressful conditions. However, to our knowledge, no study has evaluated if trait ability to remember faces could be related to witnesses' ability to describe faces during the course of an investigative interview (e.g., ECI), nor if trait ability to remember other types of stimuli (e.g., locations), measured through memory capacity tests, could be related to subsequent performance on an investigative interview (e.g., when describing the location of the perpetrator at the crime scene). Therefore, we believe that the relation between witnesses' performance on different memory tests (that access different types of memory) and witnesses' ability to provide a good report has not yet been fully addressed. We innovatively evaluated if witnesses' performance on several memory tests was related to their performance when recalling a video recording of a (mock) bank robbery in an interview setting. This topic can also have major impact on the applied field, for instance, when accessing the value of the statement in court.

Overall, the present study aimed to explore three main questions: (1) Do witnesses with a higher perception of the appropriateness of the interview, accessed on a post-interview questionnaire, have better recall and/or report more motivation to be an interviewee? (2) Are witnesses capable of performing accurate frequency judgments for their error rate? (3) Is witness performance on memory tests related to the amount/accuracy of recalled information in an interview?

We interviewed two groups of participants regarding their ability to recall a mock bank robbery 48 hours after they viewed it: one group was interviewed with the ECI, and the other group was interviewed with an SI.

## METHOD

### Participants

A total of 44 psychology students (age:  $M=21$ ,  $SD=3$ ), 36 female and 8 male, from the University of Minho (Portugal), participated in this study for course credits.

### Design

A between subjects experimental design was used with interview condition as the independent variable with two levels: (1) ECI, and (2) Structured Interview. The amount of reported information and accuracy were measured in information units and proportion, respectively.

### Materials

The participants watched a video recording of a non-violent (mock) bank robbery on a Fujitsu L7ZA LCD computer screen. The video recording, which was edited from the second episode of the first season of the 2004 Portuguese television drama 'Inspector Max' (Riccó, & Riccó, 2004), was three minutes and 11 seconds long. We used three memory tests to evaluate the following: (1) Working Memory—Working Memory Span Test (Conway et al., 2005); (2) Face Recognition—Faces (Wechsler, 1997); and (3) Spatial Span—Spatial Span (Wechsler, 1997). A post-interview questionnaire was constructed, consisting of two direct questions: (1) In your opinion, how appropriate would this interview be for a real police setting? and (2) How interesting was it for you to be interviewed during this study? These questions were answered through a seven-point Likert scale: 1—highly inappropriate and 7—highly appropriate; and 1—totally uninteresting and 7—totally interesting, respectively, for questions 1 and 2. All interviews were audio and video recorded.

### Procedure

Ethics committee approval was obtained. Participants took part in two sessions. At the first session, participants were randomly assigned to one of the two conditions (ECI or SI), having signed a consent form after reading general information about the study. Immediately after, they were shown the video recording and then administered the three memory tests. The administration order of the memory tests was varied. The second session took place approximately 48 hours later, and each participant was interviewed with either the ECI or the SI and asked to give an error rate estimate immediately after each interview phase. The error estimate for the 'Summary' phase of the interview was not asked, because many participants did not provide any new detail at this phase. After the interview, all participants were asked to give an overall error rate estimate and completed the post-interview questionnaire.

### Interview conditions

The interview protocols employed were adapted from Milne and Bull (2003) for the Portuguese language. Both interview protocols involved seven main phases: (1) preliminary phase; (2) free report; (3) open-ended questioning; (4) second retrieval; (5) third retrieval (for new information only); (6) summary; and (7) closure.

During phase 1 (preliminary phase), procedures like greeting, establishing rapport, explaining the instructions and purpose of the interview to the witness and asking not to guess were followed for both interview groups. The ECI condition included the transfer of control instruction and the report everything instruction.

During phase 2 (free report), participants were asked to recall what they could remember about the video in any order and pace they wished. In the ECI condition, they were reminded to report everything they could remember with as much detail as possible, and mental reinstatement of context was applied.

During phase 3 (open-ended questioning), three open-ended questions were asked to each participant according to his/her free report (e.g., Please describe the crime scene—if the participant previously reported seeing the crime scene). For the ECI condition, mental imagery instructions were used—e.g., *you told me that you looked at the robber when he entered the bank. Can you please close your eyes ..., think about everything that you can remember concerning him ..., his face ..., his clothes ..., his actions ..., and when you have a full picture of him in your mind, describe everything that you can remember about him.*

During phase 4 (second retrieval), participants were asked to report what they could remember about the video once again. In both conditions, participants were encouraged to give this second report, and the importance of such procedure was explained. In the ECI condition, participants were asked to recall the video in the reverse order.

During phase 5 (third retrieval), participants were asked to focus one last time on the video and report any new detail they could remember, if possible. In both interview conditions, the importance of such a procedure was explained, and participants were encouraged to do their best. In the ECI condition, participants were asked to adopt a different internal perspective in order to try to remember new details: (...) *please focus on the event as if it was a common event at the bank, instead of a robbery, as you probably assumed before seeing the robber entering the bank (...).*

On phase 6 (summary), the interviewer summarized what he understood of the witness account and asked her to correct him if he misheard, or misinterpreted, any part of the statement. He also told her/him to interrupt him if she/he could remember any new detail.

On the last phase (closure), appreciation for participants' cooperation was acknowledged, and neutral topics were again discussed. These last two phases were exactly alike for both interview groups.

Overall, the differences between the ECI and SI protocols were the four cognitive mnemonics and the transfer of control instruction and mental imagery. Both interview protocols included procedures such as rapport building and appropriate questioning (e.g., witness-compatible questioning) because they are considered an essential aspect of any investigative interview. Thus, we wanted to focus on the effect that the remaining components, only applied in the ECI condition, would have on recall. Fisher and Geilseman's (1992) guidelines for conducting the ECI were followed, and all the cognitive, social and communicative components they described were included in the ECI protocol.

### Coding

Recordings of each interview were coded using the template scoring technique from Memon, Holley, Wark, Bull, and Köhnken (1996). A comprehensive list of details in the video recording was compiled, and items of information were

categorized as referring to the following: (1) person; (2) action; (3) object; (4) location; (5) conversation; and (6) sound, resulting in 378 items of information. Recalled information was classified as either correct, incorrect (e.g., saying the shirt was brown when it was black) or confabulation (mentioning a detail or event that was not present or did not happen). The phase within the interview in which an item of information was recalled was also coded. If an item of information (correct or not) was repeated during the same, or a subsequent, phase, that information was scored only the first time (Prescott *et al.*, 2011). Subjective statements or opinions were disregarded (e.g., 'He was really good looking!').

### Inter-rater reliability

To assess inter-rater reliability, 11 (25%) interviews were selected randomly and scored independently by a researcher who was naive to the aims of the experiment and hypothesis, but familiar with the template method of scoring interviews and had access to the crime video. Intraclass correlation coefficients (ICC) were calculated for correct information, incorrect information and confabulations and for the six information categories (person, action, etc.). High inter-rater reliability was found for all measures in that the values of the ICC ranged between .979 and 1.000, with an overall ICC of .992.

## RESULTS

Exploratory data analysis was used to decide whether to conduct parametric versus nonparametric statistical tests when interval or ratio scale variables were included in such statistical tests. Bonferroni corrections were applied when multiple statistical tests were conducted on a single data set, to avoid type 1 error (Field, 2009).

### Accuracy and items of information

Participants in the ECI condition did recall more items of information ( $M=76$ ,  $SD=24.71$ ) in comparison with the control group ( $M=58$ ,  $SD=13.91$ ),  $t(42)=2.96$ ,  $p=.005$ ,  $d=.89$ , 95% CI [-30.11, -5.71]. Therefore, as expected, the ECI protocol ( $M=35.32$ ,  $SD=10.69$ ) took longer to conduct (measured in minutes) than the SI protocol ( $M=22.76$ ,  $SD=6.58$ ),  $t(42)=-4.69$ ,  $p<.001$ ,  $d=1.41$ , 95% CI [-17.96, -7.15]. As seen in Table 1, no differences were found between the two interviews regarding the proportion of (i) correct recall (ratio between the amount of correct items of information recalled over all the items of information),  $t(42)=.96$ ,  $p=.343$ ,  $d=.29$ ; (ii) errors (ratio between the amount of errors produced over all items of information),  $t(42)=1.12$ ,  $p=.269$ ,  $d=.34$ ; and (iii) confabulations (ratio between the amount of confabulated information over all

Table 1. Proportion values (mean and standard deviation) for correct recall, errors and confabulations, according to the interview condition.

	Correct recall	Errors	Confabulations
ECI	.86 (.07)	.09 (.04)	.05 (.04)
SI	.87 (.05)	.08 (.05)	.05 (.03)

items of information),  $t(42) = .80$ ,  $p = .431$ ,  $d = .24$ . Thus, participants interviewed with the ECI were able to provide more information without increasing the proportion of errors and confabulations in their reports.

We first conducted a one-way within-subjects ANOVA to see if accuracy, measured in correct recall proportion, was stable across all information categories (person vs. object vs. action vs. location vs. conversation vs. sound). We found no differences in correct recall proportion according to the information categories,  $F(2.61, 80.92) = .93$ ,  $p = .421$ ,  $\eta^2 = .03$ . Regarding the interview phases, we found that only during free recall participants interviewed with the ECI ( $M = 37$ ,  $SD = 16$ ) recalled significantly more information than SI participants ( $M = 24$ ,  $SD = 8$ ),  $t(42) = 3.37$ ,  $p = .002$ ,  $d = .54$ , 95% CI  $[-20.64, -5.18]$ . As seen in Table 2, both groups produced most information at free recall and questioning phase, recalling only a few new details at the subsequent phases.

### Interview appropriateness

We found that participants who evaluated the interview appropriateness as higher also recalled more information units,  $r_s = .32$ ,  $p = .037$ , 95% CI  $[.02, .61]$  and reported more interest in being an interviewee,  $r_s = .38$ ,  $p = .011$ , 95% CI  $[.08, .66]$ .

### Error estimate

Participants estimated that their error rate was higher during the questioning phase (27% error rate), followed by the second recall attempt (20%), the third recall attempt (19%), and, lastly, free recall phase (17%). However, no differences in accuracy across different interview phases was found,  $F(2.70, 64.88) = 2.23$ ,  $p = .099$ ,  $\eta^2 = .09$ . No correlation between participants' error estimate for the interview and their real error/confabulation rate (ratio between the amount of errors and confabulations produced over all items of information) was found,  $r = .23$ ,  $p = .123$ . Similar non-significant results were found for each individual interview phase. Lastly, we found that participants' error estimate for the interview ( $M = 27\%$ ,  $SD = 12\%$ ) was significantly higher than their real error/confabulation rate ( $M = 13\%$ ,  $SD = 6\%$ ),  $t(43) = 7.46$ ,  $p < .001$ ,  $d = 1.39$ , 95% CI  $[-.17, -.10]$ .

### Memory performance

No correlations were found between participants' scores on the memory tests and their subsequent performance on the interview, in terms of correct recall proportion and amount

Table 2. Mean ( $M$ ) and standard deviation ( $SD$ ) values for the amount of recalled items of information in each interview phase, according to the interview condition.

Interview phases	Enhanced Cognitive Interview		Structured Interview	
	$M$	$SD$	$M$	$SD$
Free report	36.50	2.71	23.59	2.71
Questioning	29.32	2.04	25.82	2.04
Second recall	4.41	.66	4.46	.66
Third recall	3.23	.51	2.64	.51
Summary	2.60	.47	1.60	.47

of recalled information. Performance on each memory test was also not associated with performance on the remaining memory tests. Because working memory could influence the performance of the ECI group at the Reverse Order and Change Perspective mnemonics, we tested if there was a correlation between participants' performance on the Working Memory Span Test and their performance on these two mnemonics. No correlation was found. We also tested if participants' results on Faces Recognition Test were correlated to the amount, and accuracy, of recalled person details. No correlation was found. Lastly, we tested if participants' results on Spatial Span test were correlated to the amount, and accuracy, of produced location details. Again, we found no correlation between these variables. Therefore, participants' results on the memory tests do not seem to be related to their subsequent performance at the interview.

## DISCUSSION

This study found that a Portuguese version of the ECI (versus a Portuguese version of the SI) produced more items of information without compromising accuracy. Major differences regarding the amount of elicited information according to the interview condition were found for free recall. Report accuracy was stable across different interview phases and information categories. We found that a higher perception of the interview appropriateness was associated with more detailed reports and more interest in being an interviewee. Performance on memory tests was not related to witnesses' subsequent recall, and witnesses' error estimate was not associated to their real error rate.

Previously published studies suggest that the ECI superiority effect could be consistent across different countries (Stein & Memon, 2006). We found a Portuguese version of the ECI to increase the amount of recalled information without compromising accuracy, that is, without increasing the proportion of errors and confabulations. Professionals have now available a Portuguese version of the ECI that has been tested and found to maximize the amount of elicited information. Such findings are crucial for Portuguese police forces and other relevant professionals because gathering more details from a crime witness, or suspect, might determine the outcome of the investigation (Fisher & Geiselman, 1992).

We found that free report was largely responsible for the ECI superiority effect, because only at free report the ECI elicited significantly more details than the SI. Thus, mental reinstatement of context, report everything and transfer of control procedures that are exclusive to the ECI are essential to obtain more information during free report. Our study supports previous findings (Dando, Wilcock, & Milne, 2009; Dando, Wilcock, Milne, & Henry, 2009; Davis, McMahon, & Greenwood, 2005; Luca, Raffaella, Elisa, & Fiorella, 2011) that suggest that shortened ECI variants can be applied when time constraint is a major investigative issue. However, the interviewer should make sure to take the time to always include mental reinstatement of context, report everything and transfer of control instructions during the free report phase.

Accuracy, that is, the amount of correct items of information elicited by a participant divided by all the items of information reported by such participant, was similar across the whole interview and across different categories of information (person, action, etc.). Such findings have two major implications. First, they support that when appropriate open-ended questions are used, accuracy can be maintained during the whole interview. Accuracy values for the free report phase, where no questions were asked, and the questioning phase, where open-ended questions were used, were similar. Therefore, although the use of inadequate questions, such as leading questions, can have a negative impact on accuracy (Goodman & Melinder, 2007), we found that the use of adequate open-ended questions does not necessarily decrease this value. Second, because accuracy values were similar for all information categories, these results suggest that professionals should not consider witnesses to be more accurate when recalling details about a given category of information (e.g., person details), in comparison with another one (e.g., action details). Such results have major implications in the 'applied' field (e.g., when evaluating the credibility of a given detail in court).

Our study also supports the hypothesis that witnesses' perceptions regarding the interview can have a major impact on their report (Ballardin *et al.*, 2013; Fisher & Geiselman, 2010). To our knowledge, this is the first study to assess whether witnesses' perceptions could be related to the amount of recalled information, as suggested by Fisher & Geiselman (2010). We found that a higher perception of the interview appropriateness was correlated with a higher amount of reported information units and higher interest in being an interviewee. Such findings are supported by previous research that suggests that communicative and social factors, such as rapport and working alliance, are very important (Ballardin *et al.*, 2013; Vanderhallen & Vervaeke, 2014; Vanderhallen *et al.*, 2011) and associated with better recall (Fisher & Geiselman, 1992; Read *et al.*, 2009; Walsh & Bull, 2012). Our results are also consistent with findings from other fields of Psychology (e.g., Clinical Psychology), which repeatedly demonstrated that the clients' perception of the therapist attributes and the therapy itself can influence therapeutic alliance as well as the outcome of the therapy (Ackerman & Hilsenroth, 2003; Martin *et al.*, 2000). This has major implications for real-life investigations. For instance, the interviewer should always explain to the witness why every procedure is being used during the interview (e.g., explaining why it is important to close her/his eyes). If the interviewer fails to do this, even what could be the most effective procedure can be perceived as inappropriate by the interviewee and have a negative impact on recall.

Previous research suggests that witnesses are able to use metacognitive techniques to monitor their own report, such as report option (Koriat & Goldsmith, 1996), confidence judgments (Allwood *et al.*, 2005; Roberts & Higham, 2002), adjusting report precision (Evans & Fisher, 2010) or frequency judgments (Sniezek & Buckley, 1991). Furthermore, metacognitive techniques can be effectively used in very different situations and contexts, such as selections from lineups (Lindsay *et al.*, 2013), cued recall tasks (Luna & Martín-Luengo, 2012) and investigative interviews, such

as the ECI (Allwood *et al.*, 2005; Paulo *et al.*, 2015; Roberts & Higham, 2002). However, to our knowledge, this was the first study to evaluate if witnesses are able to provide an accurate error estimate for their report, and we found that witnesses were unable to estimate their error rate for their overall report and for each individual interview phase. Such results are supported by previous literature that suggests that when confidence judgments are accurate, frequency judgments usually result in underconfidence (Sniezek & Buckley, 1991). Using a very similar methodology to the one we have adopted for this study, several authors (Allwood *et al.*, 2005; Paulo *et al.*, 2015; Roberts & Higham, 2002) found that witnesses were able to use different types of confidence judgments (e.g., numerical vs qualitative; instructed vs. spontaneous) to evaluate report accuracy. According to Sniezek and Buckley (1991), this realism for confidence judgments performed during, or after, the course of investigative interviews might explain why we found that participants over-evaluated their error rate when frequency judgments were asked, believing that they committed more errors and confabulations than they actually did. Such results support Sniezek and Buckley's (1991) dual-process account theory stating that confidence judgments and frequency judgments have very different natures and different accuracy values can be expected when both judgments are used in similar situations. Furthermore, both interviews (ECI and SI) remarkably achieved very high levels of accuracy, which might further explain why participants were underconfident and over-evaluated their error rate. Therefore, our study does not support using frequency judgments to evaluate report accuracy in an interview setting. Other methods to evaluate and enhance report accuracy, such as accounting for witnesses' spontaneous verbal confidence judgments (Paulo *et al.*, 2015), have shown to be much more effective for this purpose.

Lastly, our results suggest that there is no relation between witnesses' performance on memory tests and their subsequent performance during the interview. Although Morgan *et al.* (2007) found that trait ability to remember human faces was related to witnesses' accuracy when selecting a target person at a sequential photo presentation, we could not replicate such findings on an interview setting. Many factors can explain why we found different results. First, we asked participants to describe with as much detail as possible the target subject (recall task), instead of identifying her/him at a photo presentation (recognition task). Many authors (e.g., Bower, 2000) suggest that recall tasks differ considerably from recognition tasks, not only in terms of memory capacity but also in terms of how they affect different phenomena's (e.g., false memories production). Second, our study was not conducted under highly emotional circumstances during the encoding phase, and highly emotional events are remembered differently from neutral events (Reisberg & Heuer, 2004).

Furthermore, it is quite possible that report quality cannot be estimated by memory test performance because variables such as witnesses' motivation (Paulo *et al.*, 2015), perceptions about the interview and interviewer (Ballardin *et al.*, 2013), as well as interviewers' performance (among others), can have a major impact on witnesses' report on the course of an investigative interview and need to be accounted for

evaluating witnesses' performance (Fisher & Geiselman, 2010). This might explain why none of the memory capacity tests scores was related to witnesses' subsequent performance. These results have a major impact for our judicial system. Judges, attorneys, police officers, or other relevant professionals should not consider a witness's report to be 'poor' because she or he had low results on a psychological memory evaluation or self-reported to have probably committed many mistakes during recall. Instead, these professionals should account other factors to evaluate report quality, such as the type of questioning used during the interview (Fisher & Geiselman, 1992).

## ACKNOWLEDGEMENTS

We express our gratitude to the Portuguese governmental institution 'Fundação para a Ciência e a Tecnologia' (grant number: SFRH/BD/84817/2012) for funding this study, as well as Dr Becky Milne (of the University of Portsmouth) for her help and support.

## REFERENCES

- Ackerman, S. J., & Hilsenroth, M. J. (2003). A review of therapist characteristics and techniques positively impacting the therapeutic alliance. *Clinical Psychology Review, 23*, 1–33. doi: 10.1016/S0272-7358(02)00146-0
- Allwood, C., Ask, K., & Granhag, P. (2005). The Cognitive Interview: Effects on the realism in witnesses' confidence in their free recall. *Psychology, Crime & Law, 11*, 183–198. doi: 10.1080/10683160512331329943
- Aschermann, E., Mantwill, M., & Köhnken, G. (1991). An independent replication of the effectiveness of the Cognitive Interview. *Applied Cognitive Psychology, 5*, 489–495. doi: 10.1002/acp.2350050604
- Baddeley, A. D. (2002). Is working memory still working? *European Psychologist, 7*, 85–97. doi: 10.1027//1016-9040.7.2.85
- Ballardin, M., Stein, L., & Milne, R. (2013). Além das técnicas de entrevista: características individuais em entrevista investigativa com testemunhas. [Beyond the interview techniques: Individual characteristics in investigative interviews with witnesses]. *Revista Brasileira de Segurança Pública, 7*, 6–16.
- Bower, G. (1967). A multicomponent theory of the memory trace. In K.W. Spence, & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 1, pp. 229–325). New York, NY: Academic Press.
- Bower, G. (2000). A brief history of human memory. In E. Tulving, & F.I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 3–32). New York: Oxford University Press.
- Brewer, N., Weber, N., Wootton, D., & Lindsay, S. (2012). Identifying the bad guy in a lineup using confidence judgments under deadline pressure. *Psychological Science, 23*, 1208–1214. doi: 10.1177/0956797612441217
- Campos, L., & Alonso-Quecuty, M. (2008). Language crimes and the Cognitive Interview: Testing its efficacy in retrieving a conversational event. *Applied Cognitive Psychology, 22*, 1211–1227. doi: 10.1002/acp.1430
- Colomb, C., & Ginet, M. (2012). The Cognitive Interview for use with adults: An empirical test of an alternative mnemonic and of a partial protocol. *Applied Cognitive Psychology, 26*, 35–47. doi: 10.1002/acp.1792
- Colomb, C., Ginet, M., Wright, D., Demarchi, S., & Sadler, C. (2013). Back to the real: Efficacy and perception of a modified cognitive interview in the field. *Applied Cognitive Psychology, 27*, 574–583. doi: 10.1002/acp.2942
- Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12*, 769–786. doi: 10.3758/BF03196772
- Dando, C. J., Wilcock, R., & Milne, R. (2009). The Cognitive Interview: The efficacy of a modified mental reinstatement of context procedure for frontline police investigators. *Applied Cognitive Psychology, 23*, 138–147. doi: 10.1002/acp.1451
- Dando, C. J., Wilcock, R., Milne, R., & Henry, L. (2009). A modified cognitive interview procedure for frontline police investigators. *Applied Cognitive Psychology, 23*, 698–716. doi: 10.1002/acp.1501
- Davis, M., McMahan, M., & Greenwood, K. (2005). The efficacy of mnemonic components of the Cognitive Interview: Towards a shortened variant for time-critical investigations. *Applied Cognitive Psychology, 19*, 75–93. doi: 10.1002/acp.1048
- Evans, J. R., & Fisher, R. P. (2010). Eyewitness memory: Balancing the accuracy, precision and quantity of information through metacognitive monitoring and control. *Applied Cognitive Psychology, 25*, 501–508. doi: 10.1002/acp.1722
- Field, A. P. (2009). *Discovering statistics using SPSS*. London, Middx: Sage.
- Fisher, R. P., & Geiselman, R. E. (1992). *Memory-enhancing techniques for investigative interviewing: The Cognitive Interview*. Springfield, IL: Charles C. Thomas.
- Fisher, R. P., & Geiselman, R. E. (2010). The Cognitive Interview method of conducting police interviews: Eliciting extensive information and promoting therapeutic jurisprudence. *International Journal of Law and Psychiatry, 33*, 321–328. doi: 10.1016/j.ijlp.2010.09.004
- Geiselman, R. E., Fisher, R. P., Firstenberg, I., Hutton, L., Sullivan, S. J., Avetissian, I. V., & Prosk, A. L. (1984). Enhancement of eyewitness memory: An empirical evaluation of the Cognitive Interview. *Journal of Police and Science Administration, 12*, 74–80.
- Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review, 98*, 506–528. doi: 10.1037/0033-295X.98.4.506
- Higham, P. A., Luna, K., & Bloomfield, J. (2010). Trace-strength and source-monitoring accounts of accuracy and metacognitive resolution in the misinformation paradigm. *Applied Cognitive Psychology, 25*, 324–335. doi: 10.1002/acp.1694
- Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review, 103*, 490–517.
- Goodman, G. S., & Melinder, A. (2007). Child witness research and forensic interviews of young children: A review. *Criminal and Criminological Psychology, 12*, 1–19. doi: 10.1348/135532506X156620
- Larsson, A. S., Granhag, P. A., & Spjut, E. (2002). Children's recall and the Cognitive Interview: Do the positive effects hold over time? *Applied Cognitive Psychology, 17*, 203–214. doi: 10.1002/acp.863
- Liberman, V. (2004). Local and global judgements of confidence. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 729–732. doi: 10.1037/0278-7393.30.3.729
- Lindsay, R., Kalmel, N., Leung, J., Bertrand, M., Sauer, J., & Sauerland, M. (2013). Confidence and accuracy of lineup selections and rejections: Postdicting rejection accuracy with confidence. *Journal of Applied Research in Memory and Cognition, 2*, 179–184. doi: 10.1016/j.jarmac.2013.06.002
- Luca, B., Raffaella, N., Elisa, G., & Fiorella, G. (2011). The Enhanced Cognitive Interview: A study on the efficacy of shortened variants and single techniques. *Journal of Cognitive Psychology, 23*, 311–321. doi: 10.1080/20445911.2011.497485
- Luna, K., & Martín-Luengo, B. (2012). Confidence-accuracy calibration with general knowledge and eyewitness memory cued recall questions. *Applied Cognitive Psychology, 26*, 289–295. doi: 10.1002/acp.1822
- Martin, D. J., Garske, J. P., & Davis, M. K. (2000). Relation of the therapeutic alliance with outcome and other variables: A meta-analytic review. *Journal of Consulting and Clinical Psychology, 68*, 438–450. doi: 10.1037/0022-006X.68.3.43
- Memon, A., Holley, A., Wark, L., Bull, R., & Köhnken, G. (1996). Reducing suggestibility in child witness interviews. *Applied Cognitive Psychology, 10*, 503–518. doi: 10.1002/(SIECI)1099-0720(199612)10:63.0.CO;2-R
- Milne, R., & Bull, R. (2003). Does the Cognitive Interview help children to resist the effects of suggestive questioning? *Legal and Criminological Psychology, 8*, 21–38. doi: 10.1348/135532503762871219
- Morgan, C. A., Hazlett, G., Baranoski, M., Doran, A., Southwick, S., & Loftus, E. (2007). Accuracy of eyewitness identification is significantly associated with performance on a standardized test of face recognition. *International Journal of Law and Psychiatry, 30*, 213–223. doi: 10.1016/j.ijlp.2007.03.005
- Paulo, R. M., Albuquerque, P. B., & Bull, R. (2013). The Enhanced Cognitive Interview: Towards a better use and understanding of this procedure.

- International Journal of Police Science & Management*, 15, 190–199. doi: 10.1350/ijps.2013.15.3.311
- Paulo, R. M., Albuquerque, P. B., & Bull, R. (2015). The Enhanced Cognitive Interview: Evaluating accuracy by accounting metacognitive monitoring and motivation. Manuscript submitted for publication.
- Prescott, K., Milne, R., Clark, J. (2011). How effective is the Enhanced Cognitive Interview when aiding recall retrieval of older adults including memory for conversation? *Journal of Investigative Psychology and Offender Profiling*, 8, 257–270. doi: 10.1002/jip.142
- Reisberg, D., & Heuer, F. (2004). Memory for emotional events. In D. Reisberg, & P. Hertel (Eds.), *Emotion and memory* (pp. 3–41). Oxford, England: Oxford University Press.
- Riccó, A. (Director), & Riccó, R. (Director) (2004). O Assalto [The robbery] [Television series episode]. In V. Castelo (Producer), *Inspector Max*. Lisbon: Produções Fictícias.
- Rivard, J. R., Fisher, R. P., Robertson, B., & Mueller, D. H. (2014). Testing the Cognitive Interview with professional interviewers: Enhancing recall of specific details of recurring events. *Applied Cognitive Psychology*. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1002/acp.3026/full>
- Read, J. M., Powell, M. B., Kebbell, M. R., & Milne, R. (2009). Investigative interviewing of suspected sex offenders: A review of what constitutes best practice. *International Journal of Police Science & Management*, 11, 442–459. doi: 10.1350/ijps.2009.00.0.143
- Roberts, W.T., & Higham, P.A. (2002). Selecting accurate statements from the Cognitive Interview using confidence ratings. *Journal of Experimental Psychology: Applied*, 8, 33–43. doi: 10.1037/1076-898X.8.1.33
- Snizek, J. A., & Buckley, T. (1991). Confidence depends on level of aggregation. *Journal of Behavioral Decision Making*, 4, 263–272. doi: 10.1002/bdm.3960040404
- Stein, L. M., & Memon, A. (2006). Testing the efficacy of the Cognitive Interview in a developing country. *Applied Cognitive Psychology*, 20, 597–605. doi: 10.1002/acp.1211
- Tulving, E. (1991). Concepts of human memory. In L. R. Squire, N. M. Weinberger, G. Lynch, & J. L. McGaugh (Eds.), *Memory: Organization and locus of change* (pp. 3–32). New York: Oxford University Press.
- Tulving, E., & Thomson, D. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 80, 352–373.
- Vanderhallen, M., Vervaeke, G., & Holmberg, U. (2011). Witness and suspect perceptions of working alliance and interviewing style. *Journal of Investigative Psychology and Offender Profiling*, 8, 110–130. doi: 10.1002/jip.138
- Vanderhallen, M., & Vervaeke, G. (2014). Between interviewer and suspect: The role of the working alliance in investigative interviewing. In R. Bull (Ed.), *Investigative interviewing* (pp. 63–90). New York: Springer.
- Verkampt, F., & Ginet, M. (2009). Variations of the Cognitive Interview: Which one is the most effective in enhancing children's testimonies? *Applied Cognitive Psychology*, 24, 1279–1296. doi: 10.1002/acp.1631
- Vrij, A., Mann, S. A., Fisher, R. P., Leal, S., Milne, R., & Bull, R. (2008). Increasing cognitive load to facilitate lie detection: The benefit of recalling an event in reverse order. *Law and Human Behavior*, 32, 253–265. doi: 10.1007/s10979-007-9103-y
- Walsh, D., & Bull, R. (2012). Examining rapport in investigative interviews with suspects: Does its building and maintenance work? *Journal of Police and Criminal Psychology*, 27, 73–84. doi: 10.1007/s11896-011-9087-x
- Wechsler, D. (1997). *WMS-III: Wechsler Memory Scale Administration and Scoring Manual* (3rd ed.). San Antonio, TX: PsychCorp.
- Wright, A., & Holliday, R. (2006). Enhancing the recall of young, young-old and old-old adults with cognitive interviews. *Applied Cognitive Psychology*, 21, 19–43. doi: 10.1002/acp.1260