The influence of lawyers' questions on witness accuracy, confidence and reaction times and on mock-jurors' interpretation of witness accuracy.

Mark R Kebbell

ARC Centre of Excellence in Policing and Security, Griffith University

Laura Evans

University of Liverpool

Shane D. Johnson

University College London

Abstract

Two studies demonstrate the influence of lawyers' complex questions on mock-witness accuracy, confidence and reaction times and on the interpretation of witness accuracy by mock-jurors. In study one, 32 mock-witnesses were shown a short film and then questioned either with lawyers' complex questions or simple alternatives. In study two, 20 mock-jurors viewed video footage of the mock-witnesses assigned to each of the two previous conditions and were asked to rate their confidence in the witnesses' answers. The findings of the two studies indicated that lawyers' use of confusing questions reduce not only accuracy but also speed of response and both witnesses' and jurors' ability to determine accuracy. The implication of these findings is straight-forward, lawyers should ask simple questions wherever possible.

The confidence that eyewitnesses express in information seems to heavily influence 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; Brewer and Burke, 2002). In addition, the time taken for an eyewitness to respond to a question may be another factor employed by witnesses themselves and jurors in deciding witness accuracy. For example, Dunning and Peretta (2002) relate accuracy to automatic and fast decisions and Shaw (1996) relates a quick response to an easily accessed memory that leads to greater confidence in the accuracy of that response. Robinson, Johnson and Herndon (1997) concur, suggesting that a faster response correlates with greater confidence and accuracy of that response.

However, the speed with which a memory is retrieved might not necessarily always correlate with the veracity of that memory trace (Shaw, McClure & Wilkens, 2001) because the way in which a question is asked may have an impact on the speed of retrieval. For example, Loftus and Zanni (1975) have shown that using the definite article generates more false recognitions. For example, using the question "Did you see the car? Instead of, "Did you see a car? This appears to be because of the implied associations of the language and thus interference with memory. Therefore, whilst confidence and reaction time may well predict accuracy if simple questions are asked if the questions are phrased in a complex way this relationship may become weaker or non-existent.

One study that has illustrated the difference language can make to recall memory in a courtroom is that of Kebbell and Johnson (2000). In this study participants viewed a videotaped film and were individually questioned about the incident. Half the participants were asked questions using six categories of confusing questions often asked by lawyers in court (negatives, double negatives, leading, multiple questions, complex syntax, and complex vocabulary) whilst the remaining half were asked for the same information using simply phrased equivalents. Confusing questions reduced participant-witnesses' accuracy and suppressed confidence–accuracy relationships compared with the condition where simplified alternatives were asked.

In light of this literature we were interested to determine what influence the complex questions that lawyers ask might have on the speed with which mockwitnesses respond to questions and their confidence in the answers they give. This prompted Study One. Later on, in Study Two, we investigated how lawyers' questions impact on mock-jurors' perceptions of witness accuracy because ultimately jurors base decisions on their *perception* of witness accuracy.

STUDY 1

METHOD

Participants

Thirty-two participants (30 female, 2 male) with a mean age of 19.63 (SD=3.71, range = 18-35) took part in the study. These were first year undergraduate psychology and combined honours students who took part to fulfil part of their course requirements. Each participant was randomly assigned to one of two conditions. Design

A between-subjects design was used with the independent variable being the style of question asked. Participants were asked either all complex or all simple questions. The dependent variables were participants' accuracy on the recall test, their

Jurors' perceptions 5

confidence ratings in the veracity of their responses and their response times when answering the questions.

Materials

The stimulus video used by Kebbell and Johnson (2000) was also employed here. This lasted approximately 6 minutes and showed a man following and subsequently attacking a woman. Two different sets of questions were selected to test participant's memory for the content of the film. The first consisted of 'complex' questions which were similar in form to those used by lawyers in courtroom situations, whereas the second were simplified equivalents. The questions were taken from Kebbell and Johnson (2000). In that study six different types of question forms were used. Four of these different types were selected for use in this study, are summarised below:

Negatives

Complex: "Did the attacker not grab the woman around the waist?"

Simplified: "Did the attacker grab the woman around the waist?"

Double Negatives

Complex: "Is it not true that the woman did not go into the house?"

Simplified: "Is it true that the woman went into the house?

Leading

Complex: "It is correct to say that the woman ran across the baseball ground,

isn't it?"

Simplified: "Did the woman run across a baseball ground?"

Multipart

In the complex condition an example would be two questions asked at the same time, "Was the brother called Billy? Was he wearing gloves?" In the simplified condition the questions were asked separately, "Was the brother called Billy?" then there would be a pause for the participant to answer. Next the second part of the question would be asked, "Was he wearing gloves?"

For each of the four different question formats (negative, double negative, leading, multipart) there were a total of eight questions. Counting each of the multipart questions as two separate questions, this means that there were 40 questions in each condition (complex or simple). Questions were asked in the same order across the two conditions.

Procedure

All participants were tested individually and were shown the stimulus video. After watching the film they completed a phonemic fluency filler task. Participants were randomly assigned to one of the two experimental conditions (complex/simple) and questioned by one of two experimenters. Which of the two experimenters tested each participant was counterbalanced across participants and conditions. To generate materials for the second study, and to allow participant's response times to be accurately measured, each interview was recorded using a video camera.

Every question was read aloud by the experimenter and participants were required to respond yes or no in a clearly audible voice. Additionally, they were asked to indicate their confidence rating for each answer using a 10-point scale, ranging from 1 (pure guess) to 10 (absolutely sure). This confidence scale was not visible in the video recording and participants were asked not to say their rating aloud so mock-jurors in Study 2 would be unaware of their confidence judgments. No time limit was applied and hence participants could take as long as they liked to answer the questions, or to record their confidence ratings.

RESULTS

An outlier was identified and excluded from all subsequent analyses. Participant's mean accuracy, confidence ratings and reaction times are shown as Table One. Average accuracy rates were computed by dividing the number of correct responses by the total number of questions, and mean confidence ratings in a similar way. To calculate response times, using the video footage, the time elapsed between the experimenter asking each question and the participant responding was recorded. This could have been the required 'yes' or 'no' answer, or in some cases the first word of their explanation for their answer. If a participant asked for a question to be repeated, the response time calculation started after the experimenter repeated the question. In the complex condition, if the question was answered inaccurately, for example a 'yes' response where 'no' would have been correct but participants backed up their answer accurately with other details, the answers was rated as correct.

Table One. Average accuracy, confidence and response times for all participants

	Simple (N=16)		Complex (Complex (N=15)	
	Mean	SD	Mean	SD	
Accuracy	0.84	0.08	0.43	0.07	
Confidence	7.30	0.85	6.74	0.88	
Response Time	2.11	0.65	3.20	1.31	

Independent samples *t*-tests confirmed that questions were answered more accurately in the simplified condition, t(29) = 15.43, p<.001, and that response times were also faster for this condition, t(29) = 2.91, p<.01. However, there was no significant difference in the confidence ratings expressed between the two conditions, t(29) = 1.74, p > .05).

To examine the relationship between participant's confidence and accuracy and response times in more detail, a series of within-subjects Goodman-Kruskal gamma correlations were computed. This particular test was used because, unlike other measures of association, it is unaffected by overall memory performance and therefore is not influenced by the fact that different participants will have answered different questions correctly and incorrectly (for a discussion see Nelson, 1984). Therefore three correlation coefficients were computed for each participant. For each condition the average correlation coefficient was calculated and compared against

zero using a one-sample t-test. The average correlations and standard deviations are shown in Table Two.

Table Two Mean within-subject gamma correlation coefficients and standard deviations.

		C-A	C-RT	A-RT
Simple	Mean	.60*	56*	49*
	SD	.20	.13	.19
Complex	Mean	.19	40*	19
	SD	.38	.16	.13

^{*} *p* < .05

For the simplified condition the confidence-accuracy (CA), confidencereaction time (C-RT), and accuracy-reaction time (A-RT) correlations were statistically significant from zero. However, for the complex condition, whilst the C-RT correlation was significant, the C-A and A-RT associations were not. To examine the differences between the two conditions, the mean within-subjects gamma correlations were contrasted using a series of t-tests. The results indicated that relative to the complex condition the correlations were higher for the simplified condition for C-A t(29)=3.81, p<.01, C-RT t(29)=3.10, p<.01, and A-RT t(29)=7.00, p<.001.

DISCUSSION OF STUDY 1

To summarise the results of Study 1, in line with previous research (Kebbell and Johnson, 2000) relative to those in the simple condition, participants in the

complex condition were less accurate, and yet expressed similar levels of confidence in their answers. In addition, participants assigned to the complex condition tended to take longer to answer the questions. Study 1 also shows that relative to the complex condition, the correlations were higher for the simplified condition for C-A, C-RT, and A-RT indicating that the negative impact of confusing questions extends to reaction time. If question form distorts the relation between reaction time and witness accuracy, this may have an important potential impact on jurors if they use reaction time as an indicator of accuracy. For this reason we conducted Study 2 to look at what factors mock-jurors might take into consideration.

STUDY 2

METHOD

Participants

The recruitment procedure employed in Study 1 was mirrored for the second study. Twenty participants were recruited (17 Female, 3 Male), with a mean age of 18.85 years and range of seven years.

Design

This study involved a within-subjects repeated measures design. The within-subjects factor was the type of questions asked and mock-witnesses' responses (simple/complex). The dependant variables were the mock-juror's confidence in the accuracy of the mock-witness viewed, the actual accuracy and confidence of the mock witness and the time taken for the mock witness to respond (the latter three measurements obtained through the first study).

Materials

The recordings produced during Study 1 were edited for the second study so that each mock-juror watched one witness who had responded to the complex

questions and then another who had responded to the simply phrased questions. Of the 32 participant recordings previously obtained in Study 1 (16 complex, 16 simple), 10 were randomly selected from each condition. Each selected recording from one condition was then randomly paired with one from the other, producing 10 randomly generated simple-complex pairings. For each pairing, two videos were produced. For the first, the witness assigned to the simple condition appeared in the first half of the video, and the witness assigned to the complex condition appeared in the second half. For the purposes of counterbalancing the materials, for the second video of each pairing, the order with which the witnesses from the two conditions appeared was reversed. A series of statistical comparisons were made between the abbreviated dataset and the complete data set to see if there was any bias in the sampling or mockwitnesses. No bias was found and so for brevity we do not report these analyses here.

Mock-juror response forms were created to obtain mock-jurors' confidence in how accurate they believed the witness to be. Each form was identical to those used by mock-witnesses in Study1 to obtain comparable confidence levels although arguably this is more a measure of believability. There was also a second mock-juror response form, asking participants to write, in general, reasons why their particular confidence ratings were given.

Procedure

Each participant saw witnesses from Study 1 answering the simple or complex questions. After each excerpt, a pre-recorded pause occurred allowing time for the participant juror to note their confidence rating on the scales provided (again from 1 to 10). This pause lasted approximately 10 seconds and was extended further by the

experimenter if necessary. At the end of the film, participants were asked to report what factors they had considered when assessing witness accuracy.

RESULTS

On average, mock-jurors reported having higher confidence in witness accuracy for those in the simple condition (M=6.06, SD=1.01) than complex (M=5.10, SD=0.99), t(18)=3.39, p<.01. Gamma correlations were calculated to explore the associations between variables within-subjects. Witness and juror confidence were first correlated with witness accuracy (see Table 3) in order to see if witnesses and jurors could correctly distinguish between accurate and inaccurate witness responses. A 2 X 2 ANOVA (simple/complex X witness rating/juror rating) was conducted on the correlations with the witness/juror ratings as a within-subjects factor. The results showed that there was a significantly higher confidence-accuracy relation for both witnesses and jurors in the simplified condition than in the complex condition, F(1,36)=7.20, p<.01, η ²=.17. There was no significant difference between the confidence-accuracy ratings of jurors or witnesses, F(1,36)=2.51, p>.05, η ²=.01 and the interaction was not significant, F(1,36)=3.17, p>.05, η ²=.08.

Finally, for each condition, mock-juror confidence was correlated with witness reaction time (C-RT) and the differences compared with a t-test. The mean gamma correlations, displayed in Table 3, were not significantly different from one-another, t(38)=1.42, p>.05. All gamma correlations were also tested for significance against zero with t-tests, all were significant.

Table 3.

Mean within-subject gamma correlation coefficients and standard deviations

		Witness C-A	Juror C-A	Juror C-RT
Simple	Mean	.61*	.41*	61*
	SD	.34	.41	.20
Complex	Mean	.23*	.24*	55*
	SD	.32	.39	.20

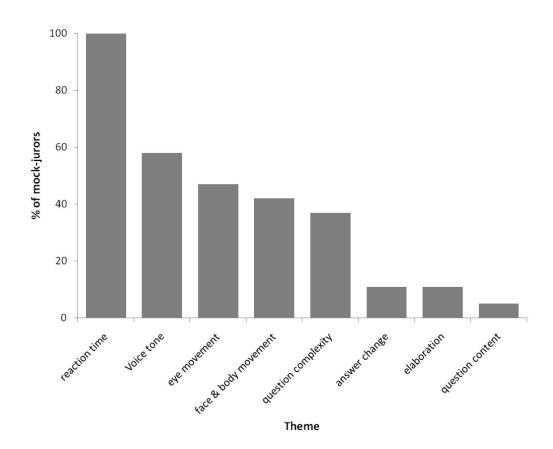
^{*} p < .05

Content Analysis

Analysis of jurors' reasons for their confidence ratings revealed the eight themes displayed in figure one. The most frequently mentioned reason for a confidence judgment mentioned by all respondents was response time, for example, 'my decisions were based on how long the subjects took to respond to the question. If they paused for a long time, I assumed they were inaccurate'. The next most frequently mentioned reason was tone of voice, for example, 'The tone of voice – if they sounded more confident, perhaps louder – I was more confident of their accuracy. Eye movement was also frequently mentioned, for example, 'I also thought they were guessing when they seemed to be looking round the room, instead of focusing on a fixed point'. Less frequently mentioned factors were face and body movement, question complexity, whether the witness changed their answers, if the witness elaborated on their answers, and the perceived difficulty of the questions,

Figure 1.

Bar Chart to show the percentage of jurors to mention each reason for their confidence judgment.



DISCUSSION OF STUDY TWO

The results of study two were clear. There is a relationship between mockjuror confidence in witness responses and their accuracy, but the strength of that relationship is a function of the way that witnesses are questioned; complex question forms appear to suppress the confidence-accuracy relation. One possible explanation for this is that while the type of question asked influences witness accuracy, it does not affect the attention that mock-jurors pay to the speed with which witnesses answer questions. Irrespective of the complexity of the question asked, the faster a mockwitness responds, the more likely a mock-juror is to perceive that response as correct.

GENERAL DISCUSSION

To reiterate, the aim of the studies was to investigate the influence of the complex questions that lawyers ask on the accuracy and speed with which mockwitnesses respond to questions, their confidence and how these factors influence mock-jurors' perceptions of witness accuracy. The results confirm the adverse impact of confusing questions. The accuracy of witnesses was significantly hampered by being questioned with confusing questions with the accuracy rate for confusing questions nearly half that for the simple equivalents. Whilst this decrease in accuracy would be a problem in itself, the potential for a miscarriage of justice based on inaccurate evidence would be reduced if witnesses, and in turn, jurors were not confident about inaccurate answers. To satisfy this condition, witnesses and jurors alike would have to be able to distinguish between accurate and inaccurate responses.

The results of the two studies suggest that both mock-witnesses and mockjurors can perform quite well in determining accuracy if the questions are simple. However, mock witnesses assigned to the complex condition were as confident in the accuracy of their answers as those assigned to the simple condition despite the fact that they were much less accurate. Moreover, whilst there was a moderate to strong correlation between witness accuracy and the confidence they expressed in their answers for those assigned to the simple condition, this was not the case for those asked confusing questions. Similarly, whilst the confidence that mockwitnesses expressed was related to the speed of their responses, it was only for those asked simply phrased questions that the relation between reaction time and accuracy was reliable. Simply put, when witnesses are asked simple questions a swift response tends to equate to an accurate response. However, with complex questions, a swift response does not equate to an accurate one, although it would appear that mock-witnesses may believe that it does.

Considering the mock-jurors, they were unaware of the confidence ratings provided by the mock-witnesses and hence remained oblivious to the fact that, as a consequence of the complex questions inaccurate witnesses remained overall fairly confident in the witnesses' responses. When asked what they had considered when evaluating witness accuracy, all nineteen jurors said they used witness reaction time in making their decisions (along with lower usage of various other cues including fidgeting, eye movement). This concurs with the results of the correlation analyses which showed that the quicker a witness responded to a question, the more confident that juror would be in the answer considered (overall across conditions, and individually in both conditions). Although the mock-jurors were correct to have more confidence in the accuracy of a fast responding witness when that witness was asked simple questions, for those asked complex questions, the speed with which a witness responded did not predict how accurate they were. Thus, using reaction time to estimate witness accuracy may be flawed, depending on how the witness is questioned. Therefore, the findings of this study are that if lawyers use the complex questions that we used here accuracy rates will decrease for witnesses, they are

likely to be overconfident, and jurors' ability to discriminate between accurate and inaccurate information will decrease.

Some legal scholars have the assumption with regards to cross-examination is that 'it is difficult to mislead a witness who is honest' (Wheatcroft, 2002, p.20). The findings from this study have shown, however, that this may not necessarily be the case. Nor may it be the case that jurors are aware of this fact. Robinson, Johnson and Herndon (1997); Dunning and Peretta (2002); and Shaw (1996) are all correct in asserting that reaction time predicts accuracy, and those arguing for the existence of a positive confidence-accuracy correlation (see Ebbesen & Konecni, 1975; Lindsay, Read & Sharma, 1998; Shaw, McClure & Wilkens, 2001; Kebbell, Wagstaff, & Covey, 1996) may similarly be right in their stance. What has now become evident, however, is that such assertions may only hold true under certain conditions — conditions that most likely do not pertain in court. As Kebbell and Johnson reported, confusing questions lead to 'adverse' effects on witness accuracy, such effects now evidently filtering through to juror perception.

Of course an issue here is whether lawyers are using complex questions as a deliberate strategy. In his popular guide to advocacy, Evans (1995) identifies a critical role of cross-examination as discrediting the evidence, he points out, 'it is not a procedure which is aiming to find out the truth' (p. 150). For example, the use of multiple questions may be used deliberately to unfairly discredit the witness because of the confusion they create. It is difficult to see how justice is served by asking witnesses complex or confusing questions. Many of these problems of language can be minimised through appropriate intervention by the judge who is obliged not only to have regard to the need to ensure a fair trial for the defendant, but also to the

reasonable interests of other parties to the court process (O'Kelly, et al., 2003). This is particularly true of vulnerable witnesses who are obliged to relive the ordeal to which they have allegedly been subjected (see Carson, 1995; Davies & Noon, 1991; Sanders, et al., 1997; Home Office, 2002). It is the Judge's duty to do everything possible to minimise the trauma suffered by other participants (Murphy, 1997). The Court of Appeal has also sanctioned the stopping of cross-examination which is repetitive and in which the witness becomes extremely distressed (R v. Brown 1998). The judge has a great deal of power. The following examples show how judges can intervene to ensure the 'best' evidence is elicited from witnesses, in these instances involving people with learning disabilities.

Lawyer All right, but my question is a slightly different one. Did you feel upset when you arrived at the discotheque? Well let me put this to you. You appeared your normal, happy self when you got there and in no way distressed because nothing had happened.

Judge You must separate these questions. You cannot have a multiple-question (O'Kelly, et al., 2003; p.233).

The implication of this is that judges should be advised of the issues concerning confusing questions we have outlined here, to ensure simple questions are asked in language the witness understands.

There are a number of improvements that could be made in a replication or extension of this study. As this was a laboratory based study, participants were aware that their responses had no 'real' implications, and so may have been less motivated to perform as well as they would in real life (see for example, Clifford & Lloyd-Bostock, 1983). However, in using 'real' participants, one perhaps compromises the tight

controls (such as equal viewing conditions across participants) permissible under such conditions as the present, which allow the impact of an independent variable to be isolated in the absence of other confounds. Future work, however, could perhaps replicate actual cases to a greater degree by varying juror conditions, perhaps investigating juror ability one week after having seen witnesses on screen since not all jury decisions are made immediately after witnesses have appeared in court. Other possibilities include the use of repeated questions that are also prevalent in cross-examination and may influence the perception of accuracy (O'Kelly, et al., 2003).

To conclude, the findings of the two studies indicate that lawyers' use of confusing questions reduce not only witness accuracy and speed of response but also the ability of both witnesses' and jurors' to distinguish between accurate and inaccurate responses. The implication of these findings is that in order for justice to be achieved lawyers should ask simple questions.

References

- Brewer, N., and Burke, A. (2002). Effects of testimonial inconsistencies and eyewitness confidence on mock-juror judgements. *Law and Human Behavior*, 26, 353-364.
- 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.
- Carson, D. (1995). Regulating the examination of children. *Expert Evidence*, 4, 2–9.
- Clifford, B. R., & and Lloyd-Bostock, S., (1983). Evaluating Witness Evidence:

 Recent Psychological Research and New Perspectives. New York: John Wiley
 & Sons Ltd.
- Cutler, B.L., Penrod, S.D., & Dexter, H.R. (1990). Juror sensitivity to eyewitness identification evidence. *Law and Human Behavior*, *14*, 185-191.
- Davies, G.M. and Noon, E. (1991). An Evaluation of the Live Link for Child Witnesses. London: Home Office.
- Dunning, D., & Peretta, W. (2002). Automaticity and eyewitness accuracy: A 10 12 second rule for distinguishing accurate from inaccurate positive identifications. *Journal of Applied Psychology*, 87, 951 962.
- Ebbesen, E. B., & Konecni, V. J. (1975). The decision making and information integration in the courts- The setting of bail. *Journal of Personality and Social Psychology*, 32, 805-821.
- Evans, K. (1995). *Advocacy in Court: A beginner's guide* (2nd edn). London: Blackstone.

- Home Office (2002). Achieving Best Evidence. Guidance for Vulnerable or Intimidated Witnesses, including Children. Home Office: London.
- Kebbell, M. R., & Johnson, S. D. (2000). Lawyers' questioning: the effect of confusing questions on witness confidence and accuracy. *Law and Human Behaviour*, 24, 629 – 641.
- 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.
- Loftus, E.F., & Zanni, G. (1975). Eyewitness testimony: The influence of the wording of a question. *Bulletin of the Psychonomic Society*, *5*, 86-88
- Murphy, P. (ed.) (1997) *Blackstone's Criminal Practice*, 7th Edn. London: Blackstone.
- Nelson, T.O. (1984). A comparison of current measures of the accuracy of feeling of knowing predictions. *Psychological Bulletin*, *95*, 109-133.
- O'Kelly, C.M.E., Kebbell, M.R., Hatton, C. and Johnson, S.D. (2003). When do Judges intervene in cases involving people with learning disabilities? *Legal* and *Criminological Psychology*, 8, 229–40.
- Robinson, M.D., Johnson, J. T., & Herndon, F. (1997). Reaction Time and

 Assessments of Cognitive Effort as Predictors of Eyewitness Memory

 Accuracy and Confidence. *Journal of Applied Psychology*, 82, 416 425.
- Sanders, A., Creaton, J., Bird, S. and Weber, L. (1997). Victims with Learning

 Disabilities: Negotiating the Criminal Justice System. Oxford: Centre for

 Criminological Research, University of Oxford.
- Scott, J. (1978). Individual and Situational Factors in Eyewitness Testimony.

- *Journal of Applied Psychology.* 63, 352 359
- 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., & Wilkens, C. E. (2001). Recognition Instructions and Recognition Practice Can Alter the Confidence-Response Time Relationship.

 *Journal of Applied Psychology, 86, 93 103.
- Wheatcroft, J. M. (2002). Some influences of courtroom questioning styles on eyewitness confidence and accuracy. Unpublished Thesis, University of Liverpool, England.