

Running head: CONFIDENCE-ACCURACY IN SOCIAL SITUATIONS

Witness confidence and accuracy: is a positive relationship maintained for recall under interview conditions?

AUTHOR

Mark R. Kebbell

ARC Centre of Excellence in Policing and Security Griffith University

\*Correspondence to: Mark Kebbell, Ph.D., ARC Centre of Excellence in Policing and Security, Mt Gravatt Campus, Griffith University, Australia, 4121.

E-mail: [m.kebbell@griffith.edu.au](mailto:m.kebbell@griffith.edu.au)

## Abstract

A large positive correlation between eyewitness recall confidence and accuracy (C-A) is found in research when item difficulty is varied to include easy questions. However, these results are based on questionnaire responses. In real interviews, the social nature of the interview may influence C-A relationships, and it is the interviewer's perception of the accuracy of a witness that counts. This study was conducted to investigate the influence of these factors for recall of a video. Three conditions were used; the same questions were used in each. Participants in condition one (self-rate questionnaire condition, N=20) were given a questionnaire that required them to answer questions and rate confidence on a scale. Pairs of participants in condition two (self-rate interview condition, N=40) were given the role of eyewitness or interviewer. Eyewitnesses were asked questions by an interviewer and responded orally with answers and confidence judgements on a Likert scale. Participants in condition three (interviewer-rate interview condition, N=40) were tested in the same way as condition two but provided confidence judgements in their own words. Interviewers independently rated each confidence judgement on the Likert scale. The experiment showed high C-A relationships, particularly for 'absolutely sure' responses. The main effect of the social interview condition was to increase confidence in correct answers but not incorrect answers. However, the advantage of this effect was tempered by the fact that although observers can differentiate between confident and less confident answers, less extreme confidence judgements were ascribed.

Key words: eyewitness; confidence; confidence-accuracy; interviewing;

## INTRODUCTION

Research has shown that the confidence that eyewitnesses express in information heavily influences both the investigative process and the credence which jurors give eyewitness testimony. If eyewitnesses express certainty that their answers are correct, their responses are likely to be perceived as correct (Brigham & Wolfskeil, 1983; Cutler, Penrod & Dexter, 1990; Cutler, Penrod & Stuve, 1988; Fox & Walters, 1986; Leippe, Manion & Romanczyk, 1992; Wells, Lindsay & Ferguson, 1979; Lindsay, Wells & O'Connor, 1989). Furthermore, surveys of the general public reveal a substantial, cross-cultural, belief that confidence predicts accuracy (Brigham & Bothwell, 1983; Deffenbacher & Loftus, 1982; Noon & Hollin, 1987; Sporer, 1983; Yarmey & Jones, 1983). These perceptions are consistent with the literature in experimental cognitive psychology that shows a moderate, yet robust, positive C-A relationship (see Nelson, 1988 for a review).

However, despite the intuitions of the public, investigators, jurists and the findings of the cognitive literature that there is a positive relationship between eyewitnesses' confidence and their accuracy, much research in eyewitness testimony appears to contradict this assumption. Reviews suggest that there is either no relationship, or only a small positive relationship between eyewitnesses' confidence and their accuracy (Bothwell, Deffenbacher, & Brigham, 1987; Cutler & Penrod, 1989; 1995; Deffenbacher, 1980; Fruzzetti, Tolland, Teller, & Loftus, 1992; Wells and Lindsay, 1985; Wells & Murray, 1984; but see also, Sporer, Penrod, Read, & Cutler, 1995; although this view may be changing, see, for example, Read, Lindsay, & Nicholls, 1998). Nevertheless, the consensus that eyewitness accuracy is, for practical purposes, unrelated to eyewitness confidence requires qualification for two reasons. First, most research has focussed on C-A relationships for identifications. Second, most

research has used 'between-participants' rather than 'within-participants' analyses. The importance of these two points is discussed next.

Previous work on eyewitness C-A relations has concentrated on identifications. For instance the recognition of a perpetrator from a line-up. Clearly, this is an important area of research. Being able to identify offenders is essential for the Criminal Justice System, and as Huff, Rattner and Sagarin (1996) point out, the evidence of an eyewitness who is confident in a false identification is likely to lead to a miscarriage of justice. However, eyewitness confidence judgements are not only used for assessing the accuracy of identifications. Kebbell and Wagstaff (1997a) suggest that the three principal reasons for eyewitness testimony are: 1) to discover whether a crime has been committed and if so, what crime; 2) to find evidence to identify the individual responsible; and 3), to produce evidence that prevents a *guilty* criminal from using an inappropriate defence (for example, false claims of self-defence). Consequently, in many situations eyewitness confidence is critical to determining the accuracy of information that does not involve identifications. This is an important distinction. Essentially it is a distinction between memory that relies on recognition and memory that relies on recall. Recognition and recall tasks are relatively independent of one-another (Flexser & Tulving, 1978; Morris & Gruneberg, 1994; Pezdek & Greene, 1993; Raaijmakers & Shiffrin, 1992). Indeed, individuals with severely impaired recall abilities can show high recognition abilities (Hanley, Davies, Downes & Mayes, 1994; Moscovitch, 1989). Therefore, a low C-A relationship for eyewitness identifications does not necessarily mean a low C-A relationship for eyewitness recall.

A second reservation concerning eyewitness C-A relationships is that researchers have concentrated on the confidence-accuracy (C-A) relationships 'between-participants',

comparing the accuracy of confident witnesses to less confident witnesses, rather than the relationship within-participants' own statements (Smith, Kassin & Ellsworth, 1989). In the latter case, an eyewitness says that he or she is absolutely certain of some details but not of others and the relationship between these confidence judgements and accuracy is calculated. The concentration on between-participants C-A relationships is not surprising given that most previous work has concentrated on identifications. In real situations eyewitnesses are usually required to identify only one perpetrator (Christianson & Hubinette, 1993; Yuille, 1986; Yuille & Cutshall, 1986). Consequently, experimental analogues of these events have relied on stimuli that typically involve only one target individual. Therefore, within-participants C-A relationships are not calculated because of a lack of data points and because the analysis would have little or no ecological relevance for identifications.

However, for recall information, what Smith, Kassin, and Ellsworth (1989) term 'within-participants' C-A relationships are more likely than between-participant C-A relationships to have practical implications for three reasons. First, criminal investigations and trials often have only one eyewitness (Yuille, 1986), thus the only C-A relationships that can be assessed are within-participants' own testimony. Second, where more than one eyewitness is available, they are interviewed by police officers or examined in court one at a time. Consequently, the relative accuracy of parts of their testimony is assessed in a within-participants manner. Third, within-participants analysis is a more realistic approximation of how jurors make decisions than a between-participants analysis. For example, a critical piece of evidence could be that an eyewitness is very confident that he or she saw a gun and that he or she saw the robber threaten a bank teller with the gun. However, he or she might be very unsure about other details such as the colour of the robber's belt and the time of the incident.

Here jurors are most likely to rely on the witness' relative confidence in specific details to assess the accuracy of those details rather than on the assessment of confidence in general.

Nevertheless, even if within-participants factors are considered, C-A relationships are not necessarily great. Consideration of within-subject C-A accuracy relationships by Smith, Kassin, and Ellsworth (1989) found a correlation of  $r=.17$  whilst Perfect, Watson and Wagstaff (1993) found a higher Goodman-Kruskal Gamma of .49 and Perfect found a still higher Goodman-Kruskal Gamma of .58 (Experiment 2, Perfect, 2004; and for a more detailed review, Perfect, 2002). Kebbell, Wagstaff, and Covey (1996) suggest that a critical aspect determining C-A relationships is the issue of item difficulty. Typically in work on C-A relationships, researchers have attempted to select items to avoid floor and ceiling effects, i.e. they have tried to avoid items that are either very easy or very hard to remember. Still, in real-life forensic investigations some questions that eyewitnesses are asked may be easier to answer than others. For example, in the case of the robbery mentioned earlier, an eyewitness may be asked 'did the robber have a gun?' Details concerning the presence or absence of a gun appear to be well remembered (Stebly, 1992) and so this question is very likely to be answered accurately. Furthermore, most eyewitnesses are likely to be very confident that their memory of whether a gun was used is correct. Alternatively, if eyewitnesses are asked, for example, 'what was the colour of the attacker's belt?' This question might be more difficult to answer and eyewitnesses may be less confident about their accuracy. Kebbell, Wagstaff, and Covey (1996) argue that in an attempt to avoid ceiling and floor effects, previous researchers may have chosen unrealistic and overly homogeneous pools of items, thus reducing the variance necessary for high correlations (see also, Brewer, Caon, Todd, & Weber, 2006; and

Weber & Brewer 2004; for a discussion of manipulation of item difficulty with regards identifications).

To assess this hypothesis Kebbell, Wagstaff, and Covey (1996) conducted two experiments that measured C-A relationships in response to information seen in video films. Like Smith, Kassin, and Ellsworth (1989) and Perfect, Watson, and Wagstaff (1993) participants were required to answer questions using a questionnaire. Participants were required to rate their confidence on a Likert scale. In each case, participants were required to answer a number of questions that ranged in difficulty, from easy to hard. Higher C-A correlations than typically reported were found in both experiments (overall  $r=.54$  and  $r=.78$ ) and this was particularly the case when easy questions were included. These findings have since been replicated (Kebbell & Wagstaff, 1997b; Perfect & Hollins, 1996; Pike, Towell, & Kemp, 1995).

Even so, these findings should be interpreted with caution. The way an eyewitness is interviewed can distort C-A relationships. For instance, suggestive information, repeated questioning, and preparation for cross-examination have all been shown to have an adverse influence on C-A relationships (Luus & Wells, 1994; Shaw, 1996; Shaw & McClure, 1996; Wells, Ferguson, & Lindsay, 1981). However, the above factors are what Wells (1978) terms 'system variables'. They are factors that can be influenced by the Criminal Justice System and so the above problems could potentially be minimized. Many researchers suggest that these steps should be taken (Luus & Wells, 1994; Shaw, 1996; Shaw & McClure, 1996; Wells, 1978; Wells, Ferguson, & Lindsay, 1981). However, there seems little point working on ways to reduce the adverse impact on system variables on C-A relationships if still no practical C-A relationship exists even in non-leading social situations and even without the problem of the

system variables that have been mentioned, two other important problems are likely to remain.

First, real interviews occur in a social environment in which questions are asked verbally during an interaction between interviewer and eyewitness. However, most research into C-A relationships has relied on questionnaires administered to groups (Kebbell, Wagstaff, and Covey, 1996; Perfect & Hollins, 1996; Perfect, Watson, & Wagstaff, 1993; Shaw, 1996; Shaw & McClure, 1996; Smith, Kassin, and Ellsworth, 1989; but see Gwyer & Clifford, 1997, although strictly speaking they do not use within-participants analysis). This social interaction might influence C-A relationships. In an interview it seems reasonable to assume that an eyewitness wishes to appear 'a good witness' (Kebbell, Wagstaff, and Covey, 1996; Wagstaff, 1981); that is to say, to appear accurate. As pointed out earlier, people believe that accuracy is indicated by confidence. Therefore, the social dynamics of an interview may encourage a witnesses to be confident so that he or she appears accurate (Moston, 1990; Velicer, Prochaska, Rossi, & Snow, 1992; Wagstaff & Mercer, 1993). This might encourage overconfidence and suppress C-A relationships.

Second, in real situations, if a positive within-participants C-A relationship exists, it is important that confidence judgements are communicated to those who need them, for instance investigators and jurors. However, real eyewitnesses do not communicate confidence on a Likert scale. Instead confidence is communicated through eyewitnesses' statements of confidence using their own words and through additional factors such as: tone of voice, eye contact and perceived nervousness (Kebbell, Wagstaff, and Covey, 1996; Kovera, Gresham, Borgida, Gray, & Regan, 1997; Murray & Wells, 1982). Agreement among participant-jurors in ascribing confidence to eyewitnesses is extremely high (Lindsay, Wells, & Rumpel, 1981;

Wells & Leippe, 1981; Wells, Lindsay, & Ferguson, 1979), even in the coding of open-ended statements (Murray & Wells, 1982). However, inter-judge agreement is not so high when one judge is the eyewitness and the other the observer (e.g. juror). In general, Wells and Murray (1984) suggest that observers' ratings of eyewitness confidence seem to regress on the eyewitness's self rated confidence. That is, eyewitnesses who rate themselves as either extremely high or extremely low in confidence will have less extreme confidence judgements ascribed to them by participant-jurors. This suppresses the C-A relationship between eyewitness and observer to .55, using a between-participants correlation. Such a finding could limit the usefulness of confidence judgements, in particular if jurors have to decide 'beyond reasonable doubt' the guilt of an individual. If they rate witnesses who are absolutely certain about critical details as less than absolutely certain, incorrect decisions could be made. However, again, previous work on individuals' ability to rate confidence has only concerned confidence in identifications and between-participants comparisons. Within-participants evaluations have been ignored, but these might be particularly high because of the ability of the observer to calibrate C-A relationships over successive questions. Calibration appears to be an important factor for determining C-A relationships (see Perfect, Watson, & Wagstaff, 1993; Wells, Lindsay, & Ferguson, 1979). Thus, C-A relationships might be sustained, even for observers, for within-participants analyses.

Consequently, the following experiment was conducted in what can be termed 'ideal but naturalistic', circumstances (e.g. short delay, non-suggestive questioning). If a strong C-A relationship was found, future work might seek to investigate how that relationship could be maintained in criminal procedures. The purpose of the following experiment, therefore, was to investigate the influence of a non-leading interactive social interview and of witnesses'

ability to communicate their confidence to an interviewer using a within-participants design. Participant-eyewitnesses were compared across three conditions. Each condition was asked the same questions. The first condition used a conventional questionnaire with Likert scales for confidence judgements. The second used an interview in which participant-eyewitnesses gave confidence judgements, using a Likert scale, to an interviewer. The third condition used an interview in which participant eyewitnesses orally gave confidence judgements in whatever way they wanted (i.e. without a Likert scale). Their confidence was rated by the interviewer.

## Method

### *Participants*

One hundred undergraduates (65 female and 35 male) participated. Mean age was 26.71 years (range 16-57,  $SD=7.79$ ). Participants were recruited and tested in pairs. Pairs of participants were randomly assigned to one of three conditions; 'self-rate questionnaire', 'self-rate interview' or 'interviewer-rate interview'.

### *Materials and Procedure*

The procedure for each of the three conditions was as follows:

*Self-rate questionnaire condition.* For the self-rate questionnaire condition both participants were shown a five and a half-minute video film that concerned the implied murder of a male

by a female. After the film participants were given a 'filler' reading-task for five minutes. Next participants were given a questionnaire to complete. This was a 33 item questionnaire (similar to that used by Kebbell, Wagstaff, and Covey, 1996) devised to test their recall of the film. Questions were open-ended (i.e. not forced-choice), but participants were instructed to provide an answer for each question, even if this were only a guess. The questionnaire was devised to include items of varied item difficulty (i.e. easy, medium and hard questions were included). To reduce the influence of being correct by chance, each question was devised so that a range of plausible answers was possible. For example, an easy question was 'where was the dead body found?', A medium difficulty question was 'what was on the dish next to the television set?', And a hard question was 'what was behind the Tabasco sauce bottle in the kitchen?' After answering each question participants were required to rate their confidence in their answer on a ten-point Likert scale calibrated as 'pure guess' (1), slightly confident (3), fairly confident (between 5 and 6), very confident (8) and 'absolutely sure' (10).

For the remaining conditions participants were randomly assigned the role of 'interviewer' or 'eyewitness'. Interviewers waited in a separate room while eyewitnesses watched the film and completed the filler task. Once this task had been completed, witnesses' recall was tested in one of the following conditions.

*Self-rate interview condition.* Participants given the role of eyewitness were interviewed by the participant given the role of interviewer. Similar instructions were given to participants as for the self-rate questionnaire condition concerning the need to answer all questions and to provide confidence judgements for each answer. However, in this condition questions were

asked orally by the interviewer who then recorded witnesses' answers. Witnesses were asked how confident they were in the accuracy of each answer. The ten-point Likert scale was provided for them to communicate their confidence judgements to the interviewer. For example, if a participant-witness was very confident of an answer he or she would tell the interviewer '8'.

*Interviewer-rate interview condition.* The interviewer-rate interview condition was identical to the self-rate interview condition with one exception. Participant-witnesses were asked to communicate confidence judgements in whatever way they thought appropriate (i.e. participants communicated confidence judgements to the interviewer without using the Likert scale). For example, if a participant-witness was absolutely sure of an answer, he or she might say 'absolutely certain' or '100% sure' or whatever way he or she wished. Participant-interviewers rated the confidence of eyewitnesses in each answer in terms of how confident the interviewer *believed the eyewitness was* in each answer (i.e. not how confident the *interviewer was* that the answer provided by the eyewitness was correct). Interviewers were asked to use whatever cues they wanted to decide confidence, for instance verbal expression, eye-contact and tone of voice. Interviewers were required to make their confidence judgements by circling a number on the same ten-point Likert scale as described in the previous conditions.

## Results

To prevent the influence of interview condition being masked by floor and ceiling effects, a strict criterion was adopted for inclusion of questions in the analysis. Each question had to be answered correctly or incorrectly by at least 10% of the participants in each condition. Thus, questions that were answered correctly or incorrectly by all participants, or almost all participants in a condition, were excluded. This left 16 questions for analysis.

Correct answers for each condition were compared with a one-way ANOVA (self-rate questionnaire/self-rate interview/interviewer-rate interview), this was not significant,  $F(2,57)=2.56$ ,  $p>.05$ . The means and standard deviations were as follows: self-rate questionnaire condition  $M=8.20$  ( $SD=2.09$ ), self-rate interview condition  $M=9.30$  ( $SD=1.81$ ); and, interviewer-rate interview condition  $M=7.90$  ( $SD=2.27$ ).

A 'within-participants' C-A correlation was calculated for each subject across the 16 questions. These correlations assess the relationship between an individual's confidence and his or her accuracy. These were averaged for each condition and are displayed in Table 1. As with previous work in this area because of the number of ties and the non-parametric nature of the data, a Goodman-Kruskal gamma correlation was calculated for each participant (see Nelson, 1984; Perfect & Hollins, 1996), and these correlations were tested for significance against zero with  $t$ -tests. All were large and significant. A one-way ANOVA (self-rate questionnaire/self-rate interview/interviewer-rate interview) showed no significant difference between the three conditions ( $F(2,57)=0.49$ ,  $p>.05$ ).

Each participant's average accuracy was correlated with his or her average confidence rating. These data were suitable for analysis using Pearson product-moment correlations (see Howell, 1992). These 'average' correlations are also shown in Table 1. Only the correlation for the self-rate questionnaire was significant.

-----  
insert Table 1 about here  
-----

As another way of understanding the data, participants' average confidence in correct answers was compared with their average confidence in incorrect answers for the three conditions with a 3 X 2 ANOVA (self-rate questionnaire/self-rate interview/interviewer-rate interview X confidence in correct/incorrect answers) with repeated measures on the second factor. These data are displayed in Table 2. A significant difference was found between conditions,  $F(2,57)=5.00, p<.01$ . Follow-up  $t$ -tests ( $p<.05$ ) indicated that participants in the self-rate interview, condition expressed higher overall confidence (in correct and incorrect answers) than participants in the self-rate questionnaire and interviewer-rate interview conditions. Participants in the self-rate questionnaire and interviewer rate interview conditions did not differ from one-another on overall confidence. The ANOVA also indicated that participants expressed significantly higher confidence in correct answers than in incorrect answers,  $F(1,57)=622.62, p<..001$ . There was no significant interaction between condition and confidence in correct and incorrect answers,  $F(2,57)=2.03, p>.05$ .

-----  
insert Table 2 about here  
-----

Finally, absolutely sure responses were considered independently. In the self-rate questionnaire condition participants made an average of 3.05 (SD=1.85) *absolutely sure* responses of which 94 % were correct; for the self-rate interview condition 5.40 (SD=2.41) absolutely certain responses were made of which 96% were correct; and, for the interviewer-

rate interview condition 1.70 (SD=1.92) absolutely certain responses were made of which 98% were correct. A Kruskal-Wallis test (self-rate questionnaire/self-rate interview/interviewer-rate interview) was used to compare the number of absolutely sure responses (correct and incorrect combined) across the three conditions. This was significant,  $\chi^2(2) = 21.02, p < .001$ . Follow-up Wilcoxon-Mann-Whitney tests ( $p < .05$ ) showed that the self-rate interview condition made more absolutely sure responses than the self-rate questionnaire condition. In turn, the self-rate questionnaire condition made more absolutely sure responses than were recorded for the interviewer-rate interview condition.

## Discussion

The results show that within-participants C-A correlations were very high, positive and significant. No difference in this C-A correlation or accuracy was found between the self-rate questionnaire, self-rate interview and interviewer-rate interview conditions. Using the example provided in the introduction, this means if an eyewitness is absolutely sure that a robber had a gun but only guesses that the robber's belt was brown, the eyewitness is more likely to be correct about the robber having a gun than having a brown belt.

However, high C-A correlations do not necessarily mean that a forensically useful C-A relationship exists. For example, with a Likert-scale a perfect C-A relationship (i.e. a correlation of 1) can be calculated from ten correct 'absolutely certain' responses and ten incorrect 'very confident' responses. The same correlation can also be calculated from ten correct 'absolutely certain' responses and ten incorrect 'pure guess' responses. Although both give the same overall correlation the second example would be more useful in real

investigations as discrimination between correct and incorrect information is more likely (see Gruneberg & Sykes, 1993). Consequently, the actual values of average confidence in correct and incorrect answers are important.

The ANOVA on confidence in correct and incorrect answers showed that participants were far more confident in correct answers (the average was approximately very confident for a correct response) than in incorrect answers (the average was approximately slightly confident), again suggesting that a witness's confidence can be used to infer accuracy. A significant difference between the three conditions was found with respect to confidence. Participants in the self-rate interview condition were significantly more confident than participants in the self-rate questionnaire condition. Thus, it appears that the demand characteristics of the interview situation may encourage higher confidence judgements (Moston, 1990; Velicer, Prochaska, Rossi, & Snow, 1992; Wagstaff & Mercer, 1993). It may be because eyewitnesses wish to appear to be 'good' witnesses and hold the belief that a good witness is a confident witness (Brigham & Bothwell, 1983; Deffenbacher & Loftus, 1982; Noon & Hollin, 1987; Sporer, 1983; Yarmey & Jones, 1983) Therefore, for eyewitnesses to appear to interviewers that they are good eyewitnesses they increase their confidence regardless of accuracy. Consequently, presence of an interviewer, at least in self-rated situations seems to increase confidence. However, in the interviewer-rate condition, interviewers rated the witnesses as significantly less confident than in the self-rate interview condition and no different to the self-rate questionnaire. This cannot be explained by the social nature of the interview as the same dynamics were also present for the self-rate interview condition. Therefore, concluding that interviewers are conservative in their ratings of witness confidence seems reasonable. The importance of this is discussed later.

Importantly, though unsurprisingly, the above findings were reflected to some extent in absolutely certain responses (absolutely sure responses were also part of the above data set). More absolutely sure responses were made in the self-rate interview condition than the self-rate questionnaire condition. Again this seems to reflect the demand characteristics of the social situation to increase confidence. However, in the interviewer-rate interview condition interviewers were far more conservative about rating witnesses as absolutely sure than the other two conditions. This seems to indicate a conservatism concerning eyewitness evidence found in between-participants designs (Wells, Lindsay & Ferguson, 1979). This finding may be very important as in most trials jurors must believe that an eyewitness is absolutely sure to produce a conviction 'beyond reasonable doubt'.

When the average accuracy rate and average confidence were correlated for each participant, a significant correlation was found only for the self-rate questionnaire condition. That is to say, participant eyewitnesses who were more confident about their answers overall were also more likely to be more accurate overall. However, this was not so for the other two, social interview conditions. This may be due to individual variations in personality and responses in the social interview condition. Some individuals might have increased confidence while others did not and accuracy remained the same. Consequently, the C-A relationship for this measure might have been suppressed.

## CONCLUSIONS

In sum, this experiment shows that: 1) participants are more confident about their correct answers than their incorrect answers; 2) when participants are 'absolutely sure' that an

answer is correct that are almost always accurate; 3) presence of an interviewer increases witness confidence, and 5) interviewers appear to rate witnesses as less confident than witnesses themselves. The results suggest a strong relationship can be present for recall. However, a realistic, social interview influences C-A relationships. The main effect in this experiment appears to be to increase confidence in correct answers and incorrect answers. However, the advantage of this effect is tempered by the fact that although observers can differentiate between confident and less confident answers they ascribe less extreme judgements to eyewitnesses who rate themselves as extremely high in confidence. In this respect the results agree with those of Wells, et al. (1979) for identifications.

Nonetheless, some caution is required in the interpretation of the results, in real situations the C-A relationships might be impaired for a variety of reasons. First, in real events question difficulty might not be so varied and in particular might not include easy questions that appear essential for high C-A relationships (Kebbell, Wagstaff, and Covey, 1996). However, this is unlikely. Many of the most important details of a crime (e.g. if a robber had a gun, if a mugger threatened his or her victim) are likely to be well remembered, easy and eyewitnesses are likely to express high levels of confidence in their answers (Christianson, 1992; Easterbrook, 1959). Still, some empirical consideration of C-A relationships with forensically important details seems worthwhile (Kebbell & Wagstaff, 1997a). Also, the results of this experiment would benefit from replication in more ecologically valid situations, for example, using a naturally occurring staged event and a longer time before testing.

Therefore, mapping out the boundary conditions of these positive C-A relationships is essential for future work. In particular, now that it is clear positive C-A relationships can

occur for recall information, understanding how procedures adopted by the Criminal Justice System influence C-A relationships seem to have renewed significance. Work on factors that can adversely effect C-A relationships may be even more critical than previously thought (Luus & Wells, 1994; Shaw, 1996; Shaw & McClure, 1996; Wells, 1978; Wells, Ferguson, & Lindsay, 1981).

## References

- Bothwell, R.C., Deffenbacher, K.A., & Brigham, J.C. (1987). Correlation of Eyewitness Accuracy and Confidence: Optimality Hypothesis Revisited. *Journal of Applied Psychology*, 72, 691-695.
- Brewer, N., Caon, A., Todd, C., Weber, N. (2006). Eyewitness identification accuracy and response latency. *Law and Human Behavior*, 30, 31-50.
- Brigham, J.C., and Bothwell, R.K. (1983). The ability of prospective jurors to estimate the accuracy of eyewitness identifications. *Law and Human Behavior*, 7, 19-30.
- Brigham, J.C., & Wolfskeil, M.P. (1983). Opinions of attorneys and law enforcement personnel on the accuracy of eyewitness identifications. *Law and Human Behavior*, 7, 337-349.
- Christianson, S.A. (1992). Emotional stress and eyewitness memory: A critical review. *Psychological Bulletin*, 112, 284-309.
- Christianson, S.A., & Hubinette, B. (1993). Hands up! A study of witnesses' emotional reactions and memories associated with bank robberies. *Applied Cognitive Psychology*, 7, 365-379.
- Cutler, B.L., & Penrod, S.D. (1989). Forensically relevant moderators of the relation between eyewitness identification accuracy and confidence. *Journal of Applied Psychology*, 74, 650-652.
- Cutler, B.L., & Penrod, S.D. (1995). Witness confidence and witness accuracy: Assessing their forensic relation. *Psychology, Public Policy, and Law*, 1, 817-845.

- Cutler, B.L., Penrod, S.D., & Dexter, H.R. (1990). Juror sensitivity to eyewitness identification evidence. *Law and Human Behavior, 14*, 185-191.
- Cutler, B.L., Penrod, S.D., & Stuve, T.E. (1988). Juror decision making in eyewitness identification cases. *Law and Human Behavior, 12*, 41-55.
- Deffenbacher, K.A. (1980). Eyewitness accuracy and confidence. *Law and Human Behavior, 4*, 243-260.
- Deffenbacher, K.A., & Loftus, E.F. (1982). Do jurors share a common understanding of eyewitness behaviour? *Law and Human Behavior, 6*, 15-30.
- Easterbrook, J.A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review, 66*, 183-201.
- Flexser, A.J. & Tulving, E. (1978). Retrieval independence in recognition and recall. *Psychological Review, 85*, 153-177.
- Fox, S.G., & Walters, H.A. (1986). The impact of general versus specific expert testimony and eyewitness confidence upon mock juror judgement. *Law and Human Behavior, 10*, 215-228.
- Fruzzetti, A.E., Tolland, K., Teller, S.A., & Loftus, E.F. (1992). Memory and eyewitness testimony. In M.M. Gruneberg & P.E. Morris (Eds) *Aspects of Memory: The Practical Aspects*, (pp. 18-50). London: Academic Press.
- Gruneberg, M.M., & Sykes, R.N. (1993). The generalisability of confidence-accuracy studies in eyewitnessing *Memory, 1*, 185-189.
- Gwyer, P. & Clifford, B.R. (1997). The effects of cognitive interview on recall, identification, confidence and the confidence/accuracy relationship. *Applied Cognitive Psychology, 11*, 121-145.

- Hanley, J.R., Davies, A.D.M., Downes, J.J. & Mayes, A.R. (1994). Impaired recall of verbal material following rupture and repair of an anterior communicating artery aneurysm. *Cognitive Neuropsychology*, *11*, 543-578.
- Howell, D.C. (1992). *Statistical Methods for Psychology*, 3rd Ed. Belmont, CA: Duxbury Press.
- Huff, R.C., Rattner, A., & Sagarin, E. (1996). *Convicted but Innocent: Wrongful Conviction and Public Policy*. Thousand Oaks, CA: Sage.
- Kebbell, M.R. & Wagstaff, G.F. (1997a). Why do the Police interview eyewitnesses? Interview objectives and the evaluation of eyewitness performance. *The Journal of Psychology*, *131*, 595-601.
- Kebbell, M.R. & Wagstaff, G.F. (1997b). An investigation into the influence of hypnosis on the confidence and accuracy of eyewitness recall. *Contemporary Hypnosis*, *14*, 157-166.
- Kebbell, M.R., Wagstaff, G.F. & Covey, J.A. (1996). The influence of item difficulty on the relationship between eyewitness confidence and accuracy. *British Journal of Psychology*, *87*, 653-662.
- Kovera, M.B., Gresham, A.W., Borgida, E., Gray, E., & Regan, P.C. (1997). Does expert psychological testimony inform or influence juror decision making? A social cognitive analysis. *Journal of Applied Psychology*, *82*, 178-191.
- Leippe, M.R., Manion, A.P., & Romanczyk, A. (1992). Eyewitness persuasion: How and how well do fact finders judge the accuracy of adults' and children's memory reports? *Journal of Personality and Social Psychology*, *63*, 181-197.

- Lindsay, R.C.L., Wells, G.L., & Rumpel, C. (1981). Can people detect eyewitness identification within and across situations? *Journal of Applied Psychology, 66*, 79-89.
- Lindsay, R.C.L., Wells, G.L., & O'Connor, F.J. (1989). Mock-juror belief or accurate and inaccurate eyewitnesses: A replication and extension. *Law and Human Behavior, 13*, 333-339.
- Luus, C.A.E. & Wells, G.L. (1994). The malleability of eyewitness confidence: Co-witness and perseverance effects. *Journal of Applied Psychology, 79*, 714-723.
- Morris, P.E. & Gruneberg, M.M. (1994). The major aspects of memory. In P.E. Morris & M.M. Gruneberg (Eds.), *Theoretical Aspects of Memory*, (pp. 29-49). New York: Routledge.
- Moscovitch, M. (1989). Confabulation and the frontal systems: Strategic versus associative retrieval in neuropsychological theories of memory. In H.L. Roediger III & F.I.M. Craik (Eds.), *Varieties of Memory and Consciousness: Essays in Honor of Endel Tulving*, (pp. 132-160). Hillsdale, NJ: Erlbaum.
- Moston, S. (1990). How children interpret and respond to questions: Situational sources of suggestibility in eyewitness interviews. *Social-Behaviour, 5*, 155-167.
- Murray, D.M. & Wells, G.L. (1982). Does knowledge that a crime was staged affect eyewitness accuracy? *Journal of Applied Social Psychology, 12*, 42-53.
- Nelson, T.O. (1984). A comparison of current measures of the accuracy of feeling of knowing predictions. *Psychological Bulletin, 95*, 109-133.
- Nelson, T.O. (1988). Predictive accuracy of feeling of knowing across different criterion tasks and across different subject populations and individuals. In M.M. Gruneberg, P.E.

- Morris, & R.N. Sykes (Eds.), *Practical Aspects of Memory: Current Research and Issues* (Vol. 1, pp. 190-196). Chichester, England: Wiley.
- Noon, E., & Hollin, C.R. (1987). Lay knowledge of eyewitness behaviour: A British survey. *Applied Cognitive Psychology, 1*, 143-153.
- Perfect, T. J. (2002) When does eyewitness confidence predict performance?. In T. J. Perfect & B. Schwartz (Eds.). *Applied Metacognition* (pp 95-120). Cambridge: Cambridge University Press.
- Perfect, T.J. (2004). The role of self-rated ability in confidence judgements for eyewitness memory and general knowledge. *Applied Cognitive Psychology, 18*, 157-168.
- Perfect, T.J. & Hollins, T.S. (1996). Predictive feeling of knowing judgements and postdictive confidence judgements in eyewitness memory and general knowledge. *Applied Cognitive Psychology, 10*, 371-382.
- Perfect, T.J., Watson, E.L. & Wagstaff, G.F. (1993). Accuracy of confidence ratings associated with general knowledge and eyewitness memory. *Journal of Applied Psychology, 78*, 144-147.
- Pezdek, K. & Greene, J. (1993). Testing eyewitness memory: Developing a measure that is more resistant to suggestibility. *Law and Human Behavior, 17*, 361-369.
- Pike, G.E., Towell, N.A. & Kemp, R.I. (1995). Eyewitness testimony: Confidence of accuracy and accuracy of confidence. Paper presented to the Fifth Annual Conference of Psychology and Law (August), Budapest, Hungary.
- Raaijmakers, J.G.W. & Shiffrin, R.M. (1992). Models for recall and recognition. *Annual Review of Psychology, 43*, 205-234.

- Read, J. D., Lindsay, D. S., & Nicholls, T. (1998). The relation between confidence and accuracy in eyewitness identification studies: Is the conclusion changing? In C. P. Thompson, D. J. Herman, J. D. Read, D. Bruce, D. G. Payne, & M. P. Toglia (Eds.), *Eyewitness memory: Theoretical and applied aspects* (pp. 107–130). Hillsdale, NJ: Erlbaum.
- Shaw, J.S. (1996). Increases in eyewitness confidence resulting from postevent questioning. *Journal of Experimental Psychology: Applied*, 2, 126-146.
- Shaw, J.S. & McClure, K.A. (1996). Repeated questioning can lead to elevated levels of eyewitness confidence. *Law and Human Behavior*, 6, 629-653.
- Smith, V.L., Kassin, S.M., & Ellsworth, P.C. (1989). Eyewitness accuracy and confidence: Within versus between-subjects correlations *Journal of Applied Psychology*, 74, 356-359.
- Sporer, S.L. (1983). Allgemeinwissen zur psychologie der zeugenaussage. In H.J. Kerner, H. Kury & K. Sessar (Eds), *Deutsche Forschungen zur Kriminalitaetsentstehung und Kriminalitaetskontrolle*, pp. 1191-1234. Germany, Koeln: Heymanns.
- Sporer, S.L., Penrod, S.D., Read, J.D., & Cutler, B.L. (1995). Choosing, confidence, and accuracy: A meta-analysis of the confidence-accuracy relation in eyewitness identification studies. *Psychological Bulletin*, 118, 315-327.
- Stebly, N.M. (1992). A meta-analytic review of the weapon-focus effect. *Law and Human Behavior*, 16, 413-424.
- Velicer, W.F., Prochaska, J.O., Rossi, J.S., & Snow, M.G. (1992). Assessing outcome in smoking cessation studies. *Psychological-Bulletin*, 111, 23-41.
- Wagstaff, G.F. (1981). *Hypnosis, Compliance and Belief*. Brighton, Sussex, UK: Harvester.

- Wagstaff, G.F. & Mercer, K., (1993). Hypnotic hypermnesia or experimental demands? In search of an explanation. *Contemporary-Hypnosis, 10*, 70-71.
- Weber, N., & Brewer, N. (2004). Confidence-Accuracy Calibration in Absolute and Relative Face Recognition Judgments. *Journal of Experimental Psychology: Applied, 10*, 156-172.
- Wells, G.L. (1978). Applied eyewitness testimony research: System variables and estimator variables. *Journal of Personality and Social Psychology, 36*, 1546-1557.
- Wells, G.L., Ferguson, T.J., Lindsay, R.C.L. (1981). The tractability of eyewitness confidence and its implications for triars of fact. *Journal of Applied Psychology, 66*, 688-696.
- Wells, G.L. & Leippe, M.R. (1981). How do triers of fact infer the accuracy of eyewitness identifications? Memory for peripheral detail can be misleading. *Journal of Applied Psychology, 66*, 682-687.
- Wells, G.L., & Lindsay, R.C.L. (1985). Methodological notes on the accuracy-confidence relation in eyewitness identifications. *Journal of Applied Psychology, 70*, 413-419.
- Wells, G.L., Lindsay, R.C.L., & Ferguson, T.J. (1979). Accuracy, confidence, and juror perceptions in eyewitness identification. *Journal of Applied Psychology, 64*, 440-448.
- Wells, G.L., & Murray, M. (1984). Eyewitness confidence. In Wells G.L., Loftus E.F. (Eds), *Eyewitness Testimony: Psychological Perspectives*, pp. 155-170 New York: Cambridge University Press.
- Yarmey, A.D., & Jones, H.P.T. (1983). Is the psychology of eyewitness identification a matter of common sense?. In S. Lloyd-Bostock & B.R. Clifford (Eds), *Evaluating Eyewitness Evidence*, pp. 13-40 Chichester: John Wiley & Sons.

Yuille, J.C. (1986). Meaningful research in the police context. In J.C. Yuille (Ed.), *Police selection and training: The role of psychology*, (pp. 225-243). Dordrecht, Germany: Martinus Nijhoff.

Yuille, J.C., & Cutshall, J.L. (1986). A case study of eyewitness memory of a crime. *Journal of Applied Psychology*, 71, 291-301.

Table 1.

Confidence-accuracy correlations for each interview condition and combined.

| Correlations                                                                                            | Combined        |                        |                               |                 |
|---------------------------------------------------------------------------------------------------------|-----------------|------------------------|-------------------------------|-----------------|
|                                                                                                         | self-rate       | Self rate<br>interview | Interviewer<br>rate condition | combined        |
| within-participants<br>correlation (gamma)                                                              | .80***<br>(.18) | .85***<br>(.20)        | .84***<br>(.13)               | .83***<br>(.17) |
| each participants' average<br>confidence correlated with<br>his or her average accuracy<br>( <i>r</i> ) | .50*            | .31                    | .39                           | .46*            |

*Note.* Standard deviations are in brackets. \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$ .

Table 2.

Average confidence expressed in correct and incorrect answers with respect to condition.

| confidence                         | condition                  |                        |                                |                        |
|------------------------------------|----------------------------|------------------------|--------------------------------|------------------------|
|                                    | self-rate<br>questionnaire | self-rate<br>interview | interviewer-<br>rate interview | conditions<br>combined |
| confidence in<br>correct answers   | 7.10<br>(1.20)             | 8.16<br>(1.25)         | 6.87<br>(0.97)                 | 7.38<br>(1.26)         |
| confidence in<br>incorrect answers | 2.79<br>(0.86)             | 3.32<br>(1.73)         | 2.89<br>(0.83)                 | 3.00<br>(0.83)         |
| confidence<br>overall              | 4.95<br>(0.96)             | 6.19<br>(1.19)         | 4.88<br>(0.81)                 | 5.34<br>(1.15)         |

*Note.* Standard deviations are in brackets.

### Acknowledgments

This research was supported by an ARC Grant to the author and the ARC Centre of Excellence in Policing and Security. I would also like to gratefully acknowledge the contribution of Professor Tim J. Perfect at the University of Portsmouth for his advice on an earlier version of the manuscript and the statistical analysis.