Helping jurors to evaluate eyewitness identifications: the role of expert evidence and judicial instruction

Author:
Martire, Kristy Anne

Publication Date:
2008

DOI:
https://doi.org/10.26190/unsworks/17850

HELPING JURORS TO EVALUATE EYEWITNESS IDENTIFICATIONS: THE ROLE OF EXPERT EVIDENCE AND JUDICIAL INSTRUCTION

Kristy A. Martire

A thesis submitted in fulfilment
of the requirements for the degree of
Doctor of Philosophy

School of Psychology University of New South Wales

February 10, 2008

Abstract

Psychologists, legal practitioners and scholars share the knowledge that honest eyewitnesses can err in their attempts to identify the perpetrator of a crime. This thesis reports an experimental investigation of the extent to which expert evidence and judicial instruction can improve juror ability to discriminate between accurate and inaccurate identifications. Special attention is also paid to the logic of inferences which have been made by psychologists regarding the efficacy of expert evidence, and compares methodologies adopting direct measures of participant Sensitivity to Eyewitness Accuracy (SEA) with those that can only indirectly assess this construct.

Study 1 surveys the knowledge and opinions of legal professionals regarding eyewitness identification issues (n = 35), showing that respondents expressed doubts that judicial instructions would exert an effect equivalent to that of eyewitness expert evidence. Accordingly, Experiments 1 to 4 (Experiment 1, n = 104; Experiment 2, n = 238; Experiment 3, n = 228; Experiment 4, n = 297) were conducted to directly assess the relative impacts of judicial instruction and expert evidence on participant-juror SEA.

The methodology utilised in these investigations incorporated the testimony of real eyewitnesses to a staged crime scenario in order to assess the impact of instruction on juror ability to discriminate between known accurate and known inaccurate eyewitnesses. Overall, little evidence was found to support the notion that expert evidence is more effective than judicial instruction, as no significant association was identified between instruction type and SEA. This result was found to hold irrespective of the objective quality of the expert’s testimony (accurate or erroneous).

In light of the results from Experiments 1 to 4, Experiment 5 was designed to investigate why the experts were not able to improve the discrimination accuracy of the jurors. This study focused on the extent to which participants of varying levels of expertise could correctly classify eyewitness accuracy. The results of Experiment 5 (n = 145) suggest that experts were no better able to discriminate between accurate and inaccurate eyewitnesses than novice laypeople. Overall, the evidence reported in this thesis raises serious questions regarding the utility of eyewitness expertise in the completion of eyewitness discrimination tasks.

Table of Contents

ABSTRACT II

ACKNOWLEDGEMENTS VIII

THESIS OVERVIEW X

SECTION 1: INTRODUCTION 1

CHAPTER 1 1

FORENSIC CONTEXT 1

CHAPTER 2 11

PUBLIC DEFENDERS' SURVEY 11

METHOD 13

RESULTS 16

DISCUSSION 21

SECTION 2: CRITICAL REVIEW 28

CHAPTER 3 28

EYEWITNESS EXPERT EFFECTS - LITERATURE REVIEW 28

A HIERARCHY FOR THE ASSESSMENT OF EXPERT EFFECTS 31

INDIRECT MEASURES 32

DIRECT MEASURES 60

SUMMARY 67

CHAPTER 4 79

JUDICIAL INSTRUCTION EFFECTS – LITERATURE REVIEW 79

SUMMARY 89

CHAPTER 5 93

INFERENCES EVALUATED 93

AIMS 99

SECTION 3: EXPERIMENTAL STUDIES 100

CHAPTER 6 100

EXPERIMENT 1 - JURY STUDY 100

METHOD 102

RESULTS 107

DISCUSSION 123

CHAPTER 7 134

EXPERIMENT 2 – IMPACT OF ACCURATE EXPERT EVIDENCE ON JUROR SENSITIVITY TO EYEWITNESS ACCURACY 135
METHOD 135
RESULTS 138
DISCUSSION 146
CHAPTER 8 **150**
EXPERIMENT 3 – THE IMPACT OF ERRONEOUS EXPERT EVIDENCE ON JUROR SENSITIVITY TO EYEWITNESS ACCURACY 150
METHOD 151
RESULTS 152
DISCUSSION 160
GENERAL DISCUSSION – EXPERIMENTS 2 AND 3 162
CHAPTER 9 **170**
EXPERIMENT 4 – BOUNDARY EFFECTS OF EXPERT EVIDENCE 170
METHOD 170
RESULTS 173
DISCUSSION 181
CHAPTER 10 **189**
EXPERIMENT 5 - EXPERTISE STUDY 189
METHOD 191
RESULTS - PART I 195
RESULTS PART II 199
DISCUSSION 209

SECTION 4: DISCUSSION **189**

CHAPTER 11 **217**
GENERAL DISCUSSION 217

REFERENCES **242**

APPENDICIES **256**

APPENDIX A **256**
APPENDIX B **256**
APPENDIX C **257**
APPENDIX D **260**
APPENDIX E **261**

APPENDIX F **261**

APPENDIX G **261**

APPENDIX H **261**

APPENDIX I **262**

APPENDIX J **279**

APPENDIX K **279**

APPENDIX L **280**

APPENDIX M **282**

List of Tables

Table 1.1 : Judicial Instruction. 4
Table 2.1: Original and Reversed Statements from the ETQ 14
Table 2.2: Original and Reworded Statements from the ETQ 15
Table 2.3: Distribution of Judgments for the 30 statements 17
Table 2.4: Comparison of Agreement Rates for Kassin et al. (2001) Experts and Legal Professionals 19
Table 3.1 : Summary of Observed Expert Effects 69
Table 4.1 : Summary of Observed Judicial Instruction Effects 91
Table 6.1 : Jurors' Mean (SE) Evaluations of Instructors 108
Table 6.2 : The Impact of Instruction and Belief type on Participant-juror Priorities 111
Table 6.3 : β (S.E.), Significance and Exp(β) for Predictor Variables in Control Condition 112
Table 6.4 : β (S.E.), Significance and Exp(β) for Predictor Variables in Expert Condition 113
Table 6.5 : β (S.E.), Significance and Exp(β) for Predictor Variables in Judicial Condition 113
Table 6.6 : Judgement Type as Proportion Within Instruction Condition, Observed *d'* and *C* Values. 115
Table 6.7 : Topics of Deliberation Within Juries and Across Instruction Conditions. 116
Table 6.8 : Juries' Mean (SE) Evaluations of Instructors 117
Table 6.9 : Frequencies of Observed Jury Outcomes 119
Table 6.10 : The Impact of Instruction and Verdict type on Jury Priorities 122
Table 6.11 : Correct Jury Verdicts by Instruction Condition 123
Table 7.1 : Jurors' Mean (SE) Evaluations of Instructors 139
Table 7.2 : β (S.E.), Significance and Exp(β) for Predictor Variables in Control Condition 143
Table 7.3 : β (S.E.), Significance and Exp(β) for Predictor Variables in Expert Condition 144
Table 7.4 : β (S.E.), Significance and Exp(β) for Predictor Variables in Judicial Condition 145
Table 7.5 : Judgement Type as Proportion Within Instruction Condition, Observed *d'* and *C* Values. 146

Table 7.6 : Jurors' Mean (SE) Ratings of Influence 148
Table 8.1 : Jurors' Mean (SE) Evaluations of Instructors 152
Table 8.2 : Jurors'Mean (SE) Ratings of Influence of Eyewitness Manner and Witnessing Conditions by Juror Decision and Instruction Condition 156
Table 8.3 : β (S.E.), Significance and Exp(β) for Predictor Variables in Control Condition 157
Table 8.4 : β (S.E.), Significance and Exp(β) for Predictor Variables in Expert Condition .. 158
Table 8.5 : β (S.E.), Significance and Exp(β) for Predictor Variables in Judicial Condition 159
Table 8.6 : Judgement Type as Proportion Within Instruction Condition, Observed *d'* and *C* Values. 160
Table 9.1 : Jurors' Mean (SE) Evaluations of Instructors 173
Table 9.2 : Jurors' Mean (SE) Ratings of Influence of Eyewitness Manner and Witnessing Conditions. 178
Table 9.3 : β (S.E.), Significance and Exp(β) for Predictor Variables in Accurate Expert Condition 179
Table 9.4 : β (S.E.), Significance and Exp(β) for Predictor Variables in Erroneous Expert Condition 180
Table 9.5 : β (S.E.), Significance and Exp(β) for Predictor Variables in Judicial Regression 180
Table 9.6 : Judgement Type as Percent Within Instruction Condition, Observed *d'* and *C* Values. 181
Table 10.1 : Statements for Evaluation (based on Kassin et al., 2001) 194
Table 10.2 : Percent of Participants Endorsing Each Statement 196
Table 10.3 : Reported Primary Areas of Research. 199
Table 10.4 : β (S.E.), Significance and Exp(β) for Predictor Variables in Community Group 205
Table 10.5 : β (S.E.), Significance and Exp(β) for Predictor Variables in Undergraduate Group 206
Table 10.6 : β (S.E.), Significance and Exp(β) for Predictor Variables in Postgraduate Group 206
Table 10.7 : β (S.E.), Significance and Exp(β) for Predictor Variables in Expert Group 207
Table 10.8 : Summary of Belief and Accuracy Rates for Each Judgement 208
Table 10.9 : Comparison Between "First Trial Only" and "All Trial" Analyses 216
Table 11.1 : Summary of the Effects of Judicial Instruction 222
Table 11.2 : Summary of Direct Comparisons between Expert Evidence and Judicial Instruction 226
Table 11.3 : Comparison Between Direct and Indirect Measures 230
Table I.1 : Factorial Structure of Witnessing Factors in Recorded Crime Videos 263

Table I.2 : Identification Type by Version of Crime 265
Table I.3 : Confidence-Accuracy Correlations for Total Sample, Choosers and Court-Choosers in Good and Poor Witnessing Conditions 266
Table I.4 : Lineup Member Selected by Lineup Type 267
Table I.5 : Identification Type by Version of Crime 270
Table I.6 : Confidence-Accuracy Correlations for Total Sample, Choosers and Court-Choosers in Good and Very Poor Witnessing Conditions 271
Table I.7 : Lineup Member Selected by Lineup Type 272

List of Figures

Figure 6.1 : Mean pre-deliberation confidence ratings by instruction type and eyewitness accuracy 109
Figure 6.2 : Mean pre-deliberation confidence ratings by instruction type and juror belief decision 110
Figure 6.3 : Mean post-deliberation confidence ratings by instruction type and eyewitness accuracy 120
Figure 6.4 : Mean confidence ratings by instruction type and unanimous verdict type. 121
Figure 7.1 : Mean confidence ratings by eyewitness accuracy. 140
Figure 7.2 : Mean confidence ratings by juror belief decision. 141
Figure 7.3 : Ratings of the influence of eyewitness confidence by belief decision 142
Figure 8.1 : Mean confidence ratings by eyewitness accuracy. 154
Figure 8.2 : Mean confidence ratings by juror belief decision. 154
Figure 8.3 : Ratings of the influence of eyewitness confidence by belief decision 155
Figure 9.1 : Mean confidence ratings by eyewitness accuracy. 175
Figure 9.2 : Mean confidence ratings by juror belief decision. 176
Figure 9.3 : Ratings of the influence of eyewitness confidence by belief decision 177
Figure 10.1 Mean confidence ratings by eyewitness accuracy. 201
Figure 10.2 Mean confidence ratings by juror belief decision. 201
Figure 10.3 Ratings of the influence of eyewitness confidence by belief decision. 202
Figure 10.4 Ratings of the influence of eyewitness manner by belief decision. 203
Figure 10.5 Ratings of the influence of witnessing conditions by belief decision. 204
Figure 10.6 Percent of participant accuracy scores across levels of expertise. 209
Figure M.1 Mean confidence ratings by juror belief decisions 283

Acknowledgements

To Richard, without your unwavering support I think I would have lost my faith in the ideals of scientific endeavour within the first six months of this project. You helped me when under-fire and you encouraged me to meet adversity head-on. More than that, you gave me the freedom to have a dissenting voice, and the confidence to believe that it was worth hearing. Thanks.

To Jill, Gary and Mehera from the Law School, thank you for your willingness to guide the way as I blundered into unfamiliar territory.

To Jacquelyn and Sue – you guys gave me my “n”; in return you got my soul for two years. I think I got the better end of that deal. Thank you.

To my *apostrophe wrangler’[e]s’*: Dad and Jodie – I can’t thank you enough for the time and attention you gave to editing my thesis, especially since you agreed to do so in full knowledge of my pre-existing punctuational impairments. You guys are tops!

To Alexa, Erin, Jade, Joanna, Ju, Justine, Kim, Maggie, Melinda, Phoebe, Shaina, Shirley and Tamara, who volunteered their time to help me make materials, interview witnesses and run participants, you helped make this all possible.

To Gut and all my family, thank you for always believing that I could do this, and for telling me so when I needed to hear it.

Finally, to Debs, Diane, Juju, David, Suzanne, Brooke and the Brown Brothers (and stockists) – I think Engrish.com said it best: *friendship make a happiness double times and divide sadness half and make us remove our unhappiness*. I just can’t make it any clearer to you than that.

There is one great difficulty with a good hypothesis. When it is completed and rounded, the corners smooth and the content cohesive and coherent, it is likely to become a thing in itself. A work of art. It is then like a finished sonnet or a painting completed. One hates to disturb it. Even if subsequent information should shoot a hole in it, one hates to tear it down because it once was beautiful and whole

(Steinbeck, 1958)

Thesis Overview

Psychologists, legal practitioners and scholars share the knowledge that honest eyewitnesses can err in their attempts to identify the perpetrator of a crime. This thesis reports an experimental investigation of the extent to which expert evidence and judicial instruction can improve juror ability to discriminate between accurate and inaccurate identifications.

Beginning with a brief introduction to the forensic context surrounding eyewitness testimony, expert evidence and judicial instruction, a survey of the knowledge and opinions of legal professionals in NSW is then reported. This survey provides evidence to suggest that legal professionals and eyewitness experts have similar concerns about the equivalency of eyewitness expert evidence and the relevant pattern judicial instruction. Accordingly, Chapters 3 and 4 provide a thorough review of the literature pertaining to the effects of eyewitness expert evidence and judicial instruction, followed in Chapter 5 by a critical analysis of the experimental methods used to investigate these issues.

Chapters 6 to 10 report a series of experimental investigations of the relative effects of eyewitness expert testimony and judicial instruction (Experiments 1 to 4), as well as the role of expertise in eyewitness discrimination performance (Experiment 5). These studies all adopt a "real eyewitness" design, and as such, allow the direct measurement of the effect of expert or judicial instruction on the ability of participant-jurors to discriminate accurate from inaccurate eyewitnesses. Experiments utilising this design are composed of two distinct stages. Stage 1. pertains to the collection of real eyewitness testimony and involves: a) the construction of a crime video; b) eyewitness identifications; and c) eyewitness interviews. In Stage 2. participant-jurors are presented with these eyewitness interviews and are asked to evaluate the accuracy of the identification made by the eyewitness. For ease of reading, only the second of these two stages is reported in the body of the thesis. Information relating to the collection of eyewitness testimony is reported in relevant Appendices. The final chapter, Chapter 11, presents a general discussion of experimental results with regard to the aims of this thesis, and summarises the significance and innovation of the work together with its limitations. It also makes suggestions for future directions for research in this field.

SECTION 1: INTRODUCTION

Chapter 1

Forensic Context

The "Eyewitness Problem"

According to Volokh (1997, p. 146), the maxim that it is "[b]etter that ten guilty persons escape than one innocent suffer" (Blackstone, 1979, p. 358) expresses an ideal that is integral to the notions of fairness in Western criminal law. Specifically, over the past decades, a fear of wrongful convictions and the resultant tragic consequences has led forensic psychologists to focus on the contribution of mistaken eyewitness identifications to these errors. This research, although originally ridiculed and dismissed (Wigmore, 1909) has gained momentum for three key reasons: 1) eyewitnesses have been shown to be unreliable; 2) eyewitness evidence is compelling in the trial context; and 3) eyewitness errors have been demonstrably linked to erroneous convictions. Accordingly, research psychologists are as motivated today to seek a resolution to the "eyewitness problem" as were their colleagues over 100 years ago (Munsterberg, 1908).

Psychologists, legal practitioners and scholars share the knowledge that honest eyewitnesses can err in their attempts to identify the perpetrator of a crime. In the case of psychologists, this knowledge has been acquired through extensive empirical research into human performance on applied memory tasks. The results of these investigations have led researchers to conclude that eyewitnesses frequently offer inaccurate recollections, believing them to be true (see Wells & Olson, 2003 for a detailed review). As a consequence, eyewitness identifications are widely considered to be one of the least reliable forms of evidence admitted in the courtroom (Devlin, 1976; Munsterberg, 1908; Watson, 1924 cited in Woller, 2004). This knowledge on its own, however, is not sufficient to conclude that eyewitness testimony represents a threat to the liberty of the innocent accused. That conclusion can only be reached if it can be demonstrated that the testimony of these unreliable eyewitnesses is as compelling to fact-finders as other more reliable forms of testimony. Unfortunately, the research conducted on this issue does indicate that jurors find eyewitness testimony to be a particularly influential form of evidence (e.g., Leippe, 1995; Loftus, 1979; Saunders, Vidmar, & Hewitt, 1983; Wells, Lindsay, & Tousignant), irrespective of the probable accuracy of the identification made (Wells & Olson, 2003). Thus, the fear that unreliable identifications are made by eyewitnesses, and subsequently believed by jurors, is supported by the available empirical evidence.

Moreover, analyses of known wrongful convictions further suggest that this tendency to be influenced by eyewitness testimony, irrespective of the quality of that evidence, can lead to the conviction of innocent defendants. Although estimates of both the number of erroneous convictions and the role of eyewitnesses in those convictions vary widely in their magnitude and methods of calculation (Gross & O'Brien, 2007), many researchers have concluded that mistaken identifications, (inappropriately evaluated by jurors), have contributed to a non-trivial number of erroneous convictions (Huff, 1987; Leippe, 1995; Penrod & Cutler, 1999; Scheck, Neufeld, & Dwyer, 2001). In fact, the most compelling evidence for the role of eyewitness testimony in mistaken convictions is emerging day-by-day through the efforts of the New York–based organisation, the Innocence Project. This group provides a means for inmates to gain access to DNA analyses which can potentially produce exculpatory evidence in their cases. As of February 2008, the Innocence Project had secured 212 exonerations for their clients based on these DNA tests (Innocence Project, 2008). The most recent analysis of these cases indicates that mistaken eyewitness identifications contributed to some 75% of the wrongful convictions. Although it is unreasonable to suggest that 75% of *all wrongful convictions* involve a mistaken identification (these exonerations pertain to a specific subset of all cases which likely over represent the role and influence of eyewitness evidence, Gross & O'Brien, 2007; Innocence Project, 2008), it is reasonable to conclude that erroneous identifications and their persuasive nature contribute significantly to erroneous convictions. Unsurprisingly, the acceptance of this fact shifts the focus of researchers from the nature of the problem to it possible solution.

Legal Safeguards

Adversarial systems around the world have adopted various safeguards to ensure the fairness of trials involving eyewitness evidence. Some of these safeguards include: voir dire, cross-examination of the eyewitness, judicial instruction and expert evidence. This thesis focuses on the last two of these safeguards as a result of the role they play in the Australian criminal justice system; specifically, while eyewitness experts were not permitted to testify in court[1], an absence of the appropriate instruction from the judge is sufficient to constitute grounds for an appeal by the defence.

Judicial Instruction

The term *judicial instruction* refers to a set of verbal directives presented to jurors by the presiding judge before the commencement of deliberations. These instructions (also known as “directions” or “warnings”) are issued in order to ensure that the decisions made by jurors are fair, having rendered “verdicts which represent a finding by the jury under the law upon the evidence presented” (Warren v. Parks, 1976 cited in Ogloff & Rose, p. 408). One subclass of judicial instruction provided at trial specifically addresses the issue of eyewitness reliability and is designed to aid jurors in their evaluations of eyewitness identification evidence.

In Australian courts the uniform *Evidence Act* (Cth, 1995) sections 116 and 165 determine the criteria for the provision of a judicial instruction regarding identification evidence. Section 116 specifies that the judge is to inform the jury that “there is a special need for caution before accepting identification evidence” (*Evidence Act* (Cth), 1995, s. 1a) and to provide reasons for that caution, both in general and in the specific circumstances of the case (*Evidence Act* (Cth),1995, s. 1b). Section 165 provides essentially the same guidance, but is applied strictly to criminal cases. Although no specific wording of the instruction is stipulated, the New South Wales Judicial Commission has formulated a pattern direction which judges may opt to use in identification cases (JCNSW, 2006, s3-020, see Table 1.1 below).

[1] The first Australian eyewitness expert was permitted to appear before the jury in the case of *R v Skaf* (2006).

Table 1.1 : Judicial Instruction.

Members of the jury, in this case you have heard evidence that the accused has been identified by a witness and that that evidence has been disputed by the accused. Now, whenever disputed identification evidence is given I, as the judge, am required to direct you that you must approach this evidence with special caution before deciding whether to accept it as reliable. This caution is necessary, even though you may be satisfied that the witness has been completely honest in their evidence. This direction relates to the reliability of the identification evidence, not to the honesty with which it was given.

Special caution is necessary before accepting identification evidence because of the possibility that even a completely honest witness may have been mistaken in their identification. I am not suggesting that the evidence of such a witness must be regarded as unreliable. My task is no more than to draw your attention to the possibility that the evidence of such a witness may be unreliable, and to explain why that is so, so as to enable you to exercise the special caution which is required in determining whether to accept that evidence as reliable and what weight is to be given to it.

The common experience of criminal courts over the years, both here in Australia and overseas, has demonstrated that identification evidence, however honestly given, may turn out to be unreliable. There have been some notorious cases over the years in which completely honest evidence of identification has been demonstrated to be wrong after innocent people have been convicted.

The reliability of an identification of a person depends upon the circumstances in which the witness observed the person whom he or she identifies as the accused, and any one of these circumstances may possibly lead to error. For example, how long was the period of observation? In what light was it made? And from what distance was it made? Was there anything about the person observed which would impress itself upon the witness? Was there any special reason for remembering the person observed? How long afterwards was the witness asked about the person seen? How did the description then given compare with the appearance of the accused? Each of these matters must be considered in every identification case.

Eyewitness Expert Testimony

The testimony of an eyewitness expert has been proposed as an alternative to the judicial instruction safeguard. Defined by Leippe (1995, p. 910), *eyewitness expert testimony* refers to “the delivery to a jury by a qualified research psychologist of information about research and theory on eyewitness behaviour”. Leippe goes on to suggest that expert testimony has been given “when a psychologist admitted by the judge as an expert authority on ‘eyewitness testimony’ takes the stand in a jury trial and presents information about research and theory concerning memory and the variables know to influence memory and memory reports” (Leippe, 1995, p. 910). This testimony typically consists of a general account of some of the factors known to influence the likely accuracy of human perceptions and memory, and is supported by the citation of relevant research literature (Konecni & Ebbesen, 1986). While we have a relatively clear understanding of who the expert is, when expert testimony has been provided, and what it likely contains, it remains unclear what precise *role* the eyewitness expert is expected to fulfil. That is, why is expert testimony admitted? And how do we know when it has served its purpose?

The role of the eyewitness expert

What then, do courts hope to achieve through the introduction of eyewitness evidence? While this seems a simple question, many different researchers have expressed many different opinions on this subject – often confusing the correlates of the outcome with the outcome itself.

At least three distinct impressions of the role of the expert can be garnered from the psychological literature on the subject. The first suggests that the role of the expert is simply to reduce the weight attributed by jurors to eyewitness evidence (Deffenbacher, 1984; Geiselman, 1994; Leippe, 1995; Leippe, Eisenstadt, Rauch, & Seib, 2004; Pezdek, 2007; Pezdek, Avila-Mora, & Sperry, *in press*; Wells et al., 1980);

> *If courts would allow expert testimony regarding scientific research on components of eyewitness accuracy, it might be possible to counter the unjustified reliance placed upon eyewitness testimony by jurors and by the courts themselves* (Hosch, Beck, & McIntyre, 1980, p. 295).

This account therefore implies that the efficacy of the expert can be measured in terms of reductions in levels of belief in eyewitnesses, and makes no reference to the appropriateness of this shift in belief. That is, the eyewitness expert has had a beneficial influence if jurors believe all eyewitnesses less, regardless of the accuracy of the eyewitness's identification.

The second stated aim of expert testimony is to educate jurors in the shortcomings of eyewitness evidence (Cutler, Dexter, & Penrod, 1989a; Cutler, Penrod, & Dexter, 1989b, 1990b; Hosch, 1980; Katzev & Wishart, 1985; Leippe, 1995; Leippe et al., 2004; Maass, Brigham, & West, 1985; Pezdek, 2007);

> *…the expert's role is to inform the judge or jury about the processes of observing, remembering, and recalling memories, to dispel common misconceptions concerning those processes, and to discuss factors that would likely affect the reliability of eyewitness identifications* (Geiselman, 1994, p. 26).

Here the efficacy of the expert can be measured by the increase he or she causes in jurors' knowledge of the factors affecting the reliability of eyewitness testimony. Although the expert has been invoked in response to concerns over erroneous convictions, here the worth of the expert is measured independently of the extent to which such errors are prevented.

The third purpose attributed to the provision of expert testimony is one also expressed in legal circles. That is, that the testimony of the expert should "aide the jury in reaching an accurate resolution of a disputed issue" (*United States v Downing* cited in Cutler & Penrod, 1995, p. 27) or to "help determine a fact in issue" (*Federal Rules of Evidence*, 1975, 1984 s702) and thereby "contribute to the administration of justice" (Honourable Justice A.R.Abadee, *Federal Rules of Evidence*, 1975, 1984). Thus, here the role of the expert is conceptualised as a means of assisting jurors to reach *just* decisions through the correct differentiation between accurate and inaccurate eyewitness identifications i.e., to "improve[s] the match between verdicts rendered and the actual guilt or innocence of the defendant" (Wells, 1986, p. 90). McCloskey and Egeth (1983) termed this the "discrimination rationale" for the inclusion of eyewitness expert evidence, and it is a position adopted widely in the psychological literature (Ainsworth, 1998; Cooper & Hall, 2000; Cutler et al., 1990b; Devenport, Stinson, Cutler, & Kravitz, 2002; Doyle, 1998; Konecni & Ebbesen, 1986; Lempert, 1986; Pezdek, 2007; Yarmey, 2001):

> *The psychologists' role as expert is to help the trier of fact reach an accurate conclusion about whether the eyewitness's testimony can be trusted in the instant case* (Lempert, 1986, p. 172).

Specifically, this position suggests that expert evidence should be evaluated in the same context as it was invoked, i.e., with regard to the prevention of erroneous convictions. A variant on this position maintains that the outcome desired as a result of expert testimony is indeed *to differentiate* between accurate and inaccurate eyewitnesses in the pursuit of justice. It also suggests that this will be achieved through the education of jurors regarding the limitations of eyewitness testimony, an education that will assist in improving juror discrimination. Education is therefore proposed as the mechanism through which reductions in erroneous convictions will be achieved rather than an end in its own right. Of the three, this is the most widely voiced opinion in the existing legal and psychological literature[2];

> …*expert testimony could serve to educate them* [jurors] *about the factors that influence eyewitness memory. The presumption here is that the expert would aid the jury so that neither too much nor too little credibility will be given to the eyewitness evidence in any particular case* (Greene & Loftus, 1984, p. 400).

Yet there need not be a "presumption" that the expert will successfully aid the jury to appropriately evaluate the testimony of an eyewitness, since the effect of the expert evidence can be empirically evaluated in terms of its influence on juror knowledge, or in terms of its effect on juror ability to discriminate between accurate and inaccurate eyewitness identifications.

Overall, researchers who believe that the expert's role is to act as a "myth buster" or "educator" tend to measure the value of an expert's testimony against the goals of increased knowledge or decreased belief, without directly considering the experts effect on juror discrimination accuracy. In doing so, these researchers are essentially prevented from addressing the legal expectations associated with eyewitness expert evidence, as they can only *infer likely performance on discrimination tasks* from the observed performance on tasks which indirectly measure this variable. These investigations of expert effects are hence divorced from the goals invoked regarding erroneous convictions, as knowing about *how* an expert influences juror education levels does not necessarily tell you about a juror's actual ability to *apply* that information in order to reach an accurate resolution. That is not to say, however, that one cannot be interested in an expert's ability to educate jurors or to change their beliefs. Indeed, it can be valuable to know about the learning process itself, yet this knowledge will not necessarily help to achieve the ultimate objective of preventing

[2] (Benton, Ross, Bradshaw, Thomas, & Bradshaw, 2006; Brigham & Wolfskeil, 1983; Cutler, Penrod, & Stuve, 1988; *United States v. Rincon*, 1994 cited in Cutler & Penrod, 1995; Cutler et al., 1989b; Daubert v. Merrell Dow Pharmaceuticals, Inc., 1993; Devenport & Cutler, 1997; Devenport et al., 2002; Fox & Walters, 1986; Goodman & Loftus, 1992; Hoffheimer, 1989; Hosch, 1980; Hosch et al., 1980; Kargon, 1986; Leippe, 1995; Leippe et al., 2004; Lindsay, 1994; Loftus, 1980; McKenna, Treadway, & McCloskey, 1992; Memon, Hope, & Bull, 2003; Park, 2003; Penrod & Cutler, 1992; Pezdek, 2007; Rahaim & Brodsky, 1982; Seelau & Wells, 1995; R v D.D, 2000 cited in Steusser, 2005; Vidmar & Schuller, 1989; Yarmey, 2001)

jurors from believing eyewitnesses who have made mistakes, while preserving the credibility of eyewitness who made accurate identifications. Thus, psychologists must take care that their conceptualisation of the role of the eyewitness expert has not unintentionally obscured their ability to measure the construct of interest: *the accurate resolution of a dispute in issue* (*United States v Downing* cited in Cutler & Penrod, 1995, p. 27).

Expert Evidence in Adversarial Systems

United States

Over time, and across judicial systems, the admissibility of the eyewitness expert has been determined according to many varied criteria. The *Frye* test, formulated in 1923 as a result of *Frye v. United States* (1923, p. 1014) established the admissibility of "novel" scientific evidence according to the "general acceptance test". Here, expert testimony was deemed admissible when its scientific foundation, either theoretical or practical, was generally accepted within the field to which it belongs. Following this same rationale, the court in *United States v. Amaral* (1973) proposed four specific criteria in the consideration of expert admissibility: 1) whether the witness was a qualified expert; 2) whether the testimony was the proper subject for an expert; 3) whether it conformed to a generally accepted explanatory theory; and 4) whether the probative value of the evidence outweighed its likely prejudicial effect (*United States v. Amaral* cited in Woller, 2004, p. 329). With the introduction of the *Federal Rules of Evidence* in 1975, the position in *Frye* and *Amaral* was subsequently softened. Specifically, the *Rules* replaced the "appreciable help" standard found in *Amaral* with an "assist the jury" standard, and removed all mention of the "general acceptance" rule found under *Frye*.

Thus, although some courts continued to apply the *Frye* test after the introduction of the *Rules*, the ruling in *Daubert* (*Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 1993) held that *Frye* had be supplanted by the *Rules*, including the removal of the criterion of "general acceptance". Instead, the court under *Daubert* focused on defining "scientific knowledge" as cited in section 702 of the *Rules*, and established that where there was scientific expert evidence before the court, its evidentiary reliability and admissibility should be based upon its *scientific validity* rather than its "general acceptance". The court then went on to specify four factors for consideration in assessing the validity of the science: 1) whether the expert's theory or technique is falsifiable and has been tested; 2) the estimated reliability of a procedure and its potential rate of error; 3) whether the results have been published; and 4) whether the expert's methods and reasoning enjoy general acceptance in the relevant scientific community. Essentially, the shift amongst U.S. courts from *Frye* to *Daubert* reflects a recasting of the role of "gatekeeper" over scientific expert evidence. In *Frye* the admissibility of an expert opinion was governed by the scientific community, while under *Daubert*, the U.S. Supreme Court expressed faith that the judiciary could take on the role of the critical and analytical evaluator of expert scientific opinion.

England & Wales

Under the common law in *R v. Turner* (1975) expert scientific evidence is only admissible in England and Wales when it meets "common knowledge" and "experience" criteria; thus "if on the proven facts a judge or jury can form their own conclusions without help, then the opinion of an expert is unnecessary" (*R v. Turner,* 1975, p. 841). This, together with the suggestion that opinions regarding human nature are not helpful to jurors, largely rendered expert psychological testimony inadmissible in English courts unless it specifically addressed mental abnormality or the defendant's own state of mind (Coleman, 1993 cited in Kapardis, 1997). This position was softened somewhat in *R v. Emery* (1993), and since that time expert psychological evidence has at times been admitted (Kapardis, 1997).

Australia

In Australia, the admissibility of expert evidence is not governed by *R v. Turner*, even though Australian courts do share a common legal heritage with England and Wales. Over the years, the common law in Australia had set down six preconditions controlling the admissibility of expert evidence: 1) the expert must be appropriately qualified; 2) the expertise must qualify as a "recognised field or area of expertise"; 3) the evidence must not be related to matters of "common knowledge"; 4) the expert must testify to a matter relevant to their expertise; 5) there must be a verifiable "basis" to their opinion; and finally 6) the expert may not testify to the "ultimate issue" (Hunter, Cameron, & Henning, 2005, pp. 1421-1422). Despite this attention to the admissibility of expert evidence, the application of these rules was "fraught with difficulty and uncertainty" (Hunter et al., 2005, p. 1422), with some courts adopting tests akin to *Frye* requiring “general acceptance” and others utilising a criterion derived from a combination of *Frye* and *Daubert* requiring “general acceptance” and “reliability”.

This confusion was addressed with the introduction of the uniform *Evidence Act* (Cth, 1995) under which the “common knowledge” rule, the “expertise” rule, the “basis” rule and the “ultimate issue” rule were replaced with a “helpfulness” standard: i.e., would the evidence materially assist the trier of fact? In the event that this criterion was met, the *Act* still provided for the exclusion of the expert on the basis that the probative value of the evidence was outweighed by the danger that: a) it was unfairly prejudicial to a party; b) it was misleading or confusing; or c) it would cause or result in undue waste of time (*Evidence Act* (Cth), 1995, s. 135). Although at first excluded under the *Act* at section 135(c) (*R v. Smith*, 2000), eyewitness identification evidence has now twice been successfully brought before jurors in New South Wales (*R v Sarago*, 2006; *R v Skaf*, 2006, 394 NSWSC).

Irrespective of the specific admissibility criteria in adversarial jurisdictions, the inclusion of eyewitness expert evidence as a safeguard against erroneous eyewitness identifications is based at least in part upon the belief that expert testimony is useful to jurors over and above other available safeguards. Thus, it is valuable to begin investigations of judicial instruction and expert evidence safeguards by first establishing if the effects of expert testimony are indeed superior to those of the alternative. This investigation is conducted in the broader context of a survey of the knowledge and opinions of legal professionals regarding eyewitness issues, as it is these professionals, whose beliefs about the relative efficacy of expert opinions and judicial instructions, often determine if expert evidence will be sought in any particular case.

Chapter 2

Public Defenders' Survey

Over at least the last 25 years, research psychologists have conducted several survey and questionnaire studies to investigate the knowledge and opinions of legal professionals regarding eyewitness identification issues. Surveys of law students (McConkey & Roche, 1989; Noon & Hollin, 1987; Yarmey & Jones, 1983), law enforcement personnel (Benton et al., 2006; Potter & Brewer, 1999; Wogalter, Malpass, & McQuiston, 2004), lawyers (Potter & Brewer, 1999; Rahaim & Brodsky, 1982; Yarmey & Jones, 1983) and judges (Benton et al., 2006; Wise & Safer, 2004; Yarmey & Jones, 1983) have provided estimates of the extent to which the opinions of the eyewitness evidence gatekeepers correspond with the evidence-based knowledge of those who research the topic. The consistent finding which emerges from these studies has been that, compared to the opinions of eyewitness experts, legal professionals demonstrate a limited appreciation for the factors known to influence the reliability of eyewitness identification evidence (Benton et al., 2006; Penrod & Cutler, 1999). This has in turn fuelled doubts that legal professionals can adequately defend the accused when faced with eyewitness identification evidence. The investigation presented here focuses specifically on the perceived disparity between the knowledge held by Australian legal professionals and eyewitness experts, touching only briefly on the legal professionals' opinions regarding the efficacy of judicial instructions.

Before expert opinion evidence can be deemed admissible in a New South Wales court, it must first be established that the person testifying *has specialised knowledge* based, for example, on the person's training, study or experience, and that *the opinion of that person is wholly or substantially based on that knowledge* (*Evidence Act* (Cth) 1995, s. 79). In addition to these criteria, general rules of admissibility also apply: if the prejudicial value of the expert's evidence is considered to outweigh its probative value, the judge has the discretion to exclude the entirety, or elements, of expert's evidence (*Evidence Act* (Cth), 1995, s. 135). This can place considerable restrictions on the admissibility of eyewitness expert evidence, and to date, eyewitness expert testimony has been ruled admissible in only two cases in New South Wales; *R v*

Sarago (2006) and *R v Skaf* (2006, 394 NSWSC), leaving it to Australian legal professionals to address eyewitnessing issues, and assist jurors in the evaluation of eyewitness evidence in the vast majority of cases.

The knowledge of Australian trainee lawyers and legal professionals relating to eyewitness issues has been surveyed on two separate occasions, first by McConkey and Roche (1989) and then again 10 years later by Potter and Brewer (1999). Using the Knowledge of Eyewitness Behaviour Questionnaire (KEBQ; Deffenbacher & Loftus, 1982), McConkey and Roche (1989) compared the performance of 60 advanced law students to that of 47 advanced and 124 introductory psychology students. It was concluded that although all three groups showed limited knowledge of eyewitness memory issues, psychology students who had been lectured on human memory displayed significantly greater knowledge than introductory psychology students who had not yet received this training and legal students familiar only with eyewitness evidence law. These findings were consistent with research measuring the attitudes and knowledge of students in the U.S. and U.K. (Deffenbacher & Loftus, 1982; Noon & Hollin, 1987). Subsequently, Potter and Brewer (1999) asked 67 detectives, 41 legal practitioners (both prosecution and defence) and 119 undergraduate psychology students to estimate how well 12 different witness behaviours – displaying excessive confidence, fidgeting and providing testimony inconsistent with other witnesses or with earlier statements – predict the accuracy of a witness's testimony. Across all groups, a number of these witness behaviours were commonly interpreted as indicators of likely testimonial *inaccuracy* (as in the cases of too much confidence or recalling items not previously remembered), even though there is no empirical data supporting the existence of such a relationship. More recently, Benton, Ross, Bradshaw, Thomas and Bradshaw (2006) replicated the approach adopted by Kassin and Barndollar (1992) by comparing the knowledge and opinions of jurors, judges and law enforcement personnel with those eyewitness experts surveyed by Kassin, Tubb, Hosch, & Memon (2001). They found that judges and law enforcement personnel displayed similar overall accuracy rates and were significantly more likely to concur with expert opinion (40%) than were potential jurors (13%). Encouragingly, this pattern of results suggests an increase over the decade in question in the psychological knowledge held by legal professionals.

The present study surveys an opportunity sample of Australian legal practitioners, using a modified version of the questionnaire employed by Kassin et al. (2001) in order to ascertain if this trend continues in the Australian context. The responses obtained will be compared to those of eyewitness experts surveyed by Kassin and colleagues, and observed differences and similarities between these samples will be discussed.

Method

Participants

A questionnaire designed to measure knowledge and opinions regarding eyewitness testimony was administered to an opportunity sample of 130 legal professionals attending the Annual Public Defenders' Conference in New South Wales. The questionnaire was completed and returned by 35 conference delegates with an average of 16 years experience in legal practice. Most respondents identified themselves as "public defenders" (49%)[3], followed by those identifying more generally as "criminal lawyers" (26%), and "barristers" (6%). Three respondents (9%) chose not to specify their professional status.

Materials

Eyewitness Testimony Questionnaire

A questionnaire was developed for the conference delegates based on the surveys of eyewitness experts conducted by Kassin and colleagues in 1989 and 2001 (Kassin, Ellsworth, & Smith, 1989; Kassin et al., 2001). The Eyewitness Testimony Questionnaire (referred to hereafter as the ETQ) differed from the revised Kassin et al. (2001) survey in three ways. Firstly, the scale on which conference delegates were asked to rate statements about eyewitness testimony was changed from a 7-point Likert scale (where 1 = The reverse is probably true; 2 = No support; 3 = Inconclusive; 4 = Tends to favour; 5 = Generally reliable; 6 = Very reliable; and 7 = I don't know) to a 5-point Likert scale (where: 1 = Definitely true; 2 = Probably true; 3 = Probably false; 4 = Definitely false; and 5 = I hadn't considered this an issue). Secondly, the direction of some items from the Kassin et al. (2001) survey (numbers

[3] All percentages are rounded to the nearest integer value.

2, 8, 11, 13, 14, 23 and 25) were reversed in order to reduce the likelihood of a response bias (see Table 2.1 below). Finally, in line with recommendations from Kassin and Barndollar (1992) three items (5, 7 and 28) were reworded in an attempt to clarify the propositions for a population of respondents unlikely to be familiar with the terminology associated with psychological research in eyewitness testimony (see Table 2.2 below). All statements from the original 30 item questionnaire appear in the ETQ in their original or an amended form (see Appendix A for the complete ETQ).

Table 2.1: Original and Reversed Statements from the ETQ

Item Number and Original Statement	**Reversed Statement**
2. The presence of a weapon impairs an eyewitness's ability to accurately identify the perpetrator's face.	The presence of a weapon does not impair an eyewitness's ability to accurately identify the perpetrator's face.
8. An eyewitness's confidence is not a good predictor of his or her identification accuracy.	An eyewitness's confidence is a good predictor of his or her identification accuracy.
11. An eyewitness's testimony about an event can be affected by how the questions put to that witness are worded.	An eyewitness's testimony about an event remains invariant no matter what the wording of the questions asked.
13. Police officers and other trained observers are no more accurate as eyewitnesses than is the average person.	Police officers and other trained observers are more accurate as eyewitnesses than the average person.
14. Hypnosis increases suggestibility to leading and misleading questions.	Hypnosis decreases suggestibility to leading and misleading questions.
23. Memories people recover from their own childhood are often false or distorted in some way.	Memories people recover from their own childhood are usually highly accurate.
25. Young children are less accurate as witnesses than are adults.	Young children are more accurate as witnesses than are adults.

Table 2.2: Original and Reworded Statements from the ETQ

Item Number and Original Statement from Kassin et al., 2001	**Reworded Statement**
5. Police instructions can affect an eyewitness's willingness to make an identification.	Police instruction can influence whether or not an eyewitness makes a selection from a lineup.
7. The rate of memory loss for an event is greatest right after the event and then levels off over time.	Memory for an event declines most rapidly immediately after its occurrence and more slowly thereafter.
28. Witnesses are more likely to misidentify someone by making a relative judgment when present with a simultaneous (as opposed to sequential) lineup.	In a lineup, the way in which photographs are presented to witnesses (e.g. simultaneously or sequentially) affects the accuracy of identifications.

Respondents were also asked a series of questions regarding their experience with eyewitness expert evidence, its admissibility, and their perceptions of the effectiveness and clarity of the required judicial instruction (*Evidence Act* (Cth), 1995, s. 116).

Procedure

The questionnaire was issued to each conference delegate as part of the conference pack. Two announcements were made on the first day of the conference, drawing delegates' attention to the questionnaire and asking them to complete it. All questionnaires were collected immediately before a keynote address, "The psychology of identification evidence", given by a research psychologist on the morning of the second day of the conference. This served to ensure that the content of the lecture did not influence the answers given by respondents.

Results

Respondents

The legal respondents in this sample had an average of 16 years experience in the law in various capacities (public defenders, $\overline{x}$ = 15yrs; criminal lawyer, $\overline{x}$ = 16.9yrs; other, $\overline{x}$ = 21yrs). Ninety-four percent of these respondents indicated that they had been involved in cases which included disputed eyewitness identification evidence; however, only three had commissioned an eyewitness expert report. Sixty percent of respondents indicated that disputed identification cases represented between 5% and 10% of their cases, with estimates ranging from zero to 50% ($\overline{x}$ = 14.91%, $\overline{\sigma}$ = 12.5%).

Judgements of Eyewitness Phenomena

Each of the 30 ETQ statements were investigated to establish how the respondents, as a group, understood the phenomenon described, and to compare this understanding with that expressed by eyewitness experts in the 2001 Kassin et al. survey. In order to aid this comparison, scores on those statements where the direction of the proposition was changed (i.e., statements 2, 8, 11, 13, 14, 23 and 25 on the ETQ), were reverse-coded to maintain consistency with the Kassin et al. data.

Table 2.3 presents the distribution of responses to each of the 30 items. *All* delegates responded either *definitely true* or *probably true* to the items regarding post-event information, attitudes and expectations, confidence malleability and alcoholic intoxication (items 9, 15, 17 and 20). A clear majority of those surveyed responded *definitely* or *probably true* to items relating to mug-shot induced bias (97%), lineup instructions (94%), stress (91%), child suggestibility (91%), unconscious transference (89%), accuracy of child witnesses (86%), cross-race bias (86%), showups (77%), exposure time (74%), weapon focus (74%), presentation format (74%), colour perception (71%), trained observers (69%) and description-matched lineups (69%) (items 21, 5, 1, 26, 12, 25, 18, 3, 6, 2, 28, 10, 13 and 27 respectively). Most delegates responded *definitely* or *probably false* to the statements relating to discriminability (80%), long-term repression (79%) and the forgetting curve (77%) (items 24, 7 and 22 respectively). However little, if any, consensus was evident with regard to the following statements: lineup fairness (57%), wording of questions (49%), event violence (49%), false childhood memories (47%), accuracy-confidence (40%), identification speed (35%) and elderly witnesses (34%) (items 4, 11, 16, 23, 8, 30 and 29 respectively). Hypnotic suggestibility and hypnotic accuracy were not considered to be relevant issues by approximately 43% and 37% of delegates respectively, and *probably* or *definitely false* by the majority of the remaining respondents in both instances.

Table 2.3: Distribution of Judgments for the 30 statements

Topic	Definitely True	Probably True	Probably False	Definitely False	Not an Issue
1. Stress	15	17	3	0	0
2. Weapon focus*	15	11	8	1	0
3. Showups	17	9	4	3	1
4. Lineup fairness	3	17	9	6	0
5. Lineup instructions	24	9	1	0	1
6. Exposure time	12	14	7	1	1
7. Forgetting curve	0	7	7	20	1
8. Accuracy-confidence*	6	8	13	5	3
9. Post-event information	15	20	0	0	0
10. Colour perception	8	17	2	2	6
11. Wording of questions*	0	17	14	4	0
12. Unconscious transference	8	23	1	0	3
13. Trained observers*	8	16	8	2	1
14. Hypnotic suggestibility*	6	11	2	1	15
15. Attitudes and expectations	12	23	0	0	0
16. Event violence	4	13	13	1	4
17. Confidence malleability	13	22	0	0	0

Topic	Definitely True	Probably True	Probably False	Definitely False	Not an Issue
18. Cross-race bias	16	14	0	2	3
19. Hypnotic accuracy	0	5	9	8	13
20. Alcoholic intoxication	22	13	0	0	0
21. Mug-shot induced bias	22	12	1	0	0
22. Long term repression	0	6	20	6	1
23. False childhood memories*	0	16	14	2	2
24. Discriminability	1	2	16	12	3
25. Child accuracy*	16	14	3	0	1
26. Child suggestibility	17	15	3	0	0
27. Description-matched lineup	6	18	7	4	0
28. Presentation format	10	16	3	0	6
29. Elderly witnesses	1	11	18	2	3
30. Identification speed	2	10	14	7	1

* Indicates topics where reverse-coding is reported (i.e. 1 (original value) = 4 (reported value), 2 = 3, 3 = 2, 4 = 1and 5 = 5).

Using the same methodology as Kassin and Barndollar (1992), and Benton et al. (2006), Table 2.4 compares the proportion of participants agreeing with a statement to the proportion of experts surveyed in Kassin et al. (2001) who endorsed the same statement. Chi-square values and significance for each comparison are reported.

Table 2.4: Comparison of Agreement Rates for Kassin et al. (2001) Experts and Legal Professionals

Item Number and Topic		% Expert Agreement (n = 64)	% Lawyer Agreement (n = 35)	$\chi2_{(1)}$ and (p)
5.	Lineup instructions	98	94	1.33 (*ns**)
11.	Wording of questions	98	49	36.28 (.000)
21.	Mug-shot induced bias	95	97	0.2 (*ns*)
17.	Confidence malleability	95	100	1.69 (*ns*)
9.	Post-event information	94	100	2.28 (*ns*)
26.	Child suggestibility	94	91	0.19 (*ns*)
15.	Attitudes and expectations	92	100	2.88 (*ns*)
14.	Hypnotic suggestibility	91	49	21.79 (.000)
20.	Alcoholic intoxication	90	100	3.49 (*ns*)
18.	Cross-race bias	90	86	0.55 (*ns*)
8.	Accuracy-confidence	87	40	24.65 (.000)
2.	Weapon focus	87	74	2.78 (*ns*)
7.	Forgetting curve	83	20	37.39 (.000)
28.	Presentation format	81	74	0.66 (*ns*)
6.	Exposure time	81	74	0.66 (*ns*)
12.	Unconscious transference	81	89	0.9 (*ns*)
3.	Showups	74	77	0.34 (*ns*)
27.	Description-matched lineup	71	69	0.03 (*ns*)
4.	Lineup fairness	70	57	1.74 (*ns*)
25.	Child accuracy	70	86	2.92 (*ns*)
23.	False childhood memories	68	47	3.98 (.046)

Item Number and Topic		% Expert Agreement (n = 64)	% Lawyer Agreement (n = 35)	$\chi2_{(1)}$ and (*p*)
10.	Colour perception	63	71	0.8 (*ns*)
1.	Stress	60	91	11.22 (.001)
29.	Elderly witnesses	50	34	2.26 (*ns*)
19.	Hypnotic accuracy	45	14	9.66 (.002)
30.	Identification speed	40	35	0.27 (*ns*)
13.	Trained observers	39	69	7.88 (.005)
16.	Event violence	37	49	1.14 (*ns*)
24.	Discriminability	32	9	7.03 (.008)
22.	Long term repression	22	19	0.18 (*ns*)

* *ns* = not significant

The responses of the legal professionals did not significantly differ from those of the experts for 21 out of the 30 items. For the nine items where there was a difference in responses between the two groups, the legal professionals were significantly less likely than the experts to agree with propositions relating to: the effects of question-wording on eyewitness testimony, the impact of hypnosis on eyewitness suggestibility, the relationship between eyewitness confidence and identification accuracy, patterns of memory decay, the likelihood that recovered memories from childhood are false and the accuracy of hypnotically induced statements. Legal professionals were significantly more likely to believe that: there was a relationship between high stress and impaired recall, that trained observers were more skilled than untrained observers, and that true and false memories could be differentiated. For four of these nine statements, the experts and lawyers showed a very marked difference in their level of agreement (i.e., by greater than 40%). These items were: that the rate of memory loss is greatest right after the event then plateaus (83% vs. 20% for experts and lawyers respectively), that eyewitness testimony can be influenced by question wording (98% vs. 49%), that eyewitness confidence is a poor predictor of eyewitness accuracy (87% vs. 40%) and that hypnosis increases suggestibility to certain types of questions (91% vs. 49%).

An "accuracy" score was computed by describing a response as "correct" when it was in agreement with more than 75% of expert opinions (the same criterion used by Benton et al., 2006 and Kassin & Barndollar, 1992). Using this criterion it was found that, of the17 items on which experts reached a consensus (items 1-16 and 30), the majority of legal professionals gave the "correct" response for all but four items (i.e., 75% correct). That is, the majority opinion of the legal professionals was the same as that of the experts 75% of the time.

The Efficacy of Judicial Instruction

The majority of legal professionals expressed the belief that jurors *definitely* (9%), or *probably* (52%) did not understand the judicial instructions regarding the limitations of eyewitness evidence, while only approximately 39% indicated that juror's *probably* did understand the instruction.

When asked about their perceptions of the influence of the judicial instruction on jury decision-making, approximately 45% of respondents indicated that the instruction would have an unbiased effect, 36% of respondents believed it would have an influence which would favour the prosecution, almost 13% believed in would introduce a pro-defence bias, and around 10% suggested that the instruction would have no effect at all.

The final question asked delegates if they thought that the relevant judicial instruction could replace the testimony of an eyewitness expert. Of the 26 participants who answered this question, significantly more than half (73%) indicated that they did not think a judicial instruction was equivalent to the evidence given by an eyewitness expert ($\chi^2_{(1)} = 5.54$, $p < 0.05$).

Discussion

Knowledge of Eyewitnessing Issues

This sample of legal professionals was seen to exhibit a substantial degree of consensus (i.e. greater than 80% agreement) on 12 of the 30 tested statements (40% of the survey). That is, the surveyed legal professionals were consistent in their opinions on the effects of post-event information on eyewitness testimony, the role of attitudes and expectations in eyewitness memory, the malleability of eyewitness confidence estimates, the effect of alcoholic intoxication, the role of mug-shot selection on identification rates, the impact of lineup instructions on identification rates, the impact of stress on recall, the suggestibility of child witnesses, unconscious transference of memories across contexts, the accuracy of child witnesses, the role of cross-race witnessing on identification accuracy and the extent to which true and false memories can be discriminated. Penrod and Cutler (1999) reported a high level of agreement (> 75%) among attorneys on only one question from the Brigham and Wolfskeil (1983) and Rahaim and Brodsky (1982) surveys combined. The level of agreement reported here, therefore, suggests that a consensus may be emerging among legal professionals on eyewitness issues that did not previously exist.

Furthermore, the opinions of this sample of legal professionals corresponded with those expressed by eyewitness experts (Kassin et al., 2001) on 21 out of the 30 items (70% of the survey). Chi-squared analyses revealed significant differences in the opinions of these two groups only on items relating to the impact of question-wording on eyewitness testimony, the role of hypnosis in increasing suggestibility, the relationship between eyewitness confidence and accuracy, the pattern in which memory for an event decays, the likelihood that recovered memories from childhood are false, the impact of stress on recall, the impact of hypnosis on recall accuracy, the relative skills of trained and untrained observers and the discriminability of true and false memories. In two of these cases, legal professionals felt that the issue was not relevant (hypnotic suggestibility and hypnotic accuracy were not considered an issue by 43% and 37% of respondents respectively), while one instance represents a significant difference on an issue where neither legal professionals or experts could be considered to have reached consensus (69% vs. 39%, agreement respectively on the issue of trained observers). The remaining six items for which there were differences between experts and legal professionals (wording of questions, accuracy-confidence, the forgetting curve, false childhood memories, stress, and discriminability), appear to be the only areas which reflect significant discrepancies between evidence-based opinions of experts and the opinions of legal professionals.

When the responses of this sample of legal professionals are compared with those of judges from Benton et al. (2006), both similarities and differences can be noted.

Firstly, the degree of correspondence between legal professionals and experts observed here (70%) appears to be substantially greater than that observed between judges and experts (40%), suggesting that the respondents from this sample are more knowledgeable regarding eyewitness identification issues. However, an examination of the propositions on which legal professionals disagree with experts reveals consistencies across surveyed samples. Of the nine items observed here to reveal discrepancies between legal professionals and eyewitness experts (including those issues legal professionals had not previously considered, i.e., hypnotic accuracy and hypnotic suggestibility), five also produced differences between the responses of judges and experts in Benton et al. (2006); wording of questions, hypnotic suggestibility, accuracy-confidence, forgetting curve, and the accuracy of recovered childhood memories. This pattern suggests not only that there may be some similarity in the knowledge and opinions of these two samples of legal professionals, but also that there may be some consistent gaps in the knowledge of legal professionals. These gaps could be targets for the future education and training of legal professionals involved with eyewitness evidence.

It is also important to consider the *validity* of the opinions expressed by legal professionals, independent of their correspondence with those of the experts. Specifically, since the 2001 survey (Kassin et al.), it has been noted that research in some topic areas has continued to develop, and as a result of recently published research it is likely that the consensus view of psychologists will undergo some change (McCullough, 2002; Shaw, Garcia, & McClure, 1999). That is, some of the views expressed in Kassin et al. (2001) survey may not now be "correct". The relationship between eyewitness confidence and identification accuracy is one example of an area of research which has "proven to be fluid over time" (Shaw et al., 1999). In fact, the consensus expressed by experts in 1989 (Kassin et al.) and 2001 (Kassin et al.) that confidence is not a predictor of eyewitness identification accuracy, may be in need of revision in light of the mounting evidence suggesting that under certain conditions, witness confidence *may* be a useful predictor of accuracy (Brewer & Wells, 2006; Juslin, Olsson, & Winman, 1996; Leippe, 1995; Lindsay, Nilsen, & Read, 2000; Olsson, Juslin, & Winman, 1998; Read, Lindsay, & Nicholls, 1998; Sporer, Penrod, Read, & Cutler, 1995; Weber & Brewer, 2003, 2004). Thus the discrepancy here between the responses of legal professionals and eyewitness experts does not *necessarily* reflect a lack of knowledge on the part of legal professionals, rather it may suggest that, in this area at least, the responses of legal professionals are closer to the current perception of the truth than were the eyewitness expert opinions expressed in 1989 and 2001.

Overall, it appears that the knowledge of eyewitness issues demonstrated by legal professionals in this sample, although still somewhat limited in comparison with eyewitness experts, reflects an improvement on that of a sample of judges assessed with a similar questionnaire (Benton et al., 2006). Thus, while it is not possible to suggest that accurate knowledge necessarily equates with an empirically sound approach to the handling of eyewitness evidence in courts, it is possible that descriptions of attorneys' "poor effectiveness" (Penrod & Cutler, 1999) may warrant revision in light of the current findings. Encouragingly, Wise and colleagues (Wise, Pawlenko, Meyer, & Safer, *under review*) report almost identical accuracy rates (71%) for their much larger sample of 1184 U.S. defence attorneys, and also note significant differences between the opinions of attorneys and experts on the issues of the forgetting curve and the effects of stress on identification accuracy. This cross-validation of the results reported here provides some evidence that psychologist's attempts to educate legal professionals have been effective although some consistent gaps in legal knowledge remain.

Perceptions of legal safeguards

Although the majority of legal professionals in the survey (60%) expressed the belief that jurors probably or definitely would not understand a judicial direction given by a judge, almost 94% indicated that they believed the direction would influence jury decision-making (45% without bias, 36% with a prosecution bias and 13% with a defence bias). In addition, the majority (73%) of those who responded indicated that the effect of the judicial direction was not equivalent to evidence of an eyewitness expert. These views are largely consistent with the available empirical evidence relating to the effects of judicial direction on juror decision-making (See Chapter 4, for a review of this literature). Greene (1988) investigated the effect of a standardised eyewitness cautionary instruction (*U.S. v Telfaire*, 1972), and compared its effect on jury verdicts with that of a revised set of directions. Consistent with the beliefs of legal professionals reported here, these results showed jurors had poor comprehension of the instruction, and overall it was concluded that the instruction had no effect on jury decisions. Cutler, Penrod and Dexter (1990) investigated the effects of the same *Telfaire* instruction and characterised it as ineffective at improving juror sensitivity to eyewitness identification issues. Although revised versions of the *Telfaire* instruction have found more support among research psychologists (Greene, 1988; Ramirez, Zemba, & Geiselman, 1996), there is no empirical evidence available regarding the efficacy of the pattern instructions adopted in other jurisdictions, in this case the specific pattern instruction recommended by the Judicial Commission of NSW (2002; s 3-020) for use in New South Wales Courts. Thus, in the absence of specific empirical knowledge on the subject, the opinions of legal professionals appear consistent with the body of information available regarding juror comprehension of pattern eyewitness instructions, if not their influence.

While there has also been some support for the suggestion that eyewitness expert evidence has beneficial effects on jury decision-making when compared with the pattern judicial instruction (Leippe, 1995; Penrod & Cutler, 1999), there has only been one study which has directly compared the influence of these two sources of information; Cutler, Dexter and Penrod (1990) showed that the expert's testimony reduced juror belief in eyewitness evidence, while the judicial instruction had no systematic effect on juror decisions. Thus, it seems that legal professionals and eyewitness experts agree that the impact of judicial instruction and expert evidence are not equivalent. This view is largely consistent with the available evidence.

Limitations

The legal professionals in this sample were recruited as part of an opportunity sample on the basis of their attendance at the Annual Public Defenders' Conference. This may have had an effect upon the generalisability of the observed results in two ways. Firstly, all surveyed individuals were attending a conference aimed at developing professional skills and providing information relevant to the role of a public defender, therefore it is likely that these individuals could represent a more knowledgeable group of legal professionals than a random sample from varied sources would have revealed.

Secondly, by virtue of being sampled at a meeting of *public defenders* it is likely that the respondents in this study are not representative of legal professionals in general. Although Wise and Safer (2004) concluded that a judge's background (either prosecution or defence) was unrelated to their knowledge of eyewitness testimony, Brigham & Wolfskeil (1983) and Lindsay, MacDonald and McGarry (1990) found that, when compared with prosecutors, defence attorneys were significantly more favourably disposed toward expert psychological testimony on eyewitness identification. Given this, it is possible, that the high degree of consensus between respondents in this sample reflects the homogeneity of their professional roles, rather than an increasing consensus among legal professionals in general. This is not to say that it is inappropriate to investigate the knowledge and opinions of defence attorneys. Kassin et al. (2001) reported, for example, that 89% of requests for expert assistance came from criminal defendants: this illustrates the fact that it is often the role of a defence attorney to challenge eyewitness evidence, and in particular, to ensure that unreliable identification evidence does not go unchallenged. Thus, by gauging defence counsels' knowledge of eyewitness identification we can assess their capacity to adequately fulfil this role

Finally, it is possible that the changes made to the questionnaire to enhance the comprehension of propositions may have changed the interpretation of some statements. Reassuringly, of the three questions with altered wording, two produced responses consistent with the opinion of the experts (presentation format and lineup instructions), and one (the forgetting curve) was consistent with the responses provided by judges (Benton et al., 2006). Thus it does not appear that modifications made to the statements in order to increase comprehension inappropriately altered the intent of the propositions themselves.

Conclusions

This survey provides insight into the knowledge and opinions held by a group of Australian legal professionals. The results suggests not only that their opinions demonstrate higher levels of internal consistency than other surveyed groups of legal professionals, but also that they exhibit a moderate to high degree of correspondence with the opinions and knowledge of eyewitness experts. This survey also revealed that legal professionals and experts share doubts about the efficacy of the judicial eyewitness direction relative to eyewitness expert testimony. These findings highlight the need to further investigate the impact of expert evidence in the field of eyewitness identification, and its relative effect when compared with judicial instruction. Although, as revealed by our survey, it seems that the lawyers share opinion of psychologists: that experts and judges are not equivalent, at present there is little empirical evidence available to test this hypothesis.

SECTION 2: CRITICAL REVIEW

Chapter 3

Eyewitness Expert Effects - Literature Review

Methodological Approaches

Since 1980 a total of 23 studies which experimentally varied the presence or absence of eyewitness expert evidence have been published in book chapters and peer-reviewed journals. This research was undertaken in the hope of gaining an understanding of the effect of expert evidence on juror decision-making. During that time, several researchers have highlighted the methodological diversity within this literature as a means by which to account, in some way, for some of the divergent effects observed. After only three studies had been published addressing the impact of eyewitness expert evidence (Hosch et al., 1980; Loftus, 1980; Wells et al., 1980), Hosch (1980) noted that researchers had conducted their investigations with operationally different approaches, and argued that replications of this type contribute to the generalisability and validity of the results obtained because the differing techniques had produced some consistent results. The author then went on to suggest that the evidence indicating that eyewitness expert evidence increased juror deliberation time and reduced juror tendency to believe eyewitness evidence showed that “expert testimony has a reliable effect on jurors’ beliefs and behaviours” (p. 300).

Nine years later Cutler, Penrod and Dexter (1989b) discussed the impact of observed methodological differences on the conclusions being made by researchers. Specifically, these authors proposed three types of effects likely to result from eyewitness expert testimony: Confusion, Sensitivity and Skepticism, and concluded that certain experimental designs confounded these distinct effects. To define the first effect: *Confusion* refers to the situation where the testimony of the expert “affects the jury in some unanticipated and undesirable way” (McCloskey, Egeth, & McKenna, 1986, p. 6); including but not limited to the absence of main or interaction effects, poor memory for the expert’s testimony, and counter-intuitive judgements by jurors (Cutler et al., 1989b). Juror *Sensitivity* is comprised of two components, Knowledge and Integration, and describes a process where jurors show the ability to weigh and combine factors likely to influence the quality of an identification, allowing them to differentiate between better and worse quality identifications. Consequently, those jurors showing Sensitivity will be more likely to believe eyewitness identifications made under relatively good witnessing conditions, and less likely to believe eyewitness identifications made under relatively poor witnessing conditions. The final effect anticipated is *Skepticism*, and refers to a "tendency to doubt or to disbelieve an eyewitness's testimony" (Devenport, Penrod & Cutler, 1998, p. 354) as a direct result of the expert's evidence. Different from Sensitivity, this type of juror response is indicated by a significant effect of expert testimony, where those jurors who heard an expert are less likely to believe an eyewitness's identification than those jurors who did not hear the expert's evidence irrespective of the quality of the evidence. Importantly, Cutler et al, (1989b) realised that certain experimental designs directly confounded the latter two of these effects: Sensitivity and Skepticism. More precisely, experiments which did not simultaneously and independently vary witnessing factors and expert evidence made it impossible for researchers to determine if observed decreases in guilty verdicts (or identification belief decisions) amongst jurors who heard expert testimony were the result of an increased awareness of the quality of the witnessing conditions, or simply a reflection of the fact that the expert had caused them to doubt the accuracy of *all* eyewitnesses. This distinction can only be made where the quality of the witnessing conditions are varied independent of the presence or the absence of expert testimony. By the time Cutler and colleagues offered their insights, nine articles had been published (including two others by Cutler and colleagues: Cutler, Dexter, & Penrod, 1990a; Cutler et al., 1989b) which reported data on the effects of eyewitness expert evidence, only some of which applied the design necessary to differentiate Sensitivity from Skepticism. Thus, although many of these studies reported effects consistent with juror Skepticism (Fox & Walters, 1986; Hosch et al., 1980; Loftus, 1980; Maass, Brigham, & West, 1987; Wells et al., 1980), Cutler et al. (1989b) reasoned that some of these results may actually be the result of the confounded study design, rather than actual Skepticism (i.e., Fox & Walters, 1986; Hosch et al., 1980; Maass et al., 1987)[4]. Accordingly, the authors suggested that only

[4] It is interesting to note that although Fox and Walters, 1986 did not independently vary witnessing and identification conditions, they did vary the confidence expressed by the eyewitness from high to

those studies systematically varying both expert evidence and witnessing and identification conditions, could speak precisely to the nature of eyewitness expert effects. However, even focusing only on these remaining studies (with the addition of Fox and Walters, 1986, see footnote 4), it is not clear what effect expert testimony has on juror decision-making. Two of these studies found Sensitivity in the absence of Skepticism (Cutler et al., 1989b; Wells & Wright, 1983 cited in Wells, 1986), suggesting that the expert testimony caused jurors to evaluate the eyewitnesses testimony rather than doubt it; three found Skepticism in the absence of Sensitivity (Cutler et al., 1990a; Fox & Walters, 1986; Loftus, 1980), which indicates that the expert caused jurors to doubt rather than to evaluate; and the final two found evidence for both Skepticism and Sensitivity (Cutler et al., 1989a; Wells et al., 1980), leading to the conclusion that expert testimony causes jurors to both evaluate and doubt eyewitnesses. Overall then, although these authors unquestionably provided a useful framework in which to analyse expert effects, the precise nature of these effects remains unclear.

In 1995 Leippe proposed another series of methodological distinctions which could be applied to the eyewitness expert effects literature. The structure introduced in this instance categorised studies in terms of their ability to differentiate Sensitivity from Skepticism, as well as on the trial context in which the experts' evidence was presented. This author differentiated three categories of trial: 1) an eyewitness description within a brief written trial, which included written descriptions of the eyewitnessing conditions (as seen in Blonstein & Geiselman, 1990; Loftus, 1980; Maass et al., 1985); 2) video eyewitness testimony that is presented in the absence of surrounding case information (Fox & Walters, 1986; Wells & Wright, 1983 cited in Wells, 1986; Wells et al., 1980); and 3) video eyewitness testimony which is

low. This variation is sufficient to separate Skepticism from Sensitivity, yet Cutler et al., (1989b) did not include this article amongst those which could separate the two effects. The rationale behind this may relate to the fact that eyewitness confidence is a witness variable rather than a situation variable, and therefore does not fit with the original definition of Sensitivity which referred to factors "which *influence* the likely accuracy of an evaluation" (Cutler et al., 1989b, p. 313). Even so eyewitness experts will testify to the issue of eyewitness confidence, and jurors can show Sensitivity to it; thus, it seems more appropriate to classify this study as one of those adopting a non-confounded design.

embedded in a trial scenario (Cutler et al., 1989a; Cutler et al., 1989b, 1990b[5]; Hosch et al., 1980; Lindsay, 1994; Wells & Wright, 1983 cited in Wells, 1986).

Some years later Pezdek Avila-Mora and Sperry (*in press*) suggested that consideration of the levels of "mundane realism" present in the trial materials could clarify eyewitness expert effects. Here two distinctions were suggested separating: 1) brief summaries with poor mundane realism (Blonstein & Geiselman, 1990; Loftus, 1980; Maass et al., 1985) from 2) richer assessments with relatively better mundane realism (Cutler et al., 1989a; Cutler et al., 1989b; Devenport & Cutler, 2004; Devenport et al., 2002; Hosch et al., 1980; Leippe et al., 2004). In spite of these various ways of partitioning the eyewitness expert effects literature, little has been gained through the distinctions imposed, and no clear pattern regarding the effects on juror decision-making has emerged. Instead, rather than clarifying the literature, it appears that these distinctions have actually served to conceal more fundamental methodological issues which have remained unaddressed in the literature.

A Hierarchy for the Assessment of Expert Effects

Cutler, Penrod and Dexter (1989) rightly highlighted a fundamental methodological issue in the eyewitness expert literature. They illustrated how a specific experimental design (which did not simultaneously vary expert evidence and witnessing and identification conditions) served to confound two outcome measures, Skepticism and Sensitivity. Although this is clearly an important distinction, the following review presents the argument that the analysis by Cutler et al. (1989b) did not go far enough, and as a result left another fundamental methodological distinction unaddressed in the expert effects domain.

Building upon the analysis of the experts' role presented in Chapter 1, studies are classified here on the basis of the outcome, or expert role, assessed by the design. Specifically, the critical distinction lies between those studies which *directly* measure an expert's influence on the accuracy of jury resolutions (henceforth Sensitivity to Eyewitness Accuracy (SEA)), and those which *indirectly* estimate this ability to discriminate through the measurement of the expert's ability to influence belief

[5] It appears that the original author erroneously referenced Cutler, Penrod and Dexter (1990) rather than Cutler, Dexter and Penrod (1990), as a study which presented participant-jurors with pattern judicial instructions.

(henceforth Response to Expert Evidence (REE)) or educate (Sensitivity to Expert Opinions (SEO)) juror belief decisions. Furthermore, under this conceptualisation, the use of *real* eyewitness testimony, rather than *fictional* or fabricated eyewitness testimony[6], represents the key design feature that determines if a particular study is to be classified as a *direct* or an *indirect* investigation of juror SEA. In particular, those studies using *real* eyewitness testimony *directly* measure expert effects on juror SEA, while those using *fictional* eyewitnesses can only *indirectly* estimate SEA on the basis of either juror Responses to Expert Evidence or their Sensitivity to Expert Opinions. It is in the context of these distinctions that the eyewitness expert effects literature will now be reviewed.

Indirect Measures

Response to Expert Evidence (REE)

Studies at the lowest level of the hierarchy offer the least amount of information regarding the effect of expert evidence on juror ability to render accurate evaluations by discriminating between accurate and inaccurate eyewitnesses. By failing to independently vary witnessing and identification conditions and expert evidence, these studies must *indirectly* estimate likely juror SEA from the experts' ability to moderate juror belief in eyewitness evidence. These studies, considered by Cutler et al. (1989b) to have confounded Sensitivity and Skepticism effects, can only describe the effect of the expert in terms of his or her ability to *change the rate* at which jurors believe an eyewitness. That is, using this methodology we can determine whether guilty verdicts occur significantly more or less frequently after the testimony of an eyewitness expert, but we cannot determine if the expert has improved juror Sensitivity. This is because there is no way of knowing whether the change (in observed guilty verdicts) was warranted. Thus, this method can only be used to establish *whether or not* participant-jurors have been influenced by the presence of expert evidence, leaving researchers to guess at the likely accuracy of those decisions actually made.

[6] The precise meaning of the term "fictional eyewitness" will be discussed in detail below.

Five studies using this approach have been described in the literature. Four of these were (or will be) published in peer-reviewed journal and will be reviewed in detail here. The fifth was referred to only briefly by McCloskey and Egeth (1983) in an article published in *American Psychologist*; as a result it is reviewed equally briefly here.

1. Hosch, Beck & McIntyre (1980)

Hosch, Beck and McIntyre (1980) constructed and enacted a complete trial scenario in order to assess the impact of expert testimony and mock-trial conditions on jury verdicts. The mock-trial was enacted before participant-jurors either in a court room or in a psychology lab and was presented either with or without expert testimony. When present, the expert gave general information about the accuracy of eyewitness testimony. They also provided information about some of the methodologies employed in psychological research, the influence of duration of view on identification accuracy, and specific information regarding their own research.

Results

Although all juries acquitted the defendant and no effect of expert testimony could be detected in jurors' evaluations of the reliability and accuracy of the *particular eyewitness* in the case, participant-jurors who heard expert testimony perceived eyewitnesses *in general* to be significantly less reliable than did those participants who did not hear eyewitness expert evidence. Moreover, those jurors who heard expert evidence were also found to place significantly less importance on the testimony of the eyewitness when reaching their verdicts, even though they did deliberate for significantly longer than their "expert-absent" counterparts. Overall then, the evidence gathered in this study is consistent with the expert having changed juror belief in eyewitness evidence, with the authors concluding that:

> *…the data revealed that expert testimony regarding the many factors that can influence eyewitness accuracy does significantly influence jurors' beliefs in the general reliability and accuracy of eyewitness testimony* (p. 294).

Strengths

This study paid close attention to the trial materials constructed for jurors and made a strong attempt to introduce consequentiality amongst their empanelled "jury" members. They did this by allowing jurors to believe that the case was real and by telling them that their decisions would be relied upon in the resolution of future settlement negotiations. Moreover, participants in this study came from both community and undergraduate samples and deliberated in juries. Subsequently, the ecological validity of this study was very high, and it seems likely that jurors watching the trial would have taken a conscientious approach to their decision-making, thus adding to the reliability of the evidence gathered.

Limitations

Although, overall, the trial presented for jurors can be described as being high in ecological validity, there were some limitations to this. In particular, the testimony of the eyewitness was not challenged by the defence through cross-examination, while the expert's testimony was challenged by the prosecution. This may explain why the prosecutions case did not appear to be compelling to jurors, even in the control conditions, and therefore did not provide a good test for the moderating effect of expert evidence on juror belief decisions.

2. Loftus (1980) – Experiment 2

The written trial materials employed by Loftus (1980) were modelled on an actual military court-martial which had included the testimony of an eyewitness expert. Experiment 2 investigated the impact of the presence or absence of eyewitness expert testimony on jury verdicts regarding a violent crime. The expert in the case testified regarding the mental processes involved in eyewitness identification, as well as some of the factors present in the case including cross-racial identification, stress, weapon focus and alcohol.

Results

In the expert-absent condition, three juries convicted the defendant, two acquitted and one failed to reach a verdict. In the expert-present condition, three convicted, four acquitted and three failed to reach a verdict. Loftus suggested that the difference observed between these conditions was likely due to increased juror scrutiny of the evidence as a result of expert testimony. This conclusion was supported by increased jury deliberation times observed in the expert-present conditions.

Strengths

This study provides a more valid investigation of jury decision-making than many other studies by virtue of the fact that jurors were required to deliberate to reach their verdicts, and through the use of real, although adapted, trial materials. This provides the researcher with greater confidence regarding the generalisability of the results than could be justified in studies without these features.

Limitations

A clear trade-off has been made in this study, between ecological validity and analytical power. By measuring jury rather than individual juror verdicts, Loftus significantly reduced the power available to detect any differences that might emerge as a result of expert testimony. Thus, the absence of differences between expert conditions in this study may simply reflect the fact that there were not enough data-points, rather than an absence of real differences.

3. McKenna, Mellott & Webb (1981)

A thorough search of the expert effects literature has located only one, very brief, reference to this study. Cited by McCloskey and Egeth (1983, p. 552), these authors suggested that the study had been conducted to examine the impact of expert psychological testimony regarding eyewitness accuracy. The case described the robbery of a bank teller who identified the suspect from a lineup two days after the incident, with the expert testifying “concerning factors that may lead to inaccurate eyewitness identifications”.

Results

Guilty verdicts were obtained from 8% of jurors in the expert-absent condition and from 6% in the expert-present condition. McCloskey and Egeth (p. 553) suggest that these results had been replicated “using adults from the Baltimore community as subjects”, but no citation for this additional study was provided.

Limitations

It is obviously not entirely fair to criticise a study with so little specific information available, however it does appear that this study suffers from some serious limitations. Firstly, there are clear problems with power, as there were only 72 participants in total, 24 of whom were in the control condition. As a result, it is unlikely that this data could be subjected to a robust statistical analysis. Moreover, there may be serious concerns regarding the trial materials themselves, as 92% of jurors acquitted the defendant in the control condition. As in the Loftus study described above, this has serious implications for the ability to identify an influence of the expert because juror acquittals were virtually at ceiling levels in the control condition.

4. Maass, Brigham & West (1985)

In this study psychology students read one of two court cases (either a burglary or a convenience-store robbery) involving the same basic facts: a) the identification was cross-race; b) no alibi was provided; and c) the eyewitness identification was the only non-circumstantial evidence. They then deliberated to reach a verdict. There were five expert evidence conditions which differentiated between sample- and person-based expert evidence: 1) expert-absent; 2) expert providing sample-based evidence with causal explanations; 3) expert providing sample-based evidence with non-causal explanations; 4) expert providing person-based evidence with causal explanations; and 5) expert providing person-based evidence with non-causal explanations.

Results

Analysis revealed that expert testimony led participants to attribute significantly less weight to the testimony of the eyewitness, and in one condition (causal, person-based expert evidence) the expert’s testimony led subjects to discount the eyewitness evidence completely. Participants who heard the expert were significantly more lenient and uncertain in their verdicts than were participants who had seen an eyewitness but did not receive expert testimony. Expert testimony was not found to significantly increase deliberation times, although it did result in a “relatively high” (p. 255) percentage of failed or “hung” deliberations (18%).

Strengths

This study attempted to prevent the loss of statistical power that tends to accompany an increase in ecological validity by obtaining both juror and jury evaluations; asking jurors to provide an estimate of the defendant’s guilt before commencing deliberations. These individual judgements permitted statistical analyses which were not possible at the jury level post-deliberation; however, overall the results were consistent with those observed on post-deliberation measures.

Limitations

It is interesting to note that researchers in this study aimed to test "not only whether the expert was at all influential" but also "whether his impact was strong enough to compensate fully for the eyewitness testimony" (p. 211). Given the results observed, the researchers then go on to lament the fact that the expert was "not wholly successful in counteracting eyewitness identifications" (p. 223). Thus Maass and colleagues appear to have adopted the standpoint that expert testimony should circumvent eyewitness evidence altogether, without having given due diligence to the eyewitness's likely, or actual, accuracy. As discussed in Chapter 1, this goal is clearly not consistent with legal expectations of the experts' effect in the trial context.

***5. Pezdek, Avila-Mora & Sperry* (in press)**

In this study researchers investigated the impact of expert testimony, expert timing, and the trial presentation medium on participant-juror evaluations of eyewitness evidence. The trial materials were adapted from those developed by Devenport et al. (2002) involving an armed robbery and a murder where the eyewitness evidence was considered "dubious" on the basis of the eyewitness's opportunity to see the perpetrator prior to the identification, and the lineup procedure adopted by the police. The expert testimony was provided in one of four forms: 1) expert-absent; 2) defence expert after the eyewitness; 3) defence expert after the eyewitness with prosecution rebuttal; and 4) defence expert before the eyewitness. This testimony began with the expert outlining their credentials, before describing basic memorial processes and describing the impact of viewing time, weapon focus, lineup construction, lineup instructions and eyewitness confidence on eyewitness accuracy.

Results

Participant-jurors in the expert-absent condition were not found to differ from participants who heard expert testimony on ratings eyewitness persuasion, alibi persuasion, defendant perceptions or defendant empathy. The introduction of expert testimony did reveal within-subjects differences (before and after expert testimony) on ratings of guilt and lineup suggestiveness such that there was a significant decrease in certainty of guilt and a significant increase in perceptions of lineup suggestiveness. The author argues that these results do not demonstrate Skepticism because the:

> *…defense expert testimony effectively increased respondents' sensitization to the suggestibility of the lineup procedure relative to the no expert condition* (p. 17).

Strengths

This study attempted to experimentally investigate the extent to which variations in expert testimony presentation could account for conflicting results in the literature. This study found compelling evidence to suggest that the order of the expert's testimony, relative to the eyewitness, did not significantly influence participant-juror evaluations.

Limitations

The authors have attempted to qualify the observed Skepticism effects (i.e., the significant decline in guilty verdicts and increase in perceived lineup suggestiveness) on the basis that they found evidence indicating that participant-jurors showed: a) "sensitization" to the lineup suggestiveness (rather than Skepticism); and b) rated the impact of expert testimony on perceptions of eyewitness credibility as neutral; the validity of these inferences seem questionable.

Taking the latter result first, neutral responses on a rating scale do not indicate that the expert had no effect on participant-juror ratings of the credibility of the eyewitness. It simply indicates that the jurors *report* that it did not have an effect. These data provide no means to independently verify if the expert's testimony did actually have an impact on their ratings of eyewitness credibility because the participant-jurors did not directly rate the eyewitness's credibility either before or after expert testimony.

The other result relied upon to moderate the observed Skepticism was described as "sensitization to the suggestibility of the lineup" (p. 17). Since this study did not systematically vary elements of the identification procedure, it is not possible to ascertain if the increase in ratings of suggestiveness are simply a response to expert evidence which would have occurred irrespective of the quality of the lineup, or a valid response to a biased lineup procedure. This can only be established where participant-jurors also have the opportunity to rate the suggestiveness of an unbiased lineup. As that did not happen in this study, the most that can be said about this result is that participant-juror ratings of suggestiveness *increased* after hearing expert evidence, it is not possible to attribute this to participant-juror Sensitivity to the stimuli provided. Given this, it appears that a safer interpretation of the results of this study is to suggest that expert evidence resulted in the utilisation of a harsher decision criterion.

Of the five REE studies reviewed above, four provided evidence indicating that the eyewitness expert significantly moderates juror perceptions of eyewitness evidence. In fact, the experts testimony appears to have influenced the weight and reliability attributed to eyewitnesses, as well as both the amount of time given to the evaluation of the evidence in the case, and the degree of confidence in resulting verdicts (see Table 3.1 (p. 69) for a summary of the expert effects detected using the REE methodology). Because of the limitations of the designs employed, none of this data speaks directly to the accuracy of the decisions reached by the jurors, nor to the experts' ability to improve their decision-making in this regard. This is a fact acknowledged by Loftus (1980):

> *The result is to increase the chances that a reasonable doubt about the defendant's guilt is raised. Is this a good thing? If it reduces the likelihood of a mistaken conviction based on faulty eyewitness identification, then it is* (p. 14)

However, other researchers seem content to infer that expert evidence would result in better decision-making by virtue of the fact that it reduced jurors' reliance on eyewitness evidence:

> *If courts would allow expert testimony regarding scientific research on components of eyewitness accuracy, it might be possible to counter the unjustified reliance placed upon eyewitness testimony by jurors and by the courts themselves*. (Hosch et al., 1980, p. 295)

Sensitivity to Expert Opinions (SEO)

A better experimental design uses fictional eyewitness statements which vary in terms of the witnessing and identification conditions described. This method allows researchers to identify the extent to which, following expert evidence, jurors are sensitive to the manipulation of variables psychologists have identified as correlates of eyewitness accuracy. For example, the experimenter may construct relatively "good" and "poor" witnessing scenarios by manipulating features of the event such as the lighting at the scene, the duration of witness exposure to the perpetrator, the presence or absence of disguise, and the extent to which the lineup shown to the eyewitness was constructed and administered in an unbiased fashion. This method can be used to determine whether jurors are sensitive to the manipulations in witnessing conditions, and if so, whether this Sensitivity is altered by expert evidence. If, on the other hand, jurors are found to be insensitive to the expert's opinions and the relative quality of the witnessing scenarios, researchers can determine whether they are being Skeptical (tend to disbelieve all eyewitnesses) or overly trusting (believing all eyewitnesses).

6. Loftus (1980) – Experiment 1

The study by Loftus (1980: Experiment 1) involved two versions of a case based on the same military court-martial reported above, adapted to include either a violent or a non-violent version of the offence (i.e., good or poor witnessing scenario), as well as the presence or absence of eyewitness expert evidence. The expert testified generally to human memorial processes, as well as to some of the eyewitnessing factors directly relevant in the case (i.e., cross-racial identification, stress, weapon focus and alcohol).

Results

Without expert testimony jurors found the defendant guilty in the violent scenario 21% more often than in the non-violent condition. No statistical analysis of this difference was presented, so it is unclear whether jurors considered the eyewitness's testimony to be less reliable in violent rather than the non-violent context. The introduction of expert psychological testimony significantly reduced the percentage of guilty verdicts of jurors by 25% in the violent condition, and 12% in the non-violent condition. Loftus went on to conclude:

> *That expert testimony had a greater impact, that is, produced a greater reduction in likelihood of conviction, in the face of violent rather than non-violent crime* (p. 13).

Strengths

As in Experiment 2 by Loftus (described above), this study endeavoured to test the effects of expert evidence in an ecologically valid context, using adapted materials from a real trial. This study also required jurors to deliberate in order to reach a verdict that reflected the group's decision. In addition to these positives, this study also appears to have adopted a more balanced trial scenario than previous studies, as guilty verdicts in the violent and non-violent expert-absent conditions were 68% and

47% respectively. This experiment therefore avoided floor and ceiling effects and provided ample opportunity to observe expert effects if present.

Limitations

At times it appears that Loftus makes inferences beyond those that the data can support. Specifically, although the rates of guilty verdicts are reported for each condition, it does not appear that any statistical analysis was conducted to test the association between the level of crime violence and the presence of expert testimony. This is the precise interaction which needs to be tested in order to be able to make conclusions about expert effects on juror Sensitivity to crime violence. Even so, the author concludes that “expert testimony …produced a greater reduction in the likelihood of conviction, in the face of a violent rather than non-violent crime” (p. 14), an inference which is not justified on the basis of an inspection of guilty rates alone. Interestingly, the article does report sufficient information to allow the reviewer to conduct a chi-squared test-of-independence test. This analysis reveals no significant association between crime violence and expert evidence on guilty verdicts ($\chi^2 = 0.19$, $p = .661$). Thus it appears that statements implying an interaction between the quality of the witnessing scenario and expert evidence may be overstating the importance of the differences observed.

7. Fox & Walters (1986)

Fox and Walters (1986) took a novel approach in their attempt to vary elements of the witnessing and identification conditions presented to their participant-jurors. In this case researchers had an actor play the role of either a high- or low-confidence eyewitness. In the low-confidence condition the eyewitness acted hesitant and uncertain, vacillating and raising the tone of their voice at the end of declarative statements. In the high-confidence condition, the eyewitness did not vacillate in their statements and did not raise the tone of voice at the end of sentences. The experimenters also exposed participant-jurors to one of three levels of expert testimony: either specific, general or expert-absent. General expert testimony comprised of a warning that eyewitnesses may choose the wrong person, explained that a stranger identification is different to the identification of a friend, and provided information about general memorial processes. The specific expert testimony condition included a description of 13 different factors known to influence or be associated with the likely accuracy of an eyewitness identification, including but not limited to: eyewitness confidence (instructing jurors to disregard confidence), duration of view, quality of view and weapon focus.

Results

Participants in this study were sensitive to the manipulation of eyewitness confidence and were significantly more likely to believe a confident eyewitness than one lacking in confidence. There was also a main effect of expert evidence such that participants in expert-present conditions were significantly less likely to believe the eyewitness than those who did not hear an expert. Had the participant-jurors shown Sensitivity, they would have been expected to believe all eyewitnesses equally, irrespective of their level of confidence. Instead, expert testimony resulted in lower levels of belief for unconfident eyewitnesses compared with those who were confident. This indicates that jurors did not disregard eyewitness confidence after hearing expert testimony, but instead exaggerated their pre-existing tendency to disbelieve low-confidence witnesses. This pattern was also significantly more pronounced when the expert provided specific rather than general evidence. Moreover, relative to the effect of general evidence, the specific expert testimony significantly reduced jurors' perception of the likely number of accurate identifications eyewitnesses would be able to make under the witnessing conditions described. This may be in part due to the fact that participants who heard the specific expert evidence reported relying on the expert's testimony significantly more when making their decisions than did those who received only general expert evidence.

Strengths

This study represents the first instance where the interaction between expert evidence and features of the eyewitness's evidence, in this case confidence, were experimentally investigated. Consequently this study provides the first opportunity to test juror Sensitivity to Expert Opinion (in this case regarding eyewitness confidence).

Limitations

Originally the authors concluded that:

> *The results obtained support McCloskey and Egeth's (1983a, 1984) position that the main effect of expert testimony is an increase in jurors' Skepticism* (p. 227).

However, this is not an entirely thorough description of the effects observed. While Skepticism was clearly evident in the fact that subjects in the expert conditions were significantly less likely to believe that the eyewitness made an accurate identification, jurors also “continued to use eyewitness confidence as a guide to assessing eyewitness accuracy after viewing expert psychological testimony” (p. 225). Thus, in addition to causing jurors to become more Skeptical of eyewitness identifications, the expert also caused jurors to become more critical of those identifications made with low-confidence rather than high-confidence, magnifying the tendency to use eyewitness confidence to estimate eyewitness accuracy. This reaction from jurors was not consistent with the expert’s opinion, which instructed jurors to disregard confidence when making their evaluations. Thus, in addition to Skepticism, there was some evidence that the expert’s testimony had an opposite effect to that which was intended.

8. Cutler, Penrod & Dexter (1989)

This complete trial study presented student-jurors with prosecution and defence opening statements, the direct and cross-examination of four to five witnesses (depending on condition), closing arguments and standard instructions from the judge. Participant-jurors were then asked to evaluate the materials presented. Within the trial each participant-juror was presented with one of two eyewitness identifications made under either “good” or “poor” witnessing and identification conditions. In the “good” condition the perpetrator appeared undisguised, his weapon was hidden, the identification by the eyewitness was made two days after the event, and the lineup was not suggestive. In the “poor” conditions, the perpetrator wore a hat concealing his hair, he was brandishing a gun, the eyewitness made their identification after a 14 day delay and the lineup instructions presented to them were suggestive. Half of the time participant-jurors saw an eyewitness who was 80% confident in the accuracy of the identification they made, and the other half of the time the eyewitness was 100% confident. Participant-jurors were also assigned to one of three levels of the expert evidence factor: 1) absent; 2) expert evidence regarding general memorial factors and specific witnessing and identification factors together with a verbal description of the magnitudes of effects; or 3) the same expert evidence with a numerical magnitude of effect. Furthermore, in half of the expert-present conditions the expert also offered the opinion either that the identification of the eyewitness was rated 7/25 when they came from the poor witnessing conditions or 20/25 in the good witnessing conditions. In all expert testimony conditions, the expert established their academic credentials before describing general memorial processes relating to storage, retention and retrieval. The expert then went on to outline some factors that can influence the accuracy of an eyewitness identification addressing the effects of stress and violence, disguises (only where relevant), weapon focus, the passage of time, suggestive lineups and the relationship between eyewitness confidence and identification accuracy. Under cross-examination the expert conceded points relating to limitations of some psychological research, as well as other limitations specific to the effects described by the expert relating to stress and violence.

Results

Participant-jurors in the control condition demonstrated that they were able to remember the witnessing and identification conditions and the expert evidence they were presented. Furthermore, jurors in the expert-absent condition demonstrated Sensitivity to the role of disguise, retention interval and lineup instructions; however, they also showed a tendency to believe that eyewitness confidence was a good predictor of identification accuracy. Descriptive expert evidence, offering no opinion, was found to significantly improve participant-juror Sensitivity to weapon visibility. The addition of an opinion to the expert testimony significantly improved juror awareness of lineup instructions. Moreover, expert evidence was also found to sensitise jurors to the weak confidence-accuracy correlation such that confidence was given less weight in determining witness credibility if the expert testified. The authors went on to state that:

> *expert testimony improved juror sensitivity to witnessing and identification conditions in comparison to the control condition* (p. 325), concluding that: *jurors exposed to expert testimony, as compared to those not exposed to expert testimony, relied more on witnessing and identification conditions when rating the probability that an identification was correct and when rendering verdicts* (p. 328).

Thus, this experiment provided considerable evidence that participant-jurors were originally Sensitive to various factors of the witnessing and identification scenario, and that expert evidence induced this sensitivity in some instances, and magnified it in others.

Strengths

In addition to providing the rationale for differentiating between studies which varied witnessing and identification conditions and those that did not, this study provided a thorough and systematic analysis of expert effects on many variables including: juror knowledge of witnessing and identification conditions, juror perceptions of eyewitness credibility, juror verdicts, and defence and prosecution case strength. This study also provided a means to differentiate Sensitivity from Skepticism and found "considerable evidence for the former" (p. 328). Thus, this study provided excellent evidence that expert testimony could not only educate jurors regarding previously unknown factors influencing the likely accuracy of an identification, but also showed that these jurors could integrate this knowledge in order to render decisions consistent with the expert's advice.

Limitations

Despite the strong evidence that expert testimony improved juror Sensitivity with respect to some witnessing and identification factors, it appears that the authors treated some non-significant results as though they had met the accepted $\alpha = .05$ criterion. This point is noted only as a minor concern regarding this study, given that there is clearly a convergence of evidence suggesting that jurors were Sensitive to the opinions offered by the eyewitness expert, and therefore the likely accuracy of the eyewitness identification.

9. Cutler, Dexter & Penrod (1989)

The design and materials adopted in this study replicate those described above in Cutler, Penrod and Dexter (1989). In this study, however, a group of experienced jurors are added to the original group of student participant-jurors.

Results

The ratings of eligible jurors regarding the strength of the prosecution's case, the strength of the defence case and verdict were significantly mediated by the witnessing and identification conditions in the crime scenario. This Sensitivity was, however, counterbalanced somewhat by the fact that eyewitness confidence also significantly moderated judgements relating to witness credibility, the strength of prosecution case and the verdict. Consequently, the authors concluded:

> *That witnessing and identification conditions (WIC) and witness confidence produced comparable main effects on juror decision making further supports the contention that jurors are not sensitive to factors complicating eyewitness evidence* (p. 222).

The introduction of expert testimony had “trivial main effects” on measures of eyewitness credibility, the strength of the prosecution case, and the verdict, leading authors to suggest that the “evidence for the skepticism hypothesis was strikingly small”(p. 222). However, one significant main effect of expert testimony was found, such that the jurors rated the strength of the defence case more highly when the expert was present than when they were absent. Thus, it appears that the expert did substantially strengthen the defence’s case, although not by reducing the credibility of the eyewitness. Moreover, the introduction of expert evidence was found to reduce juror reliance on eyewitness confidence when evaluating eyewitness credibility and the strength of the defence case. This was in addition to significant increases in participant-juror Sensitivity to witnessing and identification conditions on measures of strength of the prosecution and defence cases.

Strengths

The key strength of this study can be found in its attempts to both replicate and validate previous results for a community sample, thus adding to the ecological validity of the research and the generalisability of the results.

Limitations

One possible limitation of this study lies in the conclusions drawn from the reported results. Specifically, the authors may have underestimated the ability and knowledge of jurors in the control condition. The authors suggest that the absence of Sensitivity to eyewitness confidence nullifies the observed Sensitivity to WIC;

> *That witnessing and identification conditions (WIC) and witness confidence produced comparable main effects on juror decision making further supports the contention that jurors are not sensitive to factors complicating eyewitness evidence* (p. 222).

They go on to suggest:

> *The results provide justification for expert psychological testimony in eyewitness cases. Without such testimony, jurors appear unknowledgeable of eyewitness problems* (p. 223).

Based on the data, this may be a somewhat overly enthusiastic recommendation for the expert or an overly harsh indictment of juror knowledge. While expert testimony did induce significant Sensitivity to eyewitness confidence (i.e., by reducing the weight given to it), it did not induce significant Sensitivity to WIC; rather expert testimony significantly *improved* pre-existing levels of Sensitivity. Thus, jurors were not “unknowledgeable”.

10. Cutler, Dexter & Penrod (1990)

In this study experimenters compared court-appointed expert advice with the standard instruction issued by judges in U.S. courts in cases where eyewitness evidence must be considered (*U.S. v Telfaire*). This instruction asks participant-jurors to consider the eyewitness’s capacity and opportunity to observe the defendant, the independence of the identification, and the credibility of the witness by appraising his or her testimony. Aside from this manipulation of instruction type, this study employed the same eyewitnessing materials and eyewitness evidence as described in previous studies by these authors (see section sub-headings 8 and 9).

Results

This study found no evidence to suggest that jurors who had not heard expert evidence were either knowledgeable about which factors to consider in a witnessing scenario, or aware of the questionable predictive validity of eyewitness expressions of confidence. This situation was not improved by expert evidence which was seen to result in participant-juror Skepticism on three different measures, such that guilty verdicts, perceptions of the accuracy of eyewitnesses in general, and the strength of the prosecution’s case were all decreased as a result of expert evidence. Moreover, no significant evidence of an interaction was found, indicating that expert evidence failed to induce Sensitivity to either witnessing and identification conditions, or witness confidence. The *Telfaire* instruction had no significant impact on the participant-jurors’ responses on any dependent measures. The authors subsequently concluded that the *Telfaire* instruction was “completely ineffective at sensitizing jurors to eyewitness evidence…[and] did not reliably produce skepticism toward the identification” (p. 1205), while the “court-appointed expert produced substantial skepticism and no sensitization”.

Strengths

This study is the first and only attempt to directly compare the testimony of an eyewitness expert with a pattern judicial instruction. This is a very important comparison to make because in practice judicial instruction is often provided instead of expert evidence, while psychologists hold that the expert's evidence is superior to the judicial instruction (this assertion will be discussed further in Chapter 5). Thus, this is the only study which provides empirical evidence which can speak to the validity of these practices and preferences.

Limitations

This study is limited by the fact that there is no means by which to interpret the validity of the Skepticism effect observed amongst participant-jurors in the expert evidence condition. Thus, it is difficult to make value comparisons between the expert evidence and the judge's instruction. There is an argument in favour of juror Skepticism in instances where jurors can be shown to give too much credence to witness testimony, yet in the event that this can not be established it is very difficult to ascertain if the tendency to disbelieve eyewitnesses is a desirable outcome or not. This was a point which these same authors had made a year earlier:

> *Without further research on jurors' overall belief levels pertaining to identification accuracy, few conclusions can be reached about the desirability of a skepticism effect* (p. 314).

Given this, we cannot know if the judicial instruction or the expert evidence resulted in better outcomes, or indeed if the expert testimony had a positive impact at all. The most that can be said is that the judicial instruction changed little and the expert evidence encouraged a harsher decision criterion amongst participant-jurors.

11. Blonstein & Geiselman (1990)

Blonstein and Geiselman (1990) explored the impact of different types of eyewitness expert evidence on juror decision-making by varying the quality of the conditions surrounding the eyewitness's identification (good vs. poor), and manipulating the content of the expert evidence to either support or discredit the eyewitness's evidence. Participant-jurors evaluated the eyewitness and the case both before and after they read the expert evidence, thereby allowing a within-subjects evaluation of the impact of expert evidence.

In this written trial, the “good” witnessing and identification conditions differed from the “poor” on five variables. In the good condition: a) the weather at the time of the event was clear and sunny; b) the temperature was 80°F; c) the perpetrator was viewed from a distance of 20 feet; and d) there were no obstructions to the eyewitness’s view. In the poor condition: a) it was foggy at the time of the event; b) it was night time; c) 40ºF; d) the eyewitness saw the perpetrator from a distance of 100 feet; and e) the view was partially obscured. Little information is provided regarding the precise content of the expert evidence, however, it is known that when the expert provided supportive evidence they stated that the eyewitness had given highly confident answers and that they thought the eyewitness was credible. In the unsupportive condition, the expert concluded that the eyewitness was unsure of their recollection and had provided inconsistent testimony, and therefore was not very credible.

Results

Participant-jurors showed significant Sensitivity to the witnessing conditions in the absence of expert testimony. Supportive expert testimony was seen, however, to significantly increase participant-jurors perceptions of the credibility of the eyewitness, while unsupportive expert evidence decreased participant-juror perceptions of credibility.

Strengths

The study was the first to employ a within-subjects analysis of expert effects, thus these data control for individual differences in participant-juror ratings of credibility. This removes any noise in the dataset which might result from having unevenly distributed raters across conditions with a tendency to assign either “high” or “low” values.

Limitations

This study did not statistically evaluate the interaction between the presence and absence of expert testimony and the quality of the witnessing and identification conditions. Thus, this study provides no empirically validated data regarding how the different types of expert testimony influenced participant-juror Sensitivity to the conditions described. However, an inspection of means suggests that no significant interaction was likely, and that the expert testimony had a simple additive effect on pre-existing participant-juror sensitivity to witnessing conditions. Put another way, unsupportive expert evidence decreased credibility irrespective of WIC, thereby resulting in Skepticism, while supportive expert testimony increased credibility ratings irrespective of WIC, thereby resulting in juror overbelief. In keeping with this interpretation, the authors concluded:

> *…that, given the particular case presented in the experiment, conviction rates would go down if unsupportive expert testimony is presented* (p.18).

Lindsay (1994)

Lindsay (1994) reports three experiments which systematically vary the presence and absence of expert testimony, but provides very little information regarding each. What follows is the most thorough account possible given the information reported.

12. Experiment 2

In this study all participant-jurors were provided with video testimony in which an eyewitness described the identification he or she made. In addition to this, participant-jurors heard the testimony of the "experimenter" who ran the lineup procedure for this eyewitness. The experimenter stated that they had conducted a six-person lineup and provided the (un-biased) instructions for the procedure. Although it is stated that the eyewitness used in this study had actually viewed a staged crime and subsequently made an identification, the accuracy of that identification plays no role in the analyses subsequently reported. Instead, participant-jurors are shown one of two lineups purportedly seen by the eyewitness, one version of the lineup is composed of similar foils and is therefore considered to be "fair", while the other version of the lineup includes only one member who matches the description of the suspect and is therefore "biased". Expert evidence was also manipulated. The expert provided either: 1) a general discussion of the limitations of eyewitness reliability; 2) specific evidence highlighting the effects of high-versus low-similarity lineup foils: or 3) was absent.

Results

The results of this study are difficult to interpret. Not only did participant-jurors rate the biased lineup as significantly fairer than the un-biased lineup, but they were also more likely to find the defendant guilty in the biased lineup condition (which I refer to as "Counter-Sensitivity"). This unexpected pattern of verdicts was maintained in response to general expert evidence; however, a decline in the belief of eyewitnesses from both biased and fair conditions was observed when expert testimony was present. Surprisingly, the inclusion of specific expert evidence only served to reduce guilty verdicts in the fair condition, while levels of belief in the biased condition remained unchanged, thus the expert exacerbated the original counter-intuitive response of jurors to unbiased lineups, but did nothing to change perceptions of biased lineups. The significance of these effects and indeed pair-wise comparisons between expert conditions are not reported in the original chapter, making interpretation of the results very difficult.

13. Experiment 5

In this study expert testimony was orthogonally varied along with instructional bias and foil bias. The specific content of the expert evidence was not described.

Results

Participant-jurors appeared to be more likely to convict where the foil or the instructions were biased, however, no statistical analysis was reported. More worryingly, expert testimony significantly interacted with identification conditions such that those participant-jurors who heard that the lineup was biased in terms of both its construction and execution were significantly more likely to find the defendant guilty than jurors who presented with any other combination of lineup factors. The author offers no explanation for this perplexing and counter-intuitive finding.

Strengths

This study is the first to independently vary factors of the witnessing and identification condition such that a participant-juror may be faced with a combination of factors indicating both the likely accuracy and inaccuracy of the identification. That is, some participants heard from witnesses who had seen a lineup which included a good selection of foils but was administered with biased instructions. Thus, these jurors were required to weigh and combine these factors, in the same manner as real jurors, in order to determine the overall likelihood of an accurate identification. This feature of the design may in part account for some of the strange results observed, as it appears participant-jurors were not provided with any guidance regarding how to weigh competing predictors of accuracy. As such, these jurors may have been defeated by the difficulty of the task facing them. However, this does not account for the fact that, where all signs were consistent with an inaccurate identification, jurors chose to believe it *more often* than they did in conditions where all indicators were consistent with an accurate identification. That result remains perplexing.

14. Experiment 6

In this study jurors heard the testimony of: a) the court-appointed; b) partisan (either prosecution or defence); or c) competing experts, all of whom deliberately invaded the province of the jury by making conclusions regarding the credibility of the eyewitness. In cases where the lineup presented to the eyewitness was biased, the expert concluded that the eyewitness was not a credible source of information and should be ignored. When the lineup was fair, the expert indicated that under the conditions described the identification could be highly reliable and should be considered as strong evidence for the guilt of the accused.

Results

As in Lindsays Experiment 5, analysis revealed that participant-jurors were significantly more likely to find the defendant guilty in the biased lineup condition than the fair lineup condition. Furthermore, *all types* of expert testimony were seen to cause a significant decrease in guilty verdicts (i.e., the eyewitness was believed less). The author concluded that:

> *Expert testimony did not help (and possibly hurt). This pattern of results raises serious concerns about the ability of juries to appropriately evaluate testimony of eyewitness identifications* (p. 378).

15. Devenport, Stinson, Cutler & Kravitz (2002)

As in Lindsay (1994, Experiment 5) this study also orthogonally varied expert testimony along with a range of identification factors. Expert testimony was either present or absent and described basic memorial processes, some general factors known to influence likely identification accuracy and more specific information regarding lineup foils, instructional and presentation biases. These lineup factors were manipulated such that participant-jurors (both students and real jurors) saw either a biased or unbiased lineup, with biased or unbiased instructions and biased or unbiased presentation (i.e., sequential or simultaneous procedures).

Results

Participant-jurors who received no expert testimony were found to be Sensitive to foil, instruction and presentation bias, such that suggestiveness, overall suggestiveness, fairness, culpability and verdict were all rated lower when biased procedures, of one or other of the three types, were used. Consistent with a Skepticism effect, the introduction of expert testimony caused participant-jurors to rate the lineup instruction and presentation style as significantly more suggestive than participants in the control condition. Yet, those jurors who heard expert evidence rated the overall lineup as being significantly *less* suggestive than those jurors who did not hear expert testimony. Even so, expert evidence was found to significantly increase juror Sensitivity to instruction bias amongst jurors, such that more biased lineup instructions were rated higher on all but one measure. This effect was, however, moderated by a significant three-way interaction with foil bias which meant that this Sensitivity to instruction bias was ameliorated by the presence of foil bias. The expert did not significantly improve participant-juror Sensitivity to either foil or presentation biases. The authors concluded that:

> *…current research findings suggest that expert testimony does enhance juror sensitivity to factors influencing witnessing and identification conditions without overly increasing juror scepticism of the witness' identification* (p. 1052).

Strengths

This study builds upon Lindsay's (1994, Experiment 5) attempt to orthogonally vary the factors present in an identification. Specifically, this design requires that jurors weigh and combine the likely influence of three different biases (foil, instruction and presentation) in order to determine culpability of the defendant and reach a verdict (amongst other things). This is a considerable improvement on the designs adopted in other studies which do not combine factors in this manner. Although this will be illustrated in more detail in coming chapters, it is important to note that participant-jurors in this task were not asked to evaluate either "good" or "bad" witnessing scenarios, but were given a combination of the two such that some participants were shown biased foils, but were told the instructions and presentation style were fair, while others were shown biased foils presented in a biased manner, but were given unbiased instructions. It is clear that participant-jurors in this experiment were presented with a very difficult task, while participant-jurors in other studies who are required only to differentiate between "good" and "poor" scenarios are presented with a far simpler, and perhaps, qualitatively different task. More precisely, since it is highly unlikely that any real eyewitnessing scenario brought before jurors will include factors which suggest *only* that the identification is accurate or *only* that it is inaccurate, this experiment significantly improves the ecological validity of the task presented to participant-jurors. It requires them, instead, to consider various factors, some of which are consistent with accuracy and others that are inconsistent with accuracy. This change is not simply a minor design improvement, but a step toward the valid assessment of expert effects on juror decision-making (see Chapter 5 for further discussion of this issue).

Limitations

Though the general conclusion reached by the authors (and reported above) seems appropriately qualified given the data, other statements by the authors lack these necessary qualifications. For example:

> *For expert testimony to be an effective safeguard, it must serve the purpose of educating the jury. The results of this study suggest that expert testimony can educate jurors and assist them with the decision-making process. Thus it is important that experts are allowed to testify regarding factors that influence eyewitness identification accuracy in order to enhance juror knowledge and their verdict decisions"* (p. 1053).

This conclusion should be qualified. In particular, the statement that "the expert can educate jurors" is only true for some of the biases introduced, and any effect of education regarding one factor is eradicated by the presence of another. Moreover, this conclusion makes no reference to jurors' baseline performance, which was quite strong, or the observed Skepticism effect and the worrying reduction in perceptions of the suggestiveness of the overall lineup. Specifically, the authors adopted a definition of Skepticism as a "significant main effect of expert testimony" (p. 1047), yet despite having reported three such main effects, the authors concluded that:

> *…these data provide little evidence that expert testimony increases juror skepticism of an eyewitness identification* (p. 1051).

The authors rationalised this inference on the basis that no significant main effect of expert testimony was found with regard to juror verdicts or culpability ratings, and go on to suggest that:

> *…the absence of a main effect of expert testimony on verdict and defendant culpability supports the absence of a skepticism effect* (p. 1051).

Although, it is appropriate to reason that the direct measures of verdict and culpability may be more informative, more interesting, or more meaningful than the indirect measures which did demonstrate Skepticism, it is unwise to discount these indirect measures altogether. Two main effects consistent with Skepticism, and one main effect in the opposite direction were observed. These findings are not explained

Geiselman, Putman, Korte, Shahriary, Jachimowicz & Irzhevsky (2002)

16. Experiment 1

In the first experiment reported, three levels of expert testimony (general, specific or none) and two levels of eyewitnessing conditions (“good” and “poor”) were systematically varied. The testimony of a court-appointed expert either generally addressed human memorial processes and factors likely to influence eyewitness identification accuracy, or specifically addressed a number of the factors which varied between the good and poor witnessing conditions (i.e., lighting, time of day, distance of view, duration of view, eyewitness composure, race of witness compared to perpetrator, presence of disguise, suggestiveness of lineup instruction, type of lineup, delay before identification and the eyewitness’s visual acuity). Given the detail provided it appears that in the poor condition the expert spoke about five of these predictors: duration of view, eyewitness composure, the race of the witness compared to perpetrator, the suggestiveness of lineup instruction and the type of lineup. It is unclear if the same cues or number of cues were referred to by the expert in the good condition.

Results

Participant-jurors showed significant Sensitivity to the quality of the witnessing and identification conditions described without expert testimony. This Sensitivity was significantly enhanced by specific eyewitness expert testimony although not by the general expert evidence. The authors concluded that this result was encouraging, both with regard to juror baseline performance and the expert effect.

Limitations

Compared with the design innovation displayed in Devenport et al. (2002) where that study increased the ecological validity of the evaluation task by requiring participants to weigh and combine probabilistic predictors, this study appears to have substantially decreased ecological task validity through the introduction of specific expert testimony. This testimony tells jurors how the quality of the witnessing and identification conditions have been varied by the experimenters from “good” to “poor”, instructing them with regard to *five* predictors they can use to differentiate between the two. Thus, to show Sensitivity, all jurors need to do is correctly identify *one of these five features* and respond accordingly. This is a very simple matching task and it is hardly surprising that expert testimony of this kind significantly improved juror Sensitivity in response to the eyewitnessing stimuli provided.

17. Experiment 2

Experiment 2 replicated Experiment 1 with the addition of adversarial closing arguments which specifically linked the five factors raised in the expert’s evidence to the facts of the case. Here, expert evidence was found to result in participant-juror Skepticism such that expert testimony significantly reduced guilty verdicts irrespective of the quality of the witnessing and identification conditions.

18. Devenport & Cutler (2004)

Devenport and Cutler (2004) utilised the same trial materials as Devenport et al. (2002) to investigate the impact of different types of expert evidence on juror decision-making. In this case expert testimony was either: 1) provided only by an expert for the defence; 2) by both opposing prosecution and defence experts; or 3) was absent. The defence expert testified to general memorial processes as well as various factors known to influence identification accuracy. Where foil or instruction biases were present, the defence expert also explained their likely impact on eyewitness identifications. The prosecution expert in the “opposing” expert’s condition criticised the reliance on crime simulations, the use of college students in psychological research and the low ecological validity of eyewitness research.

Results

Participant-jurors who did not hear expert evidence displayed significant Sensitivity to both instruction bias and foil bias as measured by ratings of suggestiveness. These jurors also showed significant Sensitivity to the impact of instruction bias on likely identification accuracy. Further statistical analysis revealed no evidence of either a main effect of expert evidence, or an interaction between expert evidence and the quality of the identification conditions. Although opposing expert evidence caused participant-jurors to regard the defence expert's testimony as significantly less credible, influential, and useful, believing it to have a more negative effect on the credibility of psychology in general than did the defence-only expert, the authors note the absence of Skepticism and Sensitivity effects in response to expert evidence.

Strengths

This study has the same strengths as the work presented by Devenport and colleagues (Devenport et al.) with the addition of an investigation of different types of expert evidence, either opposed or unopposed.

Leippe, Eisenstadt, Rauch & Seib (2004)

19. Experiment 1

This investigation by Leippe and colleagues (2004) aimed to explore the impact of expert evidence in the context of relatively strong and weak prosecution cases. Though case facts varied according to case strength (i.e., prevalence of the defendant's blood type in the population, the quality of the alibi, and other physical and circumstantial evidence), in both cases a single eyewitness viewed a robbery and murder from their second-floor window at night. Participant-jurors were either provided no expert evidence or one of four different types of expert testimony which varied according to the timing of the expert's presentation (either before or after the eyewitness's testimony), and the presence of a judicial reminder (present or absent) regarding the expert's evidence. The expert provided information regarding his or her educational and occupational background, and then went on to address the reliability of eyewitness evidence, general theories of memory, limitations of experimental methodologies, the weak link between identification confidence and accuracy, the relationship between stress and memory, and the importance of unbiased identification procedures.

Results

Without the assistance of expert evidence participant-jurors showed significant Sensitivity to the strength of the prosecutions case as measured by "guilt-certainty" ratings. However, the introduction of expert evidence after the testimony of the eyewitness, with an additional reminder caused participant-jurors to become Skeptical of both the defendant's guilt and the eyewitness's believability. That is, post-evidence/reminder expert testimony reduced jurors' ratings of guilt and eyewitness believability, irrespective of case strength. The authors termed this effect the "exonerating impact" of expert evidence and went on to express their concern, concluding that the results:

> *…suggest that general educational testimony about eyewitness psychology may work in the defense's favour (i.e., prodefense asymmetry), even when it perhaps should not (i.e., an otherwise strong prosecution case)* (p. 535).

Strengths

This study focuses on the impact of expert evidence in the trial context. It does this by holding witnessing and identification conditions constant, and changing the strength of the case around these facts. Thus, this experiment provides the first indication of an expert's impact on the prosecution's case. This is obviously a very interesting area of investigation, and worthy of research attention, particularly given the finding that arguably the most ecologically valid form of the expert's evidence – post-evidence/reminder – appears to have an exonerating impact, even in instances where the prosecution's case is objectively strong.

Limitations

As this study did not vary any elements of the witnessing and identification conditions, it cannot speak to the issue of expert effects on juror Sensitivity to Expert Opinions.

20. Experiment 2

In a modified version of Experiment 1, researchers compared the effect of post-eyewitness expert evidence and a judicial reminder with a no expert control condition in the context of a new stronger case.

Results

As in Experiment 1, those participant-jurors who heard the expert rated the eyewitness as significantly less believable than in the control condition. The authors concluded that:

> *… it appears that expert testimony may have an impact under a limited set of conditions, and when it does have an impact, that impact serves to reduce the advantage to the prosecution of having eyewitness evidence* (p. 536).

21. Geiselman and Mendez (2005)

Using the eyewitnessing scenarios, specific expert testimony and attorney's closing arguments from Geiselman et al. (2002), Geiselman and Mendez (2005) investigated the effect of different trial scenarios on participant-juror ability to discriminate between "good" and "poor" witnessing scenarios. Eyewitness testimony was presented in all four versions of the trial plus: 1) attorney closing arguments; 2) judges instruction and closing arguments; 3) expert testimony plus judge's instruction and closing arguments; and 4) eyewitness-only baseline. The judicial instruction referred to stated "it is now time for the attorneys for the prosecution and the defence to summarize the evidence in this case as they see it" (p. 7).

Results

Participant-jurors in the baseline condition showed significant Sensitivity to the witnessing and identification conditions described in the scenario. This Sensitivity was not improved by any of the trial scenarios, however Sensitivity did significantly increase when participant-jurors heard the expert, the judge and closing arguments compared with the other trial conditions. The authors concluded that:

> *Jurors showed essentially no discrimination between good and poor eyewitnessing conditions when given the closing arguments and judge's instructions without expert testimony* (p. 10).

Strengths

This study shares some of the methodological limitations present in Experiment 16 by Geiselman et al. (2002), but does show that expert testimony can induce juror Sensitivity in cases where attorney's closing arguments cannot. This is despite the fact that both of these sources provided jurors with five cues (duration of view, eyewitness composure, race of witness compared to perpetrator, suggestiveness of lineup instruction and type of lineup) to discriminate between good and poor witnessing conditions. It is important not to understate the role of judicial instruction in this Sensitivity effect, however, as expert testimony plus only attorney closing arguments in Geiselman (2002) resulted in significant Skepticism. Thus, the instruction to consider closing arguments as *opinions* rather than *evidence* appears to have preserved the expert Sensitivity effect in the face of Skepticism-inducing closing arguments.

In summary the evidence relating to Sensitivity to Expert Opinions and Skepticism is very mixed (see Table 3.1 on pp. 70-73 for more detail). When analysed at the level of reported effects, most fall under the category of Skepticism (n = 18), followed by Sensitivity *induced* by the expert (n = 8), then Sensitivity *increased* by the expert (n = 4), Counter-Sensitivity (n = 3) and finally, Overbelief (n = 2). By this count Skepticism and Sensitivity effects occur with almost equal frequency (18 vs. 12). However, this only tells part of the story, as Skepticism moderated by Sensitivity is generally considered to be a more desirable outcome than Skepticism alone; eight studies found the least desirable outcome (Skepticism in the absence of a Sensitivity effect; Experiments 6, 7, 10, 11, 14, 17, 19 and 20); three found the generally desired outcome (Sensitivity (induced or increased) in the absence of Skepticism; Experiments 8, 16 and 21); two studies found both of these effect types (Experiments 9 and 15), while only one study found neither Sensitivity or Skepticism (Experiment 18). Thus, this analysis, albeit somewhat coarse, indicates that the least desirable outcome of expert testimony is observed more than twice as often as Sensitivity to Expert Evidence without an associated Skepticism effect.

Direct Measures

Sensitivity to Eyewitness Accuracy (SEA)

A third kind of experimental design utilises the testimony of *real* eyewitnesses who have actually viewed an event (usually a staged or recorded incident) to directly measure the effect of expert evidence on participant-juror Sensitivity to Eyewitness Accuracy. Unlike SEO studies which must *infer* likely identification accuracy from juror knowledge of, and response to, the correlates of accuracy, SEA studies measure the construct directly. As a result, researchers employing this methodology may or may not orthogonally vary witnessing and identification conditions, since the accuracy of the actual eyewitness provides the “verifiable criterion” (Wells, 1986, p. 90) against which juror judgments can be evaluated. Only studies employing this design permit the researcher to compare the objective accuracy of the eyewitness’s identification with the jurors' evaluation of the accuracy of that identification. That is, using this method experimenters can determine whether or not jurors believe those witnesses who made objectively accurate identifications and disbelieve those eyewitnesses who made objectively inaccurate identifications. Analysis at this level, unlike the designs considered above, permit tests of hypotheses which are in keeping with the objectives espoused in the literature and the expectations of the legal system: that the expert should assist jurors to reach an accurate resolution of a dispute in issue. These studies evaluate the extent to which expert testimony can protect the innocent accused in eyewitness cases, without compromising the integrity of the evidence provided by those accurate eyewitnesses. Moreover, when these studies are constructed to include orthogonally varied witnessing and identification conditions, this methodology can also be used to evaluate the educational impact of the expert in the same way as SEO studies described above. Therefore, this experimental design allows us to discriminate Sensitivity to Eyewitness Accuracy from Skepticism, while also permitting evaluations of the expert's educational impact. Finally, this design does not assume, as the indirect measures do, that SEO and SEA are dependent constructs, rather it provides for participant-jurors to be *either* Sensitive to Expert Opinions, *or* Sensititive to Eyewitness Accuracy, *or* both of these, or neither of these. Thus, this methodology incorporates a means by which to empirically validate the assumption that one can infer Sensitivity to Eyewitness Accuracy from indirect measures of the construct (this issue is discussed in detail in Chapter 5). To date, this SEA method has been employed in just two studies.

22. Wells, Lindsay and Tousignant (1980)

In this study, experimenters investigated the impact of court-appointed expert advice (which preceded the eyewitness's testimony) on participant-juror evaluations of witnessing and identification conditions *and* eyewitness identification accuracy.

The eyewitness testimony evaluated by the participant-jurors in Wells et al. (1980) represented a sample from the 108 eyewitnesses who made identifications after watching a staged crime scenario[7] . These eyewitnesses attempted to make identifications from a target-present lineup, after seeing a crime which occurred at one of three levels of situational accuracy. In the low-situational-accuracy condition the

[7] The sample of eyewitnesses used in this study were taken from Lindsay, Wells and Rumpel (1981).

eyewitness had 12 seconds to view the perpetrator of a theft. The perpetrator wore a hat covering all hair and part of their ears, and said, “Hey, is this your calculator?” Actual eyewitness identification accuracy in this condition was 33%. In the moderate-situational-accuracy condition, the perpetrator’s behaviour was the same as in the condition described above; however, the cap was worn higher on the head so as to reveal the colour and texture of the perpetrator’s hair. Identification accuracy in this condition was 50%. In the high-situational-accuracy condition the perpetrator was present for 20 seconds, wore no hat and asked, “Has the experimenter told us what to do yet?” Here, identification accuracy was 74%. These differences in participant-witness identification accuracy were significant. Moreover, a statistically significant difference was observed between the confidence expressed by those accurate and inaccurate eyewitnesses who consented to complete a subsequent cross-examination style interview. A subset of 48 of these interviews (16 from each level of situational-accuracy, half accurate and half inaccurate) were shown to participant-jurors in the Wells, Lindsay and Tousignant study (1980).

In this study each participant-juror viewed the testimony of four eyewitnesses from the same situational-accuracy condition, making evaluations after each one. Half of these participants also heard the testimony of a court-appointed expert who provided general testimony. This testimony began with a statement of the expert’s credentials, then moved on to discuss the potential unreliability of eyewitness identifications and the fact that an eyewitness’s expression of confidence may bear little or no relationship to the accuracy of their identification.

Results

Each participant-juror’s performance was given a value of 0%, 25%, 50%, 75% or 100% to indicate their accuracy score over the four eyewitness evaluation trials. A 2 (eyewitness accuracy) x 2 (expert testimony) x 3 (witnessing condition) ANOVA analysing these values provided evidence indicating that participant-jurors were sensitive to the relative difference between levels of situational-accuracy. The authors suggested:

> *…this effect is due to the general tendency for subject-jurors to discount the testimony of the eyewitnesses who had poor or moderate witnessing conditions relative to those who had good witnessing conditions* (p. 282)[8].

There was, however, no evidence to suggest that participant-jurors were sensitive to the actual accuracy of the eyewitness's identification. That is, participant-jurors were not more likely to believe those eyewitnesses who made accurate identifications than they were to disbelieve those eyewitnesses who made inaccurate identifications.

Consistent with a Skepticism effect, the introduction of expert testimony significantly reduced participant-jurors' tendency to consider the eyewitness to be accurate. However, there was no evidence of any interaction between witnessing and identification conditions and expert evidence, indicating that the expert did not influence or improve participant-juror Sensitivity to the differences between witnessing conditions. As a result, the authors concluded that:

> *If we assume that it is desirable to obtain juror-belief rates that closely correspond to probable witness accuracy rates across the witnessing conditions, then we must admit that expert advice did not produce an improvement in that regard* (p. 283).

Subsequent analysis of the correlations between eyewitness confidence and the proportion of belief decisions made by participant-jurors did indicate that expert testimony had reduced the association between confidence and belief from a significant relationship to one of a non-significant magnitude. This result is consistent with participant-juror Sensitivity to the Expert Opinion regarding the confidence-accuracy relationship.

Finally, no significant interaction was identified between expert evidence and witness accuracy, indicating that expert testimony did not assist participant-jurors to discriminate between accurate and inaccurate eyewitnesses. Overall, the authors conclude:

> *We are encouraged by the results of our first attempt to experimentally assess the influence of expert advice on subject-jurors' performance in judging the validity of eyewitness testimony. At the very least our results show that people are able to use such advice to change their decision criteria even when the advice runs counter to their intuitions (as with the confidence-accuracy issue)* (p. 285).

---

[8] No post-hoc tests are reported to substantiate this interpretation of the witnessing condition main effect.

Strengths

The main strength of this study lies in its use of real eyewitness testimony, thereby providing participant-jurors with an ecologically appropriate eyewitness evaluation task. Moreover, the methodology used here permits researchers to investigate the extent to which participant-jurors can be educated by expert testimony, as well as the expert's effect on their ability to discriminate between accurate and inaccurate eyewitness evidence. Thus, this study represents the first investigation of the impact of expert testimony on participant-juror Sensitivity to Eyewitness Accuracy and thereby of the juror's ability to reach an accurate resolution to a dispute in issue.

Limitations

Despite the obvious strength of this study, this experiment was limited with regard to several minor procedural elements that could readily be modified to further increase the ecological validity of this research. Firstly, the participant-jurors never heard the direct evidence of the participant-witnesses, as only a cross-examination was conducted. This may have undermined participant-juror perceptions of the eyewitness from the beginning, as the latter was never given an opportunity to state his or her version of events without being challenged.

Secondly, Wells and colleagues adopted a philosophical approach to the structure of their trial. Specifically, the experimenters opted to have a court-appointed expert provide participant-jurors with expert advice *before* the presentation of the eyewitness's testimony even though they acknowledged that this would rarely happen in real life. Thus, the data gathered in this experiment may speak only to expert effects in this somewhat artificial context, rather than those which more closely approximate the norm.

Thirdly, all participant-jurors made multiple evaluations of eyewitnesses who had made their identifications from the same level of situational-accuracy. The first part of the problem here relates to the fact that it is likely that participant-jurors made relative comparisons between eyewitnesses when making their judgements. That is, participant-jurors were provided an opportunity to cross-validate the version of the crime offered by latter witnesses with earlier versions of the same crime. This may have served to artificially improve or impair juror discrimination accuracy over trials. Moreover, since the accuracy of juror evaluations was not reported on a trial-by-trial basis, little is known about how participant-jurors performed on their first evaluation, without the benefit of collateral eyewitness testimony, or how that performance changed over trials. A second associated consequence of the repeated measures design relates to the amount of information available for analysis. Participant-jurors were each given a score reflecting their percentage of accurate evaluations (0% to 100%). These percentage ratings were then averaged across all participants in a particular condition to provide a mean percentage correct value for the condition. These means were then compared using an ANOVA. Thus, as the number of participants in each group is not reported, it is not possible to calculate the accuracy rate within each condition.

The final limitation of this study relates to the identification procedure itself. Specifically, participant-witnesses in Lindsay et al. (1981), used by Wells et al. (1980), all made their identifications from a target-present lineup. As such, all those participants who made incorrect identifications: a) failed to identify the perpetrator who was present in the lineup; and b) chose a known innocent foil. In reality, those eyewitnesses most likely to be brought before a jury will have either correctly selected the perpetrator from a target-present lineup, or have incorrectly chosen a police suspect from a target-absent lineup. That is, it is unlikely that the prosecution would call a witness who failed to make an identification or who chose a known foil. Thus, those eyewitnesses who appear in court after unknowingly having made incorrect identification, will not have *passed over* the perpetrator in the lineup to select an innocent person. In such cases their identification would immediately be identified as inaccurate because all other lineup members are (or should be) known to be innocent. It is possible, therefore, that there are qualitative differences between the inaccurate eyewitnesses used by Wells et al. who *passed over* the perpetrator (and went on to pick an innocent foil), and real world inaccurate eyewitnesses who didn't *exclude* the perpetrator, may have altered participant-juror ability to discriminate between accurate and inaccurate eyewitnesses in some non-trivial way. Thus, even after this impressive investigation, many questions still remain about the effect of expert evidence on participant-juror Sensitivity to Eyewitness Accuracy.

23. Wells and Wright (1983)

Similar to Wells et al. (1980), Wells and Wright (1983, cited in Wells, 1986) evaluated the effect of expert evidence using real eyewitness testimony from one of three witnessing conditions.

In this study the eyewitnesses from the poor condition saw the perpetrator for five seconds at a distance of 10-12 meters; the perpetrator glanced at the eyewitness once, but covered their face with a coat for most of the time. In the moderate condition the eyewitness viewed the perpetrator for 15 seconds from a distance of 6-7 meters; the perpetrator looked at the eyewitness twice and covered their face with a coat for only part of the interaction. In the good condition the perpetrator was in view for 25 seconds at a distance of 3-4 meters, the perpetrator looked directly at the witness five times and their face was never obscured. Unlike the Lindsay et al. study (1981) however, the identification accuracy rates for these conditions were not found to differ from each other significantly.

Participant-jurors viewed the testimony of one of 48 eyewitnesses (8 accurate and 8 inaccurate from each of the witnessing conditions), who were cross-examined in the same manner as described in Lindsay et al. (1981). The expert in this study encouraged jurors to pay attention to the witnessing conditions such as the eyewitness's opportunity and duration of view of the perpetrator. The expert warned participant-jurors against assuming that a good memory for trivial details increased the likelihood of an accurate identification, informed them that the confidence or certainty of an eyewitness's identification testimony bears little or no relationship to eyewitness accuracy, and highlighted the potential for unreliable eyewitness evidence.

Results

Without the aid of expert advice, participant-jurors were not Sensitive to differences in the witnessing and identification conditions, or to eyewitness accuracy. With the addition of expert testimony, however, jurors showed significant Sensitivity to both of these things. Moreover, there was no evidence to suggest that expert evidence caused participant-jurors to become more Skeptical of eyewitnesses in any way.

Strengths

This study addressed two of the limitations highlighted in Wells et al. (1980). Specifically, participant-witnesses made their identifications from target-present and target-absent lineups, and participant-jurors evaluated only one eyewitness rather than four. The differences between the findings of these two studies may in part be attributed to these methodological changes.

Limitations

As in the previous study, the testimony of the eyewitness expert preceded the testimony of the eyewitness, despite the fact this scenario is unlikely to be replicated in real courtrooms. Moreover, the expert in this study was not cross examined, and so his or her testimony remained largely unchallenged in the context of the trial. Either of these variations from general courtroom procedures may have influenced the observed results in a non-trivial manner, and therefore ought to be considered in future research. Finally, this Experiment is unlikely to have undergone peer review, as the experiment itself was never published in its own right. Instead, it was referred to in detail in an article by Wells in 1986. Thus, it may be that this study was limited in other ways which were not immediately obvious.

Overall then, only two studies have ever investigated the effect of expert evidence on juror ability to reach an accurate resolution to disputed identification evidence, and only one of those studies has been subjected to peer-review (Wells et al., 1980). Moreover, the study that was vetted by the research community only presents the group accuracy average, not accuracy rates; thus valuable information about individual juror performance and practice effects, which may have further clarified results, is not available. Despite this, there is encouraging evidence to suggest that expert testimony *can* educate participant-jurors regarding the relative quality of witnessing and identification scenarios, and that it can improve participant-juror Sensitivity to the accuracy of an eyewitness identification (see Table 3.1 for a detailed summary). Further replication of these results is necessary to ascertain if these findings are reliable and generalisable.

Summary

Taken as a whole (see Table 3.1, pp. 69-74), three methods of evaluating eyewitness expert effects have been adopted in the literature; one investigating the expert's ability to *influence* jurors (REE), one investigating the expert's ability to *educate* jurors (SEO), and lastly, one investigating the expert's ability to *aid jurors in discriminating* between accurate and inaccurate eyewitnesses (SEA). The first method has provided clear evidence that the expert can significantly alter, that is decrease, participant-juror perceptions of eyewitness credibility and believability. This finding has also been replicated several times by research conducted using the SEO approach (n = 8), however, only three of these studies also showed evidence that the expert could educate jurors regarding the relative quality of the witnessing and identification conditions. Moreover, only three studies of the 16 adopting the SEO method found evidence that the expert could educate jurors without also inducing significant Skepticism. Finally, only two studies have employed the most useful methodology, which directly assesses the impact of the expert on the accuracy of participant-juror discriminations. Of these, only one was subject to the peer review process (Wells et al., 1980), and together they provide a very mixed impression of expert effects; one identifying Skepticism and providing evidence of an educational impact, while the other found no Skepticism, clear evidence for education *and* Sensitivity to Eyewitness Accuracy.

Table 3.1 : Summary of Observed Expert Effects

Outcome Measured (Expert Role)	**Experiment**	**Expert-Absent Response to Manipulated Variables**	**Expert Effects Observed**		
			REE	**SEO**	**SEA**
Response to Expert Evidence (Moderating Effect)	1. Hosch, Beck & McIntyre, 1980		< EW* Reliability (Sig.)		
	2. Loftus, 1980 (Experiment 2)		> EW Scrutiny (Sig.)		
	3. McKenna, Mellott & Webb, 1981		< Guilty Verdicts (untested)		
	4. Maass, Brigham & West, 1985		< EW Weight (Sig.)		
	5. Pezdek, Avila-Mora & Sperry, 2007		< Guilt Certainty (Sig.) > Suggestive Lineup (Sig.)		

Outcome Measured (Expert Role)	**Experiment**	**Expert-Absent Response to Manipulated Variables**	**Expert Effects Observed**		
			REE	**SEO**	**SEA**
Sensitivity to Expert Opinion (Educating Effect)	6. Loftus, 1980 (Experiment 1)	Counter-Sensitivity to Violence (untested)	< Guilty Verdicts (Sig.)	(untested)	
	7. Fox & Walters, 1986	> Confident EW's > Believed (Sig.)	< EW Belief (Sig.)	> Confident EW's > Believed (Sig.)	
	8. Cutler, Penrod & Dexter, 1989	> Confident EW's > Believed (Sig.) Sig. Sensitivity to Disguise, Retention Interval and Lineup Instructions	None	Sig. Sensitivity to Confidence & Weapon Focus > Sensitivity to Lineup Instructions (Sig.)	
	9. Cutler, Dexter & Penrod, 1989	> Confident EW's > Credible EW's, > Strong Pros. Case, > Guilty Verdicts (Sig.) Sig. Sensitivity to WIC on Strength of Pros. Case, Strength of Def. Case & Verdicts	> Strength Def. Case (Sig.)	Sig. Sensitivity to Confidence on Credibility & Strength Def. Case. > Sensitivity to WIC on Strength of Pros. Case & Strength of Def. Case (Sig.)	

Outcome Measured (Expert Role)	**Experiment**	**Expert-Absent Response to Manipulated Variables**	**Expert Effects Observed**		
			REE	**SEO**	**SEA**
Sensitivity to Expert Opinion (Educating Effect)	10. Cutler, Dexter & Penrod, 1990	> Confident EW's > Culpability, > General Accuracy of ID's, > Strength Pros. Case & > Credibility No Sensitivity to WIC on any Variable	Sig. Skepticism on Verdict, General Accuracy of ID's & Strength of Pros. Case. From Expert	None	
	11. Blonstein & Geiselman, 1990	Sig. Sensitivity to WIC	Sig. Overbelief from Supportive Testimony Sig. Skepticism with Unsupportive Testimony	(untested)	
	12. Lindsay, 1994 (Experiment 2)	Counter-Sensitivity to Lineup Bias Pre-deliberation & Post-deliberation (untested?)	Skepticism from General Testimony (untested?)	> Counter-Sensitivity to Biased Lineup from Specific Testimony (untested)	
	13. *Ibid.* (Experiment 5)	Counter-Sensitivity to Foil Bias & Instructional Bias (untested?)	None	> Guilty Verdicts where Foil & Instruction Biased (Sig.)	
	14. *Ibid.* (Experiment 6)	None	Sig. Skepticism from Court-Appointed, Partisan & Competing Experts	None	

Outcome Measured (Expert Role)	Experiment	Expert-Absent Response to Manipulated Variables	Expert Effects Observed		
			REE	**SEO**	**SEA**
Sensitivity to Expert Opinion (Educating Effect)	15. Devenport, Stinson, Cutler & Kravitz, 2002	Sig. Sensitivity to Foil Bias, Instruction Bias on Suggestiveness Sig. Sensitivity to Foil Bias on Overall Suggestiveness Sig. Sensitivity to Foil Bias & Presentation Bias on Fairness Sig. Sensitivity to Foil Bias on Culpability	Sig. Skepticism of Instruction & Presentation Suggestiveness Sig. Overbelief of Overall Suggestiveness	Sig. Sensitivity to Instruction Bias on Overall Suggestiveness, Fairness† Culpability† & Verdict† († where no Foil Bias)	
	16. Geisleman et al., 2002 (Exp. 1)	Sig. Senstivity to WIC	None	> Sensitivity to WIC from Specific Testimony (Sig.)	
	17. Geiselman et al., 2002 (Experiment 2)	Sig. Senstivity to WIC (same sample as above)	Sig. Skepticism from General and Specific Testimony + Closing Arguments	None	

Outcome Measured (Expert Role)	**Experiment**	**Expert-Absent Response to Manipulated Variables**	**Expert Effects Observed**		
			REE	**SEO**	**SEA**
Sensitivity to Expert Opinion (Educating Effect)	18. Devenport & Cutler, 2004	Sig. Sensitivity to Instruction Bias, Foil Bias on Suggestiveness Sig. Sensitivity to Instruction Bias on Accuracy	None	None	
	19. Leippe, Eisenstadt, Rauch & Seib, 2004 (Experiment 1)	Sig. Sensitivity to **Case Strength** on Guilt-Certainty	Sig. Skepticism of Guilt & EW Believability in Post-evidence/Reminder	None (Case Strength)	
	20. *Ibid.* (Experiment 2)	(untested – no-expert sample from above + new no-expert sample)	Sig. Skepticism of EW Believability	(untested)	
	21. Geiselman & Mendez, 2005	Sig. Sensitivity to WIC	None	Sensitivity to WIC in Expert + Closing Arguments + Judge than CA or CA + J (Sig.)	

Outcome Measured (Expert Role)	**Experiment**	**Expert-Absent Response to Manipulated Variables**	**Expert Effects Observed**		
			REE	**SEO**	**SEA**
Sensitivity to Eyewitness Accuracy (Discriminating Effect)	22. Wells, Lindsay & Tousignant, 1980	Sig. Sensitivity to WIC > Confident EW's > Believed (Sig.)	Sig. Skepticism of EW Accuracy	Sig. Sensitivity to EW Confidence	None
	23. Wells & Wright, 1983	None	None	Sig. Sensitivity to WIC	Sig. Sensitivity to EW Accuracy

Skepticism and the Eyewitness Expert

The effects of the eyewitness expert were most recently reviewed 12 years ago by Leippe (1995). At that time the author concluded Skepticism was a "near ubiquitous" effect of expert evidence (p. 941), going on to state that 10 of the 12 experiments contained in the review provided some evidence for an increase in Skepticism. The author then provided a series of justifications for this Skepticism effect, suggesting that; a) Skepticism may be the result of the ordering of expert evidence in the trial rather than the expert's testimony *per se* (p. 943); b) the magnitude of the Skepticism was reasonable since "jurors were neither overwhelmed nor unmoved by the expert" (p. 948); c) expert evidence should carry weight and therefore should influence jurors (p. 922); and finally, d) that Skepticism wasn't Skepticism at all, but instead was actually a *type* of Sensitivity; "in a way then, the heightened Skepticism caused by the expert testimony reflects a heightened sensitivity" (p. 941). Thus, the evidence of the expert was considered valuable irrespective of the observed Skepticism effects. It is worth noting at this point, that these same justifications were not equally applied to rationalise the Skepticism effects resulting from other safeguards. Instead Leippe criticised cross-examination for

> *…removing the risk of believing a mistaken identification and replace*[ing] *it with a heightened risk of discounting an accurate identification* (p. 923),

and went on to suggest that:

> *…several considerations suggest that universal reliance on a judge's instruction would be unsatisfactory. First, instructions to carefully scrutinize the eyewitness testimony may imply to some jurors that the judge wants them to discount the testimony* (p. 949).

While it is unclear how these apparent contradictions were to be reconciled, just a year later, the evidence regarding the impact of expert testimony was reported to be more mixed, with researchers now suggesting that there was evidence for both Sensitivity and Skepticism effects (Devenport & Cutler, 1997; Devenport et al., 2002; Geiselman & Mendez, 2005; Geiselman et al., 2002; Ramirez, Zemba, & Geiselman, 1996; Van Wallendael, Cutler, Devenport, & Penrod, 2007). Even so, justifications for the observation of Skepticism were still felt necessary, with Leippe suggesting that "given the apparent tendency to overbelieve eyewitnesses, this [Skepticism] may not be an inappropriate outcome" (Leippe et al., 2004, p. 525). Overall then, it seems

Skepticism is not seen as a concern, even though the role of the expert is not to cause the disbelief of eyewitnesses in general, but rather to improve juror discrimination between accurate and inaccurate eyewitnesses. Thus, the frequent observation, and the justifications for Skepticism effects actually serve to illustrate another of the problems inherent in the REE and SEO methodologies. That is, in addition to the fact that they do not directly measure the dependent variable of interest, the absence of identification accuracy rate information renders Skepticism effects largely uninterpretable. If you do not know the level of belief which is appropriate in a given situation, you can never establish the presence or absence of an overbelief effect which arguably needs correcting (McCloskey & Egeth, 1983), and therefore cannot know if Skepticism is a good or bad thing (Cutler et al., 1989b; Penrod & Cutler, 1992). In response to this problem, Wells (1986) argued that a criterion for the acceptance of Skepticism was required in order to provided a means by which to evaluate observed Skepticism effects. That is, even if you have systematically varied the relative quality of the witnessing scenarios, you cannot ascertain if jurors are overly accepting of eyewitness evidence given the viewing conditions, unless you know the actual rate of eyewitness accuracy for that condition and can therefore establish the value of the Skepticism observed. Thus, although the evidence supporting the efficacy of the expert as an educator must be described as equivocal at best, few studies have actually adopted the experimental design necessary to isolate the impact of the expert with regard to juror discrimination accuracy, and therefore lack the information necessary to appropriately evaluate those Skepticism effects frequently observed. Thus, rather than seeking to excuse and justify the Skepticism effects observed, experiments in this thesis, as in the work of Wells (Wells & Wright, 1983 cited in Wells, 1986; Wells et al., 1980), will provide the criterion necessary to evaluate the absolute value of the effects observed.

Expert Effects Stimuli

The preceding review of the expert effects literature provides a good opportunity to point out some features of this literature, some of which will be addressed in more detail in coming chapters, others of which are simply noteworthy.

Unique Sets of Stimuli

It is interesting to note that although there is a substantial number of experiments reported in the eyewitness expert effects literature (n = 23), a notable proportion of those experiments have been conducted using the same crime scenarios, thereby somewhat reducing generalisability of this body of evidence. Specifically, materials were reused in 65% of the experiments reported above (n = 15), leaving only 13 truly independent experiments based on the features of the eyewitnessing stimuli.

Number & Independence of Cues in Stimuli

While this issue will be addressed in greater detail in Chapter 5, it is useful to highlight some features of the experimental stimuli now to facilitate the coming discussion. In particular, it is important to note that only five studies have either independently varied witnessing and identification conditions, or used *probabilistic* predictors, such that the predictors in the scenario do not perfectly correlate with, or act as a proxy for, eyewitness identification accuracy. Those studies were conducted by Wells and colleagues (Wells & Wright, 1983 cited in Wells; Wells et al., 1980), Devenport and colleagues (Devenport & Cutler; Devenport et al.), and Lindsay (1994: Experiment 6). The remaining studies which varied witnessing and identification factors (i.e., SEO studies) incorporated one or more cues to eyewitness accuracy, all of which were perfectly correlated with either accurate or inaccurate identifications and therefore acted as proxies for this construct. The number of such cues incorporated into the experimental stimuli ranged from one to 11 distinct differences between the good and poor scenarios; and in the latter instance (Geiselman & Mendez, 2005; Geiselman et al., 2002), five of these differences were specifically addressed in the expert's testimony. The importance of this design feature will be discussed in detail in Chapter 5.

Participant-Juror Sensitivity

The last matter worthy of mention relates to participant-juror performance *without* expert assistance. Of the 17 studies (i.e., SEO and SEA) which reported juror baseline Sensitivity to the manipulated factors, eight found significant evidence of Sensitivity to witnessing and identification conditions, and only two studies reported significant Counter-Sensitivity with regard to this manipulation. Although this is clearly not perfect performance on the part of participant-jurors, or perfect Sensitivity to WIC, it is also arguably at odds with descriptions which simply characterise participant-jurors as:

> *…insensitive to probative evidence concerning the impact of eyewitnessing factors on eyewitness performance* (Penrod & Cutler, 1992, p. 3).

It is important that researchers do not underestimate evidence suggesting that participant-jurors can evaluate the probative value of eyewitness evidence, as this information can alter the way in which expert effects are interpreted. That is, this information is necessary to establish whether the expert has *induced* Sensitivity, or *altered* its pre-existing levels.

Finally, it is interesting to note that participant-jurors were only able to discount eyewitness expressions of confidence (i.e., show Sensitivity to its poor predictive power) in five of these 17 investigations. Thus, to the extent that it is deemed appropriate for participant-jurors to ignore eyewitness confidence in their decision-making, education is clearly required on the matter.

Chapter 4

Judicial Instruction Effects – Literature Review

Eyewitness expert evidence is not the only safeguard implemented to address the fallibility of eyewitness identification evidence and its role in erroneous convictions. A judicial instruction to jurors is one such alternative provided in some adversarial jurisdictions. These instructions, which may be provided before or after eyewitness evidence, come to the fore when eyewitness identification evidence plays a significant role in a case. Generally speaking, these warnings, unlike expert evidence, are *required* in any trial where eyewitness identification evidence appears. Thus, irrespective of the eventual admissibility of an eyewitness expert, a judicial instruction may be provided to help inform jurors of the limitations of the witness identification. It is perhaps even more important, therefore, for research psychologists to investigate the extent to which participant-jurors understand and apply judicial instructions in their decision-making (Warren v. Parks, 1976 in Ogloff & Rose, 2005), and the extent to which these instructions increase the accuracy of the resolutions reached by jurors:

> *I believe that academic experts should study not only the usefulness of expert testimony, but the usefulness of less expensive alternatives. First it would be worthwhile to know more about the effect of judicial instructions…*(Park, 2003, p. 307).

The vast majority of the research relating to judicial instruction on eyewitness issues has focused on the understanding and effects of a particularised standard instruction known as the *Telfaire* instruction (*U.S. v Telfaire*, 1972), mandated for use by all judges in the United States (Ogloff & Rose, 2005). This instruction (see Appendix B) consists of a general statement highlighting the importance of identification evidence and the potential for identification errors. It requires jurors to consider three key issues when evaluating eyewitness evidence: 1) the eyewitness's capacity to observe the perpetrator with reference to the duration of view, distance of view, quality of lighting and familiarity with the perpetrator; 2) the independence and authenticity of the identification made in light of the delay between witnessing and identification and the context surrounding the identification; and 3) the credibility of the eyewitness themselves. Overall, research on the effectiveness of judicial directions, involving student and mock-juror participants, has produced a mixed picture of their effects on participant-juror knowledge and decision-making. A total of just six studies have been reported in four peer reviewed articles; these are reviewed below and are described in a modified version of the expert effects hierarchy presented in the previous chapter.

In this context, the lowest level of analysis, which investigates the moderating influence of judicial instructions, is termed Response to Judicial Instructions (RJI). The next level of analysis, investigating the judge's ability to educate jurors to the relative quality of witnessing and identification conditions is labelled Sensitivity to Judicial Instruction (SJI), and the final level of analysis investigating the impact of judicial instruction on the accuracy of a jury resolution retains the same title, Sensitivity to Eyewitness Accuracy (SEA).

Responsiveness to Judicial Instruction

1. Katzev and Wishart (1985)

The first investigation of the effects of judicial instruction relating to eyewitness issues was conducted in 1985. The judicial commentary under examination was not the standard *Telfaire* instruction. Instead this study utilises a cautionary statement regarding eyewitness identification, which is supported by psychological research on matters including the role of stress, lighting conditions and duration of observation. Moreover, in this instruction the judge informed participant-jurors that the lighting conditions in the case were poor, and that the perpetrator had a “normal everyday” face, absent any distinctive identifying features. Participant-jurors in this study were allocated to one of three instruction conditions. In the instruction-only condition participant-jurors were provided information about the charge, a definition of the legal elements of the crime, and were then informed of their responsibilities as jurors. In the summation condition, participant-jurors were given a recapitulation of the prosecution's and defence's cases in addition to the instruction presented in the instruction-only condition. In the commentary condition participant-jurors received instruction plus summation, along with the judicial cautionary statement described in detail above. In all three conditions, the judicial instruction was provided to jurors aurally as a supplement to the video of trial materials.

Results

Focusing first on pre-deliberation measures, participant-jurors rendered significantly fewer guilty verdicts in the commentary condition than in the instruction condition. This result is consistent with the judicial commentary having induced a Skepticism effect. Moreover, jurors in the commentary condition also spent a significantly shorter time deliberating over their verdicts than did participants from the other conditions. Even so, post-deliberation, participant-jurors from the three instruction conditions did not differ in the verdicts produced or ratings of the impact of instruction on their verdict. Thus the judicial commentary appeared to induce Skepticism amongst jurors only before they had deliberated upon their verdict. As a result, the authors concluded that:

> *The current study therefore refutes the claim that embedding such instructions about several issues involved in eyewitness testimony in a long list of other instructions would have very little, if any, impact on juror behaviour* (p. 742).

Strengths

Participant-jurors in this study were required to deliberate before reaching a verdict, thus experimenters were conscious of the ecological validity of their research and the generalisability of their findings. Moreover, rather than sacrifice statistical power in favour of higher ecological validity, experimenters also took pre-deliberation measures of individual juror verdicts. This design feature not only provided sufficient power for a pre-deliberation analysis, but also allowed researchers to isolate where in the decision-making process judicial instruction exerts an effect. The results in this study suggested that the deliberation process ameliorated the Skepticism induced by judicial commentary.

Limitations

Firstly, the case utilised by Katzev and Wishart did not appear to be sufficiently ambiguous for experimental purposes, as only three juries did not find the defendant guilty, and all of those were hung juries. No jury reached a unanimous agreement that the defendant had committed the crime. As a consequence of this acquittal bias, it is very difficult to ascertain if the judicial commentary actually increased Skepticism at the jury level, or simply did nothing to remove the pre-existing bias.

Secondly, the judicial commentary investigated in this study made qualitative evaluations of the lighting in the case and the perpetrator's distinctiveness. It is likely that these evaluations were treated by participant-jurors as directional predictors of the accuracy of the identification, particularly since they had been educated about the role that lighting conditions play in identification accuracy. It is therefore hardly surprising that this commentary resulted in a significant reduction in pre-deliberation guilty verdicts. Even so, this result tells us nothing about the value of this reduced belief since it is not possible to ascertain if this change in belief is a result the presence of judicial instruction or the advice provided.

Sensitivity to Judicial Instruction

Greene (1988)

Greene reported two experiments in this article:

2. Experiment 1

In the first study, Greene focused on the effect of the *Telfaire* instruction on participant-jurors' knowledge and understanding of eyewitness issues. Moreover, the author also sought to clarify the moderating effect of this instruction when combined with objectively strong or weak identification evidence. Thus, Greene sought to investigate the impact of the *Telfaire* instruction on participant-juror Sensitivity to the quality of the eyewitness's identification. Accordingly, participant-jurors were presented with either the *Telfaire* instruction, or no instruction, along with identification evidence which was either "strong" or "weak". In the strong condition, the eyewitness saw the perpetrator whilst he or she was seated directly under overhead light, with an unobstructed view, from a close distance. In the weak condition, the perpetrator was seen in a dimly lit area and the view, which was made from some distance, was obstructed.

Results

Insufficient data were provided to ascertain if participant-jurors were significantly Sensitive to case strength in the control condition. Even so, inspection of the percentages of guilty verdicts in the strong (41.9%) and weak cases (3%) suggest that participants were more likely to convict in the strong than the weak case, a pattern consistent with Sensitivity. The additional result that: participants in the weak case condition (63.6%) appear to be more likely to fail to reach a verdict than those from the strong case condition (38.7%), which is also consistent with Sensitivity in the instruction-absent condition. It remains unclear if any of these differences were statistically tested, and if the results were significant. The provision of the *Telfaire* instruction was described to decrease the likelihood of a hung verdict where the case was weak, however, no test of the effect of instruction on guilty verdicts, or interaction between instruction and case strength was reported. Thus, it is difficult to ascertain if the reduction in guilty verdicts observed in the *Telfaire* conditions reflects a significant Skepticism effect, or if the larger decrease in guilty verdicts in the strong case is consistent with something akin to Counter-Sensitivity. Despite these uncertainties, the author concluded that:

> *Jurors who heard this* [Telfaire] *instruction were no more sensitive to factors known to be problematic to eyewitness identification than were jurors who had no instruction* (p. 260).

Greene went on to attribute this result to the fact that the *Telfaire* instruction did not contain any clear instructions to help participant-jurors to assess the likely accuracy of the eyewitness identification. This rationale provided the basis for a second experimental investigation.

Limitations

Overall, the analysis of the effect of judicial instruction on individual verdicts is extremely unclear. While ANOVA's were conducted and reported on measures evaluating the ease of consensus, juror comprehension of eyewitness testimony, jurors' weighing of evidence, and finally judicial efficacy, it is very difficult to ascertain if any statistical analyses were conducted on the verdicts rendered. This is clearly a dependent variable worthy of detailed analysis, yet on the whole the information provided is not sufficient. Thus, no clear understanding of Sensitivity, or Skepticism, resulting from judicial instruction can be derived from the reported results.

3. Experiment 2

A revised judicial instruction was developed in an attempt to simplify the warning, and to provide participant-jurors with clearer directions regarding *how* eyewitness reliability is influenced by factors in the witnessing scenario. Participant-jurors in this study were allocated to either the strong or weak identification evidence condition, and heard either the *Telfaire* or the revised instruction.

Results

Despite significant Sensitivity to witnessing and identification conditions overall, those jurors who heard the revised instruction were less likely to convict than those in the other instruction conditions, both before and after deliberation. This is consistent with a Skepticism effect. Even so, fewer jurors in the revised instruction condition than the standard instruction condition selected the weak eyewitness as the person with the most important testimony. This effect was not observed when the identification was strong. Overall then, the revised instruction – despite the addition of directional predictions – led to Skepticism of eyewitness identifications, rather than Sensitivity to the likely accuracy of the eyewitness identification.

Strengths

Unlike in Experiment 1, the author provided a detailed statistical analysis of the pre-deliberation measures in Experiment 2. Accordingly, it is clear the participant-jurors in all conditions were Sensitive to the relative strength of the identifications, however, the revised instruction caused participant-jurors to become Skeptical of eyewitness identification evidence.

Limitations

As above, analysis of the post-deliberation verdicts was lacking in detail. Although the author stated that “fewer juries convicted the defendant when they heard the revised instructions” (suggesting Skepticism, p. 265) and that “the version of the identification had an effect” (suggesting Sensitivity, p. 265), no p-values or test statistics were reported to clarify if these reported effects were statistically reliable or not.

4. Hoffheimer (1989)

Using case scenarios based on a bank robbery trial (*U.S. v. Zeiler*, 1970) and a perjury case (*U.S. v. Greene*, 1979), Hoffheimer investigated the effect of the *Telfaire* instruction on participant-juror appraisals of identification testimony. After reading each of the cases, jurors were presented with one of three sets of instructions. The first group of participants received an instruction regarding: a) the presumed innocence of the defendant; b) the definition and role of the reasonable doubt standard; and c) instruction that the defendant has no burden of proof in the trial. The second group received the same three instructions described above with the addition of the statement that: d) "the government must prove identity beyond reasonable doubt" (p. 47), while those in the third group received instructions a) to c), as well as the *Telfaire* instruction.

Results

Overall, participant-jurors showed a clear tendency to acquit the defendant, which was not moderated by the introduction of the *Telfaire* instruction. Although, the *Telfaire* instruction did result in an increase in guilty verdicts (i.e., eyewitness belief) for the violent case, and a significant decrease in guilty verdicts (i.e., eyewitness disbelief) in the non-violent case, these differences were not statistically significant. Moreover, the short instruction provided in the second condition, regarding the government's responsibility to prove identity, caused the greatest decrease in guilty verdicts; resulting in significantly fewer guilty verdicts when compared with the *Telfaire* condition. Overall, these results led the author to conclude that:

> *…jury instructions on identification testimony have an influence on subjects' evaluation of identification testimony. At least in the case of violent crime, the instructions widely used by some federal courts did not have the same effect on evaluation of identification evidence as expert psychological testimony on eyewitness identification. The availability of the instructions thus does not justify the exclusion of otherwise relevant expert opinion evidence.* (p. 55).

Limitations

In the quote above, the author describes the judicial instruction as having a different effect on the evaluation of eyewitness evidence when compared with expert testimony, suggesting that the judge's effect was undesirable compared with the expert; however, it is difficult to analyse the validity of this statement given that the author makes no reference to the source of the expert effect referred to. In the context of the article, however, it seems likely that the author is making reference to Loftus's (1980) investigation of the effect of expert evidence on juror evaluations of violent and non-violent crimes. In that study (see Chapter 3, Experiment 2), Loftus suggested that expert testimony produced a greater reduction in the likelihood of conviction for violent rather than non-violent crimes, yet did not report the statistical validity of this observation. Thus, if it can be assumed that Hoffheimer is referring to this conclusion made by Loftus, which seems likely given this quote:

> *Studies of the influence of expert psychological evidence have indicated the opposite result: a narrowing of the difference between probabilities of conviction for violent and non-violent offenses* (p. 53).

The comparison apparently being made by the author, suggesting that the judge has an undesirable effect where the expert caused Sensitivity to event violence, is not borne out by either set of data. Specifically, Loftus did not report evidence of a *reliable* reduction in guilty verdicts in the violent rather than the non-violent crime, and likewise Hoffheimer reported a *non significant* increase in guilty verdicts for a violent crime and a *non significant* decrease in guilty verdicts in the case of a non-violent crime. Thus, the conclusion that the expert and the judge had different impacts on juror decision-making is not actually based upon significant effects, although the trends observed were indeed in these directions.

Ramirez, Zemba and Geiselman (1996)

5. Experiment 1

Here experimenters investigated the effects of instruction timing on participant-juror's perceptions of good and poor identification evidence. Using a modified version of the materials developed by Cutler and colleagues (1989b), participant-jurors evaluated either a good or poor eyewitness identification. The good conditions in this study were operationalised by an absence of disguise, a hidden weapon, a two day delay between viewing and identification, and an unbiased lineup instruction. The poor conditions included a disguise, an exposed weapon, a two week delay between viewing and identification, and a biased lineup instruction. In addition to this, participant-jurors were also given the *Telfaire* instruction either: 1) before and after the eyewitness evidence; 2) only before the eyewitness testimony; 3) only after the eyewitness testimony; or 4) not at all.

Results

Analysis revealed that participant-jurors only showed significant Sensitivity to the eyewitnessing conditions in the instruction-absent control condition and this was maintained in the before-only experimental condition. Moreover, when the witnessing conditions were good, participant-jurors were *less* likely to render a guilty verdict than in the other three conditions, and when the eyewitnessing conditions were poor, the subjects who heard instructions both before and after the eyewitness, were *more* likely to vote guilty than the other three conditions. The authors concluded that:

> *…sensitivity to the eyewitness evidence was minimised when the Telfaire instructions were presented at the end of the trial, and this effect was to promote either juror overbelief or skepticism depending on whether the instructions were also presented at the beginning of the trial* (p. 41).

Strengths

This study provides a replication of the Sensitivity effect observed amongst control participants by Cutler and colleagues (1989a; 1989b). This provides experimenters with a firm foundation against which to investigate the effects of order and judicial instruction.

Limitations

The conclusion reached by authors that:

> *…presenting the instructions after the evidence reduced the subject-jurors' sensitivity to the eyewitness evidence and led them to vote not guilty regardless of the nature of the eyewitness evidence* (p. 45),

does not appear to be wholly consistent with the results reported:

> *…**when the eyewitnessing conditions were good**, the subjects were less likely to vote guilty in the after-only instruction condition than in the other three conditions* [emphasis added](p. 41).

Moreover, the subsequent conclusion:

> *…presenting the instructions both before and after the evidence in the present experiment led the subjects to vote guilty regardless of the nature of the eyewitness evidence.* (p. 45),

is not wholly consistent with the associated result that:

> *…**when the eyewitnessing conditions were poor**, the subjects were more likely to vote guilty in the before-and-after instruction condition than in the other three conditions* [emphasis added](p. 41).

In each instance the result is clearly moderated by the witnessing condition, thus it is difficult to understand how the authors can conclude that jurors were either overbelieving or Skeptical "regardless of the nature of the eyewitness evidence".

Moreover, the main effects of instruction type necessary to support the conclusions of overbelief and Skepticism, quoted in the results section here, were not forthcoming anywhere in the reported statistical analysis. Instead, all that can be concluded based on the results provided is that different types of judicial instruction caused evidence of differing quality to be evaluated differently. This is not consistent with either Skepticism or overbelief.

6. Experiment 2

The *Telfaire* instruction and the eyewitnessing scenarios were both revised for Experiment 2. The eyewitnessing scenarios were modified such that the difference between good and poor conditions was determined by nine witnessing and identification factors (i.e., duration of view, distance of view, lighting, level of stress, weapon focus, delay of identification, lineup fairness, pressure to choose from a lineup and presence of a prior identification), rather than the original four. The *Telfaire* instruction was revised to include directional predictions regarding 13 factors thought by most eyewitness experts to reliably affect the performance of a significant percentage of eyewitnesses. Critically, this revised instruction directly addressed *each of the nine factors* manipulated across the witnessing scenarios, plus an additional four factors (i.e., prior exposure to the suspect, effects of post event information, cross race identification and the confidence-accuracy relationship). That is, the eyewitnessing conditions were constructed so that they almost exactly matched the empirically derived testimony of the judge.

Overall, participant-jurors evaluated either a good or a poor eyewitness identification, and were provided with: 1) the *Telfaire* instruction; 2) the authors' revised version; or 3) no instruction at all.

Results

Analysis revealed no main effect for witnessing condition or instruction type. Moreover, although an inspection of means suggests that participant-jurors in the control and revised conditions were Sensitive to the differences between witnessing conditions, and that participants provided the *Telfaire* instruction were less so, this result was not significant. This null effect was interpreted by the authors accordingly:

In summary, sensitivity to the eyewitness evidence was minimized when the Telfaire instructions were presented at the end of the trial, and this effect was to promote juror skepticism (p. 51).

Strengths

Participant-jurors in this study were asked to complete a ten-item multiple choice test measuring their knowledge of eyewitness phenomena. Results of this analysis provided clear evidence that participant-jurors attended to, and recalled the information presented by the judge; this is despite the fact that no significant differences in Sensitivity to judicial instruction were observed. Thus, this study begins to provide some information about the processes underlying juror Sensitivity, and suggests that knowledge is not necessarily sufficient to induce, or improve, Sensitivity to the relative quality of witnessing conditions.

Limitations

This study, like others, falls prey to a methodological flaw which will be discussed at some length in Chapter 5. In order to facilitate that latter discussion, it is important for the reader to note: a) that experimenters provided participant-jurors with nine different cues differentiating the good witnessing condition from the poor: b) that none of these cues were independently varied within witnessing scenarios, such that each cue in the scenario acted as a perfect proxy for the relative accuracy of the eyewitness identification; and c) that in its revised form, the judicial instruction made directional predictions regarding each of these nine cues and their association with identification accuracy.

Summary

Referring to Table 4.1 below, it can be seen that of the six experiments conducted to assess the impact of judicial instruction on juror evaluations of eyewitness evidence, five independently varied the quality of the witnessing and identification conditions, while one did not. Despite the difficulties observed in interpreting the reliability of the observed effects, the impact of judicial instruction appears to be as mixed as the effects of expert evidence, with one study providing clear evidence of Skepticism, and two providing clear evidence for Sensitivity.

It is also important to note that no study has investigated the effect of judicial instruction on participant-juror Sensitivity to Eyewitness Accuracy. Thus, although we have evidence that the judge can moderate juror belief, and indeed inform those beliefs, nothing is known at all about the influence of judicial instruction on participant-juror ability to discriminate between accurate and inaccurate eyewitness identifications.

Revising Judicial Instructions

In light of the preceding review, it is interesting to note the manner by which experimenters have chosen to revise, or indeed, construct judicial instructions. Specifically, the second experiments conducted by Greene (1988) and Ramirez et al. (1996), revised judicial instructions to include directional predictions regarding the accuracy of the eyewitness identification, in the latter case, experimenters modified the judicial instruction so that it provided jurors with information about 13 different predictor variables. This permitted researchers to evaluate the effects of judicial instruction with regard to the same outcome variable, or role, as the expert, i.e., educator to the jury. Indeed, Greene made specific reference to the fact that the *Telfaire* instruction made no directional predictions, and consequently constructed a new version of the instruction which incorporated directional opinions, thereby facilitating investigations of juror Sensitivity to Judicial *Opinions*. In essence then, the modifications made to judicial instruction have provided a common ground for the comparison of the judge and the expert as *educators* to the jury. What researchers have apparently failed to note, is that valid comparisons could also be made between unmodified judicial instruction and expert evidence in terms of their effects on participant-juror Sensitivity to Eyewitness Accuracy. This comparison is yet to be made.

Table 4.1 : Summary of Observed Judicial Instruction Effects

Outcome Measured	**Experiment**	**Control Response to Manipulated Variables**	**Judicial Effects Observed**		
			RJI	**SJI**	**SEA**
Response to Judicial Instruction (Moderating Effect)	1. Katzev & Wishart, 1985		< Pre-deliberation Guilty Verdicts from Commentary than Instruction Only (Sig.) < Deliberation in Commentary than Instruction & Summation (Sig.)		
Sensitivity to Judicial Instruction (Educating Effect)	2. Greene, 1988 (Experiment 1)	Sensitivity to WIC (untested?)	(untested – floor effect?)	> Strong EW > Believed (untested – floor effect?)	
	3. Greene, 1988 (Experiment 2)	Sig. Sensitivity to WIC	Sig. Skepticism of Pre-Deliberation Guilt from Revised Skepticism of Post-deliberation Guilt from Revised (Sig?)	Sig. Sensitivity to WIC in *Telfaire* & Revised	

Outcome Measured	**Experiment**	**Control Response to Manipulated Variables**	**Judicial Effects Observed**		
			RJI	**SJI**	**SEA**
Sensitivity to Judicial Instruction (Educating Effect)	4. Hoffheimer, 1989	None	None (Telfaire)	None	
	5. Ramirez, Zemba, Geiselman, 1996 (Experiment 1)	Sig. Sensitivity to WIC	None	Sig. Sensitivity to WIC in Before-only Conditions < Guilty Verdicts where WIC Good in After-only Condition (Sig.) > Guilty Verdicts where WIC Poor in Before+After Condition (Sig.)	
	6. Ramirez, Zemba, Geiselman, 1996 (Experiment 2)	None	None	None	

Chapter 5

Inferences Evaluated

Inferences and Advocacy

Results from the studies described above have been used by psychologists and lawyers to support arguments both for and against eyewitness expert evidence and judicial instruction as safeguards. In general, psychologists have not looked favourably upon judicial instruction, while opinions regarding expert evidence have generally been more positive.

Judicial instruction has largely been criticised for its failure to generate a Sensitivity effect (Cutler et al., 1990a; Leippe et al., 2004; Ramirez et al., 1996) and, to a lesser extent, for its tendency to create confusion among jurors (Devenport et al., 2002). These arguments, however, assume that an absence of Sensitivity to witnessing conditions, which was evident in three of the five studies (reviewed in Chapter 4), reflects an equivalent absence of Sensitivity to the *actual accuracy* of the eyewitness identification. Yet, as already discussed (in Chapter 4), there is no reason that one should expect authentic judicial instructions (seen in three studies: Greene, 1988; Hoffheimer, 1989; Ramirez et al., 1996) to make jurors more aware of the relative quality of different witnessing scenarios because they were not intended, or designed to do this. Nor is there any reason to conclude that the failure to show such Sensitivity necessarily means that the judicial instruction is ineffective with regard to SEA.

Opinions in favour of the use of expert testimony span a wide range of rationales, ranging from reliance on common belief in its utility (Kassin et al., 1989; Kassin et al., 2001) and the presence of “reliable effect[s]” (Hosch, 1980, p. 300); to rationales which emphasise: a) the educational effect the expert has on jurors (Cutler et al., 1989a; Devenport et al., 2002; Pezdek, 2007); b) resultant Skepticism (Deffenbacher, 1984; Fox & Walters, 1986; Hosch et al., 1980; Leippe, 1995); and c) resultant Sensitization effects (Cutler et al., 1989b; Hosch et al., 1980). Conversely, those opposing the inclusion of expert testimony have done so on the basis of the weak evidence attesting to: a) its utility (Ainsworth, 1998; McCloskey & Egeth, 1983); b) its tendency to evoke Skepticism (Ainsworth, 1998; Leippe, 1995; McCloskey & Egeth, 1983); and its failure to produce either c) reliable effects (Wise

& Safer, 2004) or d) sensitization (Lindsay, 1994). Thus, there is some degree of overlap in the evidence used to support each side of this debate.

Despite the equivocal evidence and arguments regarding eyewitness expert testimony and judicial instructions, many psychologists have expressed a preference for expert evidence over judicial instruction (Geiselman, 1994; Leippe, 1995; Ramirez et al., 1996; Wise et al., *under review*; Wise & Safer, 2004). Most recently:

> *The* Telfaire *instructions, the most common vehicle for educating jurors about the reliability of eyewitness evidence, are apparently ineffective in this regard* [educating jurors about the reliability of eyewitness evidence]. *Providing an eyewitness expert at the time of trial appears more promising* (Pezdek, 2007, p. 113).

This belief is even more surprising considering the fact that only one study has directly compared the effects of these two safeguards, and even in that instance, did not provide a fair test of the judicial instruction. The study in question (Cutler et al., 1990a), like all others investigating judicial instructions, investigated the effect of the warning on participant-juror *Sensitivity* to the *Opinions* expressed by the judge. Yet, since the *Telfaire* instruction does not actually provide any directional opinions as such, it is unfair to expect that jurors will be able to demonstrate any Sensitivity to them on this measure. Thus, those few voices speaking in favour of judicial instruction have done so on the basis of its practical advantages alone (Justice Major in Greene & Loftus, 1984; *R v D.D.*, 2000), having no valid evidence upon which to compare it to the alternative being advocated by most psychologists. Clearly further empirical evidence is required before valid comparisons can be made between these two safeguards, thus providing a foundation for informed advocacy on the issue. As a result of this limitation identified in the literature, this thesis aims to provide: 1) the first fair test of the effects of judicial instruction, and 2) the first valid comparison between the effects of judicial instructions and eyewitness expert evidence.

Inferences and Methodology

The following section considers the role that methodology plays in the overall inferential process, specifically its role in establishing and preserving the validity of the inferences drawn by researchers.

The hierarchy proposed in Chapter 3 grouped studies of the impact of expert evidence on the basis of the outcome measured. Studies at the lower-end of the hierarchy measure changes in rates of guilty verdicts or belief decisions, studies at the middle of the hierarchy measure the degree of Sensitivity to Expert Opinions (i.e., evidence-based directional predictions), and finally, studies at the top of the hierarchy measure juror Sensitivity to Eyewitness Accuracy. Only studies at the top of the hierarchy, which use real eyewitness designs, are able to *directly* measure expert effects with regard to the amount of aid they offer the jury "in reaching an accurate resolution of a disputed issue" (Judge Becker cited in Cutler & Penrod, 1995, p. 27). Studies lower on the hierarchy attempt to *infer* accurate resolutions from the observation of sensitivity to variables known to be *correlated* with accuracy. Therefore, these studies *indirectly* measure the effect of expert evidence on the accuracy of juror resolutions. While, generally speaking, it is valid to infer sensitivity to a construct, such as identification accuracy, from sensitivity to correlates of that construct e.g., witnessing conditions, certain assumptions must be satisfied to ensure that the inferences made are valid.

The assumption most in need of consideration takes this general form: there must be a reasonable amount of correspondence between the task one draws an inference from and the task one infers to, in order for there to be a valid basis for those inferences. Consider, for example, researchers investigating the ability to detect deception in written statements. There are two possible ways that the experimenter could study this issue, either *directly* by asking evaluators to make discriminations between true statements and lies, or *indirectly* by assessing an evaluator's abilities to identify those cues believed to be associated with truths and lies. The direct method tells researchers if their evaluators can tell the difference between truth and untruth, while the indirect measure tells researchers if their evaluators are sensitive to what we believe to be *predictors* of truths and lies.

At this point, it is appropriate to reason that Sensitivity to the features differentiating truths from lies would also result in an ability to discriminate between actual truths and lies. But consider what would happen if the tasks put before participants in the direct and indirect evaluations were to differ substantially. Imagine, then, that the lies and truths being used in the direct task were actual lies and truths generated by research participants. As such, the truths would have more characteristics of truths than lies, and the lies would have more characteristics of lies than truths. Importantly, however, because they are real they don't have *only* the characteristics of lies or truths, but represent real truths and lies which possess a combination of probabilistic indicators, each with varying predictive power. Thus, the evaluator in this task needs to estimate the likely accuracy of an account by weighing the indicators present to come to a final decision; truth or lie. This is a very difficult task, and indeed we know that evaluators generally do not perform better than chance when given real stimuli to evaluate in this way (Aamandt & Custer, 2006).

Now, compare this with the stimuli used in the indirect task, imagine that the experimenters are interested only in an evaluator's ability to differentiate between the *features* of lies and the *features* of truths. As a result, the experimenters rationalise that they do not need real truths and real lies to investigate the evaluators' ability to identify the features of each, instead, they simply need those features to be present or absent in the accounts; thus evaluators in this case are not given real lies and truths, complete with their probabilistic complexities – instead they are given fabricated transcripts which have the *features* of lies and truths. Imagine further, that these statements are constructed so that those designated as "lies" only have characteristics consistent with lies, and the statements designated as "truths" only have characteristics consistent with truths. While these stimuli clearly assess the evaluators' knowledge for the *features* of truth versus lies, this task says nothing about their ability to differentiate real truths from real lies, and doesn't assess their ability to weigh various conflicting indicators. Specifically, since the evaluator only needs to know *one feature* generally associated with truth or *one feature* generally associated with a lie in order to reliably differentiate between the two stimuli, this task provides no contingency for truths which have some features consistent with lies or vice versa. This, unlike the direct evaluation task incorporating such contingencies, is a very simple matching task; matching knowledge of the *features* of lies and truths, to the *features* of a statement, with no need to combine and weigh probabilistic predictors.

Accordingly, it is inappropriate to infer that performance on this matching task in anyway equates to performance on the direct task described earlier. Indeed, it is very difficult to imagine a situation where people researching the detection of deception would try to infer an ability to discriminate lies from truths on the basis such an unrepresentative task: as a result, this method is never used within that field. Moreover, it is clear from this example, that such a simple indirect measure will systematically overestimate evaluator performance in assessing real statements, obviously the more complex of the two tasks, therefore rendering inferences from one measure to the other inaccurate, if not invalid. Yet, as has been described in the previous chapters, this is *precisely* the methodology adopted by the majority of researchers in the expert effects (and judicial instruction) domain. Here, the direct test is to provide jurors with the testimony of real eyewitnesses who have made accurate or inaccurate identifications, and ask them to judge the accuracy of these eyewitnesses either with or without expert advice. As in the real world, although accurate identifications will probably have more features of accurate identifications than inaccurate ones, and vice versa, invariably some features of the real identifications will be consistent with accuracy e.g., there was a lengthy exposure time, while others will erroneously be consistent with inaccuracy e.g., the presence of a disguise. Thus, the evaluation of these real eyewitnesses is highly complex, requiring jurors to weigh and combine probabilistic estimates in order to make binary decisions about accuracy. Moreover, as in the detection of deception example above, participant evaluations of real eyewitness stimuli generally falls around chance levels (Wells, 1980).

Also consistent with the detection of deception example above, the indirect test of expert effects involves a situation where experimenters have determined that they do not need real eyewitness testimony in order to evaluate jurors' sensitivity to the *factors* associated with likely accurate and inaccurate identifications. Accordingly, experimenters constructed stimuli so that those identifications designated as "accurate" contained *only* elements consistent with good witnessing and identification conditions, and those identifications designated as "inaccurate" contained only features consistent with poor witnessing and identification conditions. The expert then educates jurors regarding the factors which constitute good and poor witnessing conditions, and how witnessing conditions affect the likely accuracy of an identification. Thus, the jurors were simply required to identify *one feature* of a good witnessing scenario, or *one feature* of a poor witnessing scenario, as defined by the expert who described the features of each in detail, in order to effectively discriminate between accurate and inaccurate eyewitnesses. Overall then, it is likely that performance on indirect tasks of this type will systematically overestimate the likely performance of these same jurors on the direct discrimination task by failing to provide real eyewitness testimony, complete with probabilistic predictors, for jurors to evaluate. Yet explicit and implicit inferences from performance on indirect tasks (SEO) to performance on direct tasks (SEA) characterise much of the expert effects literature[9]. These inferences are most clearly illustrated through the almost complete failure to apply direct measures in expert effects research, with only one such study ever having been published in a peer reviewed journal (Wells et al., 1980). Moreover, no previous summary of the eyewitness expert effects literature has made the distinction between direct and indirect measures, or discussed the qualitative differences between them.

Finally, it is important to understand that the issue raised here regarding direct methodologies does not relate to the fact that the expert *points out* the manipulated cues, as the expert is retained to testify to *known predictors* of accuracy. Rather, this behaviour becomes a problem when the expert is directing jurors to consider variables which have been experimentally constructed to act as *proxies* for accuracy and inaccuracy; i.e., cues that *invariably define* the difference between good and poor witnessing conditions, and therefore, accurate and inaccurate identifications. This is because it is unrealistic for an expert to be able to pinpoint every (and only) those predictors which are relevant to the accuracy of any real eyewitness. There will always be predictors that the expert does not know about and those that do not perfectly correlate with accuracy. Thus, the testimony of a real expert, which relates to a real eyewitness, will at times include redundant and irrelevant information. Participant-jurors called upon to evaluate these real eyewitnesses therefore must consider the value of each of the predictors described by the expert, and its relative weight given the other predictors in the scenario, in order to reach a final evaluation. This complex task provides a real test of the expert's efficacy by establishing a situation where it will be appropriate for jurors to follow the expert's advice regarding the specific predictors in the scenario, *most of the time*; yet, in some instances, probabilistically speaking, it will not be appropriate to do so. It is important for the ecological validity of the task that this contingency is provided for in the experimental design, and that participant-jurors are permitted to respond like real jurors, who will at times be right to disregard expert opinions and maintain their own.

[9] The studies by Lindsay (1994) and Devenport (Devenport et al., 2002; Devenport & Cutler,2004) provide notable exceptions to this statement.

To summarise, there appears to be a clear need to evaluate the validity of inferences being made from simple measures of Sensitivity to Expert Opinions, to the complex task of discriminating between real accurate an inaccurate eyewitnesses. Specifically, it is important to empirically test if, and to what extent, estimates of participant-juror performance differ from direct to indirect measures, so that inferences from SEO to SEA methodologies can be moderated accordingly.

Aims

In response to the needs identified in the preceding reviews, this thesis aims to;

1. Provide the first fair test of the effects of judicial instruction by evaluating its effects on participant-juror Sensitivity to Eyewitness Accuracy.
2. Provide the first valid comparison between the effects of judicial instructions and eyewitness expert evidence, using measures of participant-juror SEA.
3. Evaluate the correspondence between direct and indirect measures of expert effects on participant-juror SEA.

These investigations will provide an empirical foundation for the relative comparisons of the judicial instruction and expert evidence safeguards. They will also provide the first test of a judicial instruction with regard to an accuracy criterion. In addition, this research will empirically test the validity of making inferences from indirect to direct measures and perhaps provide information about how to adjust inferences when moving from one construct to the other.

SECTION 3: EXPERIMENTAL STUDIES

Chapter 6

Experiment 1 - Jury Study

The review of the literature, and results reported from the survey of public defenders, suggest that further investigation of the relative impacts of eyewitness expert evidence and judicial instruction on juror evaluation accuracy are necessary. Previous research investigating the impact of eyewitness expert evidence has focused primarily on its effects on either juror Response to Expert Evidence (Hosch et al., 1980; Loftus, 1980, Experiment 2; McKenna, Mellott & Webb, 1981 cited in McCloskey & Egeth, 1983) or their Sensitivity to Expert Opinion (Blonstein & Geiselman, 1990; Cutler et al., 1989a, 1990a; Cutler et al., 1989b; Devenport & Cutler, 2004; Devenport et al., 2002; Fox & Walters, 1986; Geiselman et al., 2002; Leippe et al., 2004; Lindsay, 1994; Loftus, 1980, Experiment 1; Pezdek et al., *in press*). This is despite the fact that at least one of the stated aims of introducing expert evidence endorsed by psychologists is to "aide the jury in reaching *an accurate resolution* of a disputed issue" (emphasis added) (*United States v. Downing*, 1985 in Cutler & Penrod, 1995, p. 27). However, the impact of eyewitness expert evidence and judicial instruction with regard to this particular outcome has gone largely unexplored, as only two studies have included the real eyewitness testimony necessary to permit such investigations (Wells & Wright, 1983 cited in Wells, 1986; Wells et al., 1980); neither of which included a judicial instruction condition.

Wells, Lindsay and Tousignant (1980) were the first to incorporate eyewitness expert evidence into experiments designed to measure participant-jurors' baseline sensitivity to eyewitness accuracy. The data gathered from this study suggested firstly, that expert evidence could induce participant-jurors to be sensitive to some witnessing and identification variables (i.e., expert evidence significantly reduced juror reliance on expressions of eyewitness confidence), although not necessarily to all of those present in the eyewitnessing scenario (i.e., expert evidence did not significantly increase the degree to which jurors evaluated the witnessing conditions). Secondly, this study provided evidence that this sensitivity to the experts opinion could coexist with

Skepticism to eyewitnesses in general. Indeed, eyewitness expert testimony caused participant-jurors to believe all eyewitnesses less, *irrespective* of the relative quality of the witnessing and identification conditions. Finally, this study demonstrated that it was possible for significant Response and Sensitivity to Expert Opinions to be observed, *without* significant Sensitivity to Eyewitness Accuracy. Wells et al. (1980, p. 283) concluded that:

> *if we assume that it is desirable to obtain juror-belief rates that closely correspond to probable witness accuracy rates across witnessing conditions, then we must admit that expert advice did not produce an improvement in that regard*.

Therefore, participant-jurors in this case were no closer to arriving at an *accurate resolution* as a result of expert evidence than they were in its absence.

The study by Wells and Wright (1983 cited in Wells, 1986) produced an equally interesting although qualitatively different pattern of results. In this case, a significant interaction between eyewitnessing conditions and expert testimony was found *in spite* of the fact that the identification rates across levels of witnessing condition did not vary significantly. Here, participant-jurors who heard expert testimony were significantly more likely to believe eyewitnesses who made their identifications under “good” conditions than they were to believe those eyewitnesses who made their identifications under “poor” conditions. Thus the experts’ testimony induced participant-jurors to be sensitive to factors (i.e., exposure time, distance and clarity of view), which although *likely* to influence identification accuracy, had not done so in this particular case. Furthermore, after expert testimony, these participant-jurors *also* showed significant Sensitivity to Eyewitness Accuracy such that accurate eyewitnesses were believed at a rate 15% higher than inaccurate eyewitnesses, without an associated Skepticism effect. Given that the accuracy of the eyewitnesses did not significantly vary between good and poor witnessing conditions, this Sensitivity to Accuracy after expert evidence is somewhat surprising, but may be explained in-part by the fact that the expert in this case did not actually make any directional predictions regarding the effect of witnessing conditions on likely accuracy.

In light of the somewhat confusing evidence regarding the co-occurrence of Sensitivity to Expert Opinions and Sensitivity to Eyewitness Accuracy, this study aims to provide the first empirical data regarding the relative impact of expert evidence and judicial instruction on the accuracy of jury verdicts. Not only is this the first *comparative investigation* of the effects of judicial instruction and expert evidence where accuracy rather than sensitivity to the expressed opinion is the dependent variable of interest, it is also the first investigation examining this accuracy both before and after group deliberations. Given the dearth of previous research varying instruction type and incorporating group deliberation, it is difficult to make robust predictions regarding the outcome of this investigation. Even so, some tenuous predictions can be made; Firstly, previous research suggests that jurors who hear expert testimony will be less influenced by the eyewitness than jurors who do not hear the expert, as both Hosch et al. (1980) and Maass et al. (1985) found that expert testimony significantly decreased the importance and weight attributed to eyewitness evidence. Secondly, it seems likely that jurors who have heard expert evidence will be less likely to convict than their control counterparts. This decline in convictions may be evident through either a significant decrease in guilty verdicts (Loftus, 1980, Experiment 1) or increases in lenient or hung verdicts (Loftus, 1980, Experiment 2; Maass et al., 1985). Finally, in light of the significant increases in deliberation time observed by Hosch et al. (1980) and Loftus (1980, Experiment 2), it is reasonable to expect that those juries provided expert evidence will deliberate longer than those juries given no instructions. In spite of these predictions, it remains unclear how performance in the judicial condition will compare with expert evidence, and indeed how instruction type will interact with any deliberation effects to influence the final accuracy of the deliberated verdicts.

Method

Overview

Experiments utilising real eyewitness designs are composed of two distinct stages. Stage 1. pertains to the collection of real eyewitness testimony and involves: a) the construction of a crime video; b) eyewitness identifications; and c) eyewitness interviews. In Stage 2. participant-jurors are presented with these eyewitness interviews and are required to evaluate the accuracy of the identification made. For ease of reading only the second of these two stages has been reported in full in the body of this thesis (see below), specific information regarding the method used in Stage 1. is reported in full in Appendix C (p. 257).

Participants

Participant-witnesses

Participant-jurors in this study viewed the four eyewitnesses obtained from Eyewitness Evidence Protocol (I) (see Appendix C for a detailed description of this protocol). These eyewitnesses had viewed a video of an office break-in before identifying either a police suspect from a target-absent lineup or the perpetrator from a target-present lineup.

Participant-jurors

One hundred and four jury eligible undergraduate psychology students (40 males and 64 females, mean age 19.8 years) from the University of New South Wales served as participant-jurors and received partial course credit for their participation.

Design

A 2 (witness type: correct vs. mistaken) x 3 (instruction type: eyewitness expert vs. judicial instruction vs. no instruction) between-subjects factorial design was employed. Six places were made available for participant-jurors in each experimental session, however in practice only 53% of juries had six members, the remainder having between 3-7 members (mean 5.47 jurors) due to unanticipated attendance or non-attendance. Each jury (n = 19) was assigned at random to view one eyewitness and hear one of the three instruction types.

Materials

Video of Participant-Witness Testimony

Undergraduate psychology students fulfilled the role of participant-witnesses in this study. Two of these four witnesses were female, and two were male. One male and one female had made accurate identifications while the other two had made inaccurate identifications. All participant-witnesses were over the age of seventeen years and under the age of 25. The confidence-accuracy correlation for this sample of eyewitnesses was not statistically significant (r_{pbi} = 0.323, p = .667, n = 4). The relevance of this information will become clear in the context of the statistical analysis to follow.

The first interview completed by these witnesses (conducted by council for the prosecution) was in the style of an examination-in-chief, during which the witness was asked to describe what they saw and to outline the details of the identification process and their resulting decision. On average the examination-in-chief lasted just under four minutes. The subsequent cross-examination style interview focused on the witness' estimate of the duration of the incident and other details provided by the participant-witness. In order increase the realism of the interview the witness was not prepared for any of the questions contained within the cross-examination. On average the cross-examination lasted for approximately three minutes. The same interview schedule was used for each of the four witnesses (see Appendix C (disc) for the complete interview schedules).

Pre-trial Instruction

A series of pre-trial instructions were read aloud to the jury. All juries were told that they were about to see the testimony of an *actual witness* to a crime who identified the police suspect and were asked to watch the footage "as though they were a juror in the trial of the accused". They were then asked to "examine and scrutinize the testimony of the witness with great care" (JCNSW, 2006, s3-610; see Appendix D). Participant-jurors were also informed in general terms about the structure and purpose of the examination-in-chief and cross-examination. Those participant-jurors in the expert evidence condition were also provided a direction regarding the purpose and evaluation of expert testimony (JCNSW, 2006, s2-1110; see Appendix E for verbatim instructions). Furthermore, jurors were told what charges the defendant was facing and were provided with information regarding the legal requirements for conviction. Finally jurors and were informed that in order for a verdict to be reached, all members of the jury must unanimously agree (as required under NSW law at the time this experiment was conducted) and that in this case believing the accuracy of the eyewitness's identification beyond a reasonable doubt was enough to satisfy the legal requirements for guilt in this matter (see Appendix F for these materials).

The Minimal Trial

Participant-jurors in all conditions were provided with the same set of case facts and general legal instructions (see Appendix G) irrespective of the eyewitness the viewed. These facts stated: 1) three hundred dollars was stolen from the victim in the theft; 2) the accused was arrested with $328 in his possession for which he did not account; 3) the accused was arrested in the vicinity of the theft; 4) the accused did not have an alibi for the time of the theft; 5) at the time of his arrest the accused was wearing clothes similar to those described by the witnesses. After being provided this information, participant-jurors in the control condition viewed approximately seven minutes of footage showing the examination-in-chief and cross-examination of one participant-witness. Participant-jurors in the expert condition next watched approximately 15 minutes of footage showing the testimony of one eyewitness as well as the examination-in-chief and cross-examination of the eyewitness expert. The participant-jurors in the judicial instruction condition saw the testimony of an eyewitness followed by the judicial instruction. This version of the trial lasted approximately 11 minutes.

Expert Testimony

A research psychologist, who has appeared in court as an eyewitness expert, acted in the role of the expert on eyewitness identification issues. In the examination-in-chief the expert outlined his credentials, his current position, his area of expertise and his research history. He then addressed three key issues regarding eyewitness testimony: 1) the nature of memory as a reconstructive process; 2) system and estimator variables including distance, lighting, disguise, race and lineup type; and 3) the limits of the confidence-accuracy relationship as endorsed by 87% of experts (in Kassin et al., 2001). The cross-examination of the expert highlighted some of the limitations of expert psychological testimony including: 1) the reliance on mock-crime paradigms and undergraduate participants in laboratory research; 2) the questionable ecological validity of mock-crimes and mock-witnesses; 3) the probabilistic nature of psychological testimony. The expert was given no prior knowledge regarding the nature or content of the questions contained within the cross-examination. The cross-examination lasted for 2.5 minutes while the examination-in-chief was five minutes in duration.

Judicial Instruction

The role of the judge was played by the same research psychologist who provided the expert evidence. The judicial instruction lasted for three minutes and was based on the direction recommended by JCNSW (2006, s 3-020). See Table 1.1 (p. 4) for a transcript of the judge's direction.

Juror Responses

After watching the minimal trial participant-jurors completed a brief pre-deliberation - questionnaire, containing 10 questions. Participants were asked to provide some demographic information before being asked if they thought the witness had made an accurate identification (yes/no) and to rate their confidence in this decision using a 7-point scale ("not at all confident" to "extremely confident"). All participant-jurors also rated the eyewitness on the dimensions of trustworthiness, attractiveness, credibility and confidence on 7-point scales ("not at all" to "extremely"). Jurors were asked to provide their individual responses to these same questions again (excluding the demographic questions) after deliberations on the post-deliberation questionnaire. They were also asked to report the jury's verdict. In the expert and judicial conditions, the post-deliberation questionnaire also required jurors to evaluate the evidence given by the expert or the judge in terms of its credibility, clarity and utility, before completing a series of cued recall items designed to test participant-jurors memory and understanding for the instruction they heard (see Appendix H for pre- and post-deliberation questionnaires).

Procedure

As a jury of 3-7 members, participants were read the pre-trial instructions specific to their condition and were then asked to watch the video materials that were about to be played for them. A slide of the facts of the case, and the elements of the charge, was projected prominently in the room to ensure that all jurors had access to this information for the duration of the experiment. Those juries in the expert evidence and judicial instruction conditions first watched the participant-witness' testimony, followed by either the expert evidence or judicial instruction as dictated by experimental condition. They were then asked to individually complete their pre-deliberation response sheet. Those in the no instruction condition watched the participant-witness' testimony and then completed their pre-deliberation response sheets. After all jury members completed the first questionnaire, they were given a maximum of thirty minutes to reach a unanimous verdict relating to the guilt of the accused. Jurors were told that their deliberations would be tape recorded, and were instructed that if no decision was reached within the allocated time period a “hung” verdict would be recorded. Once the jury had either reached a unanimous verdict, or had “hung”, each member was asked to complete the post-deliberation questionnaire.

Results

Pre-deliberation Responses

Jurors’ Qualitative Evaluations of Eyewitnesses

Jurors used 7-point scales to evaluate the eyewitnesses on four personality and performance dimensions; attractiveness, trustworthiness, credibility and confidence. Analysis of these responses, comparing ratings assigned to witnesses who had made accurate identifications and those who had made inaccurate identifications, revealed that accurate witnesses were rated as significantly more attractive ($t_{(100.61)}$ = -2.69, p = <.05) and trustworthy ($t_{(102)}$ = -2.37, p = <.05) than inaccurate witnesses. Participant-jurors did not significantly differentiate between accurate and inaccurate eyewitnesses with respect to perceived credibility ($t_{(102)}$ = -1.09, p = .280). The jurors’ ratings of eyewitness confidence will be discussed separately as confidence emerged as a noteworthy variable in this study.

Jurors’ Qualitative Evaluations of Instructors

Participant-jurors rated the expert and the judge with regard to their credibility, and the clarity and utility of the advice they provided (see Table 6.1 for mean ratings). Comparisons of the ratings assigned to the instructors revealed no significant differences in either utility ($t_{(69)}$ = -0.75, p = .455) or clarity ($t_{(69)}$ = 0.48, p = .632), however, the expert was rated as being significantly more credible than the judge ($t_{(69)}$ = 2.13, p = <.05).

Table 6.1 : Jurors' Mean (SE) Evaluations of Instructors

	Expert (SE)	Judge (SE)
Credibility	5.87 (0.13)	5.41 (0.18)
Utility	4.79 (0.19)	5.00 (0.19)
Clarity	5.46 (0.13)	5.35 (0.22)

Juror Decision Criterion

A chi-squared test-of-independence indicated that there was a significant association between belief (i.e., the participant-jurors decision to believe the eyewitness or not) and instruction condition ($\chi^2_{(2)} = 7.68$, $p < .05$). Subsequent goodness-of-fit analyses indicated that participants in the control and judicial conditions showed a significant bias (i.e., greater than 50%) towards *believing* the eyewitness viewed (control belief, 87.9%, $\chi^2_{(1)} = 18.94$, $p < .05$; judicial belief, 75%, $\chi^2_{(1)} = 8.00$, $p < .05$), while those participant-jurors in the expert condition showed no significant bias toward either belief or disbelief (59% belief, $\chi^2_{(1)} = 1.26$, $p = .262$). This result lends support to suggestions that participant-jurors will generally tend to believe eyewitnesses, and that expert evidence can address this bias by inducing Skepticism. What remains to be seen however, is whether participant-jurors decisions to believe, at whatever rate, are significantly associated with eyewitness accuracy.

Juror Sensitivity to Eyewitness Confidence

A 3x2 between groups ANOVA was conducted to explore the effect of instruction type (judicial, expert or control) and eyewitness accuracy (accurate or inaccurate) on participant-jurors' estimates of eyewitness confidence. There was a significant main effect for eyewitness accuracy ($F_{(1,98)} = 27.80$, $p < .0005$, $\eta^2_p = 0.221$), such that witnesses who had accurately identified the perpetrator from the lineup were rated by the jurors as being significantly more confident than eyewitnesses who had made a false identification (accurate eyewitnesses, $\bar{x} = 4.69$, $\bar{\sigma} = 0.16$; inaccurate eyewitnesses, $\bar{x} = 3.57$, $\bar{\sigma} = 0.14$). The main effect of instruction type was not significant ($F_{(2,98)} = 2.17$, $p = .120$, $\eta^2_p = 0.042$), and although a significant interaction was identified ($F_{(2,98)}$ = 4.58, p < .05, η^2_p = 0.086) such that control jurors, unlike judicial and expert jurors, appear to rate accurate and inaccurate eyewitnesses as being equal in confidence (see Figure 6.1 below), post-hoc tests revealed no significant pair-wise differences.

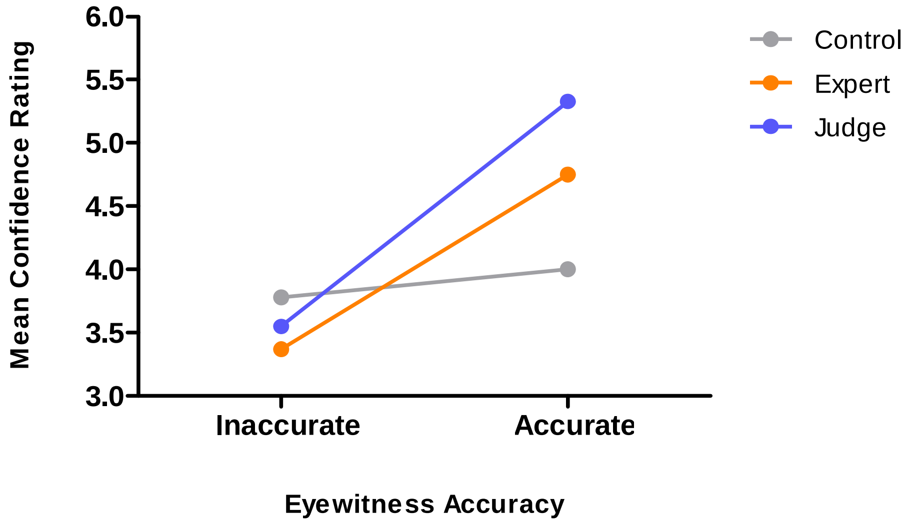

Figure 6.1 : Mean pre-deliberation confidence ratings by instruction type and eyewitness accuracy

A 3x2 between groups ANOVA was also conducted to investigate the impact of instruction type (judicial, expert or control) and belief (i.e., whether the juror believed the eyewitness or not) on participant-jurors' estimates of the confidence of the eyewitness they saw. Although instruction type had no impact on juror's perceptions of confidence ($F_{(2,98)}$ = 1.82, p = .168, η^2_p = 0.04), a significant main effect for belief was observed such that those eyewitnesses who were believed ($\bar{x}$ = 4.21, $\bar{\sigma}$ = 0.14) were rated as being significantly more confident than those who were disbelieved ($\bar{x}$ = 3.35, $\bar{\sigma}$ = 0.26; $F_{(1,98)}$ = 8.39, p < .01, η^2_p = 0.08). The interaction between instruction and belief type was significant ($F_{(2,98)}$ = 4.27, p < .05, η^2_p = 0.08), however, as before, post-hoc tests failed to identify any significant pair-wise differences. Even so, the trends present in the data suggest that those participant-jurors who heard an expert testify that confidence was not a useful predictor of identification accuracy, rated believed and disbelieved eyewitnesses as equal in confidence (see Figure 6.2 below). Thus, participant-jurors rated accurate eyewitnesses as more confident than inaccurate eyewitnesses, and rated those eyewitnesses they believed as more confident than those they chose not to believe. The significant interaction between belief and instruction type suggests that participant-jurors in the expert condition attempted to follow the advice they were given, rating believed eyewitnesses ($\overline{x}$ = 4.00, $\overline{\sigma}$ = 1.13) as equally confident as disbelieved eyewitnesses ($\overline{x}$ = 4.19, $\overline{\sigma}$ = 1.47).

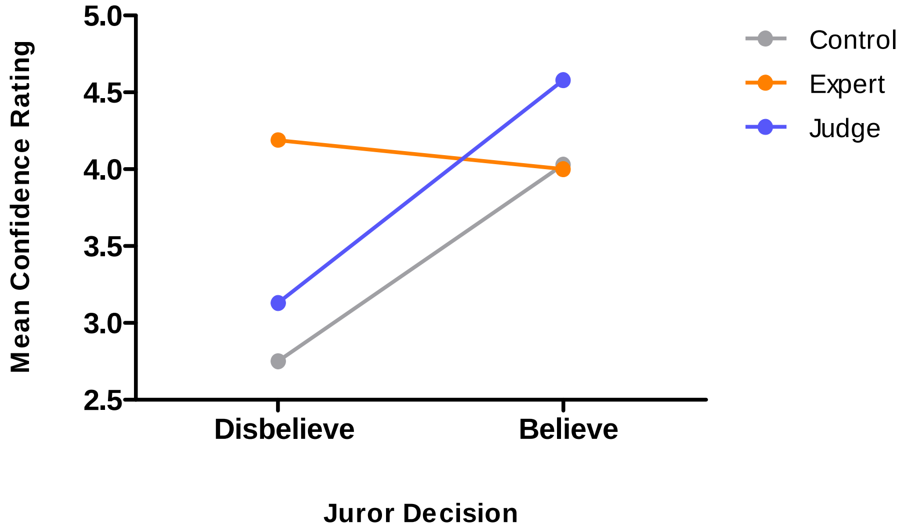

Figure 6.2 : Mean pre-deliberation confidence ratings by instruction type and juror belief decision

Considerations Affecting Belief Decisions

In order to investigate the factors considered by participant-jurors when making their decision to believe or disbelieve the eyewitness they saw, a series of (2x3) ANOVA's were conducted, each investigating the effect of belief type (whether the eyewitness was believed or not) and instruction condition (judicial, expert or control) on participant-jurors' ratings of the extent to which the three variables (eyewitness confidence, eyewitness manner and eyewitnessing conditions) influenced their judgements (see Table 6.2 for a summary of these analyses). For each of these three analyses, there were no main effects or interaction effects, indicating that participant-jurors reported being influenced by these factors to equal degrees irrespective of the instruction they were given or their belief type. This is a somewhat surprising result given that the eyewitness expert asks jurors to ignore confidence, and the judge attempted to draw the participant-jurors' attention to the witnessing and identification conditions. This analysis suggests that instruction type had no influence on the types of considerations prioritised by participant-jurors.

Table 6.2 : The Impact of Instruction and Belief type on Participant-juror Priorities

	***F*(df)**	***p***	**η_p^2**
Eyewitness Confidence			
Belief Type	3.27 (1,98)	.074	.032
Instruction Type	0.03 (2,98)	.969	.001
Interaction	0.46 (2,98)	.634	.009
Eyewitness Manner			
Belief Type	0.80 (1,98)	.374	.008
Instruction Type	0.27 (2,98)	.762	.006
Interaction	0.37 (2,98)	.694	.007
Witnessing & Identification Conditions			
Belief Type	3.29 (1,98)	.073	.032
Instruction Type	1.21 (2,98)	.303	.024
Interaction	1.40 (2,98)	.251	.028

Predictors of Juror Belief Decisions

A binary logistic regression was conducted in order to investigate whether the participant-juror ratings of the characteristics of the eyewitness predicted their decisions to believe or disbelieve the eyewitness they saw. This analysis was conducted separately for each instruction condition using the participant-jurors' rating of eyewitness credibility, confidence, attractiveness, and trustworthiness as the predictors in the model. Overall, the models created significantly predicted belief decisions in all three instruction conditions, correctly classifying 93.9% of juror decision in the control condition ($\chi^2_{(4)}$ = 14.17, p = <.01), 84.6% of decisions in the expert condition ($\chi^2_{(4)}$ = 32.00, p = <.0005), and 87.5% of decisions in the judicial condition ($\chi^2_{(4)}$ = 12.56, p = <.05, see Tables 6.3 to 6.5 for the complete models).

Significant individual predictors were identified only in the expert condition, where participant-jurors' estimates of eyewitness credibility, confidence and attractiveness were all predictive of their belief decisions. An increase in rated credibility by one response-scale unit increased the odds that a participant-juror would believe an eyewitness by a minimum of 3.6 times ($\beta = 4.88$, $p < .01$), a one unit increase in confidence decreased the odds of belief by a factor of approximately 0.21 ($\beta = -1.55$, $p < .05$), and a one unit increase in attractiveness decreasing the odds of belief by a factor of approximately 0.13 ($\beta = -2.16$, $p < .05$), when holding all other factors constant. These results provide an alternative interpretation of the significant interaction between belief and instruction type reported earlier, where it appeared that participant-jurors rated believed and disbelieved eyewitnesses to be *equal* in confidence. Rather, this analysis suggests that participant-jurors may have been confused, or misunderstood the testimony of the expert, rating believed eyewitnesses *as less* confident than disbelieved eyewitnesses. This interpretation is consistent with a "Counter-Sensitive" response by participant-jurors to expert evidence.

Table 6.3 : β (S.E.), Significance and Exp(β) for Predictor Variables in Control Condition

	β (S.E.)	***p***	**Exp(β) (95% CI)**
Credibility	4.73 (2.76)	.087	113.00 (0.50-25401.85)
Confidence	0.32 (1.20)	.788	1.38 (0.13-14.35)
Attractiveness	-1.65 (1.17)	.159	0.19 (.02-1.91)
Trustworthiness	-0.37 (1.94)	.850	0.69 (0.01-31.10)

Table 6.4 : β (S.E.), Significance and Exp(β) for Predictor Variables in Expert Condition

	β (S.E.)	***p***	**Exp(β) (95% CI)**
Credibility	4.88 (1.83)	.008	132.03 (3.66-4768.57)
Confidence	-1.55 (0.70)	.026	0.21 (0.05-0.83)
Attractiveness	-2.16 (0.94)	.022	0.12 (0.02-0.73)
Trustworthiness	0.04 (0.81)	.964	1.04 (0.21-5.08)

Table 6.5 : β (S.E.), Significance and Exp(β) for Predictor Variables in Judicial Condition

	β (S.E.)	***p***	**Exp(β) (95% CI)**
Credibility	1.14 (0.77)	.142	3.12 (0.68-14.20)
Confidence	1.46 (0.78)	.063	4.31 (0.93-20.05)
Attractiveness	-0.50 (0.32)	.315	0.60 (0.23-1.61)
Trustworthiness	-1.28 (0.76)	.091	0.28 (0.06-1.23)

Juror Sensitivity to Eyewitness Accuracy

Overall, the participant-jurors' evaluations of the eyewitness's identification were significantly less accurate than would be expected by chance alone, with participant-jurors making correct evaluations only 37.5% of the time ($\chi^2_{(1)} = 6.50$, $p < .05$). When analysed separately by condition, only participant-jurors in the expert evidence condition performed with accuracy significantly below chance levels at 25.6% ($\chi^2_{(1)} = 9.26$, $p < .001$); with control and judicial conditions performing at 39.4%, and 50.0% accuracy respectively. However, a chi-squared test-of-independence revealed that there was no significant association between instruction type and pre-deliberation evaluation accuracy ($\chi^2_{(2)} = 4.52$, $p = .104$).

Signal Detection Analysis

As noted by Wells and colleagues (1980), the task of differentiating between accurate and erroneous eyewitness identifications is analogous to a signal detection task where a signal is an accurate eyewitness and the response is the decision made by the participant-juror to believe the eyewitness or not. In accordance with Signal Detection Theory, Wells et al. (1980) went on to suggest that a “hit” represents the situation where a juror believes and accurate eyewitness, a “miss” occurs where a juror disbelieves an accurate eyewitness, a “false alarm” results where a juror believes an inaccurate eyewitness, and a “correct rejection” is the situation where a juror disbelieves an inaccurate eyewitness. However, Wells et al. (1980) do not take this analogy to its logical conclusion: that is to draw comparisons between the eyewitness constructs known as Sensitivity and Skepticism with the Signal Detection metrics of *d’* and *C*. The metric *d’* provides an index of the discriminability of the signal (in this case an accurate eyewitness), from the noise (and inaccurate eyewitness), taking into account both the hit rate and the false alarm rate. This index is directly comparable with the construct of Sensitivity which describes the situation where participant-jurors can discriminate between the relative quality of witnessing conditions (in SEO studies), or eyewitness accuracy (in SEA studies). In addition, the index known as *C* measures the response criterion adopted, that is whether participants tend to provide one of the two response types more often than the other, or both response types equally. In this context, this is a measurement of Skepticism on a continuum, where participants may show a tendency to believe all eyewitnesses (as in the case of overbelief), to disbelieve all eyewitnesses (as in Skepticism) or show no response bias by being as likely to believe as to disbelieve. Thus, Signal Detection Theory provides an alternate means by which Skepticism and Sensitivity can be measured, which incorporates both correct belief decisions, and incorrect belief decisions into overall calculations of discriminability and bias. Given this, signal detection measures were calculated for each instruction condition (see Table 6.6) in order to estimate participant-jurors’ Sensitivity to the accuracy of eyewitness identifications (*d’*) and their Skepticism (*C*).

Table 6.6 : Judgement Type as Proportion Within Instruction Condition, Observed *d'* and *C* Values.

	Judgement Type					
	Miss	**Hit**	**Correct rejection**	**False alarm**	***d'***	***C***
Control	9.1%	36.4%	3.0%	51.5%	-0.75	1.22
Expert	33.3%	17.9%	7.8%	41.0%	-1.39	0.31
Judge	6.1%	31.3%	18.8%	43.8%	0.44	0.75

When evaluating Sensitivity, greater values of *d'* indicate a greater ability to discriminate between a signal (in this case a witness who has made an accurate identification) and noise (a witness who inaccurately identified an innocent suspect). Inferential confidence intervals were calculated (Macmillan & Creelman, 2005) to test the largest observed pair-wise difference. At the $p = .05$ level, no reliable difference in Sensitivity was observed between the expert condition (*d'* 95% CI : -2.27 to -0.50) and the judicial condition (*d'* 95% CI : -0.58 to 1.47). It is important to note, however, that the negative values for *d'* estimates indicates that participant-jurors in the expert condition (and less so in the control condition) were actually responding "yes" to noise rather than signal, believing inaccurate eyewitnesses and not believing accurate eyewitnesses.

With respect to Skepticism, positive values of *C* indicate a bias toward indicating the signal is present (i.e., the jurors are biased towards believing that the eyewitness correctly identified the perpetrator – "overbelief") while negative values of *C* describe a bias toward indicating the signal is not present (i.e., jurors are Skeptical and have a bias towards disbelieving the eyewitness's identification). At the $p = .05$ level there was no significant difference in Skepticism between the control (*C* 95% CI : -1.81 to -0.62) and expert conditions (*C* 95% CI : -0.75 to 0.13).

Deliberation

Comparisons of the length of deliberations between instruction conditions revealed no significant differences in duration ($F_{(2,16)} = 0.25$, $p = .782$), with juries in the control, expert and judicial conditions deliberating for an average of 10.58, 12.31 and 14.33 minutes respectively (range : 1.58 mins to 29.98 mins).

An independent rater listened to the recordings of the juries' deliberations and coded the topics being discussed at 30 second intervals. The rater categorised the topics in the deliberation using the headings: eyewitness characteristics (e.g., eyewitness confidence, identity, accuracy and credibility), situational characteristics (e.g., lighting, timing, lineup type, disguise, exposure and quality of view), legal requirements (e.g., burden of proof and elements of the crime), evidence (e.g., non eyewitness and missing evidence) and human memory. These categories were developed after listening to a sample of deliberations. The frequency of each type of discussion was recorded for each deliberating jury. The results of this coding is summarised in Table 6.7. Inspection of the Table suggests that juries across all conditions discussed the same issues with comparable average frequency. The only exception to this general pattern indicates that participant-jurors in the expert condition were more likely to discuss human memory than participants in the control condition; however, it is unlikely that this difference is statistically reliable.

Table 6.7 : Topics of Deliberation Within Juries and Across Instruction Conditions.

	Control Condition		**Expert Condition**		**Judicial Condition**	
Topic	**% of Juries Citing**	**$\bar{x}$ (SE) Frequency Cited**	**% of Juries Citing**	**$\bar{x}$ (SE) Frequency Cited**	**% of Juries Citing**	**$\bar{x}$ (SE) Frequency Cited**
Eyewitness Characteristics	83.3	3.60 (1.44)	71.4	5.00 (1.79)	83.3	4.40 (1.44)
Evidence	83.3	3.40 (0.81)	85.7	3.33 (1.23)	100	4.33 (1.33)
Legal Requirements	100	3.00 (0.73)	71.4	3.80 (1.24)	66.7	3.50 (1.89)
Situational Characteristics	83.3	3.00 (0.63)	85.7	5.83 (1.33)	83.3	5.40 (1.86)
Human Memory	0.0	-	28.6	2.00 (0.00)	16.7	2.00 (0.00)

Post-deliberation Responses

After jurors reached a unanimous verdict, or 30 minutes had elapsed, participants were required to complete the post-deliberation questionnaire. Given that the independence of these responses has been violated by the group deliberation process, average response (means) of the members of each jury, rather than individual responses, were used in the subsequent analyses (Stevens, 2002).

Juries' Qualitative Evaluations of Eyewitnesses

Juries rated accurate and inaccurate witnesses as equal in attractiveness ($t_{(17)}$ = -2.02, p = .060), trustworthiness ($t_{(17)}$ = -0.11, p = .918) and credibility ($t_{(17)}$ = -0.58, p = .567). Jury ratings of eyewitness confidence will be discussed below.

Juries' Qualitative Evaluations of Instructors

Participant-juries in the expert and judicial conditions did not differ in their ratings of the credibility, clarity and utility of the instruction they received (credibility $t_{(11)}$ = 1.14, p = .278; clarity $t_{(11)}$ = 1.16, p = .270; utility $t_{(11)}$ = 0.05, p = .963). See Table 6.8 below for mean ratings.

Table 6.8 : Juries' Mean (SE) Evaluations of Instructors

	Expert (SE)	**Judge (SE)**
Credibility	5.53 (0.13)	5.34 (0.11)
Utility	4.58 (0.40)	4.56 (0.17)
Clarity	5.40 (0.18)	5.12 (0.16)

Juries' Recall for Instructions

On the cued recall task, jurors were asked to select one of four multiple choice options which "most accurately reflect[s]" the testimony of the expert or judge. For each question one response option was a quote taken directly from the judicial instruction or the expert's testimony (e.g. "confidence isn't necessarily a predictor of accuracy"), one option interpreted or paraphrased this quote (e.g. "nervous, uncertain witnesses can be right") and two were inaccurate accounts of what was said (e.g. "confidence is a good predictor of accuracy" or "confident witnesses are generally right"). Jurors were given one mark for an item where they selected the direct quote, half a mark for its paraphrased alternative, and a zero for either of the incorrect options. Out of a possible score of four, on average those juries in the expert condition scored 3.18 ($\bar{\sigma}$ = .06) while those from the judicial instruction condition scored on average 2.87 ($\bar{\sigma}$ = .11). An independent samples t-test revealed that jury recall for instruction was significantly better in the expert evidence condition than it was for the judicial instruction condition ($t_{(11)} = 2.59$, $p < .05$). Responses to the item relating to the confidence-accuracy relationship were of particular interest in light of participant-juror tendency to rate more confident eyewitness as believed in all but the expert condition. A total of 94.9% of *jurors*[10] accurately recalled the testimony provided by the eyewitness expert, with 87.2% of jurors selecting the verbatim quote, and 7.7% choosing its paraphrased alternative.

Jury Decision Criterion

A chi-squared test-of-independence indicated that there was no significant association between the decision reached by a jury (i.e., the jury decision to believe the eyewitness and convict; disbelieve the eyewitness and acquit; or fail to reach a unanimous decision) and instruction condition ($\chi^2_{(4)} = 7.69$, $p = .104$). Moreover, no significant association was seen between instruction type and jury decisions, when considering only unanimous verdicts (i.e., not guilty or guilty; $\chi^2_{(2)} = 3.38$, $p = .184$), or unanimous verdicts compared with undecided verdicts ($\chi^2_{(2)} = 3.89$, $p = .142$). See Table 6.9 below for the frequencies of each jury decision type in each instruction condition. It is interesting to note here that no jury reached a guilty verdict after having heard expert evidence, this is consistent with these juries having disbelieved the eyewitness, as guilt in this case was designed to hinge on the belief or disbelief of the eyewitness.

[10] This is the percent of individual jurors' responses rather than averaged jury responses.

Table 6.9 : Frequencies of Observed Jury Outcomes

	Outcome		
	Not Guilty	**Guilty**	**Hung**
Control	3	3	0
Expert	5	0	2
Judge	2	1	3

Jury Sensitivity to Eyewitness Confidence

A 3 (instruction) x 2 (accuracy) between groups ANOVA was conducted to explore the effect of instruction type and eyewitness accuracy on average jury estimates of eyewitness confidence. Identical to the pre-deliberation analysis, there was a significant main effect for eyewitness accuracy ($F_{(1,13)}$ = 7.51, p < .05, η^2_p = 0.366), such that witnesses who had accurately identified the perpetrator from the lineup were rated by juries as being significantly more confident than eyewitnesses who had identified the suspect (accurate eyewitnesses, $\bar{x}$ = 4.50, $\bar{\sigma}$ = 0.26; inaccurate eyewitnesses, $\bar{x}$ = 3.52, $\bar{\sigma}$ = 0.24). Also consistent with the pre-deliberation analysis, no significant main effect of instruction type was identified ($F_{(2,13)}$ = 0.86, p = .445, η^2_p = 0.117), but a significant interaction effect was found ($F_{(2,13)}$ = 4.61, p < .05, η^2_p = 0.415). Subsequent post-hoc analyses revealed no significant pair-wise differences, but inspection of Figure 6.3 suggests that there was a trend for juries in the control condition to rate accurate eyewitnesses as *less* confident than inaccurate eyewitnesses in contrast to juries in the expert and judicial groups.

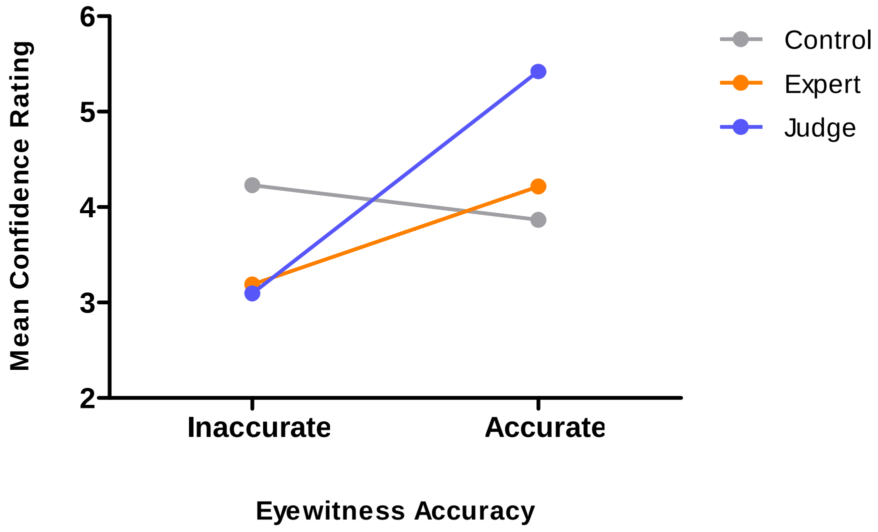

Figure 6.3 : Mean post-deliberation confidence ratings by instruction type and eyewitness accuracy

A 3 (instruction) x 2 (verdict: guilty vs. not guilty) between groups ANOVA was also conducted to investigate the impact of instruction type and unanimous verdict type (i.e., whether the jury reached a guilty or not guilty verdict) on participant-jurors' estimates of the confidence of the eyewitness they saw. Although instruction type had no impact on juror's perceptions of confidence ($F_{(2,9)}$ = 0.41, p = .675, η^2_p = 0.084), a significant main effect for verdict type was observed such that juries who convicted rated the eyewitness as more confident ($\bar{x}$ = 5.11, $\bar{\sigma}$ = 0.43) than those who did not convict ($\bar{x}$ = 3.45, $\bar{\sigma}$ = 0.25; $F_{(1,9)}$ = 10.39, p < .05, η^2_p = 0.536). The interaction between instruction and verdict type was not significant ($F_{(1,9)}$ = 4.26, p = .069, η^2_p = 0.321) and largely uninterpretable due to the fact that no jury in the expert condition chose to convict the defendant (see Figure 6.4 below). Thus, while it seems appropriate to suggest control and judicial juries were more likely to convict when they rated the eyewitness as more confident in their identification, it is difficult to interpret the relationship between confidence and verdicts amongst expert juries. Accordingly, a point-biserial correlation was conducted to ascertain if confidence was associated with hung or unanimous verdicts in the expert condition to establish if juries in this condition were simply acquitting when the eyewitness was low in confidence and failing to reach a verdict when the witness was high in confidence. The analysis revealed that there was no association between eyewitness confidence and verdict type ($\chi^2_{(1)}$ = 0.21, p = .645). Thus, it appears that expert juries did not rely on eyewitness confidence when making their decisions. This is consistent with

Sensitivity to the expert's recommendation that eyewitness confidence acts as a poor predictor of identification accuracy.

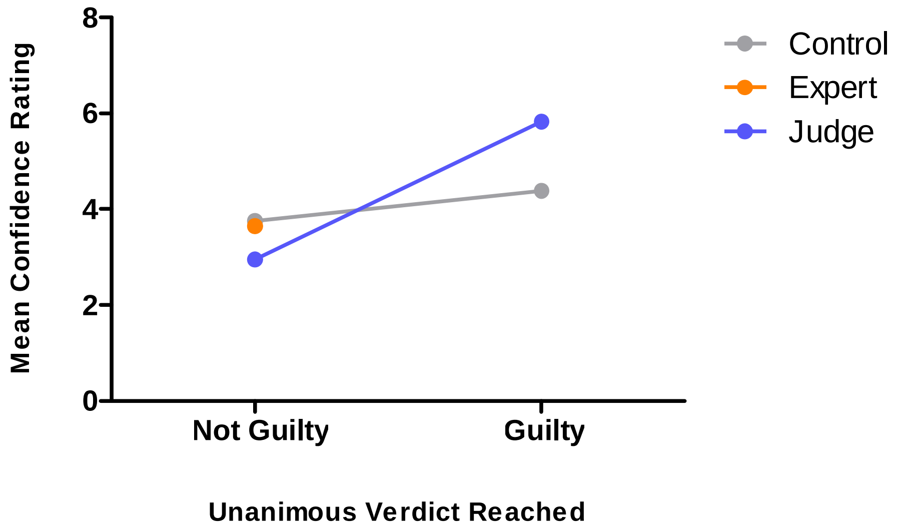

Figure 6.4 : Mean confidence ratings by instruction type and unanimous verdict type.

Considerations Affecting Jury Outcomes

Analyses investigating three factors considered by juries when formulating their verdicts (eyewitness confidence, eyewitness manner, and witnessing and identification conditions) revealed no significant main or interaction effects for any factor (see Table 6.10 for summaries of these analyses).

Table 6.10 : The Impact of Instruction and Verdict type on Jury Priorities

	***F*(df)**	***p***	**η_p^2**
Eyewitness Confidence			
Verdict Type	1.86 (2,12)	.198	.236
Instruction Type	0.53 (2,12)	.601	.081
Interaction	1.78 (2,12)	.211	.228
Eyewitness Manner			
Verdict Type	2.89 (2,12)	.094	.325
Instruction Type	1.81 (2,12)	.205	.232
Interaction	3.36 (2,12)	.069	.359
Witnessing & Identification Conditions			
Verdict Type	2.80 (2,12)	.101	.318
Instruction Type	3.36 (2,12)	.069	.359
Interaction	1.46 (2,12)	.271	.196

Jury Verdict Accuracy

Overall, the decisions reached by jurors could not be tested statistically due to the small numbers of observations of each type in each instruction condition. The verdicts reached by juries in the control, expert and judicial conditions were objectively accurate 33.3%, 40.0% and 100% of the time respectively when considering only *unanimous* verdicts (i.e., excluding hung juries). The percentages of accurate verdicts resulting from *all deliberations* were 33.3%, 28.6% and 50% for the control, expert and judicial conditions respectively (see Table 6.11 below). It is interesting to note that juries in the judicial condition reached a correct verdict each time, making no wrongful convictions or wrongful acquittals. This strong performance is likely associated with the apparent tendency for groups in this condition to hang rather than err.

Table 6.11 : Correct Jury Verdicts by Instruction Condition

	Correct Jury Verdicts	
	% All Verdicts	**% Unanimous Verdicts Only**
Control	33.3%	33.3%
Expert	28.6%	40.0%
Judge	50%	100%

Discussion

Before discussing the post-deliberation responses made by juries, it is appropriate to examine the data collected from jury members before they commenced their deliberations. These data provide us with our first impressions regarding the impact of various types of eyewitness instructions on participant-juror discrimination accuracy.

Pre-deliberation Responses

Independent Effects of Instruction

Consistent with previous research suggesting that participant-jurors tend to “overbelieve” eyewitness evidence (Cutler, Penrod & Stuve, 1988; Loftus, 1979; Loftus & Monahan, 1980; Penrod & Cutler, 1999; Wells, 1980), jurors in the control condition displayed a significant preference for accepting the testimony of eyewitnesses, with jurors believing the eyewitness they saw 87.9% of the time. There was no evidence to suggest that standard judicial instructions resulted in any change in participant-juror response, as jurors in the judicial condition were also significantly biased towards believing eyewitnesses; doing so 75% of the time. Thus, rather than *inducing* Skepticism as concluded by Katzev and Wishart (1985) and Ramirez et al. (1996), this result suggests that standard judicial instruction instead did not alter the pre-existing tendency that participant-jurors have for believing the eyewitness before them.

Some evidence was found to indicate that expert testimony altered participant-jurors use of eyewitness confidence as a determinant of accuracy. In particular, the significant interaction between belief and instruction types suggests that participant-jurors in the expert condition were less prone to this than were participants in the judicial and control conditions. Indeed, regression analyses predicting the belief decisions of participant-jurors on the basis of their estimates of the eyewitnesses' confidence, credibility, attractiveness and trustworthiness revealed that the belief decisions of members of the expert condition were *negatively* influenced by estimates of the eyewitnesses' confidence. That is, participant-jurors in the expert evidence condition were more likely to believe those eyewitnesses they rated lower in confidence than those eyewitnesses they rated higher in confidence. This indicates that expert evidence caused participant-jurors to respond to eyewitness confidence in a manner different to that seen in other conditions. Yet the effect of expert testimony is undermined somewhat by a failure to identify any significant differences in the amount of influence attributed to eyewitness confidence across instruction conditions, with participant-jurors in all conditions assigning comparable levels of influence to eyewitness confidence. Even so, further evidence of the influence of expert evidence on participant-juror decision-making can be found through an examination of the belief criterion adopted by participant-jurors. Unlike in the control and judicial conditions, members of expert evidence condition were *as likely* to believe the testimony of an eyewitness as they were to disbelieve their evidence. The inference from this result is: that expert evidence *reduced* the pre-existing participant-juror bias towards belief. This is supported by the significant association identified between instruction type and belief type. Thus, expert evidence appears to have caused *both* Sensitivity to Expert Opinion (although not in the desired direction), by altering participant-jurors use of confidence information, *and* Skepticism by changing the rate of belief decisions. While some might suggest that the Skepticism observed here poses no cause for alarm due to the fact it was observed together with Sensitivity to Expert Opinons, the ultimate value of the experts' testimony can only be ascertained with regard to participant-jurors' Sensitivity to Eyewitness Accuracy. This finding emphasizes the importance of the use of direct methodologies.

Relative Effects of Instruction

As the first direct comparison between the effects of adversarial expert testimony and judicial instruction on participant-juror discrimination performance, this analysis provides little evidence to suggest that the judicial instruction produces significantly poorer outcomes than expert evidence as has previously been suggested (Greene & Loftus, 1984; Leippe, 1995; Pezdek, 2007). All that can be said, thus far, is that judicial instruction did little to change the belief criterion adopted by participant-jurors (i.e., their tendency to overbelieve), and had no discernable influence on the use of the confidence-accuracy heuristic: Expert evidence on the other hand appears to have induced Skepticism and altered participant-juror treatment of eyewitness confidence expressions, although not in the desired way. On this basis, it is difficult to advocate for either the expert or the judge.

Instruction Type & Sensitivity to Eyewitness Accuracy

Overall, participant-jurors demonstrated a marked inability to believe those witnesses who were correct and disbelieve those witnesses who were incorrect in their identifications. That is not to say, however, that they were unable to discriminate between accurate and inaccurate eyewitnesses. The data instead suggests that participant-jurors were capable of differentiating *between* the two types of eyewitnesses, but were incapable of correctly determining which type of eyewitness they saw; regarding accurate eyewitnesses as inaccurate and vice versa. That is, they could separate the eyewitnesses into two groups (accurate and inaccurate), but were confused about how to respond to each. This was clearly evident in the fact that overall participant-jurors made accurate determinations of eyewitness accuracy only 37.5% of the time, a figure that is *significantly lower* than chance performance.

Moreover, although there was no significant association between instruction type and juror discrimination accuracy, post-hoc tests showed participant-jurors from the expert condition to be the only group performing significantly below chance with 25.6% accuracy, while those in the control and judicial condition had accuracy rates of 39.4%, and 50.0% respectively. This suggests that participant-jurors in the expert evidence condition were making systematic, rather than random errors, something which cannot be achieved in the absence of the ability to discriminate between accurate and inaccurate eyewitnesses. Irrespective of this, these accuracy rates are in keeping with previous research indicating that participant-juror evaluation accuracy may at times be as low as 25% depending on the accuracy of the eyewitness (Wells & Leippe, 1981; Wells, Lindsay, & Ferguson, 1979), but they also call into question the validity of any suggestions that expert evidence results in more accurate resolutions to criminal trials than can be achieved in the absence of this evidence or with judicial instruction.

Predicting Sensitivity to Eyewitness Accuracy

Based upon indirect tests indicating participant-juror “Counter-Sensitivity” to expert opinions and Skepticism of eyewitness testimony, one could readily anticipate that the experts’ evidence would result in less than perfect discrimination accuracy – as was observed. Indeed, it is reasonable to predict that jurors misapplying expert advice (by allowing expressions of confidence to influence accuracy evaluations), will likely perform more poorly than those correctly applying an expert’s advice (to ignore expressions of confidence). However, there is one crucial element missing from this analysis which impacts upon the quality of possible inferences; that is the fact that participant-jurors in the expert condition rated *accurate eyewitnesses* as being significantly more confident than *inaccurate eyewitnesses.* Thus, simply knowing that jurors had misapplied the experts evidence (i.e., they were Responsive to Experts Evidence but not appropriately Sensitive to the Expert’s Opinion), was not sufficient for someone to accurately predict, or infer the *magnitude* of the resulting impairment in Sensitivity to Eyewitness Accuracy demonstrated amongst jurors. Once it is known that participant-jurors were rating accurate eyewitnesses as significantly more confident than inaccurate ones, then it is possible to estimate how poor participant-juror discrimination performance would be in the expert condition. Without this full understanding of how the participant-jurors are responding to *real* eyewitnesses, inferences based upon knowledge regarding participant-jurors Sensitivity to Expert Opinion could only have grossly underestimated the impact of their failure to accurately follow the expert advice. Specifically, if one assumed what the expert said was “correct”; that there was no association between confidence and accuracy, it would not have been so concerning that participant-jurors considered these variables to be negatively associated. That is, if there is no association between confidence and accuracy, then it doesn’t matter if participant-jurors treat confidence as a negative or a positive predictor. Thus, logically, their erroneous beliefs after expert evidence would be no more concerning than their erroneous beliefs before hearing expert evidence, either way they would be using an imperfect predictor of accuracy to guide their decisions. The real magnitude of the problem only becomes clear when one considers that participant-jurors perceived accurate eyewitnesses to be significantly more confident, and that as a result of responding to the experts' evidence inappropriately (by disbelieving more confident eyewitnesses), jurors began to apply a rule which *systematically impaired their discrimination accuracy*, rather than improving it. Thus, knowing about REE and SEO did not provide sufficient information to estimate how these effects would impact upon participant-jurors ability to believe accurate eyewitnesses and disbelieve inaccurate eyewitnesses.

Deliberation Effects

Consistent with the findings of Maass et al. (1985), there was no evidence to suggest that expert evidence served to significantly increase jury deliberations times in comparison to the no instruction control condition or the judicial instruction condition. This result is in direct contrast with Loftus (1980) where a significant increase in deliberation time was associated with the presence of expert testimony. Moreover, summaries of the topics discussed by deliberating juries in each condition indicated few differences in the content covered across groups. Indeed, the only noteworthy difference related to discussions of human memory, which although raised by a few of the juries in the expert condition, and one jury in the judicial condition, was not raised at all in the control condition. Thus, it seems that judicial instruction and expert evidence induced some juries to consider the nature of human memory as a factor in their decision-making, where they otherwise might not have discussed this issue. Interestingly, there did not appear to be a difference between expert evidence and judicial instruction conditions in either the number of juries mentioning human memory, or the frequency with which this topic was raised. Thus, if one considers juror education regarding human memory to be a key feature of expert evidence, this data must give some pause, suggesting as it does that expert evidence does not induce either widespread or in-depth discussion of the topic.

Post-deliberation Responses

Participant-jurors, irrespective of instruction condition, showed themselves to be limited in their capacity as individual evaluators of eyewitness evidence. It is feasible, however, that the deliberation process might serve to improve their performance. The evidence speaking to this issue is discussed in detail below.

Independent Effects of Instruction

Juries in the judicial instruction condition were as likely to reach a unanimous verdict as they were not to. Of those juries reaching a unanimous verdict, two found the defendant not guilty, indicating that they disbelieved the eyewitness, while the remaining jury chose to convict the defendant, indicating that they believed the eyewitness's identification was accurate. Although the numbers of each verdict type are insufficient to allow statistical analysis, there is no indication that these juries showed any bias toward disbelieving the eyewitness as a result of judicial instruction. Therefore, no evidence was found to suggest that juries receiving judicial instruction show Skepticism in either their post-deliberation verdicts, or their pre-deliberation belief decisions. This type of instruction does, however, appear to have increased the likelihood that a jury would hang, as all juries in the control condition, unlike the judicial condition, were able to reach unanimous verdicts (evenly split between guilty and not guilty verdicts). Therefore, while judicial instruction did appear to increase juror uncertainty, no verdict bias was evident in the judicial and control conditions. This balance in verdicts is somewhat surprising given that these individual jurors showed a significant bias towards believing the eyewitness in their pre-deliberation responses. Thus, it appears that the deliberation process itself may act as an important corrective for the tendency to overbelieve eyewitness testimony.

Verdicts reached by juries in the expert evidence condition appear to have been substantially less balanced than those in the judicial and control conditions. Five of the seven juries in the expert condition indicated that they did not have significant faith in the accuracy of the eyewitnesses' identification to convict the defendant, while the remaining two juries failed to reach unanimous verdicts. Thus, consistent with Maass et al. (1985) expert evidence not only appears to have increased the frequency of "unsuccessful" deliberations, but also appears to have induced a bias against accepting the eyewitnesses evidence. There may however be a Sensitivity effect embedded in these apparently Skeptical responses. Specifically, jury decisions to acquit or hang were not associated with rating of eyewitness confidence. Thus, although there were no convictions, or eyewitnesses who were believed unanimously, jurors did appear to be Sensitive to the Expert's Opinion to the extent that they should ignore confidence. Moreover, this Sensitivity appears to have been gained through the deliberation process, as pre-deliberation jurors were rating those eyewitnesses believed as less confident than those disbelieved. This interpretation is further supported by post-deliberation recall for the experts' testimony regarding confidence and accuracy which was very good, with only 5.1% of all jurors misremembering the experts' evidence on the subject. Yet, despite all this, participant-jurors in the expert condition reported eyewitness confidence as equally influential as those in the other instruction conditions. It is therefore difficult to ascertain with any certainty if juries exhibited Sensitivity to the Expert's Opinion in addition to their tendency to disbelieve all eyewitnesses.

Relative Effects of Instruction

The significant pre-deliberation difference observed in the rated credibility of the expert and the judge disappeared in post-deliberation ratings. On average, juries rated the expert and the judge as equally credible, useful and clear; however, juries did show significantly better recall for the testimony of the expert than the judge. Moreover, to the extent that one considers the revised judicial instruction used by Greene (1988) to be akin to expert evidence, the verdicts obtained here are consistent with the finding that those jurors who heard the revised (expert-like) instruction were less likely to convict than those provided the standard judicial instruction. Thus, there appear to be some, although not many, qualitative differences between the influence of judicial instruction and expert evidence on jury evaluations and verdicts rendered. Although recall for the judicial instruction was significantly worse than recall for the expert evidence, this finding is only of interest: a) to the extent that the outcome of interest is jury education or knowledge for the instruction provided; or b) because it shows that any failure to find SEA was not due to a failure for the jury to understand or remember the instruction provided. This fact is uninformative with regard to the relative effects of these different instruction types on the accuracy of jury verdicts.

Instruction Type & Jury Verdict Accuracy

The accuracy of jury verdicts was analysed excluding the hung juries. When considering only unanimous verdicts, the accuracy rates (i.e., percent of juries who reached a "guilty" verdict after seeing the evidence of an accurate eyewitness and a "not guilty" verdict after seeing evidence of an inaccurate eyewitness) in the control, expert and judicial conditions were 33.3%, 40.0% and 100% respectively; with the accuracy rates overall falling equal to chance performance at 50%. Consistent with the findings of Lindsay et al. (1989) there was no association between conviction rates and eyewitness accuracy in the control condition, indicating that mock-juries were unable to discriminate between accurate and inaccurate eyewitness identification testimony. In contrast, one might be tempted to praise judicial instruction as a cure-all for erroneous convictions considering the perfect discrimination accuracy observed. A more measured analysis must, however, take account of the number of observations per condition, and the impact of the high number of hung juries in the judicial condition. Indeed, when hung juries are considered together with unanimous verdicts, performance in the judicial condition drops to chance levels. Even so, it does not feel entirely satisfactory to attribute the performance observed in the judicial condition to chance alone. Juries in this condition reached accurate verdicts both when they convicted and when they acquitted, suggesting the accuracy of their verdicts was not a consequence of a response bias. Moreover, the high rate of hung juries was largest seen in all instruction conditions, suggesting that jurors did not acquiesce to the demands of their peers, instead preferring not to reach a unanimous verdict where doubts persisted. Thus, while we must be careful not to be too swayed by jury performance in the judicial condition, we can be fairly confident that the judicial instruction did not have the negative impact psychologists have suggested in the past.

It is equally tempting to suggest that expert evidence had no positive impact on jury verdict accuracy, and indeed the low accuracy rates do little to challenge this interpretation. However, the positive argument posed in defence of the judicial instruction can be applied almost equally to the expert condition. Worthy of note is the fact that juries in the expert evidence condition appear to have taken a biased approach to the evaluation of eyewitness identification, tending to disbelieve eyewitness testimony rather than convict where the eyewitness was accurate. Thus, it is with some confidence that we can suggest that although jury accuracy in the expert condition was likely underestimated, there does appear to be clear evidence of a response bias which can inhibit jury verdict accuracy.

Limitations

The conclusions reached in this experiment may be limited by elements of both the design and analysis adopted in this study.

Design

Firstly, and most importantly, this study did not seek to systematically vary any element of the witnessing and identification conditions, and therefore was not originally designed to detect Sensitivity to Expert Opinions. This was simply because the intention of this first experimental study was to focus on the relative effects of judicial instruction and eyewitness expert testimony on participant-juror Sensitivity to Eyewitness Accuracy, as measured directly. Even so, it was possible to measure participant-juror Sensitivity to Expert Opinions because the expert did testify "that eyewitness confidence was not a good predictor of identification accuracy". Thus, it was possible to observe Sensitivity to the Expert's Opinion through analysis of the confidence ratings attributed to believed and disbelieved eyewitnesses. Specifically, participant-jurors who were Sensitive to Expert Opinion ought to have rated believed and disbelieved eyewitnesses as equal in confidence rather than relying on confidence as a predictor of accuracy (see Cutler et al., 1989a; Cutler et al., 1989b, 1990b; Fox & Walters, 1986; Wells et al., 1980). Accordingly, it was possible to investigate the stated aims of the experiment, whilst also engaging in some preliminary investigations regarding the correspondence between direct and indirect measures of SEA.

Secondly, the small number of eyewitnesses available for evaluation by participant-jurors may also have limited the quality of the inferences being made, and the validity of the tests being conducted, in this experiment. Specifically, there appeared to be an association between eyewitness confidence and eyewitness accuracy, which although not statistically significant, was responded to by participant-jurors none the less. Given the small numbers of eyewitnesses, however, it is inappropriate to suggest that these eyewitnesses, or the confidence-accuracy association observed by participants, was representative of broader, more realistic samples. Indeed, there may have been some unseen selection pressure in operation, which resulted in an artificial association between confidence and accuracy amongst these eyewitnesses. Thus, one must be cautious when generalising from the observed effects in this study to likely effects in the real world, as these eyewitnesses may differ in some, or many ways from real eyewitnesses.

That aside, these eyewitnesses were not knowingly selected on the basis of any quality apart from the identification decision made. Thus, while certainty regarding the representativeness of these eyewitnesses is lacking, these eyewitnesses are *real eyewitnesses*, who made accurate or inaccurate identifications. Thus, there is value in evaluating the effects of different instruction types with regard to these eyewitnesses. Moreover, even if it is established that these eyewitnesses are wholly unrepresentative of eyewitnesses in general, there is still value in examining the effects of instruction on their evaluation, as probabilistically speaking, jurors will at times be confronted, in real life, with eyewitness testimony which is *inconsistent* with the available empirical evidence; and it is these precise instances that provide jurors with the opportunity to show that they can correctly apply expert evidence, or judicial instruction, in the individual instance to reach an accurate determination. As previously discussed, this is a contingency which must be provided for, and evaluated, if the effects of expert evidence and judicial instruction are to be thoroughly understood.

The final limitation associated with the design of this experiment relates to the questionnaires constructed for completion pre- and post-deliberation. Specifically, due to the fact that participant-jurors were not prompted to recall the expert's testimony or judicial instruction prior to their deliberations, it is not possible to ascertain if jurors entered the deliberation process with a thorough and accurate understanding, or at least recall, for the instructions they were given. Thus, it is unclear whether participant-jurors misremembered, or misunderstood the expert's evidence before deliberations, causing them to treat confidence as a *negative* predictor of eyewitness identification accuracy rather than a *neutral* predictor. Either way, this appears to have been resolved post-deliberation.

Analysis

The analysis of post-deliberation jury responses was limited by several factors. The first relates to the varied numbers in juries both within and across conditions. Given that real juries in Australia are comprised of 12 jurors, it is difficult to ascertain if the decisions reached by juries of three to seven members are likely to be representative of those decisions reached by real juries. Thus, some caution must be exercised when generalising from the verdicts observed in this study to the likely effects of different types of instruction in forensic contexts.

Another associated limitation relates to the statistical power of the jury level analysis. In particular, the fact that there were only between five and seven juries in each instruction condition meant that it was difficult to conduct a meaningful statistical analysis of the types of verdicts rendered, and to make comparisons across conditions. Future research investigating deliberated verdicts will require larger numbers of juries in order to reach valid conclusions about reliable differences between the effects of instruction types.

The final limitation relating to the analysis conducted in this experiment is associated with the coding procedure adopted to investigate the content of jury deliberations. Instead of listening to the entire deliberation and coding each topic addressed by the jury, the topic of discussion of was coded at 30-second intervals. This practice may have misrepresented the range and frequency of the topics discussed, although there is no reason to expect that this process might have influenced juries from different instruction conditions in different ways.

Conclusions

Overall, this study provides evidence to suggest: a) that judicial instruction did not exert an undesirable influence of participant-juror or jury decision-making. Specifically, although participants showed significant overbelief in both the control and judicial conditions pre-deliberation, the deliberation process appears to have resolved this tendency in both groups, producing an even mix of guilty and not guilty verdicts; conversely, b) there is also evidence to suggest that expert testimony caused participant-jurors and juries to adopt a more stringent criterion for belief both pre- and post-deliberation; c) the apparent Skepticism evident in the expert condition did not significantly improve the accuracy of participant-juror or jury evaluations, in fact, participants in the expert condition performed with discrimination accuracy significantly below chance levels, while also showing a tendency to hang. Thus, the totality of this evidence leaves in doubt the validity of the cited conclusion that eyewitness expert evidence is superior to judicial instruction.

Chapter 7

The following two studies (reported in Chapters 7 and 8) were designed to systematically address an issue inadvertently raised in Experiment 1. In that study, the eyewitness expert provided participant-jurors with unhelpful information regarding the confidence-accuracy correlation for that sample. This occurred because participant-jurors actually perceived accurate eyewitnesses to be significantly more confident than inaccurate eyewitnesses even though the relationship between these two variables was not statistically significant. Thus, the expert's recommendation that participants ignore confidence was unhelpful given the context. Although it is inevitable that an expert's probabilistic evidence will be objectively inaccurate in some *individual cases*, it is not clear what impact such erroneous information will have on juror decision-making. Moreover, it is not known if, in Experiment 1, the expert's testimony was unhelpful simply by chance, or because there was some selection pressure in operation which generated an unrepresentative relationship in our sample. Thus, that experiment may not have provided either a fair test of expert evidence, or a fair test of its impact when given in error. Experiments 2 and 3 therefore attempt to investigate the impact of expert evidence regarding the confidence-accuracy correlation in two contexts: 1) in a sample of eyewitnesses for whom confidence and accuracy are uncorrelated, and; 2) in a sample of eyewitnesses where confidence and accuracy are significantly correlated. In these studies the accuracy of the expert's testimony will be manipulated to construct a "best case" scenario where the expert provides accurate evidence, and a "worst case" scenario where the expert is providing erroneous or objectively unhelpful evidence. This permits a direct exploration of the impact of qualitatively varied expert evidence on participant-juror Sensitivity to Eyewitness Accuracy. This approach also provides the greatest capacity to identify an effect of eyewitness expert evidence by comparing its impact when operating at functional extremes. Finally, this investigation couples the investigation of these expert evidence boundary conditions with a direct comparison of the impact of the standard judicial instruction.

It is predicted that juror Sensitivity to Eyewitness Accuracy will be significantly greater in instances where the expert provides accurate evidence than in those where erroneous evidence is provided. Furthermore, since the standard judicial instruction makes no directional predictions regarding the confidence-accuracy relationship, and therefore neither accurately or inaccurately describes the ground truth for the sample (Geiselman, 1994), the jurors in the judicial condition are expected to perform significantly worse than those who hear accurate expert evidence and significantly better than those who hear inaccurate expert evidence.

Experiment 2 – Impact of Accurate Expert Evidence on Juror Sensitivity to Eyewitness Accuracy

The eyewitness testimony used in this study was collected using a new eyewitnessing Protocol (II), which is described in full in Appendix I. This testimony was collected from participant-witnesses who had viewed a crime scenario under “poor” conditions before making either a suspect or perpetrator identification. Importantly, the expressions of confidence made by these eyewitnesses were not statistically associated with the accuracy of their identifications. Thus, this study investigates the effect of accurate expert evidence on participant-juror Sensitivity to Eyewitness Accuracy by providing participant-jurors with expert evidence, which correctly indicates that an eyewitness’s confidence is unrelated to the accuracy of their identification.

Method

Overview

Experiments utilising real eyewitness designs are composed of two distinct stages. Stage 1. pertains to the collection of real eyewitness testimony and involves: a) the construction of a crime video; b) eyewitness identifications; and c) eyewitness interviews. In Stage 2. participant-jurors are presented with these eyewitness interviews and are required to evaluate the accuracy of the identification made. For ease of reading only the second of these two stages has been reported in full in the body of this thesis (see below), specific information relating to the collection and analysis of the real eyewitness testimony is reported in full in Appendix I (p. 262).

Participant-Jurors

Two hundred and twenty-eight undergraduate psychology students (142 female, 85 male and one unspecified) ranging in age from 17 to 43 years ($\overline{x}$ = 20.1 yrs, $\overline{\sigma}$ = 0.21) from the University of New South Wales participated in the role of jurors during scheduled psychology tutorials.

Design

A 2(witness type: accurate vs inaccurate) by 3 (instruction type: judicial instruction, expert evidence, no instruction control) factorial design was employed to investigate the impact of instruction on participant-juror discrimination accuracy. Each participant-juror viewed one participant-witness who was randomly selected. Each of these participant-witnesses appeared with equal frequency in all three instruction conditions.

Materials

The pre-trial instruction, expert testimony and judicial instruction used in Experiment 1 were also used in this study (see p. 104 and Appendices D and E for details.

Eyewitness Testimony

The eyewitness evidence protocol used in this study (see Appendix I) resulted in seven eyewitness interviews from the “poor” witnessing condition (4 accurate and 3 inaccurate). Six of these seven interviews were used as stimuli for this study. The testimony of one of the accurate eyewitnesses was not included in this experiment in order to balance the number of accurate and inaccurate eyewitness identifications presented to participant-jurors. The eyewitnesses included in the study were selected on the basis of their order of participation (i.e. the testimony of the first three accurate eyewitnesses was used). The confidence-accuracy correlation for these six eyewitnesses was .63, however, this was not statistically significant (p = .117).

The Minimal Trial

Each participant-juror saw the examination-in-chief and cross-examination of one eyewitness. For the participant-jurors in the judicial instruction condition, the testimony of the eyewitness was followed by a judicial instruction. This version of the trial lasted approximately 11 minutes. Participant-jurors in the expert condition watched approximately 15 minutes of footage showing the testimony of one eyewitness followed by the examination-in-chief and cross-examination of an eyewitness expert. Participant-jurors in the control condition viewed approximately seven minutes of video showing only the eyewitness testimony.

Expert Testimony

The eyewitness expert testimony used in Experiment 1 was also used in this study (see p. 105 for complete details). Importantly, in his testimony the expert described the limits of the confidence-accuracy relationship as endorsed by 87% of experts (in Kassin et al., 2001). Given that confidence and accuracy were not correlated for participant-witnesses who viewed the “poor” crime scenario, the expert provided accurate testimony for the sample.

Juror Responses

After watching the minimal trial participant-jurors completed a brief questionnaire, containing either 31 questions (in the expert and judicial conditions) or 23 questions (in the no-instruction condition). Common to the three variants of the questionnaire, participants were asked to provide demographic information before being asked if they thought the witness had made an accurate identification and to rate their confidence in this decision using a 7-point scale (“not at all confident” to “extremely confident”). All participant-jurors also rated the eyewitness on the dimensions of trustworthiness, believability, attractiveness, credibility and confidence using a 7-point scale. In addition, participant-jurors from the expert and judicial conditions evaluated the information provided by the expert or the judge (on a 7-point scale) in terms of credibility, clarity and utility. Finally, a series of multiple choice questions assessed participant-juror memory for the testimony of the expert or the judge (see Appendix J for the complete questionnaire).

Procedure

First year psychology students were tested in groups as part of a tutorial activity. Each of the 25 tutorial groups were randomly assigned to one of the three instruction conditions (expert, judicial or control). In each group, each participant-juror sat at a computer containing the trial materials as video files. Allocation of participant-jurors to witness was random. The participant-witness videos varied both within and across tutorial groups such that all witnesses were shown under all instruction conditions and in all tutorial groups.

Participant-jurors were read the pre-trial instructions specific to their condition and were then asked to open the participant-witness video file on their computer. After participants located the file, they were asked to put on their headphones and follow the instructions provided in the video. Those in the expert evidence and judicial instruction conditions first watched the participant-witness' testimony, followed by accurate expert evidence or the judicial instruction as dictated by experimental condition. They were then asked to complete their response sheet. Those in the no-instruction condition watched the participant-witness' testimony, participated in a 5-minute filler task (participant-jurors memorised strings of letters and numbers, between seven and 14 items long, to recall after a 30 second delay), and then completed their response sheets.

Results

Jurors' Qualitative Evaluations of Eyewitnesses

Jurors were asked to evaluate the eyewitness they saw on seven personality and performance dimensions using a 7-point scale. These ratings revealed that accurate and inaccurate witnesses were not seen to differ significantly on the dimensions of credibility ($t_{(226)} = -0.58$, $p = .564$), likeability ($t_{(209)} = -0.69$, $p = .493$), trustworthiness ($t_{(209)} = -0.57$, $p = .570$) or anger ($t_{(210)} = -0.12$, $p = .902$). However, accurate eyewitnesses were rated as significantly higher in accuracy ($t_{(223.51)}$[11] $= -2.13$, $p < .05$), and attractiveness ($t_{(222.69)}$[11] $= -4.87$, $p < .0005$). The results for the remaining dimension, confidence, will be discussed ahead.

Jurors' Qualitative Evaluations of Instructors

Participant-jurors were asked to evaluate the information provided by the judge or expert in terms of its credibility, utility and clarity (see Table 7.1 for Mean qualitative evaluations of each instructor). Comparisons between the average ratings of the expert and the judge on these three dimensions revealed no significant differences for either

[11] Levene's Test for Equality of Variance was violated in this comparison. Accordingly the t-value reported above does not assume equal variances.

utility ($t_{(141.49)}$[11] = -0.44, p = .693) or clarity ($t_{(153)}$ = 1.24, p = .216), however the expert was rated as significantly more credible than the judge ($t_{(153)}$ = 2.39, p < .05).

Table 7.1 : Jurors' Mean (SE) Evaluations of Instructors

	Expert (SE)	**Judge (SE)**
Credibility	5.61 (0.12)	5.26 (0.09)
Utility	4.87 (0.13)	4.95 (0.13)
Clarity	5.72 (0.13)	5.52 (0.10)

Jurors' Recall for Instructions

As in Experiment 1, participant-jurors were asked to select which of four multiple choice options "most accurately reflects" what the expert or judge said. Jurors were given one mark for an item where they selected the direct quote, half a mark for the paraphrased alternative, and a zero for either of the two incorrect options. Out of a possible score of four, on average those jurors in the expert condition scored 3.30 ($\bar{\sigma}$ = .07) while those in the judicial instruction condition scored on average 2.78 ($\bar{\sigma}$ = .09). An independent samples t-test revealed that jurors in the expert condition had significantly higher recall accuracy than those jurors in the judicial instruction condition ($t_{(151.83)}$[11] = 4.77, p < .0005). For the specific item relating to the confidence-accuracy relationship, 100% of jurors were accurate in their recollection of the expert's testimony, with 87% of jurors selecting the verbatim quote, and 13% choosing its paraphrased alternative.

Juror Decision Criterion

A chi-squared goodness-of-fit test indicated that the proportion of participant-jurors who believed eyewitnesses did not significantly differ from 50%, either when collapsed across conditions ($\chi^2_{(1)}$ = .07, p = .791), or within instruction conditions (control, 46.6% belief, $\chi^2_{(1)}$ = .34, p = .558; expert, 44.4% belief $\chi^2_{(1)}$ = .67, p = .414; or judge, 53.5% belief, $\chi^2_{(1)}$ = .49, p = .486). Together these results suggest that neither expert evidence nor judicial instruction caused a change in the decision criterion adopted by participant-jurors.

Juror Sensitivity to Eyewitness Confidence

A 3x2 between groups ANOVA was conducted to explore the effect of instruction type and eyewitness accuracy on participant-jurors' estimates of the confidence of the eyewitness they saw. There was a significant main effect of witness accuracy ($F_{(1, 222)}$ = 40.17, p <.0005, η^2_p = 0.153) and of instruction type ($F_{(2, 222)}$ = 5.58, p <.005, η^2_p = 0.048; control $\bar{x}$ = 4.43, $\bar{\sigma}$ = 0.16; expert $\bar{x}$ = 3.73, $\bar{\sigma}$ = 0.18; judge $\bar{x}$ = 4.42, $\bar{\sigma}$ = 0.13), but no significant interaction ($F_{(2, 222)}$ = 0.29, p = .749, η^2_p = 0.003). Overall, accurate eyewitnesses ($\bar{x}$ = 4.77, $\bar{\sigma}$ = 0.13) were rated as significantly more confident than inaccurate eyewitnesses ($\bar{x}$ = 3.61, $\bar{\sigma}$ = 0.13), while those participant-jurors in the expert condition were significantly more conservative in their estimates of eyewitness confidence than were participant-jurors in the judicial condition (see Figure 7.1 below).

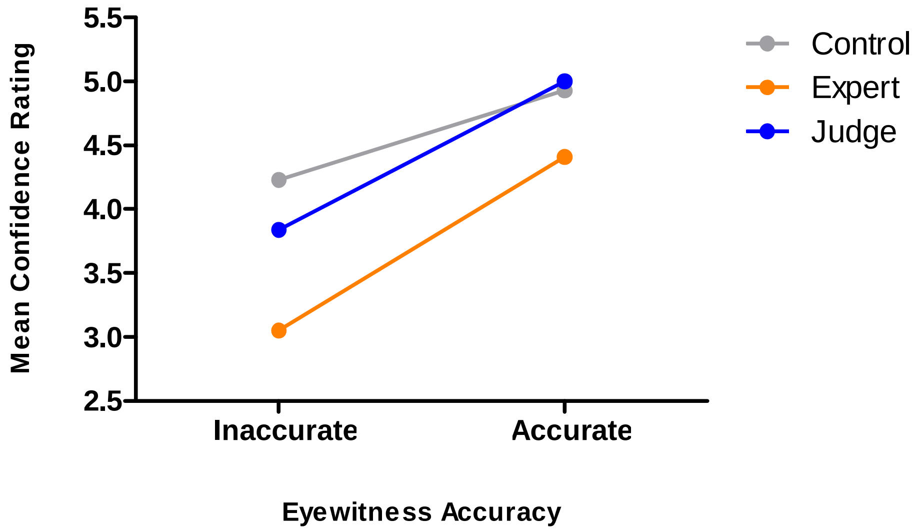

Figure 7.1 : Mean confidence ratings by eyewitness accuracy.

A second 3x2 ANOVA was conducted to investigate the extent to which instruction type and belief decision (whether a participant-juror decided to believe the eyewitness or not) impacted upon participant-juror ratings of eyewitness confidence (or vice versa). The main effects of belief type ($F_{(1, 222)}$ = 33.11, p <.0005, η^2_p = 0.130) and instruction type ($F_{(2, 222)}$ = 3.80, p <.05, η^2_p = 0.033; control $\bar{x}$ = 4.37, $\bar{\sigma}$ = 0.15; expert $\bar{x}$ = 3.87, $\bar{\sigma}$ = 0.18; judge $\bar{x}$ = 4.46, $\bar{\sigma}$ = 0.13) were significant, as was the interaction ($F_{(2, 222)}$ = 7.99, p <.0005, η^2_p = 0.067). This analysis indicated that those eyewitnesses who were believed ($\bar{x}$ = 4.75, $\bar{\sigma}$ = 0.13) were also rated as more confident than those eyewitness who were disbelieved ($\bar{x}$ = 3.72, $\bar{\sigma}$ = 0.12). As before, participant-jurors in the expert condition rated eyewitnesses as significantly less confident than those participant-jurors in the judicial condition. Inspection of the interaction term (see Figure 7.2 below) revealed that participant-jurors in the control condition attributed a larger difference in the confidence of believed and disbelieved eyewitnesses ($\Delta\bar{x}$ = 2.02) than participants from either the expert ($\Delta\bar{x}$ = 0.34) or judicial conditions ($\Delta\bar{x}$ = 0.70).

Figure 7.2 : Mean confidence ratings by juror belief decision.

Considerations Affecting Belief Decisions

All participant-jurors were asked to rate the extent to which each of three factors (eyewitness confidence, eyewitness manner and witnessing condition) affected their decision whether to believe the eyewitness or not. These ratings were analysed using three separate 2 (belief type: believe or disbelieve) x 3 (instruction type: control, expert, judge) ANOVA's.

The main effects for belief decision ($F_{(1,206)}$ = 12.41, p < .005, η^2_p = 0.057) and instruction type were both significant ($F_{(2,206)}$ = 8.14, p < .005, η^2_p = 0.073), while the interaction term was not ($F_{(2,206)}$ = 1.86, p = .158, η^2_p = 0.018). This meant that participant-jurors who believed the eyewitness they saw rated eyewitness confidence as significantly more influential ($\bar{x}$ = 3.66, $\bar{\sigma}$ = 0.09) than those participant-jurors who disbelieved the eyewitness they saw ($\bar{x}$ = 3.21, $\bar{\sigma}$ = 0.09). Moreover, post-hoc analyses revealed that eyewitness confidence was significantly more influential for those in the judicial instruction condition ($\bar{x}$ = 3.72, $\bar{\sigma}$ = 0.09) than those in the expert instruction condition ($\bar{x}$ = 3.11, $\bar{\sigma}$ = 0.12; see Figure 7.3 below).

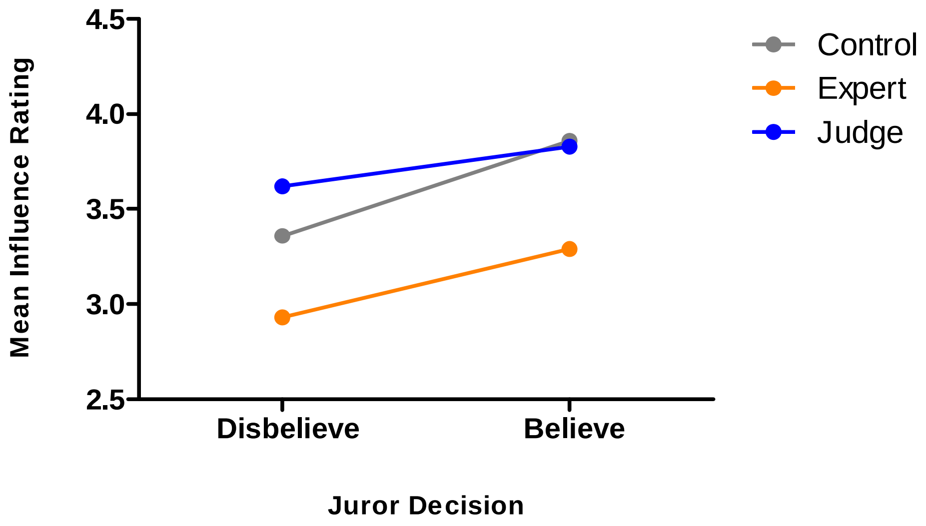

Figure 7.3 : Ratings of the influence of eyewitness confidence by belief decision

Analysis of rated reliance on eyewitness manner revealed a significant main effect of instruction type ($F_{(2,222)}$ = 11.99, p < .0005, η^2_p = 0.097), however, the main effect of belief type ($F_{(1,222)}$ = 0.67, p = .412, η^2_p = 0.003; believe $\bar{x}$ = 3.88, $\bar{\sigma}$ = 0.16; disbelieve $\bar{x}$ = 3.70, $\bar{\sigma}$ = 0.15) and the interaction ($F_{(2,222)}$ = 1.98, p = .140, η^2_p = 0.018) were not statistically significant. Post-hoc analysis indicated that reliance on eyewitness manner was significantly greater amongst participants from the control condition ($\bar{x}$ = 4.55, $\bar{\sigma}$ = 0.19), than either of those from the expert ($\bar{x}$ = 3.25, $\bar{\sigma}$ = 0.22) or judicial conditions ($\bar{x}$ = 3.57, $\bar{\sigma}$ = 0.16).

Ratings of the influence of witnessing conditions were significantly affected by instruction type ($F_{(2,222)}$ = 12.52, p < .0005, η^2_p = 0.101), but not by belief type ($F_{(1,222)}$ = 0.43, p = .512, η^2_p = 0.002; believe $\bar{x}$ = 3.68, $\bar{\sigma}$ = 0.16; disbelieve $\bar{x}$ = 3.82, $\bar{\sigma}$ = 0.16). The interaction was not significant ($F_{(2,222)}$ = 0.73, p = .482, η^2_p = 0.007). Analysis of this factor produced the same pattern of results observed for participant-jurors ratings of the influence of eyewitness manner, with control participants reporting that witnessing conditions had higher levels of influence in their decision-making ($\bar{x}$ = 4.55, $\bar{\sigma}$ = 0.19) than those in the expert ($\bar{x}$ = 3.16, $\bar{\sigma}$ = 0.23) and judicial conditions ($\bar{x}$ = 3.55, $\bar{\sigma}$ = 0.16).

In summary, participant-jurors who heard judicial instruction rated eyewitness confidence as more influential than did those in the expert condition, while those participants in the control condition reported being influenced by witness manner and witnessing conditions to a greater extent than those in the expert and judicial conditions, regardless of their decision to believe or disbelieve the eyewitness. Finally, those participants who chose to believe the eyewitness they saw also rated eyewitness confidence as being more influential in their decisions than those participant-jurors choosing to disbelieve the eyewitness.

Predictors of Juror Belief Decisions

Participant-jurors rated the credibility, accuracy, confidence, attractiveness, likeability, trustworthiness and anger of the witness that they saw. A binary logistic regression was used to investigate which, if any, of these ratings predicted participant-jurors' decisions to believe or disbelieve an eyewitness. This analysis was conducted separately for each instruction type.

Table 7.2 : β (S.E.), Significance and Exp(β) for Predictor Variables in Control Condition

	β (S.E.)	***p***	**Exp(β) (95% CI:LL-UL)**
Credibility	-0.357 (1.02)	.727	0.70 (0.09-5.19)
Accuracy	3.92 (1.81)	.031	50.51 (1.44-1768.72)
Confidence	2.22 (0.96)	.021	9.16 (1.40-59.99)
Attractiveness	-0.37 (0.58)	.524	0.69 (0.22-2.14)
Likeability	0.98 (1.24)	.432	2.66 (0.23-30.40)
Trustworthiness	-0.78 (1.44)	.585	0.46 (0.03-7.62)
Anger	-1.57 (1.22)	.199	0.21 (0.02-2.29)

In the control condition (see Table 7.2 above for all variables in the equation) the overall model was significant ($\chi^2_{(7)} = 54.15$, $p < .0005$), correctly classifying 87.7% of participant-juror decisions. Moreover, participant-juror estimates of eyewitness accuracy and eyewitness confidence each contributed significantly to the model, with a one unit increase in the participant-jurors' rating of the accuracy of the eyewitness resulting in an increase in the odds of believing that eyewitness by at least 1.4 times ($\beta = 3.92$, $p < .05$). An increase in ratings of the eyewitness's confidence of one unit also resulted in an increase in the odds of believing the eyewitness by at least the same amount ($\beta = 2.22$, $p < .05$).

For the expert condition the overall model was not found to be significant ($\chi^2_{(7)} = 11.47$, $p = .120$), although it accurately classified participant-juror decisions 70.4% of the time (see Table 7.3 below for all variables in the equation). Even so, juror estimates of eyewitness credibility significantly predicted eyewitness belief, such that an increase in the rating of credibility by one descriptive unit increased the odds that a participant-juror would believe an eyewitness by approximately 3.3 times ($\beta = 1.20$, $p < .05$).

Table 7.3 : β (S.E.), Significance and Exp(β) for Predictor Variables in Expert Condition

	β (S.E.)	***p***	**Exp(β) (95% CI:LL-UL)**
Credibility	1.20 (0.57)	.035	3.33 (1.09-10.15)
Accuracy	0.10 (0.45)	.833	1.10 (0.46-2.65)
Confidence	-0.10 (0.30)	.733	0.90 (0.50-1.62)
Attractiveness	-0.01 (0.30)	.995	1.00 (0.55-1.82)
Likeability	-0.02 (0.49)	.966	0.98 (0.38-2.55)
Trustworthiness	-0.05 (0.58)	.928	1.05 (0.34-3.31)
Anger	0.26 (0.59)	.658	1.30 (0.41-4.07)

In the judicial instruction condition participant-juror estimates of eyewitness accuracy significantly predicted belief decisions such that a one unit increase in ratings of accuracy increased the odds that an eyewitness would be believed by approximately 2 times ($\beta = 0.72$, $p < .05$, see Table 7.4 below). In addition, the overall model was significant ($\chi^2_{(7)} = 18.86$, $p < .0005$), accurately classifying 69.4% of cases.

Table 7.4 : β (S.E.), Significance and Exp(β) for Predictor Variables in Judicial Condition

	β (S.E.)	***p***	**Exp(β) (95% CI:LL-UL)**
Credibility	0.11 (0.33)	.743	1.11 (0.59-2.11)
Accuracy	0.72 (0.30)	.016	2.06 (1.14-3.71)
Confidence	0.28 (0.17)	.101	1.33 (0.95-1.86)
Attractiveness	0.20 (0.22)	.376	1.22 (0.78-1.91)
Likeability	-0.16 (0.29)	.589	0.85 (0.48-1.51)
Trustworthiness	0.13 (0.32)	.684	1.14 (0.61-2.13)
Anger	-0.26 (0.39)	.502	0.77 (0.36-1.66)

Juror Sensitivity to Eyewitness Accuracy

Overall, participant-jurors correctly evaluated eyewitnesses 60% of the time, performing significantly better than would have been expected by chance alone ($\chi^2_{(1)} = 9.28$, $p < .005$). Participant-jurors in the control, expert and judicial conditions attained 57.5%, 55.6% and 64.4% evaluation accuracy respectively. However, there was no association between instruction type and accuracy ($\chi^2_{(2)} = 1.42$, $p = .49$). Within instruction conditions, only those receiving judicial instruction performed significantly better than chance ($\chi^2_{(1)} = 8.33$, $p < .005$).

Signal Detection Analysis

Signal detection measures were also calculated for each instruction condition (see Table 7.5) in order to estimate participant-jurors' Sensitivity to the accuracy of eyewitness identifications (*d'*) and their Skepticism (*C*). At the $p = .05$ level, no reliable difference in Sensitivity was observed between the two most disparate groups: the expert condition (*d'* 95% CI : -0.34 to 1.38) and judicial instruction (*d'* 95% CI : 0.22 to 1.23). Comparison of the largest observed pair-wise difference in values of *C* also suggest there is no reliable difference in observed Skepticism between expert evidence (*C* 95% CI : -0.17 to 0.52) and the judicial instruction (*C* 95% CI : -0.30 to 0.20).

Table 7.5 : Judgement Type as Proportion Within Instruction Condition, Observed *d'* and *C* Values.

	Judgement Type					
	Miss	**Hit**	**Correct rejection**	**False alarm**	***d'***	***C***
Control	16.4%	20.5%	40.0%	26.0%	0.36	0.04
Expert	29.6%	29.6%	25.9%	14.8%	0.35	0.17
Judge	18.8%	36.6%	27.7%	16.8%	0.73	-0.05

Discussion

Participant-juror responses on numerous dependent measures indicate that they paid attention to, and followed the advice provided by, the eyewitness expert regarding the confidence-accuracy relationship. Firstly, the result that 100% of participant-jurors correctly recalled what the expert said clearly indicates that participants attended to the experts testimony and recalled the information presented. Secondly, evidence that participant jurors in the expert and judicial conditions, unlike the control condition, rated believed and disbelieved eyewitnesses as being equal in confidence further supports the interpretation that participant-jurors from these were not using eyewitness confidence to infer eyewitness accuracy. Thirdly, although those participant-jurors who believed the eyewitness they saw also indicated that eyewitness confidence was influential in their decision-making, it was reassuring to note that the levels of influence reported in this group were significantly lower than those levels reported by members of the judicial instruction condition, who were free to use eyewitness confidence in any manner they saw fit. Finally, regression models designed to predict belief or disbelief on the basis of participant-jurors ratings of the eyewitness on seven performance and personality dimensions described earlier, suggests that although no significant overall model could be reached in the expert condition, confidence was *not* a significant predictor of belief. Instead, eyewitness credibility was the most powerful single predictor of belief decisions in the expert condition. This differed from the pattern observed in the control condition; where belief decisions were significantly predicted by participant-juror ratings of eyewitness confidence and accuracy, and the judicial condition where ratings of eyewitness accuracy significantly predicted belief decisions. Thus, these results provide a strong indication that participant-jurors attended to, recalled and responded to the information provided by the eyewitness expert, and as a result, like those in Wells and Wright (1983, cited in Wells, 1986) showed Sensitivity to Expert Opinion.

Given this, it was surprising that participant-jurors in the expert and judicial conditions reported being influenced by witnessing conditions significantly *less* than participants in the control condition; this is despite the fact that both expert and judicial instruction highlighted the role of witnessing conditions in determining likely identification accuracy. It is difficult to explain why this may be, but it does appear that participant-jurors in the control condition were more generous in their attributions of the influence of each of the three considerations than were participant-jurors in the other three conditions (see Table 7.6 below). Thus, this difference may be attributed to a dampening effect of instruction on participant-juror estimates of influence, rather than a disinclination for participants in the experimental conditions to consider witnessing conditions and witness manner when making their decisions.

Table 7.6 : Jurors' Mean (SE) Ratings of Influence

	Eyewitness Confidence (SE)	Eyewitness Manner (SE)	Witnessing Conditions (SE)
Control	3.46 (0.12)	4.55 (0.19)	4.55 (0.19)
Expert	3.11 (0.12)	3.25 (0.22)	3.16 (0.23)
Judge	3.72 (0.09)	3.57 (0.16)	3.55 (0.16)

Direct comparisons between participant-jurors' perceptions of the judicial instruction and expert evidence revealed few significant differences. The expert and the judge were rated as being equally useful, and their testimony equally clear. However, participant-jurors considered the eyewitness expert to be more credible than the judge. This may be attributed in part to the fact that the person who played both roles was in fact a research psychologist and therefore may not have been able to present the judicial instruction with the same authority as the expert evidence. This difference in perceived credibility was accompanied by a significant difference in recall for the two types of instruction; participant-jurors were significantly better able to recall the testimony of the expert than they were the judicial instruction. The possibility that these two results are associated with each other is supported by theories of persuasion which suggest that perceptions of source credibility (and other heuristic cues) establish expectancies which come to influence the extent to which a message and its content are processed systematically (Chaiken & Maheswaran, 1994). It is reasonable then, in light of this proposed association, to suggest that the significantly poorer recall of the judicial instruction compared to the expert may result from the lower levels of credibility attributed to the judicial source. Even so, these differences did not appear to impact upon participant-jurors' ability to discriminate between accurate and inaccurate identifications.

Participant-jurors from the judicial instruction condition were as successful in discriminating accurate eyewitnesses from inaccurate eyewitnesses as were participants in the control and expert conditions. Indeed, a generous interpretation of the results of this study could suggest that participant-jurors in the judicial condition were actually the *best able to discriminate* between accurate and inaccurate eyewitnesses, as it was only in this condition (when the three groups were analysed individually) that participant-juror accuracy was significantly better than chance performance. This favourable account is, however, difficult to maintain in the light of signal detection analyses which suggest that the judicial instruction did not result in either greater Sensitivity or Skepticim (as measured by *d'* and *C*) amongst participant-jurors than did the experts' evidence. Moreover, the failure to detect an effect of expert evidence is not likely to be due to low power as this study had a 98.7% chance ($1-\beta = 0.99$, $\alpha = .05$) of detecting an effect of instruction that was of a medium size ($w = .03$). Thus, there is no evidence to suggest that either: a) expert testimony improved juror SEA relative to judicial instruction; or b) that either of these instructions served to increase participant-juror SEA above or beyond that which was attained in the control condition.

Chapter 8

Experiment 3 – The Impact of Erroneous Expert Evidence on Juror Sensitivity to Eyewitness Accuracy

The results of Experiment 2 suggested that the correct information provided by an eyewitness expert had no appreciable effect on participant-juror Sensitivity to Eyewitness Accuracy. The next logical step is to ascertain if inaccurate expert testimony has a negative impact on participant-juror ability to correctly evaluate eyewitness identification evidence. This investigation is important for three key reasons: 1) probabilistically speaking, it is inevitable that an eyewitness expert will provide evidence which is inaccurate in some instances; 2) there is evidence to suggest that it is possible for confidence and accuracy to be correlated under certain circumstances (Brewer & Wells, 2006; Juslin, Olsson, & Winman, 1996; Lindsay, Nilsen, & Read, 2000; Leippe, 1995; Olsson, Juslin, & Winman, 1998; Read, Lindsay, & Nicholls, 1998; Sporer, Penrod, Read, & Cutler, 1995; Weber & Brewer, 2003, 2004); 3) eyewitness experts have shown a willingness to testify that confidence is not a good predictor of eyewitness identification accuracy (Kassin et al., 1989; Kassin et al., 2001); and finally 4) by running an experiment where the eyewitness expert provides inaccurate information, it is possible to begin to estimate the *range of likely effects* an eyewitness expert may have in the real world, where the accuracy of the expert's testimony is determined by the eyewitness in each individual case.

Accordingly, in Experiment 3, participant-jurors were presented with erroneous information regarding the confidence-accuracy relationship among eyewitnesses. In this instance, participant-jurors were asked to evaluate the testimony of eyewitnesses from the "good" witnessing conditions in Protocol II (see Appendix I), where a significant association between eyewitness confidence and identification accuracy was observed. Thus, the expert's testimony that "eyewitness confidence is not a good predictor of eyewitness accuracy" was an objectively inaccurate description given the group of eyewitnesses being evaluated.

Method

The method used in this experiment was essentially the same as in Experiment 2, however, a different group of eyewitnesses were shown to participant-jurors. These eyewitnesses viewed a staged crime under “good” conditions (see Protocol II, Appendix I) and identified either a foil from a target-absent lineup, or the perpetrator from a target-present lineup. Interviews conducted with these eyewitnesses were evaluated by participant-jurors.

Participant-Jurors

Two hundred and thirty-eight undergraduate psychology students (144 females, 93 males and one unspecified) ranging in age from 17 to 48 years ($\overline{x}$ = 20.0, $\overline{\sigma}$ = 0.24) from the University of New South Wales acted in the role of jurors during scheduled psychology tutorials.

Design

A 2 (witness type: accurate v inaccurate) x 3 (instruction type: judicial instruction v expert evidence v no instruction control) between-subjects factorial design was employed.

Materials

The minimal trial, pre-trial instruction, expert testimony, judicial instruction and juror response forms are all identical to those used in Experiment 2 (see Chapter 7 and Appendices D, E, H and I for more details).

Eyewitness Testimony

The eyewitness evidence protocol produced eight eyewitnesses giving interviews who witnessed the crime under “good” conditions (four accurate and four inaccurate). Six of these eight interviews were used as stimuli for Experiment 3. The testimony of the seventh eyewitness (who was accurate) was not included in this experiment as she was a school of psychology Course Administrator and it was considered that she would be familiar to a large number of first-year psychology students. Thus, in order to balance the number of accurate and inaccurate eyewitness identifications presented to participant-jurors, the testimony obtained from the last inaccurate eyewitness was also set aside. The confidence-accuracy correlation for these eyewitnesses was both positive and significant ($r_{pbi} = 0.83$, $p < .01$).

Results

Jurors' Qualitative Evaluations of Eyewitnesses

Jurors' ratings of the eyewitnesses revealed that accurate and inaccurate witnesses were not seen to differ significantly on the dimensions of credibility ($t_{(236)} = -1.55$, $p = .121$), accuracy ($t_{(231.75)}$[12] $= -1.24$, $p = .216$), attractiveness ($t_{(235.64)}$[12] $= -1.92$, $p = .056$), likeability ($t_{(218)} = -1.30$, $p = .194$), trustworthiness ($t_{(213.15)}$[12] $= -1.67$, $p = .965$), or anger ($t_{(218)} = -0.34$, $p = .736$). Ratings of confidence will be discussed further below.

Jurors' Qualitative Evaluations of Instructors

Participant-jurors were asked to rate the credibility, utility and clarity of the judge and the expert (see Table 8.1 for mean ratings). Comparisons between the average ratings of the expert and the judge on these three dimensions revealed significant differences on all three, with the expert being rated as significantly more credible ($t_{(157)} = 3.59$, $p < .0005$) and clearer ($t_{(157)} = -3.13$, $p < .005$) than the judge. However, the judge was seen to provide more useful information than the expert ($t_{(157)} = 2.263$, $p < .05$).

Table 8.1 : Jurors' Mean (SE) Evaluations of Instructors

	Expert (SE)	**Judge (SE)**
Credibility	5.61 (0.09)	5.09 (0.12)
Utility	4.46 (0.11)	5.00 (0.14)
Clarity	5.51 (0.90)	5.15 (0.13)

[12] Levene's Test for Equality of Variance was violated in this comparison. Accordingly the t-value reported above does not assume equal variances.

Jurors' Recall for Instructions

With regard to participant-jurors' memory for the instructions they heard, on average those in the expert condition scored 3.22 ($\bar{\sigma}$ = 0.06) out of a maximum possible score of four, while those in the judicial instruction condition scored 2.89 ($\bar{\sigma}$ = 0.10) on average. Again, juror recall for instruction was shown to be significantly better in the expert evidence condition than it was for the judicial instruction condition ($t_{(157)}$ = 3.04, p < .005). For the specific item relating to the confidence-accuracy relationship 96.8% of jurors demonstrated an accurate recollection of the testimony provided by the eyewitness expert, with 91.5% of jurors selecting the verbatim quote, and 5.3% choosing its paraphrased alternative.

Juror Decision Criterion

A chi-squared goodness-of-fit test indicated that the proportion of participant-jurors who believed eyewitnesses did not significantly differ from 50%, either when collapsed across conditions ($\chi^2_{(1)}$ = 2.84, p = .092), or within instruction conditions (control, 51.9% belief, $\chi^2_{(1)}$ = 0.11, p = .739; expert, 54.3% belief, $\chi^2_{(1)}$ = 0.68, p = .409; or judge, 61.5% belief, $\chi^2_{(1)}$ = 3.46, p = .063). Together these results suggest, as in Experiment 2, that neither expert evidence nor judicial instruction caused a change in the decision criterion adopted by participant-jurors.

Juror Sensitivity to Eyewitness Confidence

A 2x3 ANOVA was conducted to assess the extent to which participant-juror estimates of eyewitness confidence differed as a function of instruction type and participant-witness accuracy. Participant-jurors rated accurate eyewitnesses ($\bar{x}$ = 4.41, $\bar{\sigma}$ = 0.12) as being significantly more confident than inaccurate eyewitnesses overall ($\bar{x}$ = 3.60, $\bar{\sigma}$ = 0.11, $F_{(1,232)}$ = 24.19, p < .0005, η^2_p = 0.094). The main effect of instruction ($F_{(2,232)}$ = 0.82, p = .444, η^2_p = 0.007; control, $\bar{x}$ = 3.91, $\bar{\sigma}$ = 0.14; expert, $\bar{x}$ = 3.94, $\bar{\sigma}$ = 0.13; judge, $\bar{x}$ = 4.16, $\bar{\sigma}$ = 0.15) and the interaction term ($F_{(2,232)}$ = 1.61, p = .202, η^2_p = 0.014) were not significant (see Figure 8.1 below).

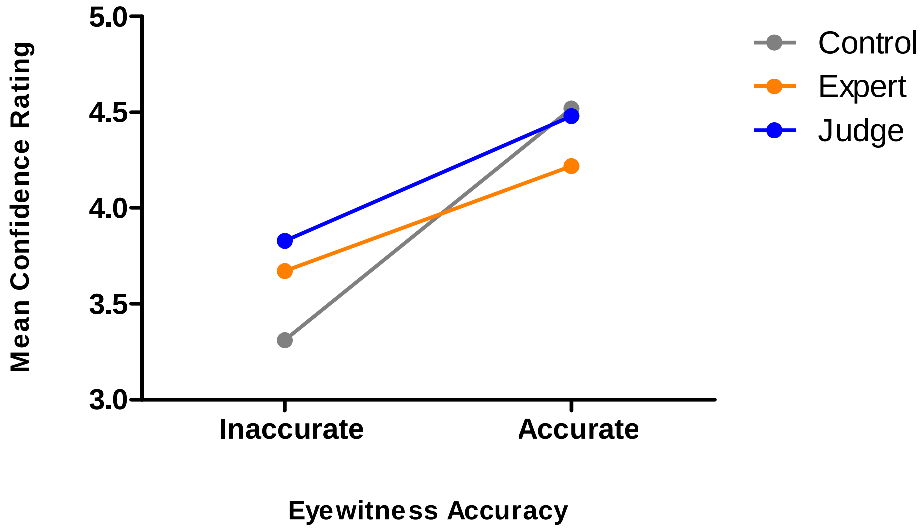

Figure 8.1 : Mean confidence ratings by eyewitness accuracy.

A 2x3 ANOVA investigating participant-juror estimates of eyewitness confidence given the instruction heard and the belief decision reached provided evidence for a significant difference in the estimates of confidence attributed to believed ($\bar{x}$ = 4.28, $\bar{\sigma}$ = 0.11) and disbelieved eyewitnesses ($\bar{x}$ = 3.55, $\bar{\sigma}$ = 0.12; $F_{(1,232)}$ = 19.82, p < .0005, η^2_p = 0.079) such that those eyewitnesses who were believed were perceived to be more confident than those who were disbelieved. There were no other significant main ($F_{(2,232)}$ = 0.83, p = .436, η^2_p = 0.007; control, $\bar{x}$ = 3.77, $\bar{\sigma}$ = 0.14; expert, $\bar{x}$ = 3.97, $\bar{\sigma}$ = 0.13; judge, $\bar{x}$ = 4.01, $\bar{\sigma}$ = 0.16) or interaction effects ($F_{(2,232)}$ = 1.36, p = .252, η^2_p = 0.012; see Figure 8.2 below).

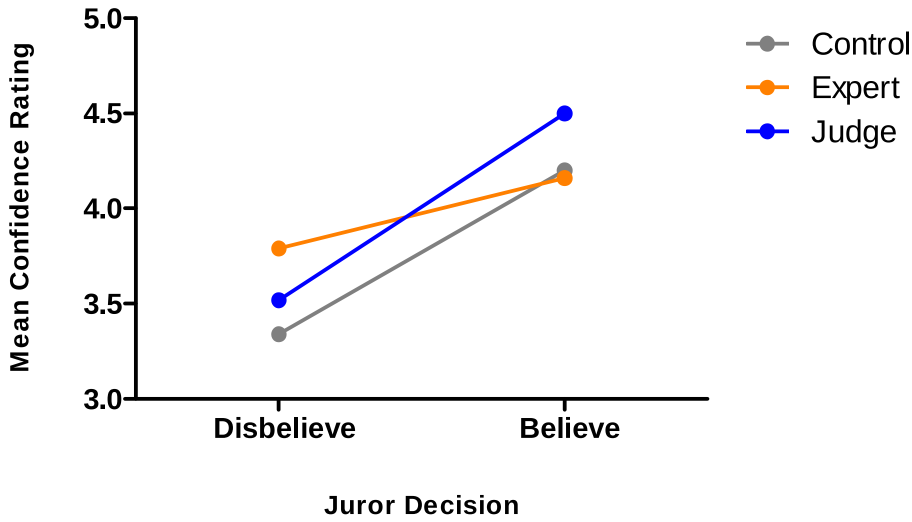

Figure 8.2 : Mean confidence ratings by juror belief decision.

Considerations Affecting Belief Decisions

Three 2 (belief type: believe or disbelieve) x3 (instruction type: control, expert, judge) ANOVA's were conducted to assess the extent to which participant-jurors considered that their decisions were influenced by eyewitness confidence, eyewitness manner and witnessing condition. Participant-jurors rated eyewitness confidence as significantly more influential when they believed the eyewitness ($\bar{x}$ = 3.24, $\bar{\sigma}$ = 0.10) than when they disbelieved them ($\bar{x}$ = 3.54, $\bar{\sigma}$ = 0.09; $F_{(1,214)}$ = 4.68, p < .05, η^2_p = 0.021). They also reported relying on confidence significantly more in the judicial ($\bar{x}$ = 3.69, $\bar{\sigma}$ = 0.12) condition than in the expert condition ($\bar{x}$ = 3.19, $\bar{\sigma}$ = 0.10; $F_{(2,214)}$ = 5.17, p < .01, η^2_p = 0.046). The interaction term was not significant ($F_{(2,214)}$ = 1.74, p = .178, η^2_p = 0.016; see Figure 8.3 below).

Figure 8.3 : Ratings of the influence of eyewitness confidence by belief decision

Those participants in the control condition reported relying on eyewitness manner ($F_{(2,232)}$ = 9.13, p < .0005, η^2_p = 0.073) and eyewitnessing conditions ($F_{(2,232)}$ = 6.57, p < .01, η^2_p = 0.054) significantly more often than participant-jurors in either the expert or control conditions. The other main and interaction effects for eyewitness manner (belief main effect, $F_{(1,232)}$ = 1.24, p = .267, η^2_p = 0.005; interaction effect, $F_{(2,232)}$ = 0.23, p = .792, η^2_p = 0.002) and eyewitnessing conditions (belief main effect, $F_{(1,232)}$ = 1.57, p = .211, η^2_p = 0.007; interaction effect, $F_{(2,232)}$ = 0.11, p = .900, η^2_p = 0.001) were not significant (see Table 8.2 below for the means).

Table 8.2 : Jurors'Mean (SE) Ratings of Influence of Eyewitness Manner and Witnessing Conditions by Juror Decision and Instruction Condition

	Mean (SE) Eyewitness Manner		**Mean (SE) Witnessing Conditions**	
	Disbelieved	**Believed**	**Disbelieved**	**Believed**
Control	4.05 (0.28)	4.51 (0.27)	4.29 (0.28)	4.71 (0.27)
Expert	3.05 (0.26)	3.29 (0.24)	3.49 (0.26)	3.67 (0.24)
Judge	3.44 (0.35)	3.50 (0.27)	3.64 (0.34)	3.90 (0.27)

In summary then, participants in the expert and judicial conditions differed significantly in their reported reliance on eyewitness confidence, with those in the judicial instruction condition reporting greater reliance on eyewitness confidence when making their decisions. Eyewitness manner and witnessing conditions were more compelling for jurors from the control condition compared to those in expert and judicial conditions. Again, this may be due to instruction having a dampening effect on participant-jurors tendency to rate the influence of a variable, as values appear to be higher overall in the control condition.

Predictors of Juror Belief Decisions

Participant-jurors rated the credibility, accuracy, confidence, attractiveness, likeability, trustworthiness and anger of the witness who they saw. A binary logistic regression was used to investigate which, if any, of these ratings predicted participant-jurors' decisions to believe or disbelieve an eyewitness. This analysis was conducted separately for each instruction type.

In the control condition the overall model was significant ($\chi^2_{(7)} = 23.93$, $p < .005$), correctly classifying 78.7% of participant-juror decisions. Participant-juror estimates of eyewitness credibility and eyewitness anger each contributed significantly to the model (see Table 8.3 below for all variables in the equation), with a one unit increase in the credibility rating being associated with an increase in the odds of the eyewitness being believed by approximately 3.6 times ($\beta = 1.30$, $p < .05$). Conversely, a decrease in the rating of eyewitness anger of one unit was seen to result in an increase in the odds of believing the eyewitness by a factor of 0.17 (β = -1.76, $p < .05$).

Table 8.3 : β (S.E.), Significance and Exp(β) for Predictor Variables in Control Condition

	β (S.E.)	***p***	**Exp(β) (95% CI:LL-UL)**
Credibility	1.30 (0.63)	.039	3.67 (1.07-12.60)
Accuracy	0.35 (0.47)	.450	1.42 (0.57-3.57)
Confidence	-0.03 (0.31)	.935	0.98 (0.53-1.80)
Attractiveness	0.34 (0.35)	.329	1.41 (0.71-2.80)
Likeability	0.14 (0.44)	.747	1.15 (0.48-2.76)
Trustworthiness	-0.69 (0.54)	.203	0.50 (0.17-1.45)
Anger	-1.76 (0.74)	.017	0.17 (0.04-0.73)

The model was also found to be significant in the expert condition ($\chi^2_{(7)} = 34.13$, $p < .0005$), accurately classifying participant-juror decisions 72.3% of the time (see Table 8.4 below for all variables in the equation). Moreover, juror estimates of eyewitness credibility and accuracy were seen to be significant predictors in this model, such that a one-point increase in ratings of either credibility or accuracy increased the odds that a participant-juror would believe an eyewitness by at least 1.2 times (β = 0.95, $p < .05$) and 1.4 times (β = 1.14, $p < .01$) respectively.

Table 8.4 : β (S.E.), Significance and Exp(β) for Predictor Variables in Expert Condition

	β (S.E.)	***p***	**Exp(β) (95% CI:LL-UL)**
Credibility	0.95 (0.42)	.022	2.59 (1.15-5.85)
Accuracy	1.14 (0.43)	.008	3.11 (1.35-7.19)
Confidence	-0.09 (0.24)	.710	0.92 (0.57-1.46)
Attractiveness	0.14 (0.25)	.562	1.15 (0.71-1.86)
Likeability	-0.10 (0.37)	.785	0.90 (0.44-1.87)
Trustworthiness	-0.30 (0.38)	.425	0.74 (0.36-1.55)
Anger	-0.62 (0.41)	.135	0.54 (0.24-1.21)

Although no significant predictors of juror belief were identified among participant-jurors who heard judicial instruction (see Table 8.5 below), the overall model was significant ($\chi^2_{(7)} = 20.61$, $p < .005$), accurately predicting belief decisions in 72.3% of cases.

Table 8.5 : β (S.E.), Significance and Exp(β) for Predictor Variables in Judicial Condition

	β (S.E.)	***p***	**Exp(β) (95% CI:LL-UL)**
Credibility	0.13 (0.42)	.752	1.14 (0.50-2.57)
Accuracy	0.95 (0.59)	.108	2.59 (0.81-8.24)
Confidence	0.32 (0.31)	.300	1.38 (0.75-2.53)
Attractiveness	0.50 (0.29)	.085	1.65 (0.93-2.91)
Likeability	-0.26(0.37)	.482	0.77 (0.37-1.60)
Trustworthiness	0.58 (0.43)	.177	1.78 (0.77-4.13)
Anger	-0.27 (0.59)	.648	0.76(0.24-2.43)

Juror Sensitivity to Eyewitness Accuracy

Overall, participant-jurors were able to correctly classify the accuracy of eyewitness identifications 62.6% of the time. This level of accuracy was found to be significantly better than would have been expected by chance alone ($\chi^2_{(1)} = 15.13$, $p < .0005$). No significant association was identified between instruction type and juror evaluation accuracy ($\chi^2_{(2)} = .806$, $p = .668$), with jurors in control, expert and judicial conditions attaining 64.5%, 63.8% and 58.5% accuracy respectively. When this analysis is performed within conditions, only those receiving judicial instruction did not perform significantly better than chance ($\chi^2_{(1)} = 1.862$, $p = .172$), while those in the control ($\chi^2_{(1)} = 7.716$, $p = .005$) and expert conditions ($\chi^2_{(1)} = 7.191$, $p = .007$) showed significant sensitivity to the accuracy of the eyewitness they viewed.

Signal Detection Analysis

At the $p = 0.05$ level, confidence intervals on *d'* (Sensitivity) revealed no reliable difference between the control condition (*d'* 95% CI : 0.23 to 1.41) and judicial instruction (*d'* 95% CI : -0.11 to 1.16) (see Table 8.6 below). The largest observed pair-wise difference in values of *C* (Skepticism) was also tested using confidence intervals derived from inferential interval estimates. Again, at the *p* = .05 level, the estimates suggest there is no reliable difference in observed Skepticism between expert evidence (*C* 95% CI : -0.32 to 0.21) and the judicial instruction (*C* 95% CI : -0.65 to -0.01).

Table 8.6 : Judgement Type as Proportion Within Instruction Condition, Observed *d'* and *C* Values.

	Judgement Type					
	Miss	**Hit**	**Correct rejection**	**False alarm**	***d'***	***C***
Control	11.4%	27.8%	36.7%	24.1%	0.82	0.14
Expert	20.2%	38.3%	25.5%	16.0%	0.69	0.05
Judge	12.3%	32.3%	26.2%	29.2%	0.53	0.33

Discussion

In Experiment 3, as in Experiment 2, participant-jurors showed some Sensitivity to Expert Opinion, as well as to the underlying confidence-accuracy correlation, and as a result, sensitivity to eyewitness accuracy.

Participant-jurors rated accurate eyewitnesses as having greater confidence than inaccurate eyewitnesses; suggesting that these jurors perceived the underlying positive correlation between confidence and accuracy. As in Experiment 2, the eyewitness expert informed jurors that, among other things, an eyewitness's confidence was not a good predictor of his or her identification accuracy, and, as in Experiment 2, participant-jurors have heard and applied this information when formulating their decisions. Firstly, 96.8% of participant-jurors' correctly recalled what the expert said about the confidence-accuracy relationship. Furthermore, in line with the advice they received, participant-jurors in the expert condition rated eyewitness confidence as significantly *less influential* in their decisions than did participants in the judicial condition. In addition, the participant-jurors decision to believe the eyewitness could be predicted by their rating of the characteristics of the eyewitness. In the expert condition participant-jurors perceptions of credibility and accuracy significantly predicted their belief decisions, while eyewitness confidence did not. Overall, belief decisions in the control condition were predicted by participant estimates of credibility and anger, while no individual variables significantly predicted belief in the judicial condition. However, the conclusion that the participant-jurors showed Sensitivity to Expert Opinion must be tempered somewhat by the finding that, irrespective of instruction condition, participant-jurors tended to rate the eyewitnesses they believed as more confident than those they disbelieved. This suggests that although participant-jurors heard and understood the expert's evidence and *reported* behaving in a manner consistent with the advice, they still appear to have either rated those eyewitnesses they believed as more confident, or believed those eyewitnesses who were seen to be more confident. Even so, binary logistic regressions indicated that perceptions of eyewitness confidence did not significantly predict participant-juror belief decisions in the expert condition. Thus, although perceptions of eyewitness confidence were associated with juror belief decisions, this was to a lesser extent than perceptions of eyewitness credibility and accuracy.

When participant-juror ratings of the expert and the judge were compared a number of significant differences revealed themselves. The participant-jurors rating of the expert and the judge differed significantly from each other in terms of their perceived clarity, utility and credibility. As in Experiment 2, the expert was seen as significantly more credible than the judge, and their testimony was recalled more accurately. In this study, however, the expert's evidence was also rated as significantly clearer than that of the judge, but significantly less useful. Although it seems somewhat counter-intuitive for the clearer, more credible source to also be rated as less useful, participant-juror ratings of the utility of the expert may reflect their frustration at being instructed to disregard a cue which they could clearly identify, and which was objectively useful to their decision-making. It may be that though the expert was able to testify clearly on the subject of confidence-accuracy, and appeared to be credible, this did not preclude participant-jurors from feeling that this unknowingly erroneous evidence was less useful than the judges more general comments.

Turning now to the decision criterion adopted by participant-jurors, belief decisions in the three instruction conditions seem to have been made on the basis of different eyewitness qualities, none of which appear to predict eyewitness identification accuracy. Despite this, participant-juror discrimination performance did not differ significantly across instruction conditions. Overall, participant-jurors were significantly better than chance at discriminating between accurate and inaccurate identifications, performing at 62.6% accuracy. This may be due to their sensitivity to the underlying confidence-accuracy correlation or some other cue to accuracy, which they were able to detect and respond to. However, even those jurors from the expert condition, who reported, and to some extent behaved in a manner consistent with having dismissed the strongest predictor of eyewitness identification accuracy, did not perform significantly worse than the remaining groups of jurors who had perceived, and were free to use, this cue. In fact, although not significantly so, it was actually those participant-jurors in the judicial condition who performed worst of all when accuracy was evaluated within conditions (58.5% accuracy). Yet, signal detection analyses suggest that this inference is likely to be unduly harsh as confidence estimates of both Sensitivity and Skepticism reveal no significant differences between the 95% confidence intervals of the most disparate groups, and power analyses suggest that this result is unlikely to represent a Type II error due to low statistical power as this experiment had 98.9% chance of detecting an effect of instruction that was moderate in size ($1-\beta = 0.989$, $\alpha = .05$, $w = 0.3$). Thus, as in Experiment 2, there is no evidence to suggest that erroneous expert testimony impairs juror SEA to a greater extent than judicial instruction, or that either of these instructions serve to influence participant-juror SEA above and beyond that which was attained in the control condition.

General Discussion – Experiments 2 and 3

These studies were designed to make relative comparisons between the effects of judicial instruction and expert evidence on the accuracy of participant-juror decision-making. They also aimed to systematically explore the impact of expert advice of differing quality (i.e., accurate or erroneous).

Independent Effects of Instruction

The results of these studies provide evidence regarding the impact of expert and judicial instruction which, in light of the mixed results already presented in the literature, both support and contradict previous research findings.

Participant-jurors in the judicial instruction condition were as likely to believe an accurate eyewitness as to disbelieve them. Multiple analyses revealed that this pattern of responding did not differ significantly from that observed in the control condition therefore there is no evidence to suggest that the judicial instruction served to induce Skepticism. This finding contrasts with previous evidence where standard judicial instructions were found to induce participant-juror Skepticism (Katzev & Wishart, 1985; Ramirez et al., 1996). Moreover, the comment made by Greene (1988), that judicial instruction cannot induce sensitivity to the opinions offered because no opinion is provided by the judge, is equally true here. Thus we shouldn't be surprised to find that no Sensitivity was observed in the judicial instruction condition either.

Experiments 2 and 3 both produced evidence of significant participant-juror Sensitivity to opinions of the eyewitness expert regarding confidence and accuracy. This was illustrated by the fact that participant-jurors who heard expert evidence significantly moderated their tendency to believe more confident eyewitnesses and disbelieve those who were less confident, seen in Experiment 2. Furthermore, regression models conducted in Experiments 2 and 3 also revealed that where participant-jurors heard the expert say that confidence did not predict accuracy, estimates of confidence were not found to be a significant predictor of participant-juror belief type. In a separate analysis confidence was rated as significantly less influential by these participants than members of the judicial instruction condition (which did not include advice regarding confidence and accuracy). In addition, participant-jurors were clearly able to recall what the expert said regarding the confidence-accuracy correlation in both studies, thereby leading to the conclusion that participant-jurors displayed Sensitivity by not only attending to, but also remembering and acting upon the evidence of the expert regarding eyewitness confidence. Encouragingly, this Sensitivity was observed in the absence of an associated Skepticism effect, as participant-jurors did not adopt a significantly more biased decision criterion as a result of hearing the expert's testimony. This combination of significant Sensitivity in the absence of Skepticism is in keeping with findings from previous research (Cutler et al., 1989a; Cutler et al., 1989b; Wells & Wright, 1983 cited in Wells, 1986) but is at odds with many more studies which have found either Sensitvity together with Skepticism (Cutler et al., 1989a; Devenport et al., 2002) or Skepticism in isolation (Blonstein & Geiselman, 1990; Cutler et al., 1990a; Fox &

Walters, 1986; Geiselman et al., 2002, Experiment 2; Leippe et al., 2004, Experiments 1 and 2; Lindsay, 1994, Experiment 6; Loftus, 1980, Experiment 1; Wells et al., 1980). Thus, the finding of significant Sensitivity to expert testimony *in the absence* of significant Skepticism, is a finding worthy of further replication.

Effect of Expert Evidence on Eyewitness Confidence

Experiments 2 and 3 differed from each other in design on the basis of the confidence-accuracy relationship represented in the sample of eyewitnesses. In Experiment 2 eyewitness expressions of confidence did not significantly predict identification accuracy, while in Experiment 3, more confident witnesses were significantly more likely to have been accurate in their identifications than were less confident eyewitnesses. This variation across studies allowed us to investigate how the presence or absence of a significant underlying predictor variable which altered the accuracy of the eyewitnesses opinions might impact upon participant-jurors' Sensitivity to Eyewitness Accuracy. In Experiment 2, where the confidence-accuracy correlation was neutral, participant-jurors in the control and judicial conditions appear to have relied on this cue significantly more than participants in the expert condition (as evidenced by logistic regression and ANOVA analyses relating to participant-juror beliefs). In Experiment 3, the same overall pattern was replicated. However, in this instance participant-jurors' tendency to believe more confident eyewitnesses was not significantly moderated by expert instruction, even though belief decisions in the expert condition were significantly predicted by perceptions of eyewitness accuracy and credibility rather than confidence. Thus, irrespective of the objective accuracy of the eyewitness expert's advice regarding the confidence-accuracy relationship, there is evidence to suggest that participant-jurors followed the advice they were given to some extent. Overall then, our results replicate those of Wells et al. (1980), who found that the introduction of expert evidence telling jurors that confidence was not a predictor of accuracy, successfully convinced participant-jurors to reduce their reliance on that variable. However, the finding that expert evidence did not significantly moderate the impact of belief on participant-juror ratings of eyewitness confidence in Experiment 3 (in spite of other indicators of Sensitivity), is in keeping with Fox and Walters' (1986) finding that participant-jurors continued to use the cue even against instruction. Together these results may suggest that although participant-jurors do try to follow the advice given, and believe they are doing so, the extent to which they actually comply may be moderated by the *objective accuracy* of this advice. That is, where the advice of the expert was inaccurate (Experiment 3), there was less evidence that the participant-jurors had followed expert advice than there was when the advice was objectively accurate (Experiment 2). This result is certainly worthy of further investigation as evidence that jurors were able to evaluate the accuracy of the advice they are provided is the best of all possible outcomes. In addition, this result also provides compelling evidence for the need to validate participant-juror self reports of their own decision-making strategies using objective measures. Participant-jurors in Experiment 3 indicated that they relied on confidence significantly less than participant-jurors in the judicial instruction condition, however, objective analysis revealed that the presence of expert evidence did not moderate the observed association between perceived eyewitness confidence and belief. Thus, to some extent, participant-jurors were simply reporting that they did what they were told to do without having actually followed through on the advice. This is an important observation given past reliance on participant-jurors' reports of the extent to which they rely on expert advice.

Relative Effects of Instruction

The direct comparison of the effects of adversarial expert testimony and judicial instruction on participant-juror discrimination performance, suggests that judicial instruction may not be inferior to expert evidence as previously asserted (Cutler & Penrod, 1995; Greene & Loftus, 1984; Leippe, 1995; Pezdek, 2007). Rather, the evidence presented here suggests that although there are differences between participant-jurors' perceptions of the two instruction types, these differences do not appear to influence participant-juror discrimination performance. The results of Experiments 2 and 3 replicate Experiment 1 which indicated that participant-jurors perceived the judge to be a less credible source than the eyewitness expert, and in all three studies, this finding was accompanied by significantly poorer recall for the judge's instruction than the experts' evidence. This result may in part be attributed to the fact that the judicial instruction was delivered by a research psychologist rather than a judge, yet even so, the observed discrepancies in participant-juror ratings and recall performance appear to be irrelevant to the subsequent eyewitness discrimination task, which revealed no systematic impact that could be attributed to these differences. Moreover, in Experiment 3 the expert was rated as being significantly clearer, yet significantly less useful to jurors in their decision-making. As above, while the difference in clarity is not entirely unexpected in light of previous research regarding jurors comprehension of judicial instruction (Greene, 1988), and the difference in utility may reflect participant-juror frustration at being directed to disregard what may appear to be a useful cue, these discrepancies did not appear to systematically influence participant-juror ability to differentiate accurate from inaccurate identifications across conditions. Indeed, it is more appropriate to suggest that participant-juror perceptions of, and recall for the source of their instruction did not serve to significantly *impair* their ability to identify mistaken and accurate identifications, which was already significantly above chance levels in the control condition, at least in Experiment 3.

Instruction Type & Sensitivity to Eyewitness Accuracy

Overall, participant-jurors in these studies performed significantly better than chance irrespective of the sample of eyewitnesses used (60% accuracy in Experiment 2; 62.6% in Experiment 3). This is a somewhat surprising and gratifying result. Previous research has reported that participant-juror discrimination accuracy is as low as 25 to 29% for inaccurate eyewitnesses (Wells & Leippe, 1981; Wells et al., 1979) and as high as 84% for accurate eyewitnesses (Wells et al., 1979) in part due to participant-jurors' tendency to believe (overbelieve) all eyewitnesses (Cutler et al., 1988; Loftus, 1979; Loftus & Monahan, 1980; Penrod & Cutler, 1999; Wells, 1980). The results from Experiments 2 and 3 present evidence, on the basis of conventional and signal detection analyses, that participant-jurors in these studies were as likely to believe as to disbelieve an eyewitness, and in light of their significant discrimination accuracy, were therefore equally likely to *believe accurate eyewitnesses* as to *disbelieve inaccurate eyewitnesses*. This finding directly contradicts extensive evidence suggesting that subject-jurors were as likely to believe accurate eyewitnesses as inaccurate eyewitnesses (Lindsay et al., 1989; Lindsay et al., 1981; Loftus, 1974; Wells & Wright, 1983 cited in Wells, 1986; Wells et al., 1979). The pattern of results is, however, in keeping with Wells & Wright (1983, cited in Wells, 1986) who found that participant-jurors in the expert evidence condition were sensitive to the actual accuracy of the eyewitnesses. Thus, although this finding is not without precedent, it is important and requires replication.

Predicting Sensitivity to Eyewitness Accuracy

In keeping with the findings of Wells et al. (1980) and Wells and Wright (1983, cited in Wells, 1986), these data suggest that Response and Sensitivity on the part of participant-jurors (i.e., indirect measures) does not necessarily result in a significant influence on participant-juror Sensitivity to Eyewitness Accuracy (i.e., the direct measure). In both experiments reported here, there is compelling evidence to suggest that participant-jurors heard and tried to apply the evidence of the eyewitness expert, with participant-jurors from the expert condition in Experiment 2 seemingly ignoring confidence when formulating their decisions as instructed, and participant-jurors in Experiment 3 relying on credibility and accuracy to a greater extent than confidence in their determinations. Moreover, binary logistic regressions predicting juror belief decisions in each experimental condition established different models to predict belief decisions for each instruction type, suggesting that participant-jurors in the control condition are basing their determinations on different factors to participants from the expert and judicial conditions, which also appear to differ with respect to each other. Even so, participant-juror discrimination accuracy is not significantly improved, nor impaired, by the participant-jurors attempts to follow the differing advice given in each instruction condition.

Overall, participant-jurors were able to differentiate accurate from inaccurate eyewitnesses at levels significantly above chance, and this performance was not influenced by advice providing an accurate description of the confidence-accuracy correlation for the eyewitness sample, as in Experiment 2, or an inaccurate description as in Experiment 3. In addition, neither the presence nor the absence of a significant cue to eyewitness identification accuracy in the eyewitness sample appears to have had an impact on participant-juror ability to discriminate accurate from inaccurate eyewitnesses. This is not necessarily as surprising a result as it may seem in light of the work by Wells and colleagues. In the first instance the data of Wells et al. (1980) suggested that jurors can be sensitive to the expert evidence without also showing sensitivity to witness accuracy. In the second instance evidence from Wells and Wright (1983 cited in Wells, 1986) showed participant-jurors to be sensitive to expert evidence, eyewitness accuracy and the manipulation of witnessing conditions, even when those manipulations did not significantly influence identification accuracy rates. Thus, there is a necessity to be very careful when extrapolating from indirect to direct measures.

Limitations – Experiments 2 and 3

The main limitation to this pair of studies relates to the fact that the eyewitnesses evaluated by participant-jurors in Experiments 2 and 3 were taken from different crime scenarios (good and poor), causing the variation between the confidence-accuracy correlations observed across experiments. Thus, although we know that these samples of eyewitnesses differ with regard to this key quality, we cannot be sure that they do not differ from one another in other important and unanticipated ways. If this proved to be true it would make it more difficult to attribute any observed differences between the results of these experiments to differences in the underlying confidence-accuracy correlation, or the accuracy of the expert evidence, as they may actually be attributable to a third variable. Despite this possibility, it is reassuring to note that there are actually very few differences between the results of Experiments 2 and 3; thus, even if there are important differences between the samples of eyewitnesses, they do not appear to be influencing the data in any obvious ways. Conversely, the similarity between the results from Experiments 2 and 3 may also be evidence of a failure to produce qualitatively different types of expert evidence. Specifically, although not statistically significant, the confidence-accuracy correlation in the “poor” witnessing condition appeared to be rather high. This may have substantially reduced any possible differences which could have emerged between accurate and erroneous expert evidence.

Conclusion

Thus, the results of Experiments 1 to 3 and those of Wells and colleagues (Wells, 1980; Wells and Wright 1983 cited in Wells, 1986) provide compelling evidence to suggest that participant-juror’s Sensitivity to Eyewitness Accuracy can operate *independently* of their Sensitivity and Response to Expert Opinion. Accordingly, we should be cautious about inferring Sensitivity to Eyewitness Accuracy on the basis of indirect measures (i.e., REE, or SEO). That is, since compliance with the expert’s advice may or may not predict performance accuracy in a particular instance, it follows that knowing whether or not a juror has complied with this advice does not *necessarily* tell us anything about their ability to differentiate accurate from inaccurate eyewitnesses. Thus, the evidence presented here suggests that it may be invalid to infer Sensitivity to Eyewitness Accuracy on the basis of indirect designs. Instead, the results suggest that SEA needs to be assessed directly through the use of real eyewitness designs, which provide a verifiable criterion for evaluation. A point made by Wells over 20 years ago (Wells, 1986), but largely ignored since.

Chapter 9

Experiment 4 – Boundary Effects of Expert Evidence

In the previous two experiments participant-jurors were shown an eyewitness either from a sample for which there was a significant confidence-accuracy correlation or from a sample where there was no significant association between confidence and accuracy. In this way it was possible to investigate the impact of erroneous and accurate expert evidence (and judicial instructions) on participant-juror decision-making. This study, however, was limited by the fact that jurors hearing accurate evidence viewed a *different* subset of eyewitnesses than those jurors who heard erroneous evidence, thus the characteristics of individual eyewitnesses were confounded with the qualitative differences in expert evidence. An alternative approach which addresses this possible confound is to vary the testimony given by the eyewitness expert with respect to the same set of eyewitnesses. Accordingly, in this study we take this second approach, by holding constant the sample of eyewitnesses viewed by participant-jurors (i.e., the ground truth), and instead altering the testimony of the eyewitness expert. This permits us to look at the impact of different forms of instruction on the same evaluation task. It also expands our understanding of the impact of how erroneous eyewitness expert evidence might impact upon juror decision-making (i.e., its relative costs and benefits). Finally, this experimental approach serves to increase the power of our analysis by allowing us to directly compare the consequences of accurate and erroneous advice, something which was not possible in Experiments 2 and 3.

Method

The eyewitnesses used in this experiment also viewed a staged crime under “good” conditions (see Protocol II, Appendix I) and identified either a foil from a target-absent lineup, or the perpetrator from a target-present lineup. Interviews conducted with these eyewitnesses were evaluated by participant-jurors.

Participant-Jurors

Two hundred and ninety-seven undergraduate psychology students (190 female and 107 male) from the University of New South Wales with ages ranging between 17 and

47 years ($\bar{x}$ = 19.55, $\bar{\sigma}$ = 3.25) acted in the role of jurors during scheduled psychology tutorials.

Design

A 2 (witness type: correct vs. mistaken) x 4 (instruction type: accurate eyewitness expert evidence vs. erroneous eyewitness expert evidence vs. judicial instruction vs. no instruction) between-subjects factorial design was employed.

Materials

The minimal trial and pre-trial instruction were identical to those used in Experiments 1-3.

Eyewitness Testimony

The testimony of six of the eight eyewitnesses from the "good" condition were used in this experiment (see Appendix I, p. 262). These six interviews were obtained from the three eyewitnesses who expressed the most confidence in their decisions and made accurate identifications, and the three least confident eyewitnesses who made inaccurate identifications. This procedure did not alter the magnitude of the confidence-accuracy correlation ($r_{pbi} =$ .827, p = .042, n = 6). Five of these witnesses were the same as those used in Experiment 3.

Accurate and Erroneous Expert Evidence

A research psychologist (a different individual from in the previous studies) acted in the role of the expert in eyewitness identification issues. During the examination-in-chief the expert outlined her credentials, her current position, her area of expertise and her research history. She then addressed three key issues regarding eyewitness testimony: 1) the nature of memory as a reconstructive process; 2) system and estimator variables including distance, lighting, disguise, race and lineup type; and 3) the confidence-accuracy relationship. The information provided by the expert on the last of these issues differed across expert conditions. In the accurate expert condition, the expert explained that "the confidence expressed by a witness was a good indicator of the accuracy of their identification", and also that it was a "strong predictor" of identification accuracy. Conversely, in the erroneous expert condition, the expert suggested that there was no relationship between a witness's confidence and their accuracy. In all other respects the examination in chief was identical across conditions. The cross-examination of the expert was also identical across expert conditions, and highlighted some of the limitations of expert psychological testimony including: 1) the reliance on mock-crime paradigms and undergraduate participants in laboratory research; 2) the questionable ecological validity of studies which employ mock-crimes and mock-witnesses; and 3) the probabilistic nature of psychological testimony. The expert was given no prior knowledge regarding the nature or content of the questions contained within the cross-examination. The cross-examination lasted for 2.5 minutes while the examination-in-chief was approximately five minutes in duration.

Judicial Instruction

The role of the judge was played by the same research psychologist providing the expert evidence for this study. The judicial instruction (see Table 1.1 p. 4) lasted for approximately four minutes.

Juror Responses

Participant-jurors completed the same response form used in Experiments 2 and 3, however, the precise wording of the multiple choice questions relating to the expert's evidence were altered to incorporate direct and paraphrased quotes from the new expert's testimony. Of the four multiple choice options relating to the confidence-accuracy relationship, one directly quoted the accurate expert's testimony, one directly quoted the erroneous expert's testimony, while the remaining two paraphrased each of these quotes (see Appendix K).

Procedure

Participant-jurors were tested in groups as part of a tutorial activity and were seated at individual computers loaded with the trial video. Tutorial groups were randomly assigned to one of the four instruction conditions (accurate expert, erroneous expert, judicial or control) and participant-jurors within these groups were randomly assigned to view one of the six witness interviews. Thus, across the experiment each witness interview was seen under each of the different instruction conditions. In all other respects, the procedure adopted for this experiment was identical to Experiments 2 and 3.

Results

Jurors' Qualitative Evaluations of Eyewitnesses

Jurors evaluated the eyewitnesses on seven personality and performance dimensions using 7-point scales. Analysis of these responses, comparing ratings assigned to witnesses who had made accurate identifications and those who had made inaccurate identifications, revealed that accurate witnesses were rated as significantly more credible than inaccurate witnesses ($t_{(293)}$ = -3.24, p = <.005). Participant-jurors did not significantly differentiate between accurate and inaccurate eyewitnesses with respect to perceived accuracy ($t_{(294)}$ = -0.03, p = .975), attractiveness ($t_{(294)}$ = -0.38, p = .707), likeability ($t_{(295)}$ = 0.17, p = .863), trustworthiness ($t_{(293)}$ = -1.71, p = .089) or anger ($t_{(295)}$ = 81, p = .419). The jurors' ratings of eyewitness confidence will be discussed below.

Jurors' Qualitative Evaluations of Instructors

Participant-jurors rated the accurate expert, the erroneous expert and the judge with regard to their credibility, and the clarity and utility of the advice they provided (see Table 9.1 for mean ratings). Comparisons of these ratings revealed no significant differences in either utility ($F_{(2,230)}$ = 0.75, p = .473) or clarity ($F_{(2,230)}$ = 0.75, p = .474). A one-way ANOVA did however identify a significant difference between instruction types for credibility ($F_{(2,230)}$ = 16.74, p = <.0005). Post-hoc analysis showed that the judge was rated as significantly less credible than either the accurate or erroneous expert.

Table 9.1 : Jurors' Mean (SE) Evaluations of Instructors

	Accurate Expert (SE)	**Erroneous Expert (SE)**	**Judge (SE)**
Credibility	5.43 (1.02)	5.65 (0.78)	4.79 (1.03)
Utility	4.78 (1.07)	4.88 (0.93)	4.67 (1.09)
Clarity	5.38 (1.11)	5.49 (1.03)	5.29 (1.00)

Jurors' Recall of Expert Evidence or Judicial Instruction

On average those jurors in the accurate expert condition scored 3.27 ($\overline{\sigma}$ = 0.64) correct out of a possible four on the multiple choice recall task. Those in the erroneous expert condition also scored an average of 3.27 ($\overline{\sigma}$ = 0.55), while those in the judicial instruction condition scored an average of 2.84 ($\overline{\sigma}$ = 0.61). A univariate ANOVA with post-hoc analyses indicated that juror recall for expert evidence (of either type) was significantly better than for judicial instruction ($F_{(2,230)} = 12.43$, $p < .0005$). For the specific item relating to the confidence-accuracy relationship, 83.8% of jurors in the accurate expert condition selected the verbatim quote, and 8.8% chose its paraphrased alternative, meaning that 92.6% of jurors had a largely accurate recollection of the evidence they had heard. A very similar pattern was evident in the erroneous evidence condition, with a total of 90.1% of jurors accurately recalling the details of the erroneous expert evidence (88.9% selecting the verbatim quote and 1.2% choosing the paraphrased alternative). The observed patterns of accuracy were not found to differ significantly across expert instruction conditions ($\chi^2_{(1)} = 0.29$, $p = .539$). These findings show that the participant-jurors had a good memory of the instruction they heard (although significantly poorer in the judicial condition than the expert conditions), and that memory for the accurate and erroneous evidence was equivalent.

Juror Decision Criterion

A chi-squared test-of-independence indicated that there was no significant association between belief (i.e., whether the participant-juror believed the eyewitness or not) and instruction condition ($\chi^2_{(3)} = 1.16$, $p = .764$). This result indicates that belief rates were comparable across all conditions, thereby showing that neither expert testimony (accurate, 42.5% belief; erroneous, 48.1% belief) nor judicial instruction (51.4% belief) caused jurors to adopt a more stringent or lenient belief criterion than that displayed by jurors in the control condition (45.5% belief). Furthermore, a Chi-squared goodness-of-fit test indicated that this decision criterion was not biased towards belief or disbelief decisions, as rates of belief (46.9% overall) did not differ significantly from 50% across all instruction conditions ($\chi^2_{(1)} = 1.10$, $p = .295$).

Juror Sensitivity to Eyewitness Confidence

A 4x2 between groups ANOVA was conducted to explore the effect of instruction type and eyewitness accuracy on participant-jurors' estimates of eyewitness confidence. There was a significant main effect for eyewitness accuracy ($F_{(1,289)}$ = 72.60, p < .0005, η^2_p = 0.201), such that witnesses who had accurately identified the perpetrator from the lineup were rated by the jurors as being significantly more confident than eyewitnesses who had identified the innocent suspect (accurate eyewitnesses, $\bar{x}$ = 4.57, $\bar{\sigma}$ = 0.10; inaccurate eyewitnesses, $\bar{x}$ = 3.35, $\bar{\sigma}$ = 0.10). The main effect of instruction type was not significant ($F_{(3,289)}$ = 0.40, p = .756 η^2_p = 0.004) and there was no significant interaction between accuracy and instruction type ($F_{(3,289)}$ = 0.69, p = .559 η^2_p = 0.007; see Figure 9.1 below).

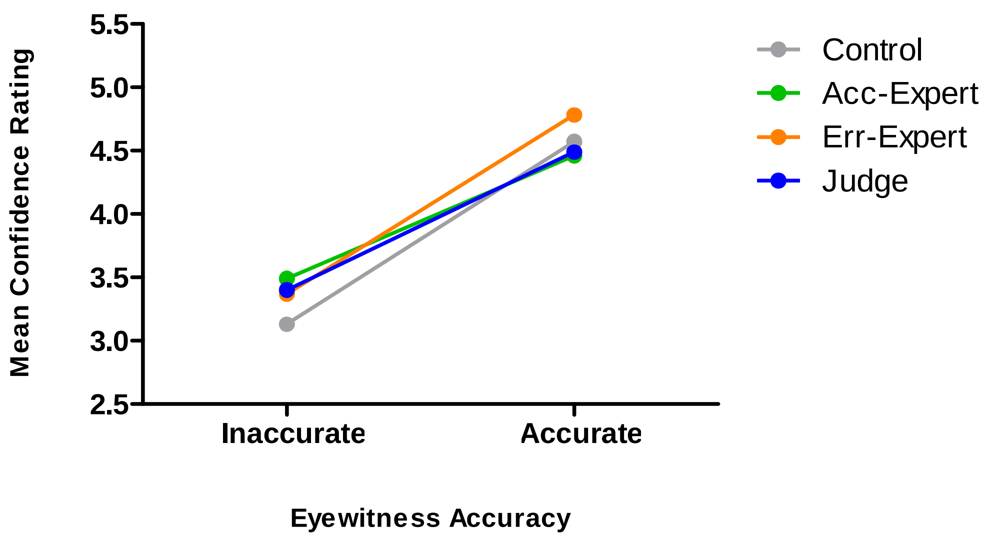

Figure 9.1 : Mean confidence ratings by eyewitness accuracy.

A 4x2 between groups ANOVA was also conducted to investigate the impact of instruction type and belief (i.e., whether the juror believed the eyewitness or not) on participant-jurors' estimates of the confidence of the eyewitness they saw. Although instruction type had no impact on juror's perceptions of confidence ($F_{(3,288)}$ = 0.29, p = .803, η_p^2 = .003), a significant main effect for belief was observed such that those eyewitnesses who were believed ($\bar{x}$ = 4.44, $\bar{\sigma}$ = 0.11) were rated as being significantly more confident than those who were disbelieved ($\bar{x}$ = 3.54, $\bar{\sigma}$ = 0.11, $F_{(1,288)}$ = 37.37, p < .0005, η_p^2 = .115). A significant interaction effect was also identified ($F_{(3,288)}$ =

2.69, p = .047, η_p^2 = .027), however post-hoc analyses revealed no significant pair-wise differences.

Figure 9.2 : Mean confidence ratings by juror belief decision.

A follow-up analysis considered only the ratings of witness confidence assigned by jurors in one of the two expert conditions (see the data represented by the green and orange lines in Figure 9.2 above). In this analysis, in addition to the main effect of belief reported above, the belief by expert evidence interaction was also significant ($F_{(1,159)}$ = 7.07, p = <.01). Participant-jurors who heard the testimony of the accurate expert showed a greater differentiation in the ratings of confidence they assigned to believed and disbelieved witnesses ($\Delta\bar{x}$ = 1.47) than jurors who heard the evidence presented by the erroneous expert ($\Delta\bar{x}$ = 0.42). Thus, jurors told to use witness confidence to help them to evaluate the accuracy of the eyewitness believed those witnesses they perceived to be more confident and disbelieved those witnesses perceived to be less confident (or alternatively rated believed eyewitnesses as more confident and disbelieved eyewitnesses as less confident). In contrast, the jurors who were told that confidence was not a useful indicator of eyewitness accuracy rated the confidence of believed and disbelieved eyewitnesses at equivalent levels. Finally, the significant, though imperfect, correlation between the participant-jurors' *perceptions* of eyewitness confidence and the eyewitness's own ratings of confidence (r = .593, p < .0005), suggest that when making their evaluations, participant-jurors were considering factors other than the eyewitness's numerical estimate of confidence. Had they been basing their confidence estimates on the eyewitness's verbal estimates alone, the correlation between the two estimates ought to have approached unity.

Considerations Affecting Belief Decisions

In order to investigate the factors considered by participant-jurors when making their decision to believe or disbelieve the eyewitness they saw, a series of (2x4) ANOVA's were conducted, each investigating the effect of belief decision (whether the eyewitness was believed or not) and instruction condition on participant-jurors' ratings of the extent to which the three variables (eyewitness confidence, eyewitness manner and eyewitnessing conditions) influenced their judgements. For each of these three analyses, there was a main effect of instruction condition, but no main effect of belief type or interaction effect. Post-hoc analysis revealed that jurors in the accurate expert condition (those told to consider confidence when evaluating the eyewitness) rated eyewitness confidence as significantly more influential than did the jurors in the erroneous expert condition ($F_{(3,288)}$ = 4.26, p = <.01, η_p^2 = .042; see Figure 9.3 below).

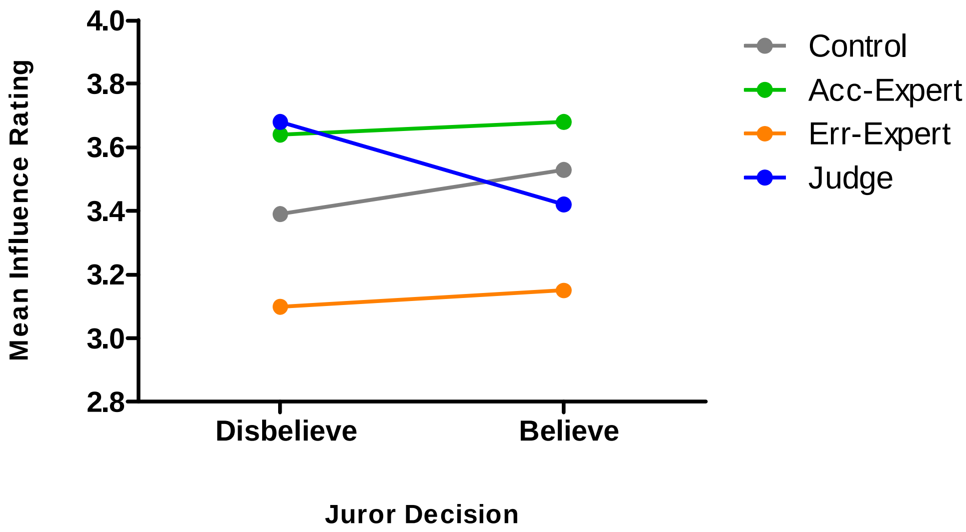

Figure 9.3 : Ratings of the influence of eyewitness confidence by belief decision

Participant-jurors in the judicial instruction condition reported relying on eyewitness manner significantly more than those from any other condition ($F_{(3,288)}$ = 5.33, p = <.005, η_p^2 = .053), while participant-jurors from the control condition were significantly *less likely* than other groups to report using information about witnessing and identification conditions when formulating their decision ($F_{(3,288)}$ = 7.24, p = <.0005, η_p^2 = .07).

Table 9.2 : Jurors' Mean (SE) Ratings of Influence of Eyewitness Manner and Witnessing Conditions.

	Mean (SE) Eyewitness Manner		**Mean (SE) Witnessing Conditions**	
	Disbelieved	**Believed**	**Disbelieved**	**Believed**
Control	3.00 (0.17)	3.53 (0.18)	3.03 (0.16)	3.07 (0.17)
Accurate Expert	3.40 (0.15)	3.47 (0.17)	3.44 (0.14)	3.65 (0.16)
Erroneous Expert	3.10 (0.15)	3.39 (0.16)	3.64 (0.15)	3.64 (0.15)
Judge	3.88 (0.17)	3.78 (0.17)	3.68 (0.16)	3.81 (0.16)

Predictors of Juror Belief Decisions

A binary logistic regression was conducted in order to investigate which, if any characteristic of the eyewitness predicted participant-juror decisions to believe or disbelieve the eyewitness they saw. This analysis was conducted separately for each instruction condition using the participant-jurors' rating of eyewitness credibility, accuracy, confidence, attractiveness, likeability, trustworthiness and anger as the predictors in the model.

In the control condition none of these factors significantly predicted belief decisions, although the overall model was significant ($\chi^2_{(7)} = 18.42$, $p = <.05$). A significant model emerged from analysis of the accurate expert condition, which correctly classified participant-juror decisions 76.9% of the time (see Table 9.3 below for all variables in the equation; $\chi^2_{(7)} = 32.55$, $p = <.0005$). Juror estimates of eyewitness confidence was a significant predictor in this model, such that when holding all other factors constant, an increase in rated confidence by one response-scale unit increased the odds that a participant-juror would believe an eyewitness by approximately 3.03 times ($\beta = 1.11$, $p < .005$).

Table 9.3 : β (S.E.), Significance and Exp(β) for Predictor Variables in Accurate Expert Condition

	β (S.E.)	***p***	**Exp(β) (95% CI)**
Credibility	0.11 (0.44)	.804	1.12 (0.47-2.65)
Accuracy	-0.11 (0.39)	.786	0.90 (0.42-1.94)
Confidence	1.11 (0.32)	.001	3.03 (1.61-5.69)
Attractiveness	0.20 0.29)	.481	1.23 (0.69-2.15)
Likeability	-0.34 (0.39)	.386	0.71 (0.33-1.54)
Trustworthiness	0.68 (0.43)	.113	1.97 (0.85-4.57)
Anger	-0.16 (0.38)	.682	0.86 (0.41-1.80)

In the erroneous expert condition analysis of participant-juror responses also revealed a significant model ($\chi^2_{(7)}$ = 33.64, p = <.0005), which accurately classified participant-juror decisions 79% of the time (see Table 9.4 below for all variables in the equation). In this model, the participant-jurors' evaluation of the accuracy of the eyewitness significantly predicted the final decision whether to believe them or not, such that an increase in participant-jurors ratings of the accuracy of the eyewitness of one scale point increased the odds that a participant-juror would believe an eyewitness by approximately 3.85 times (β = 1.35, $p < .005$) when holding all other predictors constant.

For participant-jurors who received the judicial instruction, both perceived accuracy and attractiveness of the eyewitness were shown to be significant predictors of the decision to believe (see Table 9.5 below), such that a one unit increase in perceived accuracy increased the odds of belief by approximately 2.15 times (β = 0.78, $p < .05$), while a decrease in attractiveness of one unit saw the likelihood of belief decrease by a factor of 0.51 (β = -0.67, $p < .05$) when all other predictors were held constant. Overall this model was significant ($\chi^2_{(7)}$ = 20.08, $p < .01$), accurately classifying 71.4% of cases.

Table 9.4 : β (S.E.), Significance and Exp(β) for Predictor Variables in Erroneous Expert Condition

	β (S.E.)	***p***	**Exp(β) (95% CI)**
Credibility	0.60 (0.48)	.207	1.82 (0.72-4.62)
Accuracy	1.35 (0.45)	.003	3.85 (1.59-9.32)
Confidence	-0.11 (0.25)	.646	0.89 (0.55-1.45)
Attractiveness	0.40 (0.31)	.195	0.48 (0.82-2.70)
Likeability	-0.33 (0.38)	.393	0.72 (0.34-1.53)
Trustworthiness	-0.04 (0.44)	.934	0.97 (0.41-2.27)
Anger	0.12 (0.35)	.724	1.13 (0.57-2.24)

Table 9.5 : β (S.E.), Significance and Exp(β) for Predictor Variables in Judicial Regression

	β (S.E.)	***p***	**Exp(β) (95% CI)**
Credibility	0.29 (0.32)	.365	1.34 (0.71-2.53)
Accuracy	0.78(0.33)	.021	2.15 (1.12-4.13)
Confidence	0.19 (0.24)	.428	1.21 (0.76-1.92)
Attractiveness	-0.67 (0.32)	.035	0.51 (0.28-0.95)
Likeability	0.36 (0.43)	.405	1.43 (0.61-3.35)
Trustworthiness	-0.04 (0.33)	.910	0.96 (0.51-1.83)
Anger	-0.51 (0.47)	.283	0.60 (0.24-1.52)

Juror Sensitivity to Eyewitness Accuracy

Overall, the participant-juror evaluations of the accuracy of the eyewitness were correct 63.6% of the time. This represents a level of discrimination which is significantly better than would have been expected by chance alone ($\chi^2_{(1)} = 22.09$, $p <$ .0005), this was also true within the control (71.2% accuracy, $\chi^2_{(1)} = 11.88$, $p < .005$), accurate expert (66.3% accuracy, $\chi^2_{(1)} = 8.45$, $p < .005$) and (marginally) the judicial instruction condition (61.4% accuracy, $\chi^2_{(1)} = 3.67$, $p = .058$). Performance in the erroneous expert condition was not significantly better than chance (56.8% accuracy, $\chi^2_{(1)} = 1.49$, $p = .222$). There was no significant association between instruction type and accuracy ($\chi^2_{(3)} = 3.661$, $p = .300$). This was also true when only the two expert conditions were considered ($\chi^2_{(1)} = 1.52$, $p = .271$).

Signal Detection Analysis

Signal detection measures were calculated for each instruction condition (see Table 9.6) in order to estimate participant-jurors' Sensitivity to the accuracy of eyewitness identifications (*d'*) and their Skepticism (*C*). At the $p = .05$ level, no reliable difference in Sensitivity was observed between the control condition (*d'* 95% CI : 0.51 to 1.81) and erroneous expert evidence (*d'* 95% CI : -0.21 to 0.89). A comparison of the pair of the *C* estimates showing the greatest difference (control vs. judge) suggest that there is no reliable difference in observed Skepticism between the control condition (*C* 95% CI : -0.16 to 0.51) and the judicial instruction (*C* 95% CI : -0.35 to 0.24).

Table 9.6 : Judgement Type as Percent Within Instruction Condition, Observed *d'* and *C* Values.

	Judgement Type					
	Miss	**Hit**	**Correct rejection**	**False alarm**	***d'***	***C***
Control	18.2%	34.8%	36.4%	10.6%	1.16	0.17
Accurate Expert	17.5%	26.6%	40%	16.3%	0.81	0.15
Erroneous Expert	22.2%	27.2%	29.6%	21%	0.34	0.05
Judge	17.1%	30%	31.4%	21.4%	0.59	-0.06

Discussion

The results of this experiment are consistent with those of Experiments 2 and 3, providing evidence that participant-juror accuracy did not vary significantly as a function of the objective quality of the eyewitness expert's evidence. Furthermore, also in keeping with earlier experiments, participant-jurors in the judicial instruction condition performed as well as participants in the expert and control conditions, and direct and indirect methods provided differing estimates of participant-juror SEA.

Independent Effects of Instruction

Beginning with judicial instruction, overall it was found that participant-jurors in this condition were as likely to believe an eyewitness as to disbelieve them. Moreover, as there was no significant association between instruction type and likelihood of belief, there is no evidence to suggest that the judicial instruction served to induce Skepticism among these participant-jurors.

Participant-jurors in both the accurate and erroneous evidence conditions showed responses consistent with significant Sensitivity to the testimony given by the eyewitness expert regarding confidence and accuracy. Firstly, there was a significant interaction between the type of expert evidence and participant-jurors' tendency to believe more confident eyewitnesses and disbelieve those who were less confident. Specifically, participant-jurors told not to rely on confidence as a predictor of accuracy showed a significantly smaller difference in the confidence attributed to believed and disbelieved eyewitnesses compared to jurors who were told that confidence was a useful predictor of identification accuracy. Secondly, regression analyses revealed that a different set of factors predicted participant-juror belief decisions in the accurate and erroneous evidence conditions. In fact, the belief decisions of those participants told confidence was associated with accuracy were significantly predicted by the jurors estimates of the eyewitness's confidence. While the belief decisions of those told not to rely on eyewitness confidence, appear to have been influenced by their perceptions of the eyewitness's accuracy. The model for the decisions reached in the judicial instruction condition was different again, showing participant-juror beliefs to be significantly predicted by estimates of the eyewitness's accuracy and attractiveness. Thirdly, when deciding whether to believe or disbelieve the identification made by an eyewitness, participant-jurors in the accurate evidence condition (who were told confidence was associated with accuracy) reported relying on eyewitness confidence significantly more than did members of the erroneous expert group; while participant-jurors in the judicial condition reported relying on eyewitness manner more than any other group, and members of the control condition reported relying on witnessing conditions significantly less than any of the instruction conditions. Finally, participant-jurors were able to recall what the accurate and erroneous experts said regarding the confidence-accuracy correlation, thereby leading to the conclusion that participant-jurors displayed Sensitivity to the Expert's Opinions in that they attended to, remembers and attempted to act in accordance with the advice they were provided. Moreover, this Sensitivity was found in the absence of an associated Skepticism effect, as the decision criterion adopted by participant-jurors in the expert conditions did not differ significantly from that observed in the control or judicial conditions.

This pattern of results is entirely consistent with the results from Experiments 2 and 3 despite significant differences in the materials used: a) the expert (and judge) in this study was female rather than male; b) the general testimony of the expert differed in minor ways according to the style of the speaker (although the interview schedule remained the same); and c) the expert evidence was varied in this case, so as to be accurate or erroneous, rather than remaining constant. Thus, the replication of the previous finding of Sensitivity in the absence of Skepticism does not appear to be an artefact of the expert, their individual style, or the eyewitnesses evaluated. At first sight this result may look like encouraging evidence for the inclusion of eyewitness expert testimony, however direct analyses of participant-juror Sensitivity to Eyewitness Accuracy (discussed under the subheading *Instruction Type and Sensitivity to Eyewitness Accuracy* below) provides some grounds for caution.

Expert Evidence on Eyewitness Confidence

As in the previous investigation of expert evidence regarding eyewitness confidence, we have found that participant-jurors are willing and able to follow an expert's advice. Moreover, this particular experiment demonstrates that participant-jurors are equally able to act on advice *to disregard*, as *to rely* on an eyewitness's confidence, as participant-jurors complied with the experts advice in both the accurate and erroneous conditions. Thus, irrespective of the true situation with regard to the confidence-accuracy relation, participant-jurors understand and attempt to follow the expert's advice – even when it is inappropriate to do so.

This result suggests that it is of vital importance for the ground truth to be reliably established before expert evidence is provided (Elliott, 1993), as jurors will just as readily follow accurate as erroneous advice. This result is in contrast to the hopeful interpretation of Experiments 2 and 3, where it was suggested that jurors might be able to determine the accuracy of the expert's evidence and act accordingly. Instead, there is evidence that participant-jurors are happy to defer to the expert's advice and, as a result de-emphasise a useful predictor when told to do so. Although there is no evidence that this practice resulted in Skepticism on the part of the jurors, some might argue that this adherence to erroneous expert advice illustrates the concern that jurors might be overwhelmed or subservient to expert evidence:

> *The idea that eyewitness expert testimony may prejudice jurors against the eyewitness rather than compel them to carefully apply new knowledge in their own evaluation of the eyewitness may be the strongest objection to such testimony* (Doyle, 1989 cited in Leippe, 1995, p. 922).

Of course the fear that participant-jurors are unduly influenced by expert evidence is only a concern if the demonstrated tendency to follow erroneous evidence, when given, also results in an increased tendency to err when evaluating the accuracy of an eyewitness; that is when estimates taken from indirect measures closely approximate performance accuracy values arrived at through direct means. This issue will be discussed next.

Instruction Type & Sensitivity to Eyewitness Accuracy

Overall, participant-jurors performed significantly better than chance with 63.6% of their decisions whether to believe the eyewitness proving to be correct. Despite this, there was no significant association between instruction condition and participant-juror performance accuracy, indicating that this level of accuracy was achieved irrespective of the advice given (expert or judicial) or its quality (accurate or erroneous). This is surprising given that those participants in the erroneous expert condition were following advice which required them to adopt an unhelpful decision cue, while those jurors from the accurate expert condition followed advice directing them to use a highly useful decision cue. Indeed, even analyses investigating the association between juror discrimination accuracy and instruction *in only these two groups* failed to identify any significant relationship. However, it should also be noted that the levels of accuracy, although far from perfect, are also far from discouraging particularly in the case of the control condition where accuracy was over 70%.

Moreover, the evidence reported here suggests that participant-jurors were equally likely to *believe accurate eyewitnesses* as to *disbelieve inaccurate* *eyewitnesses*. While this does seem encouraging, it is also likely given previous research (Bradfield & Wells, 2000; Brewer & Burke, 2002; Cutler, Penrod & Stuve, 1988; Lindsay, Wells & O'Connor, 1989; Luus & Wells, 1994), and the evidence collected here (that believed eyewitnesses were rated as significantly more confident than disbelieved eyewitnesses), that jurors in all but the erroneous expert condition were using eyewitness confidence to determine their decisions, at least to some extent. Given that confidence was such a strong predictor of accuracy, it is likely that this one factor accounts for much of the discrimination accuracy observed. Thus, it may not be realistic to suggest that the performance of real jurors approximates that observed in this sample of participant-jurors, unless of course similar confidence-accuracy relationships were anticipated. Equally, it could be argued that the jurors' failure to achieve perfect or near perfect discrimination performance is surprising and somewhat concerning in light of the fact that confidence was such a strong predictor of eyewitness accuracy in this sample (r_{pbi} = .827). Irrespective of whether juror discrimination performance is considered to be good compared with previous research, or poor given the specifics of this study, it remains difficult to account for the performance of those participant-jurors who were shown to be using cues other than eyewitness confidence to guide their decision-making. That is, in all respects, members of the erroneous expert condition reported and behaved consistent with having abandoned their use of eyewitness confidence, yet somehow their performance is not significantly worse than the other groups of participants. This finding together with those of previous experiments reported here suggests that eyewitnesses are, in some way, Sensitive to the *actual accuracy* of the eyewitness. This is an exciting result to have replicated particularly since it has now been found in three studies (using the testimony of 13 eyewitnesses from two different crime scenarios and the judgements of a total of 763 participant-jurors), and although it is unclear what mechanism is being used to achieve this outcome, the persistence of the effect warrants further investigation.

Relative Effects of Instruction

The results of this experiment also replicate the findings from Experiments 1, 2 and 3 regarding the relative effects of judicial instruction and expert evidence. The evidence suggests that although there are differences between participant-jurors' perceptions of the two instruction types, these discrepancies do not extend to participant-juror discrimination performance, which was statistically equivalent in all conditions. Participant-jurors perceived the judge to be less credible than either the accurate or erroneous expert, and recalled significantly fewer correct items from the expert evidence. Yet, as before, these discrepancies did not appear to translate into any discernable difference in performance accuracy.

Comparisons between participant-juror perceptions of accurate and erroneous expert evidence revealed no differences whatever either with regard to participant-juror accuracy or the subjective ratings. Participant-jurors rated each of the variants of expert evidence as being equally credible, useful and clear, and they recalled the evidence with equivalent accuracy. Moreover, the jurors who heard the evidence of an expert who correctly described a method by which they could improve their performance, were no better evaluators of eyewitness accuracy than the participant-jurors who heard evidence which, had they followed it, ought to have impaired their performance. The evidence presented here therefore indicates that accurate eyewitness expert evidence regarding the confidence-accuracy relationship does not significantly improve the accuracy of juror evaluations of eyewitness evidence even when it is compared with erroneous information designed to negatively impact performance on this issue. Given that no effect of expert evidence could be discerned when comparing the impact of helpful and accurate advice to the impact of advice deliberately designed to disrupt juror performance, this suggests that any impact of expert advice on jurors in this case is minimal at best. Moreover, the absence of a significant difference between the types of expert evidence does not appear to be due to a lack of statistical power. For this comparison there was a 71% chance ($1-\beta = 0.71$, $\alpha = .05$) of being able detect an effect that was small to moderate in size (w : between 0.1 and 0.3), and a 96% chance of detecting an effect that was moderate ($w > 0.3$). Given that this was the power to detect a difference between the two conditions which should have resulted in the greatest possible effect of this expert evidence (i.e., the difference between accurate and erroneous advice) then it is reasonable to expect to see at least a moderate effect. Thus, it is with some confidence that we conclude that in this case the expert did not exert anything other than a small, if not trivial, effect on juror discrimination accuracy.

Predicting Sensitivity to Eyewitness Accuracy

Consistent with Experiments 1 through 3, and research by Wells and colleagues (Wells & Wright, 1983 cited in Wells, 1986; Wells et al., 1980) the results from this study indicate that direct and indirect measures of participant-juror Sensitivity to Eyewitness Accuracy produce qualitatively different impressions of eyewitness expert effects. Specifically, in this experiment participant-jurors were shown to be Sensitive to Expert Opinions, of varying quality, however the nature of advice provided did not moderate participant-juror capacity to evaluate eyewitness identifications. Thus, the estimates provided by SEO methodologies do not appear to be directly related to participant-juror accuracy, with indirect measures providing a more positive impression of the expert's impact than is warranted given the pattern of the data collected using direct measures.

Limitations

The external validity of this study, and indeed Experiments 2 and 3, may have been limited somewhat by the decision to concentrate on the confidence-accuracy relationship as the key feature of the expert's evidence. It is fully accepted that a different pattern of instruction effects might have been observed if a different eyewitnessing factor had been manipulated and different advice proffered by the experts. On the other hand, however, there is substantial evidence to suggest that eyewitness experts are willing to provide testimony on this issue (Kassin et al., 1989; Kassin et al., 2001), and that there is disagreement regarding the appropriateness of an expert doing so given the empirical evidence available (Brewer & Wells, 2006; Brigham, 1988; Fleet, Brigham, & Bothwell, 1987; Sporer, Penrod, Read, & Cutler, 1995; Weber & Brewer, 2003). Thus, while the issue of confidence-accuracy may not be ideal or representative of *all* topics the expert may testify to, it is neither inappropriate to focus on this issue, nor irrelevant to an understanding of eyewitness expert effects on juror decision-making.

Conclusions

Despite the evidence that the participant jurors understood and respond appropriately to the expert evidence or judicial instruction they received, the objective accuracy of the judgements they made were not found to be significantly associated with instruction type. That is, participant jurors who were correctly informed that confidence was a good predictor of eyewitness accuracy did not perform significantly better than participant jurors who were erroneously told that confidence was a poor predictor of accuracy. Thus, this experiment provides little support for the efficacy of eyewitness expert testimony with regard to a SEA criterion, yet it does add to a growing body of evidence which suggests that juror discrimination accuracy may not actually be as poor as previously reported.

Chapter 10

Experiment 5 - Expertise Study

Results from the previous four experiments have indicated that jurors understand and are sensitive to the opinions provided by the experts, but despite this, expert evidence has not significantly improved the jurors' ability to determine the accuracy of a witness's testimony. Three possible explanations for this present themselves. The first: that participant-jurors are not receiving or acting upon expert recommendations, is not borne out by the data; as the evidence clearly indicates that participants are responsive to expert advice. The second suggests that the experimental materials used to represent the testimony of the eyewitness expert may be at fault:

> *it would be naive to expect that a teacher could be assessed by evaluating a one-page verbal summary of one of her or his lectures, and no one would do this. Similarly, it is naive to expect that the effectiveness of an eyewitness expert could be assessed by evaluating a one-page verbal summary of her or his testimony* (Pezdek, 2007, p. 112).

Accordingly, it may be that the video interviews conducted with the expert, lasting approximately eight minutes in total, were insufficient to either accurately represent expert testimony, or to provide participant-jurors with sufficient information to improve their discrimination accuracy. The third possible explanation holds that the failure to detect a significant improvement in discrimination accuracy is related to the difficulty of the actual discrimination task and the utility of the experts' expertise. Results from Experiment 4 indicate that even when participants are told to rely on a cue which strongly predicts the accuracy of an identification, discrimination accuracy, although significantly better than chance, remained below 70%. It is possible that this objectively low rate of accuracy is tied to the nature, or form, of the testimony the expert can offer. That is, the eyewitness expert can only provide testimony indicating that confident witnesses are *generally* accurate or vice versa. The expert can never state with any certainty if this particular confident witness was accurate. Thus, the participant-juror is left to integrate numerous probabilistic predictors of accuracy of each eyewitness in order to reach a final evaluation of their accuracy; it may be that the difficulty of this task overwhelms the potential benefit that can be derived from the experts' knowledge of the significance of various predictors.

To date there is no data available in the eyewitness expert effects literature which can serve to disentangle the last two of these possible explanations for the failure to observe an expert effect on SEA: a) that the trial materials were inappropriately abridged; or b) that the knowledge of significant predictors alone is insufficient to improve the integration of probabilistic predictors in the individual instance. Accordingly, Experiment 5 represents a means by which to begin to resolve this issue. Specifically, this experiment will test the eyewitness discrimination accuracy of participants with varying levels of expertise, from novice to expert.

This experimental design helps to differentiate between the two remaining explanations for expert inefficacy by recruiting a sample of participants who are experts themselves, and who therefore have all of the information that an expert could provide novice jurors regarding eyewitness identification issues. Moreover, this approach also has the potential to provide valuable information regarding the *amount* of training necessary before a novice juror can perform with equivalent accuracy to an eyewitness expert. The implications of this research for the eyewitness expert and eyewitness discrimination tasks are clear. If it can be demonstrated that the expertise of an expert (i.e., the knowledge of significant predictors of identification accuracy) significantly improves discrimination accuracy when compared with novice performance, it seems likely that the expert inefficacy observed in Experiments 1 to 4 can be attributed to the abridged experimental materials. Moreover, it will then be possible to turn our attention to the acquisition of expertise, with the study providing estimates of the amount of training necessary to attain “expert” levels of performance. On the other hand, if no evidence is found to show that the experts themselves are more likely to make accurate identifications than untrained jurors, then we must begin to question the role eyewitness experts play in the courtroom. Specifically, if the role of the expert is to educate jurors in order to improve discrimination accuracy, the absence of a significantly greater ability to do so themselves makes it difficult to expect: a) that they possess the knowledge necessary to improve task performance, or b) that they could then pass that knowledge onto others faced with the same task.

While it is difficult to make any firm predictions regarding either expert performance on the discrimination task, or the amounts of training necessary to achieve “expert” discrimination accuracy, evidence obtained in the detection of deception literature provides some guidance. In this field it is common for researchers to evaluate the expertise claimed by people in regards to the detection of deception, such as police officers and FBI agents. In their recent meta-analysis, Bond and DePaulo (2006) reviewed studies which assessed the ability of a total of 2842 “experts” to discriminate a truth from a lie. Their analysis found no evidence that “experts” were more sensitive to lies than a layperson. Thus, despite the many differences between truths and lies, and inaccurate and accurate eyewitnesses, this data suggests that it is possible, if not likely, that the expertise of the expert will not actually translate successfully into improved discrimination accuracy. If this is the case the role of the eyewitness expert in court may need to be re-evaluated.

Method

Design

This study employed a 4 (expertise level; community, undergraduate psychology student, postgraduate psychology student, expert psychologist) x 2 (witness accuracy; accurate v inaccurate) factorial design, where each participant evaluated the testimony of three eyewitnesses, each from different crime scenarios. The three witnesses seen by each participant were either all accurate, all inaccurate or a combination of accurate and inaccurate witnesses (two accurate and one inaccurate eyewitness or one accurate and two inaccurate eyewitnesses).

Participants

There were four different groups of participants:

Community Participants

Community participants were contacted through an existing memory research database. All people within that database had, at some earlier point, indicated that they would be willing to participate in memory research conducted by the School of Psychology at UNSW. An email was sent to each member of this database asking if the member would like to participate in an eyewitness evaluation task. Those people who indicated that they would be willing to complete the task were subsequently mailed the necessary experimental materials. One hundred and forty-eight jury eligible community members were sent the experiment pack; 53 participants responded and were included in the final analysis (35.8% participation rate).

Undergraduate Participants

Psychology students who were completing a third year psychology and law course participated in this experiment as part of their tutorial activities. As this experiment was run in the final teaching week of the course, these participants had been lectured extensively on issues surrounding human memory as well as the system and estimator variables known to be associated eyewitness identification accuracy. Fifty-eight students in five tutorials completed the experiment.

Postgraduate Participants

These fourteen participants were recruited via an email to all members of a mailing list containing the names of graduates from the M.Psych (Forensic) program at the University of New South Wales. Additional postgraduate participants were recruited personally at meetings of both the British Association of Forensic Psychology and the European Association of Psychology and Law in 2006.

Expert Participants

Experts in the field were identified through: a) a PsycINFO search specifying articles published between 1996 and 2006 citing the key words *eyewitness, identification* and *eyewitness expert*; b) a selective search of PsycINFO to identify an additional list of authors whose work had been influential in the development of this thesis but who had not published in the above time period; c) a Google search using the terms *expert witness* and *eyewitness expert;* and d) through personal recruitment at meetings of both the British Association of Forensic Psychology and the European Association of Psychology and Law in 2006.

In total 129 discs were distributed to those experts and postgraduates for whom mailing addresses could be located, or who could collect the disc in person. Overall, 34 completed responses (n = 20 experts) were received making the response rate for the postgraduate and expert groups 26.4% overall.

Materials

All participants in this study watched the testimony of three eyewitnesses. After seeing each eyewitness they answered a series of questions about that witness before viewing the next one. Participants at each level of experience were provided slightly different participation materials, described below.

Eyewitness Testimony

The eyewitness testimony provided to participants in this study was collected using eyewitness Protocols I, II and III (see Appendix L for details regarding Protocol III). Overall, these three protocols resulted in testimony from 20 eyewitnesses, four from Protocol I, 14 from Protocol II and two from Protocol III. This sample of 20 eyewitnesses included the testimony of seven males and 13 females and the confidence-accuracy correlation for this group was 0.56 (p <.05, n = 20).

Participation Packs

Community Participants

The participation pack sent to interested community members included an instructional letter (giving participants guidance regarding informed consent and the experimental procedure), a paper questionnaire and a CD which contained three .wmv format video files. Each video sequence showed the evidence and the cross-examination of one eyewitness. The questionnaire included a written version of the pre-trial instructions used in all previous experiments, followed by a series of 10 questions pertaining to each eyewitness. Participants were asked: 1) if they believed the eyewitness had made the correct identification; 2) how confident they were in their decision; 3) how they rated the credibility, confidence, attractiveness and trustworthiness of the eyewitness; and 4) how they rated the role of witness confidence, manner, witnessing conditions and identification conditions when deciding whether to believe or disbelieve the eyewitness. Participants also provided basic demographic information, and were tested on their knowledge of eyewitness identification issues using five statements derived from Kassin et al. (2001, see Table 10.1 below). These participants were also asked if they had ever heard an expert psychologist talk about human/eyewitness memory (participants who answered “yes” to this question were excluded from subsequent analysis). The CD contained a random selection of eyewitnesses, the first from Protocol II, the second from Protocol I and the third from Protocol III. This order was selected to ensure that the largest number of eyewitnesses from the same scenario would be viewed by these participants on their first trial.

Table 10.1 : Statements for Evaluation (based on Kassin et al., 2001)

Statements for Evaluation	
If asked, would you believe the following statements?	
1	The potential for the wording of questions put to witness to affect their testimony.
2	The confidence of an eyewitness can be influenced by factors that are unrelated to identification accuracy.
3	Instructions can influence an eyewitness's willingness to make an identification.
4	An eyewitness's confidence is not a good predictor of their identification accuracy.
5	The presence of a weapon negatively impacts eyewitness identification accuracy.

Undergraduate Participants

Undergraduates participated in this study in their tutorial groups. Participants within each tutorial group saw the same three eyewitnesses testify, however, different tutorial groups saw different combinations of eyewitnesses. These participants also watched one eyewitness from each eyewitnessing protocol in an attempt to minimize the relative comparisons that could be made across trials. Undergraduate participants completed a paper version of the questionnaire differing slightly from the community group as these participants provided demographic information before evaluating the eyewitnesses, and did not rate the impact of identification conditions on their decision-making.

Postgraduate and Expert Participants

Participants in these groups received one CD containing all of the experimental materials. This disc included a .wmv video file for each of the three eyewitnesses as well as an electronic questionnaire (see Appendix M) which could be completed and returned online. The questionnaire began with the same pre-trial instructions provided in each of the other conditions, and was followed by the same evaluative questions as in the community questionnaire. These participants were then asked to provide some demographic information as well as a description of their employer and their general role i.e., practitioner or academic. Participants were then asked to describe their training, qualifications and experience before answering questions relating to their court experience (i.e., the number of court reports and appearances in the last five years and the grounds for exclusion if relevant). The expert participants were then asked whether or not they would provide the five statements (taken from Kassin et al., 2001) to a jury as part of their expert testimony. They were also asked to describe their areas of research, and estimate their number of peer reviewed and other publications relating to eyewitness evidence and eyewitness expert evidence. The eyewitness testimony included on each CD was selected in the same manner as used for the community sample, including a random selection of eyewitnesses from each of the three Protocols represented in the order II, I, III.

Procedure

Depending on group, participants were either read the pre-trial instructions (undergraduates) or read them themselves (community, postgraduates and experts). These instructions directed participants to watch one video at a time and to answer the relevant questions on their questionnaire after seeing each video clip. Those participants in the community sample mailed their completed questionnaires back to the University, while those participants in the postgraduate and expert samples emailed their electronically completed questionnaire back to a specified email account.

Results - Part I

Descriptive Analysis

Community Sample

The community sample was composed of 53 participants (37 female and 16 male) aged between 21 and 78 years ($\bar{x}$ = 53.58). Approximately 25% of these participants were employed by a government institution, 17% were self employed, 15% worked for a private company, 6% worked at a university or research institution and 35.9% specified “other”. Overall, these participants showed high levels of endorsement for all five statements testing their knowledge of eyewitness identification issues, with the highest level of agreement for statement one and the lowest level of agreement for statement four (see Table 10.2 below).

Table 10.2 : Percent of Participants Endorsing Each Statement

Statement	**% of Community sample** (n = 51) **willing to endorse the statement**	**% of Postgraduate sample** (n = 12) **willing to testify that:**	**% of Expert sample** (n = 20) **willing to testify that:**
1. The potential for the wording of questions put to a witness to affect their testimony	100%	58.3%	80.0%
2. The confidence of an eyewitness can be influenced by factors that are unrelated to identification accuracy	98.0%	50.0%	80.0%
3. Instructions can influence an eyewitness's willingness to make an identification	92.2%	33.3%	50.0%
4. An eyewitness's confidence is not a good predictor of their identification accuracy	74.5%	41.6%	75.0%
5. The presence of a weapon negatively impacts eyewitness identification accuracy	80.4%	33.3%	70.0%

Undergraduate Sample

The ages of the 58 students in the Undergraduate sample ranged between 19 to 47 years of age ($\overline{x}$ = 22.05). Forty-five of these participants were female.

Postgraduate Sample

The 14 postgraduate participants were aged between 22 and 33 years of age ($\overline{x}$ = 26.15), and all but one person in this sample was female.

Approximately 40% of the postgraduate students had a Bachelor level degree, 29% had a Masters level degree and 14% had a doctoral degree. Almost half (42.9%) described themselves as students (14.3% as students in professional training, 28.6% practitioners and 28.6% as academic), with 57.1% of the sample indicating that their employer was a university or research institution, 21.4% indicating a government forensic institution and 7.1% describing themselves as being employed in the private sector. None of these 14 postgraduate participants had published articles relating to eyewitness identification or eyewitness experts in peer reviewed journals.

Of those participants specifying their research interests, eyewitness memory (14.3%) and eyewitness identification (7.1%) were most frequently cited; however, almost 35.7% of this group did not respond to this question. Finally, of those postgraduates responding to the questions regarding viable topics for expert testimony, the largest number showed a willingness to testify to Statement 1 (58.3%), while the least were willing to testify to Statements 3 and 5 (33.3%) (see Table 10.2, p. 196).

Expert Sample

Twenty experts (seven female and 13 male) with an average age of 44.65 years (range 25-74 years) responded to the questionnaire.

More than two-thirds of this sample had reached the Doctoral level in their education (70%), 15% had attained a Masters degree, while the remainder of the sample were at the Bachelor level. When asked to describe themselves, participants in this group classified themselves most often as academics (65%), followed by practitioner-academics (25%) and practitioners (10%). In describing their employer, 10% indicated that they were either self employed or working for government forensic institutions, while the majority (80%) worked for universities or other research institutions. Overall, these experts reported an average of 9.55 ($\overline{\sigma}$ = 15.51) peer reviewed articles relating to eyewitness identification issues and an average of 25.65 ($\overline{\sigma}$ = 54.99) other publications in the same field. Furthermore, they reported having an average of 3.25 peer reviewed articles each ($\overline{\sigma}$ = 11.23) on the topic of eyewitness expert evidence and an average of 20.35 ($\overline{\sigma}$ = 53.30) other forms of publication on this matter. Between them these experts had published over 1176[13] articles on eyewitness identification and eyewitness experts.

Twenty-five percent of this sample had been asked to prepare a report for court relating to eyewitness identification issues, and 35% had been asked to testify on these issues. Within the five year period between January 2001 and December 2005, 40% of the sample had been asked to consult with legal representatives on a case involving eyewitness evidence ($\overline{x}$ = 5.35 times, $\overline{\sigma}$ = 11.67), 15% had prepared a report for court ($\overline{x}$ = 1.60 reports, $\overline{\sigma}$ = 4.52) and 20% had actually testified ($\overline{x}$ = 1.90 times, $\overline{\sigma}$ = 4.87). These experts reported being involved in cases in the United Kingdom, the United States of America and Australia. When asked whether or not they would be prepared to testify in relation to five issue, of those surveyed, the most (80%) indicated that they would testify to Statements 1 and 2, while only half of the sample showed a willingness to testify to statement 3 (see Table 10.2, p. 196 above).

Each participant was also asked to specify three areas of research interest. Overall, members of this group conducted research most often in decision-making (11.6%), followed by eyewitness identification, eyewitness memory and face perception (each 8.3%; see Table 10.3 below).

[13] Please note that this figure likely overestimates the number of separate articles published as participants may have been co-authors on some articles.

Table 10.3 : Reported Primary Areas of Research.

Research Topic	**Overall %**	**Research Topic**	**Overall %**
Decision-making	11.7	Memory	3.3
Eyewitness Identifications	8.3	Face Composites	3.3
Eyewitness Memory	8.3	Meta-cognition	1.7
Face Memory	8.3	Social Cognition	1.7
Child Witnesses	5	Interviewing	1.7
Face Perception	5	Perception	1.7
Face Recognition	5	Recognition Memory	1.7
Eyewitness Testimony	5	Attitudes	1.7
Forensic Interventions	3.3	None	23.3

Results Part II

Comparative Analysis

The following analysis includes data from all trials for all participants. Although there may be practice effects in this analysis, as it assumes an independence of judgements which can't be guaranteed given the repeated measures design, this approach offers the largest power possible. An analysis which includes data only from the first trial for all participants is reported in full in Appendix N. A discussion of the differences between these two analyses is presented in the *Discussion* at the end of this chapter.

Participants Qualitative Evaluations of Eyewitnesses

Evaluations of eyewitnesses on four personality and performance dimensions revealed no significant differences in the ratings of accurate and inaccurate eyewitnesses for credibility ($t_{(432)}$ = 0.86, p = .390), attractiveness ($t_{(431)}$ = *1.46*, p = .144) and trustworthiness ($t_{(431)}$ = 0.52, p = .604). Ratings of eyewitness confidence will be discussed below.

Participant Decision Criterion

A chi-squared goodness-of-fit test indicated that all participants believed significantly more than half of the eyewitnesses ($\chi^2_{(1)}$ = 12.31, p = <.0005). Members of the community group were significantly more likely to believe than disbelieve eyewitnesses (72.6% belief; $\chi^2_{(1)} = 32.11$, $p = <.0005$) while the undergraduates (49.4% belief; $\chi^2_{(1)} = 0.02$, $p = .879$), postgraduates (47.6% belief; $\chi^2_{(1)} = 0.10$, $p = .758$) and experts (55% belief; $\chi^2_{(1)} = 0.60$, $p = .439$) were as likely to believe the eyewitness as to disbelieve them. A (2x4) chi-squared revealed that this pattern of results represented a significant association between expertise level and proportion of belief judgements ($\chi^2_{(3)} = 21.12$, $p = <.0005$). This significant reduction in the belief of eyewitnesses with training is consistent with a Skepticism effect.

Effect of Expertise on Decisions

A two-way between groups ANOVA was conducted to explore the effect of expertise and eyewitness accuracy on participant estimates of eyewitness confidence. A significant main effect was observed for eyewitness accuracy ($F_{(1,433)} = 8.98$, $p < .005$, $\eta_p^2 = .021$), such that accurate eyewitnesses ($\bar{x} = 3.48$, $\bar{\sigma} = 0.10$) were rated as being significantly more confident than inaccurate eyewitnesses ($\bar{x} = 3.05$, $\bar{\sigma} = 0.11$). A significant main effect was also found for expertise ($F_{(3,433)} = 2.71$, $p < .05$, $\eta_p^2 = .019$), with posthoc comparisons revealing that those participants in the community group ($\bar{x} = 3.55$, $\bar{\sigma} = 0.10$) attributed significantly higher confidence to the eyewitnesses they saw than did participants at postgraduate or expert levels (postgraduate, $\bar{x} = 3.00$, $\bar{\sigma} = 0.20$; expert, $\bar{x} = 3.16$, $\bar{\sigma} = 0.16$). A significant interaction effect was also observed ($F_{(3,433)} = 2.92$, $p = <.05$, $\eta_p^2 = .020$). Inspection of Figure 10.1 (below) suggests that the difference between ratings of confidence attributed to eyewitnesses by experts do not follow the same pattern observed in other groups where accurate eyewitnesses are rated as significantly more confident than inaccurate eyewitnesses. In the expert condition this trend was reversed with inaccurate witnesses being rated as having higher confidence than accurate eyewitnesses.

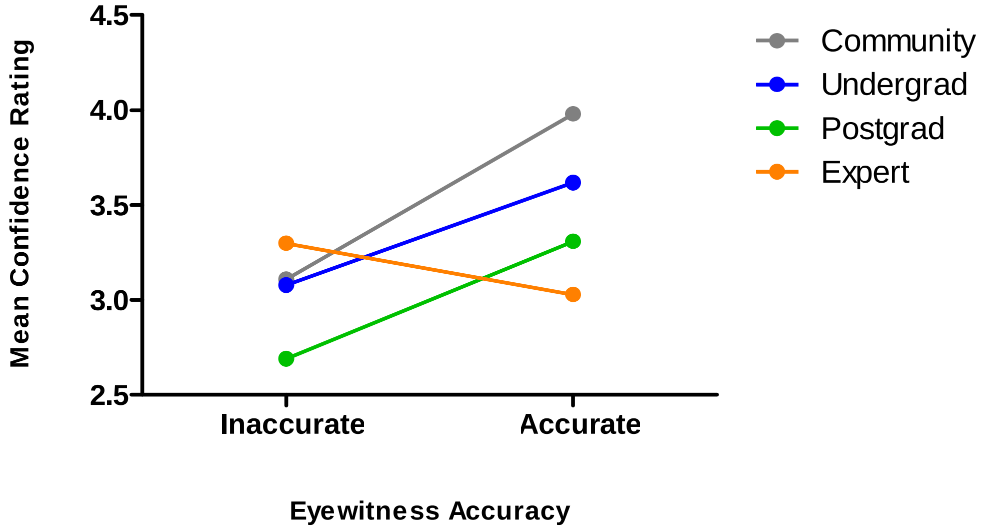

Figure 10.1 Mean confidence ratings by eyewitness accuracy.

A two-way between groups ANOVA was conducted to investigate the impact of expertise and belief on participant estimates of eyewitness confidence. Although level of expertise had no impact on participants perceptions of confidence ($F_{(3,432)}$ = 1.02, p =.386, η_p^2 = .007), a significant main effect for belief was observed such that those eyewitnesses who were believed ($\overline{x}$ = 3.75, $\overline{\sigma}$ = 1.14) were rated as being significantly more confident than those eyewitnesses disbelieved ($\overline{x}$ = 2.83, $\overline{\sigma}$ = 1.33, $F_{(1,432)}$ = 27.13, p < .0005, η_p^2 = .06). No significant interaction was observed ($F_{(3,432)}$ = 2.06, p =.105, η_p^2 = .014). See Figure 10.2 below.

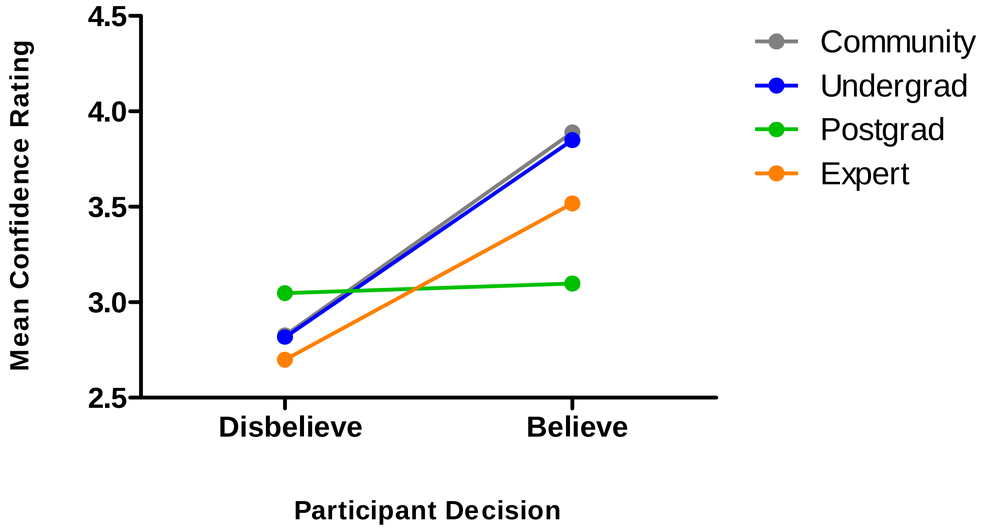

Figure 10.2 Mean confidence ratings by juror belief decision.

These results suggest that all but the expert participants rated accurate eyewitnesses as significantly more confident than inaccurate eyewitnesses. Furthermore, those eyewitnesses who were believed were rated as being significantly more confident than those eyewitnesses who were disbelieved.

Considerations Affecting Belief Decisions

A series of two-way ANOVAs were conducted, each investigating the effect of belief type and expertise level on participant-jurors' ratings of the extent to which of the four variables (eyewitness confidence, eyewitness manner, eyewitnessing conditions and identification conditions) influenced participant judgments. The latter of these analyses (identification conditions) did not include data from the undergraduate group as they had not been asked to rate the influence of this factor.

Community participants rated eyewitness confidence as significantly more influential in their decision-making ($\bar{x}$ = 2.74, $\bar{\sigma}$ = 0.83) than did the undergraduates ($\bar{x}$ = 2.42, $\bar{\sigma}$ = 1.00), postgraduates ($\bar{x}$ = 1.79, $\bar{\sigma}$ = 1.13) or experts ($\bar{x}$ = 1.83, $\bar{\sigma}$ = 1.12, $F_{(3,431)}$ = 15.5, p < .0005, η_p^2 = .099). There was no significant main effect of belief type ($F_{(1,431)}$ = 2.49, p = .116) or interaction effect ($F_{(3,431)}$ = 0.20, p = .894). Furthermore, post-hoc analysis also revealed that undergraduate participants also relied on confidence significantly more than either the postgraduate or expert participants (see Figure 10.3 below).

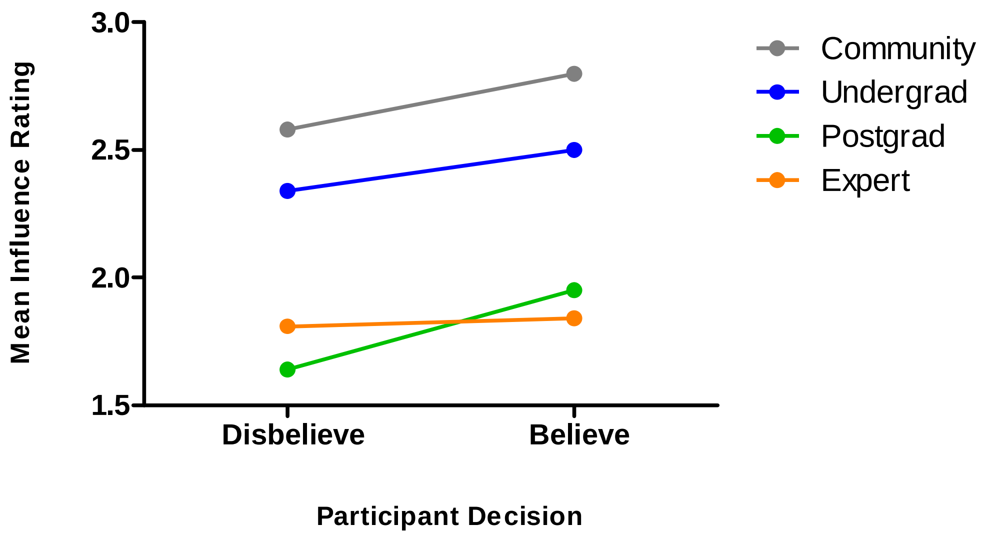

Figure 10.3 Ratings of the influence of eyewitness confidence by belief decision.

Analysis revealed that reliance on eyewitness manner also differed significantly with expertise levels ($F_{(3,430)}$ = 6.10, p = <.0005, η_p^2 = .041). Post-hoc analyses indicated that experts ($\bar{x}$ = 1.90, $\bar{\sigma}$ = 0.93) reported being influenced by eyewitness manner significantly less than the community ($\bar{x}$ = 2.54, $\bar{\sigma}$ = 0.91) and undergraduate participants ($\bar{x}$ = 2.43, $\bar{\sigma}$ = 01.03). No interaction ($F_{(3,430)}$ = 0.50, p = .680) or main effect for belief type ($F_{(1,430)}$ = 1.12, p = .294) was identified (see Figure 10.4 below).

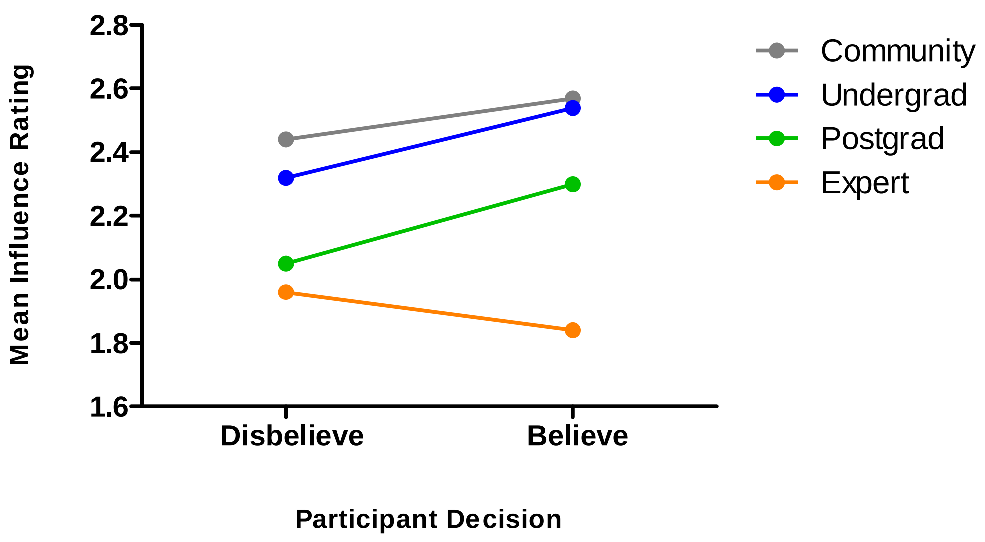

Figure 10.4 Ratings of the influence of eyewitness manner by belief decision.

Witnessing conditions were differentially influential both across levels of expertise ($F_{(3,432)}$ = 6.58, p = <.0005, η_p^2 = .044) and belief type ($F_{(1,432)}$ = 15.48, p = <.0005, η_p^2 = .035). Not only was witnessing condition rated to be more influential when participants decided to believe than disbelieve the eyewitness ($\bar{x}$ = 2.61, $\bar{\sigma}$ = 0.90 cf. disbelieve, $\bar{x}$ = 2.22, $\bar{\sigma}$ = 1.09), but post-hoc analyses demonstrated that those participants in the postgraduate group reported being significantly more affected by this variable ($\bar{x}$ = 2.90, $\bar{\sigma}$ = 0.85) than either community ($\bar{x}$ = 2.36, $\bar{\sigma}$ = 0.93) or undergraduate group members ($\bar{x}$ = 2.34, $\bar{\sigma}$ = 1.07). No significant interaction effect was identified ($F_{(3,432)}$ = 0.079, p = .971). See Figure 10.5 below.

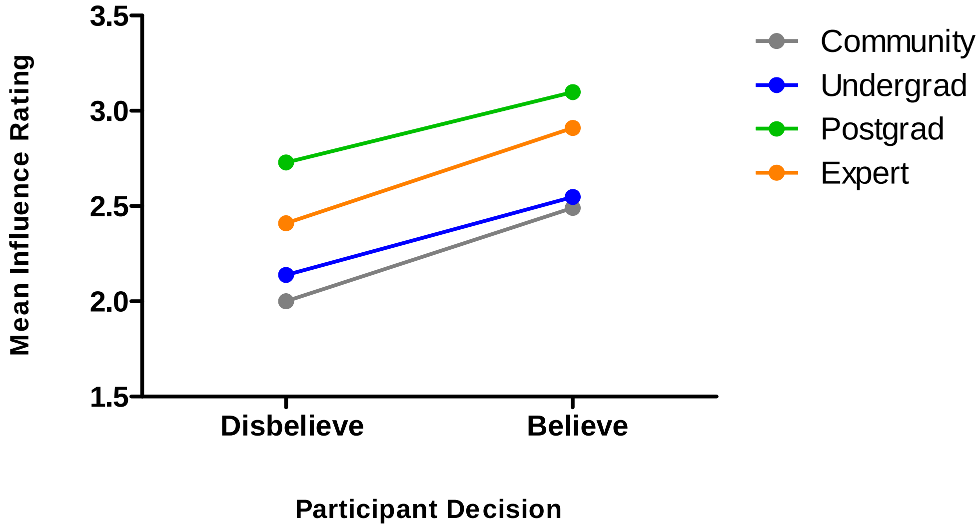

Figure 10.5 Ratings of the influence of witnessing conditions by belief decision.

The participants' ratings of the influence of the final factor, identification conditions, was equivalent irrespective of expertise level ($F_{(2,252)}$ = 1.33, p = .265), however identification conditions were estimated as significantly more influential when participants chose to believe ($\bar{x}$ = 2.52, $\bar{\sigma}$ = 0.87) rather than disbelieve the eyewitness they viewed ($\bar{x}$ = 2.19, $\bar{\sigma}$ = 1.01, $F_{(2,252)}$ = 10.03, p = <.005). There was no significant interaction effect ($F_{(2,252)}$ = 0.75, p = .472).

In summary, it was found that community and undergraduate participants relied more heavily on eyewitness confidence and manner than did experts, and that postgraduate participants relied more heavily on witnessing conditions than either community or undergraduate participants. It was also apparent that witnessing and identification conditions were more influential in belief decisions than in disbelief decisions.

Predictors of Belief Decisions

A binary logistic regression was conducted in order to investigate whether participants' ratings of the eyewitness (on the dimensions of credibility, confidence, attractiveness and trustworthiness) predicted their decision to believe the eyewitness they viewed. For community participants the overall model was significant ($\chi^2_{(4)}$ = 36.19, p < .0005), correctly classifying 80.8% of participants' decisions. Participants' estimates of eyewitness credibility contributed significantly to the model (see Table 10.4 below for all variables in the equation), with a one unit increase in perceptions of credibility resulting in decrease in the odds of disbelieving the eyewitness by approximately a factor of 0.45 times (β = -0.80, p < .0005)[14].

Table 10.4 : β (S.E.), Significance and Exp(β) for Predictor Variables in Community Group

	β (S.E.)	***p***	**Exp(β) (95% CI:LL-UL)**
Credibility	-0.80 (0.27)	.003	0.45 (0.27-0.77)
Confidence	-0.19 (0.20)	.354	0.83 (0.56-1.23)
Attractiveness	-0.04 (0.21)	.866	0.97 (0.64-1.45)
Trustworthiness	-0.08 (0.25)	.741	0.92 (0.56-1.51)

For undergraduate participants the model was found to be significant ($\chi^2_{(4)}$ = 68.08, p = <.0005), accurately classifying participant decisions 74.7% of the time (see Table 10.5 below for all variables in the equation). Here the participants' estimates of the eyewitness's credibility and confidence were significant predictors of belief, with an increase of one scale point in the participants' rating of the eyewitness's confidence decreasing the odds that a participant would disbelieve an eyewitness by a factor of 0.60 (β = -0.51, p < .005), and an increase of one scale point in ratings of credibility decreasing the odds that the eyewitness would be disbelieved by a factor of 0.31 (β = -1.18, p < .0005)

[14] The reader may note that negative β values reflect and increase in the likelihood of belief in this study, where in previous experiments negative β values reflected a decrease in the likelihood of belief. This change is an artifact of differences in data-coding between this study and previous experiments (i.e., where belief had previously been coded as 1, in this study it was coded as 0).

Table 10.5 : β (S.E.), Significance and Exp(β) for Predictor Variables in Undergraduate Group

	β (S.E.)	***p***	**Exp(β) (95% CI:LL-UL)**
Credibility	-1.18 (0.27)	.000	0.31 (0.18-0.53)
Confidence	-0.51 (0.17)	.002	0.60 (0.43-0.83)
Attractiveness	-0.35 (0.18)	.064	0.72 (0.50-1.02)
Trustworthiness	0.09 (0.24)	.690	1.10 (0.69-1.74)

A significant model was also identified for the group of participants with postgraduate training ($\chi^2_{(4)}$ = 17.60, p = <.005), accurately classifying participant decisions 76.2% of the time (see Table 10.6 below for all variables in the equation). As in the community sample, credibility alone was found to be a significant predictor of belief judgments such that an increase of one scale point in the participants' rating of the eyewitness's credibility decreased the odds that a participant would disbelieve an eyewitness by a factor of 0.21 (β = -1.55, p < .01).

Table 10.6 : β (S.E.), Significance and Exp(β) for Predictor Variables in Postgraduate Group

	β (S.E.)	***p***	**Exp(β) (95% CI:LL-UL)**
Credibility	-1.55 (0.59)	.009	0.21 (0.07-0.68)
Confidence	0.43 (0.40)	.278	1.54 (0.705-3.37)
Attractiveness	0.40 (0.49)	.408	1.50 (0.58-3.90)
Trustworthiness	-0.18 (0.43)	.671	0.83 (0.36-1.95)

The model was also significant for the expert participants ($\chi^2_{(4)}$ = 12.23, p < .05) correctly classifying 76.3% of judgments. A one unit increase in participants' ratings of eyewitness confidence resulted in a decrease in the odds that an expert would disbelieve an eyewitness by a factor of 0.52 (β = -0.65, $p < .05$). See Table 10.7 below.

Table 10.7 : β (S.E.), Significance and Exp(β) for Predictor Variables in Expert Group

	β (S.E.)	***p***	**Exp(β) (95% CI)**
Credibility	-0.34 (0.40)	.395	0.72 (0.33-1.55)
Confidence	-0.65 (0.31)	.032	0.52 (0.29-0.95)
Attractiveness	-0.02 (0.25)	.923	0.98 (0.60-1.58)
Trustworthiness	-0.37 (0.34)	.282	0.69 (0.35-1.35)

Participant Sensitivity to Eyewitness Accuracy

Overall, participants accurately classified eyewitnesses 53.3% of the time, a chi-squared goodness-of-fit test indicated that this was not significantly different from the 50% accuracy expected by chance alone ($\chi^2_{(1)} = 1.93$, $p = .164$). Participants in the community group were correct 58.5% of the time – this was significantly better than chance ($\chi^2_{(1)} = 4.59$, $p = <.05$). This was not true for participants with higher levels of expertise (undergraduate, 47.7% accuracy, $\chi^2_{(1)}$ 0.37, $p = .544$; postgraduate, 57.1% accuracy, $\chi^2_{(1)} = 0.86$, $p = .355$; expert, 53.3% accuracy, $\chi^2_{(1)} = 0.27$, $p = .606$). Even so, analysis indicated that there was no significant association between expertise and accuracy ($\chi^2_{(3)} = 4.16$, $p = .245$).

A summary of the decisions made by the participants', their decision criteria, and performance accuracy is reported below (see Table 10.8). Overall, this summary suggests that when Skepticism was low, accuracy was high as in Trial 1, and that increased Skepticism, as in Trial 3, was not associated with increased accuracy. This illustrates how the accuracy rates in the sample of eyewitnesses viewed can influence the appropriateness of a relatively more or less Skeptical decision criteria. Indeed, as can be seen in Trial 2, low Skepticism, was not associated with significant accuracy as in Trial 1.

Table 10.8 : Summary of Belief and Accuracy Rates for Each Judgement

	% Belief	$\chi^2_{(1)}$	***p***	**% Accuracy**	$\chi^2_{(1)}$	***p***
1st Witness Evaluated	64.4% (overbelief)	21.78	.000	59.3% (sig. acc)	5.03	.025
2nd Witness Evaluated	63.9% (overbelief)	11.11	.001	43.4% (no sig. acc)	2.40	.115
3rd Witness Evaluated	42.1% (no bias)	3.65	.056	57.2% (no sig. acc)	3.04	.081

Discrimination Expertise within Expertise Levels

The performance of each participant was calculated by adding the number of correct identifications made (i.e., an accuracy score out of three). The distribution of the scores was then compared between participant groups (see Figure 10.6 below). This analysis was conducted in order to clarify whether or not there was evidence of more "expert" participants (those scoring three out of three for accuracy) in the expert and postgraduate groups, and more 'novice' participants (those scoring zero out of three for accuracy) in the community and undergraduate groups. Inspection of the Figure 10.6 suggests that this is not the case. Chi-squared analyses confirmed that there was no association between expertise level and accuracy score ($\chi^2_{(9)} = 12.38$, $p = .193$).

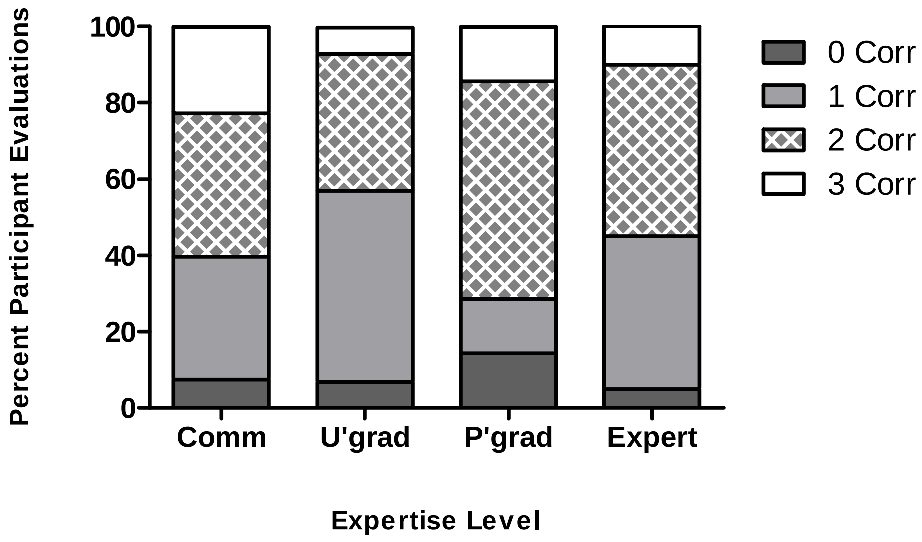

Figure 10.6 Percent of participant accuracy scores across levels of expertise.

Discussion

This study represents the first attempt to evaluate how increasing amounts of expertise in the area of eyewitness identification and human memory impact upon the accuracy with which eyewitness discrimination tasks are completed. The data gathered here offer little encouragement that jurors can be successfully trained to complete this task with greater accuracy than they otherwise would.

Expertise Manipulation

Based on the descriptive analysis of the participants recruited for this study, and subsequent analysis of their decision strategies, it does appear that there was a successful sampling of groups of people with qualitatively and quantitatively different levels of expertise relevant to eyewitness identification performance. Of the community members sampled, none had heard a psychologist speak about either human memory or eyewitness memory. Moreover, members of this group showed they were least willing to accept that confidence was not a predictor of eyewitness identification accuracy, and like other samples of novice participants (Bradfield & Wells, 2000; Brewer & Wells, 2006; Cutler & Penrod, 1988; Cutler et al., 1989b; Leippe et al., 2004; Lindsay, Wells & O'Connor, 1989; Wells et al., 1998) showed a significant tendency to rate eyewitness confidence as an influential factor in their decision-making. They also showed a significant tendency to believe the eyewitnesses seen (Cutler & Penrod, 1988; Lindsay et al., 1981; Loftus, 1979; Loftus & Monahan, 1980; Maass et al., 1985; Penrod & Cutler 1999; Wells et al., 1980). Thus, the community members' responses were consistent with what we know of novice jurors.

Members of the undergraduate sample had received several hours of education relating to eyewitness identification issues through tutorials and lectures in their psychology and law course, and were observed to exhibit greater Skepticism in their belief decisions than community members. Yet, these participants showed themselves to be similar to the novice sample in some ways, rating eyewitness manner, and eyewitnessing conditions to be as influential in their decision-making as did community participants.

The postgraduates in this sample performed differently from both the community and undergraduate groups; showing significant Skepticism in their belief decisions, rating witnessing conditions as significantly more influential than those participants who were less experienced, and rating the role of confidence and eyewitness manner similar to the expert group. In addition, the responses of postgraduate participants provided a rank order of endorsement for eyewitness identification issues which was identical to that derived from expert responses. Despite this evidence, which is consistent with increasing expertise, these participants also showed themselves to be less experienced than the Expert sample, with an absence of peer reviewed publications on the topic of eyewitness identification or eyewitness experts, and an average education at only the Bachelors level.

The members of the expert group distinguished themselves as such through their court appearances, their numerous publications in relevant domains and their high level of education. Moreover, these participants showed themselves to be relatively more Skeptical in their belief decisions; however, they were still prone to use eyewitness confidence to estimate accuracy, despite the fact that 75% of these participants indicated that they would testify that eyewitness confidence is not a good predictor of identification accuracy. It is difficult to resolve this apparent inconsistency, except to say that these participants *reported* being influenced by confidence in a way which was consistent with their professional opinions (i.e., they reported relying on confidence significantly less than community and undergraduate participants). This suggests that expert reliance on eyewitness confidence was unintentional. Even so, it is difficult to argue that this unexpected result detracts from the expertise of these individuals. Thus, overall, it is with some confidence that we can attribute differences between groups in this study (or lack there of), to expertise, as each group differed from those above and below it in meaningful ways.

Effects of Expertise Level

Focusing first on community members, these participants perceived accurate eyewitnesses as being more confident than inaccurate eyewitnesses and rated believed eyewitnesses as more confident than those who were disbelieved. Indeed, participants in this group rated all eyewitnesses as being significantly more confident than did postgraduate and expert participants. Moreover, community members displayed a significant preference for believing the eyewitnesses they viewed; this was not seen amongst other participant groups. When asked, community participants reported that they relied on eyewitness confidence in their decision-making to a greater extent than any of their more experienced counterparts. They also relied on eyewitness manner significantly more than members of the expert group, and witnessing conditions significantly less than postgraduates. Finally, the belief decisions made by community members were significantly predicted by their perceptions of eyewitness credibility, and overall, participants in this group performed with accuracy at levels significantly higher than could be expected by chance alone, with 58.5% accuracy. Thus, although participants in this group were sensitive the confidence-accuracy correlation in the sample, their reliance on credibility, a non-significant predictor (p. 199 - *Participants Qualitative Evaluations of Eyewitnesses*), may have somewhat impaired their discrimination performance.

Overall, undergraduate participants rated accurate eyewitnesses as significantly more confident than inaccurate eyewitnesses, and rated the eyewitnesses they believed as significantly more confident than those who they disbelieved. Thus it seems they were sensitive to the underlying confidence-accuracy correlation. Undergraduate participants were, however, no more likely to believe than to disbelieve an eyewitness, suggesting that their increased experience served to reduce levels of belief compared with the community sample. The undergraduate participants also reported being significantly less influenced by eyewitness expressions of confidence than community members. They were also influenced by eyewitness manner significantly more than experts, and influenced by witnessing conditions significantly less than postgraduate participants. The belief decisions of undergraduates were significantly predicted by perceived eyewitness credibility and confidence, however, discrimination accuracy in this condition was at chance levels. Thus, although participants were using confidence information when formulating their decisions, the regression model suggests that the magnitude of the influence of credibility (a non-significant predictor of accuracy for this group of eyewitnesses) was at least three times greater. This reliance on an unhelpful predictor may have obscured any positive gains achieved by utilising eyewitness confidence information.

Across all trials, postgraduate psychology students also perceived and responded to the underlying confidence-accuracy correlation, rating believed eyewitnesses as significantly more confident than those eyewitnesses who were disbelieved. At this level of expertise, participants were as likely to believe an eyewitness as to disbelieve them, and reported relying on witnessing conditions significantly more than their less experienced counterparts. The decisions made by postgraduates could be significantly predicted by their estimates of eyewitness credibility, and overall, these participants performed at chance levels with 57.1% accuracy.

Finally, expert participants were found to use confidence information in a manner unlike their less experienced counterparts. Members of this group misperceived the underlying confidence-accuracy correlation, and although they rated believed eyewitnesses as significantly more confident than disbelieved eyewitnesses, they rated *inaccurate* eyewitnesses as significantly more confident than accurate eyewitnesses. Thus, these participants were mistaking the accuracy “signal”, for the inaccurate “noise”. Even so, the Skepticism effect observed among the undergraduates and postgraduates persisted amongst the experts, with belief rates significantly lower than those observed in the community sample. In addition, experts were less influenced by eyewitness confidence than community members, and less influenced by eyewitness manner than either the community members or undergraduate students. Surprisingly, however, expert decisions were significantly predicted by ratings of eyewitness confidence, such that increases in perceived confidence resulted in a significant increase in the odds of belief. Given that participants in this group were systematically mistaking accurate eyewitnesses for inaccurate eyewitness, it is hardly surprising to find their discrimination accuracy falling at chance levels. It is somewhat puzzling, however, that their performance was not significantly *worse* than chance.

Expertise Level and Sensitivity to Eyewitness Accuracy

Bearing in mind that there was no significant association between expertise level and discrimination accuracy, the knowledge that a group performed significantly better than chance, while others did not, is limited in value. The only conclusion that can be made with any confidence is that expertise had *no reliable effect on discrimination accuracy*, despite the fact that members of some groups were making systematic errors while others were not. Thus, the decisions strategy adopted, the cues relied on, and the amount of expertise available to an eyewitness assessor, appear to have had a very minimal impact upon the ability to discriminate between accurate and inaccurate eyewitnesses. This is most clearly illustrated by the finding that there was no significant association between individual accuracy scores (out of three) and level of expertise. Moreover, when the data for all trials is treated independently ($n = 435$ observations) this experiment had a 99% chance of detecting an effect that was moderate in magnitude ($w = .03$), supporting the interpretation that expertise had only a small effect on discrimination performance, if any at all.

Practice/ Order Effects

As noted earlier, it could be argued that it is inappropriate to analyse the data in the manner reported because this assumes the observations are all independent from each other. This is not the case from the three observations made by each participant. A more valid, but less powerful, analysis involves only the first decision made by each participant. The results of such an analysis are reported in Appendix N.

Overall, these two analyses produced very similar results, and most differences (see Table 10.9 below, p. 216) are likely to be a reflection of the differences in statistical power rather than practice or order effects. Moreover, even in instances where this does not appear to be the case (see the postgraduate and expert groups), discrimination performance did not appear to differ across analyses, thus the changes observed do not appear to substantially affect participant SEA.

Finally, the “First Trial Only” analysis revealed that participants were able to discriminate between accurate and inaccurate eyewitnesses at levels significantly higher than chance, however, the “All Trials” analysis indicated that participant-jurors were as likely to believe an accurate eyewitness as to believe an inaccurate eyewitness. This result suggests that participants did not actually become better discriminators over trials; rather their sensitivity to eyewitness accuracy appears to have been blunted somewhat by repeated trials. Even so, it is important for the reader to be aware that it is unclear if this finding can be attributed to a practice effect, or simply an order effect, as participants in the community, postgraduate and expert groups all viewed an eyewitness from Protocol II, followed by Protocols I and III.

Limitations

The conclusion that expertise in eyewitness identification issues does not significantly improve eyewitness discrimination accuracy may be limited by various factors. Firstly, it could be argued that the experts in this study were unrepresentative of all eyewitness experts indicated by: a) their small number of court appearances; b) the small number in the sample; and c) the method of recruitment. While it may be valid to suggest that experts, in general, are likely to appear in court more often than experts in this sample, and to propose that the method of selection may have unintentionally contributed to a selection bias, it is very difficult to suggest that these expert participants do not have sufficient expertise to be classed as experts. Specifically, participants in this sample have undertaken research programs in relevant areas and have published extensively in peer reviewed journals on the topics of eyewitness identification and eyewitness expert evidence. Thus it is difficult to suggest that these respondents are lacking in the knowledge necessary to be considered experts and to have a level of expertise over and above lay jurors; this is the key distinction necessary to test the efficacy of expert evidence or education.

It could also be argued that these eyewitness experts were unable to utilise their expertise in *this* experiment because the sample of eyewitnesses used were atypical in some significant way, making the experts evaluations unfairly difficult. Although this is a matter which will be discussed in some detail in the *General Discussion* to follow, it is important for the reader to note that the eyewitnesses in this sample: a) were real and therefore have the characteristics of real accurate or inaccurate eyewitnesses; and b) behaved consistent with the available literature with respect to the confidence-accuracy correlation. Thus it cannot be suggested that eyewitnesses like these would never appear in court and therefore never have to be evaluated by jurors. Accordingly, it is of interest to evaluate the aid expertise can provide in the discrimination of such eyewitnesses.

Not withstanding this defence, this research could be improved upon in the future by providing participants with a larger number of eyewitnesses to evaluate from a broader range of crime scenarios. This would maximize the likelihood of gathering testimony from eyewitnesses whose identification accuracy has been significantly influenced by varied combinations of different eyewitnessing and identification factors, and would therefore provide an even more realistic challenge for expert evaluators; giving them an opportunity to utilize their knowledge with regard to numerous predictors of accuracy rather than one as in this case.

Conclusions

Analysis of the “First Trial Only” and subsequent analysis of “All Trials” provided no evidence of a significant association between a participants’ level of expertise and their ability to differentiate between accurate and inaccurate eyewitnesses. Indeed, there was no significant association between expertise level and the number of “expert” or “novice” performers in each group; as perfect performances were as likely to be seen in the community sample as the expert sample. This result suggests that there is nothing inherent in the expertise of the experts which assists them to discriminate between accurate and inaccurate eyewitnesses. In addition, there appears to be no prospect of educating novice jurors to improve their discrimination accuracy through expert testimony, as their performance was on par with that of the experts themselves. This finding will be discussed in more detail in the *General Discussion* to follow.

Table 10.9 : Comparison Between “First Trial Only” and “All Trial” Analyses

	Analysis	
	“First Trial Only”	**“All Trials”**
Community	• EW’s not rated as sig. > confident in this group *cf.* other groups. • Belief decisions not sig. predicted by credibility ratings.	• EW’s rated as sig. > confident in this group *cf.* other groups. • Belief decisions sig. predicted by credibility ratings.
Undergraduate	• EW’s rated as sig. > confident in this group *cf.* other groups. • Sig. overbelief. • Belief decisions sig. predicted by credibility and trustworthiness ratings	• EW’s not rated as sig. > confident in this group *cf.* postgraduate and expert groups. • No sig. belief bias. • Belief decisions sig. predicted by credibility and confidence ratings.
Postgraduate	• EW confidence treated as a negative predictor of identification accuracy.	• EW confidence treated as a positive predictor of identification accuracy.
Expert	• Accurate EW’s rated as sig. > confident. • No sig. predictors of belief decisions.	• Accurate EW’s rated as sig. < confident. • Belief decisions sig. predicted by confidence ratings.

SECTION 4: DISCUSSION

Chapter 11

General Discussion

Summary

This thesis reports a thorough investigation of the efficacy of eyewitness expert evidence, as well as its actual and perceived utility when compared with pattern judicial instructions. Overall, the six different studies reported here analyse data gathered from more than 1000 participants from various sub-populations including community members (n = 53); undergraduate (n = 961) and postgraduate psychology students (n = 14); legal professionals (n = 35); and eyewitness experts (n = 20). This thesis also pays particular attention to the logic of inferences which have been made by psychologists regarding the efficacy of expert evidence, and compares methodologies adopting direct measures of participant Sensitivity to Eyewitness Accuracy with those that can only indirectly assess this construct.

The first study reported in this thesis provided an overview of the knowledge and opinions of legal professionals regarding eyewitness identification issues. It showed a high degree of correspondence between the opinions expressed by research psychologists and those of surveyed legal professionals. Moreover, both groups of professionals expressed doubts that judicial instructions relating to eyewitness identification evidence would exert an effect equivalent to that of eyewitness expert evidence. Given this shared perception of inequality (discussed in Chapter 5), and the dearth of evidence supporting it (as only one study has previously compared a judicial instruction with eyewitness expert evidence Cutler et al., 1990a), Experiments 1 to 4 were conducted to directly assess the relative impacts of judicial instruction and expert evidence on participant-juror Sensitivity to Eyewitness Accuracy.

The methodology utilised in these studies incorporated the testimony of real eyewitnesses to a staged crime scenario in order to assess the impact of instruction on juror ability to discriminate between known accurate and known inaccurate eyewitnesses. This methodology has been reported only twice previously in the

literature (Wells & Wright, 1983 cited in Wells, 1986; Wells et al., 1980) and in at least one of those cases (Wells et al., 1980) participant-jurors were required to evaluate the testimony of eyewitnesses who would likely never appear before a jury (i.e., those who identified a known innocent foil from a target-present array). Specifically, Experiment 1 required participant-jurors and juries to evaluate the testimony of one of four different eyewitnesses who had identified either the perpetrator or the police suspect after watching the video of a burglary. Confidence and accuracy were perceived by participants to have been associated in this study. Accordingly the efficacy of the eyewitness expert’s testimony was assessed with regard to its ability to: a) moderate juror use of confidence information in decision-making; b) educate jurors regarding the confidence-accuracy correlation; and c) improve juror ability to discriminate between accurate and inaccurate identification evidence. In this case the provision of expert evidence stating that confidence was not a useful predictor of accuracy resulted in significantly poorer discrimination performance, with participants incorrectly using confidence as a negative predictor of accuracy, rather than ignoring the information as they had been directed to.

Experiments 2 to 4 attempted to improve upon Experiment 1 by: a) providing more eyewitnesses for evaluation; b) by varying witnessing and identification conditions; and c) by manipulating the objective quality of the experts’ testimony with regard to these witnessing and identification conditions. Consequently, the testimony of 13 new eyewitnesses, who had viewed one of two versions of a bag-snatch scenario, was presented to participant-jurors. Importantly, although attempts were made to significantly vary the witnessing conditions viewed by eyewitnesses in Experiments 2 through 4, no significant association with accuracy was identified as a result of manipulations to the lighting, angle of view, presence of disguise or distance of view in the crime scenarios (i.e., good, poor or very poor). Indeed, accurate and inaccurate eyewitnesses were only found to differ *reliably* from each other with respect to their expressions of confidence. Consequently, the efficacy of the eyewitness expert’s evidence regarding the confidence-accuracy correlation was investigated and compared with pattern judicial instruction in these studies. Overall, little evidence was found to support perceptions of the superiority of expert evidence, as no significant association was identified between instruction type and Sensitivity to Eyewitness Accuracy. This was irrespective of the objective quality of the expert testimony.

Overall, substantial evidence was collected in Experiments 2 to 4 which suggested that, although participant-jurors were both receptive to the expert and compliant with their advice, Sensitivity to Eyewitness Accuracy was unaffected. As a result Experiment 5 was designed to investigate why the experts were not able to improve the discrimination accuracy of the jurors. Specifically, it was reasoned that if it could be demonstrated that experts themselves could complete eyewitness discrimination tasks with greater accuracy than untrained laypeople, this would implicate either jurors or the experimental materials (i.e., the expert testimony) as the likely cause of the failures observed. If, however, it could not be demonstrated that expertise improved eyewitness discrimination, the validity of the expertise itself would warrant closer consideration. Thus, Experiment 5 focused on the extent to which participants of varying levels of expertise could correctly classify the testimony of three eyewitnesses from three different crime scenarios. This study utilised the testimony of 17 eyewitnesses from Experiments 1 to 4 plus the testimony of three additional eyewitnesses who had viewed either a new burglary (n = 2) or whose testimony had been collected, but not used in Experiments 2 to 4 (n = 1). The results of this investigation suggest that eyewitness experts were no better able to discriminate between accurate and inaccurate eyewitnesses than novice laypeople. This raises questions as to the utility of eyewitness expertise in the completion of eyewitness discrimination tasks.

The results of these five experimental investigations will now be discussed in detail with regard to the specific aims of this thesis.

Aims

Aim 1: To Provide a Fair Test of Judicial Instruction

Although judicial instructions have been criticised for their inability to generate Sensitivity (Cutler et al., 1990a; Leippe et al., 2004; Ramirez et al., 1996), analysis of this criticism shows it to be somewhat unfair. Specifically, it is not possible to induce Sensitivity to the quality of an identification, unless one is provided with information that specifies what feature, or features, are generally associated with better or worse quality identifications. Thus, pattern judicial instructions such as the *Telfaire* instruction and the direction recommended by the NSW Judicial Commission will not be able induce this type of Sensitivity as they do not incorporate the necessary directional predictors (Greene, 1988). As a consequence, the only valid way to assess the impact of judicial instruction is to measure its effects on participant-juror Sensitivity to Eyewitness Accuracy (as assessed here for the first time) using real eyewitness designs. Indeed, analyses of this kind provide a far richer impression of the effects of judicial instruction than previous analyses, and also offer the first valid data upon which to debate the utility of the judicial directions provided to jurors.

Response to Judicial Instruction

Analysis of participant-juror responses to judicial instruction have been investigated in 11 studies to date including the first four experiments reported in this thesis (Experiments 1a[15], 1b, 2, 3 and 4; Cutler et al, 1990a; Greene, 1988; Hoffheimer, 1989; Katzev & Wishart, 1985; Ramirez et al, 1996). Overall, just two of these 11 studies provide evidence to suggest that judicial instruction is sufficient to alter the decision criterion adopted by participants, inducing significant levels of Skepticism (Greene, 1988; Katzev & Wishart, 1985). Moreover, both of these studies utilised a revised or non-standard judicial instruction, complete with directional predictions. The remaining nine of these 11 studies have found no evidence to suggest that the presence of a judicial instruction causes participant-jurors to significantly alter their pre-existing criterion of belief, and therefore have found no evidence to suggest that these instructions induce Skepticism amongst jurors.

Sensitivity to Eyewitness Accuracy

Investigations evaluating the effect of pattern judicial instructions on participant Sensitivity to Eyewitness Accuracy provide a somewhat more mixed impression. Although no significant association between instruction type and SEA was identified for Experiments 1 to 4 using chi-squared tests-of-independence, goodness-of-fit analyses suggest that judicial instruction may have improved participant SEA, relative to chance. This was evident in Experiment 1b (by increasing accurate verdicts from 33.3% in the control condition to 100% in the judicial condition) and Experiment 2 (by increasing discrimination accuracy from chance levels in the control condition to 64.4% in the judicial condition). However, Experiments 3 and 4 also produced

[15] Please note that the term "Experiment 1a" refers to the pre-deliberation juror analysis, while "Experiment 1b" refers to the post-deliberation jury analysis.

evidence to suggest that judicial instruction may have impaired SEA when compared with controls, with accuracy rates dropping to chance levels. Lastly, Experiment 1a produced no evidence for an effect of judicial instruction on discrimination accuracy, with SEA falling at chance levels in both the judicial and control conditions. On average, across studies (1a, 2, 3 and 4), discrimination accuracy in the judicial condition was 58.6% (range: 50% to 64.4%) compared with 58.2% in the control condition (range: 39.4% to 71.2%).

Overall then, this detailed analysis found little compelling evidence that judicial instruction is likely to result in a more stringent belief criterion (i.e., Skepticism), but provided mixed evidence that judicial instruction can improve SEA (see Table 11.1 below).

Table 11.1 : Summary of the Effects of Judicial Instruction

Experiment	Control Response to Manipulated Variables	Judicial Effects Observed	
		RJI	SEA
Katzev & Wishart, 1985		< Pre-deliberation Guilty Verdicts from Commentary than Instruction Only (Sig.) < Deliberation in Commentary than Instruction & Summation (Sig.)	
Greene, 1988 (Experiment 2)	Sig. Sensitivity to WIC	Sig. Skepticism of Pre-Deliberation Guilt from Revised Skepticism of Post-deliberation Guilt from Revised (Sig?)	
Hoffheimer, 1989	None	None (*Telfaire*)	
Ramirez, Zemba, Geiselman, 1996 (Experiment 1)	Sig. Sensitivity to WIC	None	
Ramirez, Zemba, Geiselman, 1996 (Experiment 2)	None	None	
Cutler, Dexter & Penrod, 1990	> Confident EW's > Culpability, > General Accuracy of ID's, > Strength Pros. Case & > Credibility	None	

Experiment	Control Response to Manipulated Variables	Judicial Effects Observed	
		RJI	SEA
Experiment 1a (Pre-deliberation)	Sig. Overbelief of EW No Sensitivity to Accuracy (χ^2 GOF)	No Change in Bias	No Change in Sensitivity (χ^2 GOF)
Experiment 1b (Post-deliberation)	No Bias 33.3% Accuracy (untested)	No Change in Bias	100% Accuracy (untested)
Experiment 2	No Bias No Sensitivity to Accuracy (χ^2 GOF)	No Change in Bias	Sig. Sensitivity (χ^2 GOF)
Experiment 3	No Bias Sig. Sensitivity to Accuracy (χ^2 GOF)	No Change in Bias	No Sensitivity (χ^2 GOF)
Experiment 4	No Bias Sig. Sensitivity to Accuracy (χ^2 GOF)	No Change in Bias	No Sensitivity (χ^2 GOF)

Aim 2: To Make a Valid Comparison of Judicial Instruction and Expert Evidence

Given that all previous analyses of the effects of judicial instruction have either: a) used indirect outcome measures (e.g. from SEO methodologies); or b) altered the instruction in non-trivial ways (i.e., by introducing directional predictions), and only one direct comparison has ever been made between judicial instruction and eyewitness expert effects, perceptions of the superiority of expert evidence may warrant revision. Indeed, evidence gathered in this thesis suggests that the differences between the effects of judicial instruction and expert evidence may be so minor as to be inconsequential.

Moderating Effects of Expert Evidence *cf.* Judicial Instruction

As discussed above, the evidence presented here suggests that judicial instruction does little to alter the pre-existing belief criterion used by participants, with nine of the 11 studies finding no evidence for a reliable change. Similarly, only three of the six studies which directly compare expert evidence and judicial instruction, produced evidence consistent with a Skepticism effect (Experiment 1a & b[16]; Cutler et al., 1990a), while the remaining three experiments suggested that the testimony of the expert caused no marked change. Thus, while overall Skepticism does appear more likely to result from instances where directional predictions are provided to participants (Cutler et al., 1990a; Greene, 1988; Katzev & Wishart, 1985), Experiments 1 to 4 offer no evidence for a clear difference in the decision criterion adopted as a result of hearing pattern judicial instruction or eyewitness expert evidence.

Discriminating Effects of Expert Evidence *cf.* Judicial Instruction

Turning now to the effects of expert evidence on participant Sensitivity to Eyewitness Accuracy, the data from studies incorporating both expert and judicial conditions yields only five interpretable expert effects[17]. Of these five results, two provide evidence suggesting that the expert testimony impaired participant discrimination

---

[16] In the latter case this inference is based upon an inspection of conviction rates (conviction dropped from 50% to 0% while acquittals increased from 50% to 71.4%) as the jury-level analysis did not have sufficient power to permit valid statistical tests.

[17] The results from Experiment 1b are not considered interpretable given that no statistical analysis could be conducted on accuracy rates.

accuracy (Experiments 1a and 4), and three provide no evidence that the expert altered participant performance accuracy in any significant way. Although upon first inspection this does not provide compelling evidence in favour of expert testimony, it is important to note: firstly, that no significant differences between instruction conditions and participant SEA were found using tests-of-independence, thus the declines in performance noted here have been inferred from patterns of results derived from simple goodness-of-fit analyses; and secondly, one of those studies where declines in performance were noted was also an instance where the expert provided erroneous advice, and therefore were to be expected in the event that the expert's evidence was being utilised by jurors. Thus overall, although there was little evidence of improvement to participant SEA as a result of expert evidence, these results are consistent with the expert, at times, having the effect anticipated given the objective accuracy of their testimony (see Table 11.2 below).

As already mentioned, direct comparisons between expert evidence and judicial instruction (which was also shown to have a mixed impact on SEA), revealed no reliable associations between instruction type (or quality) and Sensitivity to Eyewitness Accuracy. This is reflected in the average accuracy rates across Experiments 1(a) to 4, where accuracy in the judicial condition was 58.9%, while performance accuracy in the accurate expert and erroneous expert conditions fell at 49.2% (range: 25.6% to 66.3%) and 60.3% (range: 56.8% to 63.8%, n = 2) respectively. Thus, overall the available data provides little evidence to support suggestions of the superiority of expert evidence over judicial instruction, instead suggesting that the two forms of instruction may have equivalent effects on participant performance. This issue will be discussed further below (see *Limitations)*.

Table 11.2 : Summary of Direct Comparisons between Expert Evidence and Judicial Instruction

Experiment	**Control Response to Manipulated Variables**	**Expert Effects Observed**		**Judicial Effects Observed**	
		REE	**SEA**	**RJI**	**SEA**
Cutler, Dexter & Penrod, 1990	> Confident EW's > Culpability, > General Accuracy of ID's, > Strength Pros. Case & > Credibility No Sensitivity to WIC on any Variable	Sig. Skepticism on Verdict, General Accuracy of ID's & Strength of Pros. Case.		None	
Experiment 1a (Pre-deliberation)	Sig. Overbelief of EW No Sensitivity to Accuracy (χ^2 GOF)	No Sig. bias $\therefore$ Skepticism (cf. Control)	Below Chance Sensitivity (χ^2 GOF)	No Change in Bias	None (χ^2 GOF)
Experiment 1b (Post-deliberation)	No Bias 33.3% Accuracy (untested)	No Convictions $\therefore$ Skepticism (cf. Control) Not Sig.	40% Accuracy (untested)	No Bias	100% Accuracy (untested)
Experiment 2 (Accurate Expert)	No Bias No Sensitivity to Accuracy (χ^2 GOF)	No Bias	No Change in Sensitivity (χ^2 GOF)	No Bias	Sig. Sensitivity (χ^2 GOF)

Experiment	Control Response to Manipulated Variables	Expert Effects Observed		Judicial Effects Observed	
		REE	SEA	RJI	SEA
Experiment 3 (Erroneous Expert)	No Bias Sig. Sensitivity to Accuracy (χ^2 GOF)	No Bias	No Change in Sensitivity (χ^2 GOF)	No Bias	No Sensitivity (χ^2 GOF)
Experiment 4	No Bias Sig. Sensitivity to Accuracy (χ^2 GOF)	No Bias	No Change in Sensitivity in Accurate Expert No Sensitivity in Erroneous Expert (χ^2 GOF)	No Bias	No Sensitivity (χ^2 GOF)

Aim 3: To Assess the Validity of Inferring From Indirect to Direct Methodologies

The third aim of this thesis was to empirically investigate the validity of making inferences about Sensitivity to Eyewitness Accuracy from fictional eyewitness designs, referred to here as indirect measures (see Table 11.3 below). Specifically, this thesis used the testimony of real eyewitnesses in order to provide participants with a probabilistic indicator of eyewitness identification accuracy, thus requiring participants to evaluate *individual* eyewitnesses who may or may not have behaved as predicted by the eyewitness expert. It was predicted that under these conditions participant jurors would be less capable of discriminating between accurate and inaccurate identifications than their responses to expert evidence would. That is, it was anticipated that indirect measures would overestimate performance on the more complex real eyewitness discrimination task.

Beginning with participant responses to expert evidence (i.e., the Response to Expert Evidence level of analysis), only three studies using real eyewitness designs have reported evidence that the eyewitness expert will cause participants to change their decision criterion and become more Skeptical of eyewitness identifications: Wells et al., 1980; Experiment 1a and Experiment 1b (see Table 11.3 below). However, contrary to suggestions that the testimony of an expert ought to reduce the weight attributed to eyewitness evidence (Deffenbacher, 1984; Geiselman, 1994; Leippe et al., 2004; Pezdek, 2007; Pezdek et al., *in press*; Wells et al., 1980), thereby improving participant discrimination accuracy, this Skepticism was not associated with an improvement in SEA in any instance. Instead it co-occurs with a decrease in evaluation accuracy (Experiments 1a & b[18]) or with no effect of expert evidence (Wells et al., 1980). Thus, although it is not clear if Skepticism predicts a decline in performance accuracy, or an absence of an effect, there is certainly no evidence to suggest that Skepticism is a *sufficient* analogue of Sensitivity to Eyewitness Accuracy. That is, there is little evidence that one can validly infer from performance on this indirect measure to the direct measure. This is even true in cases where it might be argued that Skepticism ought to be most beneficial (Leippe et al., 2004), namely

[18] This analysis is true *iff* the rate of performance accuracy in Experiment 1b (40%) is estimated to be significantly below chance performance. This statistical analysis was not conducted as a result of the limited sample size.

where participants showed a tendency to overbelieve the eyewitness (as in Experiment 1a). Thus it does not seem appropriate to either advocate for, or measure only, the expert's moderating effects. This is because neither the role, nor the outcome corresponds closely with the intended effect of expert evidence: to assist jurors to discriminate between accurate and inaccurate identification evidence.

Turning now to Sensitivity to Expert Opinion analyses, there is no statistically validated evidence to suggest that significant Sensitivity to Expert Opinions reliably predicts significant improvements in Sensitivity to Eyewitness Accuracy. This conclusion is based on evidence collected from six studies (four of which are presented in this thesis), where participants were found to respond to the expert opinions, thereby showing significant SEO, yet the discrimination accuracy of these participants was significantly improved, relative to other instruction conditions, in only *one* instance (Wells & Wright, 1983 cited in Wells,1986). A more generous statistical analysis of the data presented in this thesis (utilising chi-squared goodness-of-fit tests rather than tests-of-independence) suggests that there have been: a) two instances where SEO appears to be operating as an analogue for SEA (Experiments 1a & 4 (erroneous expert)), with observed SEO corresponding to appropriate changes in SEA; b) three instances where SEO studies appear to have either overestimated (Experiments 2 & 4 (accurate expert)) or underestimated SEA (Experiment 3); and c) one case where the pattern of results is unclear (Experiment 1b[19]).

In summary then, even taking the most generous interpretation, there does not appear to be any clear correspondence between this indirect measure and Sensitivity to Eyewitness Accuracy; as the evidence suggests that SEO can at times predict, overestimate and underestimate SEA. These results call into question the validity of: a) conducting studies which can only measure SEO, rather than SEA; and b) the conceptualisation of the eyewitness expert as simply an educator (see p. 5), as once again, neither the role, nor the outcome prove to be consistent with legal expectations of the impact or expert evidence.

[19] In this study participants showed significant Sensitivity to Expert Opinion, leading one to anticipate chance levels of discrimination accuracy. However, measurement of SEA *appeared* to show accuracy at levels significantly worse than chance (40%). This figure has not been statistically compared to chance performance due to limited statistical power.

Table 11.3 : Comparison Between Direct and Indirect Measures

Experiment	**Expert Effects Observed**		
	REE	**SEO**	**SEA**
Wells, Lindsay & Tousignant, 1980	Sig. Skepticism of EW Accuracy	Sig. Sensitivity to EW Confidence	None
Wells & Wright, 1983	None	Sig. Sensitivity to WIC	Sig. Sensitivity to EW Accuracy
Experiment 1a	No Sig. bias $\therefore$ Skepticism (cf. Control)	Counter-Sensitivity (Qualitative Logistic Reg.)	Below Chance Sensitivity (χ^2 GOF)
Experiment 1b	No Convictions $\therefore$ Skepticism (cf. Control) Not Sig.	No Sig. Ass'n between Confidence & Verdict $\therefore$ Sensitivity (Verdict by Confidence r_{pbi})	40% Accuracy (untested)
Experiment 2	No Bias	Sig. Sensitivity (Confidence by Belief ANOVA)	No Change in Sensitivity (χ^2 GOF)
Experiment 3	No Bias	Sig. Sensitivity (Influencing Considerations ANOVA)	No Change in Sensitivity (χ^2 GOF)
Experiment 4	No Bias	Sig. Sensitivity to Accurate Expert Sig. Sensitivity to Erroneous Expert (Confidence by Belief ANOVA)	No Change in Sensitivity in Accurate Expert No Sensitivity in Erroneous Expert (χ^2 GOF)

Explanations for the Failure for Experts to Improve Juror Discrimination Accuracy

Given the data presented, it is reasonable to ask: how can we account for the experts' failure to improve participant Sensitivity to Eyewitness Accuracy? This is of particular importance since the evidence overwhelmingly suggests that participants heard and utilised the expert advice (i.e., were Sensitive to Expert Opinions), *and* we know that that advice was objectively useful (at least in some cases) for the task at hand.

Experiment 5 was conducted in an attempt to answer this question. By investigating the role of expertise in eyewitness discrimination tasks the need to rely on participant-jurors to enact the advice expert advice was removed. It was reasoned that this provided the best possible test of the utility of the experts' expertise as the involvement of participant-jurors was no longer necessary. Moreover, this design had the potential to provide information about the amount of training necessary to improve novice SEA to expert levels. The results of this study provided less than compelling evidence for the utility of expertise in the eyewitness discrimination task. In fact, participants across all levels of expertise performed with equivalent accuracy, and individuals who performed particularly well were as likely to come from the novice community members as they were to come from trained experts. This study thereby made it very difficult to attribute the failure of expert evidence (to improve SEA) to participant-jurors, instead placing the responsibility firmly with the expert and the nature of their expertise. The question therefore becomes: how is it that those people who know the most about eyewitness identification issues are not significantly better than novice members of the community at discriminating between accurate and inaccurate eyewitness testimony?

The answer to this question may lie in the nature of the testimony itself. Consider the real world situation where multiple factors are known to influence the likely accuracy of an identification, each to varying extents. This means that some, but not all, eyewitnesses will behave consistently with probabilistically derived predictions. Yet other eyewitnesses will display *some* characteristics that are *consistent* with these predictions and *some* characteristics that are *inconsistent* with these predictions. An example of this would be the eyewitness who made an accurate identification after a long exposure to a disguised perpetrator. Given this situation, the expertise needed for this task must include an ability to determine *when, if* and *how* a particular known predictor has influenced eyewitness identification accuracy. That is, the information experts have about the *significance* of a predictor is simply not enough. To discriminate accurate eyewitnesses from inaccurate eyewitnesses, one must know how these factors operate in combination and how to weigh each in order to decide *which* of all possible predictors they need to consider in their decision-making. However, eyewitness experts currently do not have sufficient knowledge regarding the relative contributions or the interactions between predictors to even begin to provide jurors with what amounts to an actuarial tool for the discrimination of eyewitness accuracy (Seelau & Wells, 1995). This explains why simply providing jurors with education, no matter how relevant, will only ever have a small impact on discrimination accuracy, as unless the expert can instruct jurors to consider factor X, for witness Y, giving it Z weight (relative to the other factors), then the juror is still left to determine *what, when* and *how* in each individual instance. Without this capacity, there is no reason to anticipate that performance in the presence of expert testimony would be systematically better than performance without it. Indeed:

> *decision-making research in a variety of psychological domains…shows that integration is quite difficult to achieve, even with trained judges* (Cutler et al., 1989b, p. 313).

Importantly, this account also sheds some light on the results observed in Experiment 4 in particular, where participant-jurors were provided a near perfect predictor of eyewitness identification accuracy: confidence. Specifically, in that study participants were: a) not aware that confidence was the only useful cue to accuracy in the scenario; and b) they were not aware how much weight to attribute to it. Without this information it is likely that participants in the expert evidence condition attempted to weight confidence along with *other factors* raised by the expert (as well as some of their own invention), and in doing so performed less well than they could have if they had considered eyewitness confidence in isolation. In addition to this, there is also some evidence that participant-jurors were less than perfectly calibrated to the confidence information available to them, as the correlation between the eyewitness's actual numerical expression of confidence and the participants' estimate of eyewitness confidence was significant although not perfect ($r_{pbi} = .593$, $p < .0005$). Thus, there also appears to be room for the expert to assist jurors not only regarding the *when* and *how*, but also the *what* of eyewitness identification factors.

Limitations

Using Confidence as the Manipulated Variable

This thesis has largely focused on the impact of expert evidence regarding the confidence-accuracy correlation. Although this issue wasn't targeted at the outset of this research, the use of real eyewitnesses limited decisions about experimental design. Specifically, although two separate (and time consuming) pilots were conducted in an attempt to incorporate the effects of various witnessing factors on eyewitness identification accuracy, in the end these manipulations did not prove to effect the accuracy of real eyewitness identifications. The only factor seen to reliably differentiate accurate from inaccurate testimony was ultimately the confidence-accuracy correlation. Thus, the role of this variable in the current research was determined largely by the eyewitnesses themselves.

It may well be that the inferences made in this thesis are limited somewhat by this concentration on the confidence-accuracy relationship as the key feature of expert evidence. Firstly, I fully accept that a different pattern of instruction effects might have been observed if a different eyewitnessing factor had been manipulated and different advice offered by the eyewitness expert. However, as already mentioned in Chapter 9, it does not logically follow that investigations of expert testimony regarding eyewitness confidence are worthless *per se*, particularly given that: a) this factor was significantly and repeatedly associated with the accuracy of these real eyewitnesses; b) this factor has been associated with the accuracy of eyewitnesses in general (although there is disagreement in the literature regarding the nature and magnitude of that association; Brewer & Wells, 2006; Juslin, Olsson, & Winman, 1996; Lindsay, Nilsen, & Read, 2000; Leippe, 1995; Olsson, Juslin, & Winman, 1998; Read, Lindsay, & Nicholls, 1998; Sporer, Penrod, Read, & Cutler, 1995; Weber & Brewer, 2003, 2004); and c) eyewitness experts have expressed a willingness to testify to the confidence-accuracy issue (Kassin et al., 1989; Kassin et al., 2001). Thus, while the issue of confidence-accuracy may not be ideal or representative of *all* topics the expert may testify to, it is neither inappropriate to focus on this issue, nor irrelevant to an understanding of eyewitness expert effects on juror decision-making.

A second issue associated with the use of the confidence-accuracy correlation relates to the fact that this heuristic has previously been identified as one that participant-jurors naturally tend to rely on when making their identifications

(Bradfield & Wells, 2000; Brewer & Wells, 2006;Cutler & Penrod, 1988; Lindsay, Wells & O'Connor,1989). It may be the case, then, that participant-jurors in the control and judicial instructions were going to use eyewitness confidence to predict identification accuracy regardless, artificially inflating their performance where the confidence-accuracy correlation amongst eyewitnesses was significant (Experiments 3 & 4). Yet even the prospect of this assistance does not account for: a) the failure to identify a reliable *difference* between accurate and erroneous expert conditions in Experiment 4; b) the objectively low rates of discrimination accuracy in the accurate expert condition in Experiment 4 (66.3%), particularly given the extremely robust confidence-accuracy correlation, in that case $r_{pbi\,=}$ .827; c) the objectively high accuracy rates in Experiment 3 (63.8%) where jurors were erroneously told that confidence was not useful and therefore should not have been expected to perform significantly better than chance; or d) the failure for experts themselves to utilise and benefit from this cue in Experiment 5. At most, the potential artificial inflation of performance in the control and judicial conditions may account for a failure to find an association between instruction type and accuracy in cases where participants in the expert condition were told to use confidence, that is, in Experiment 4 where comparisons are made between the control, judicial and accurate expert conditions. In all other instances Sensitivity to Eyewitness Accuracy in expert evidence conditions should have been equivalent to chance, as these participants were told *not to use confidence*, while performance in the control and judicial conditions ought to have varied as a function of the predictive power of the confidence-accuracy correlation. The former prediction was not supported in Experiments 1a and 3, where participant performance differed significantly from chance. The latter prediction was violated by control participants in Experiment 1a, and judicial instruction participants in all instances, indicating that the assumed role of eyewitness confidence in participant decision-making is not sufficient to account for the performance accuracy observed in any instruction condition. Thus, overall it is unsatisfactory to attribute the failure to find significant associations between accuracy and instruction type to the use of eyewitness confidence as the key predictor of identification accuracy; some responsibility must be attributed to the nature, rather than the content, of the expert's evidence.

Representative Nature of Eyewitnesses

The issue of the representative nature of eyewitnesses is another which may serve to limit the conclusions and inferences drawn in this thesis. That is, doubts may be raised in defence of the expert and their evidence, suggesting: that the eyewitnesses used in this study were systematically different from those eyewitnesses real jurors will be asked to evaluate, therefore rendering the expert testimony less effective. If this were true, one might successfully argue that experts could not have been expected to perform better than novices in Experiment 5 because their expertise pertains to a different set of witnesses than those they were asked to evaluate; and related to this, that expert testimony could not therefore have improved participant discrimination accuracy. Yet, as already mentioned, the key feature of the identifications made by the eyewitnesses, the confidence-accuracy correlation, is hardly unexpected given the literature, so it cannot be said that these eyewitnesses were *entirely* unrepresentative. In fact, the only regard in which it is likely that these real eyewitnesses differ from real world eyewitnesses relates specifically to the number of factors influencing the accuracy of their identification. That is, it is very probable that eyewitnesses in these studies were *easier* to evaluate than real world eyewitnesses, given that eyewitness accuracy could be predicted on the basis of just one witnessing factor, while the accuracy of real world eyewitnesses is free to vary with respect to innumerable interrelated factors. Thus, while the eyewitnesses presented to participants in this thesis may not perfectly represent real world eyewitnesses, it is more likely that they provided participants with an easier, rather than more difficult evaluation task than a juror would face in the courtroom. Specifically, if one accepts the earlier explanation offered for the failure of expert evidence (see *Explanations for the Failure...* above), then participants in Experiments 1 to 5 only had to establish *which* factor to utilise, and after having done this correctly did not need to consider *how* this factor should be weighed in the context of other influential factors in the witnessing scenario. Therefore, while the representative nature of the eyewitnesses used may indeed be questioned, it does not necessarily follow that this impaired the efficacy of expert evidence; rather, it appears likely that this simplified what would otherwise be a very complex integrated evaluation. It is important to note, however, that despite the relative simplicity of this task, one fundamental feature of the design remained the same at all times: real eyewitnesses were used to provide participants with *probabilistic* predictors of identification accuracy. Thus, despite their relative simplicity, it must be taken into consideration that in comparison to other fictional eyewitness stimuli, the tasks in this research were *always* more akin to courtroom evaluations.

Null Effect of Expert Evidence on SEA

The reader will likely note that no test of the association between instruction type and discrimination accuracy revealed a significant relationship, leading to one of two possible interpretations: that the data represent a) a Type II error, i.e., the failure to detect and existing association between instruction type and discrimination accuracy; or b) the real state of the world where there is no association between expert testimony and participant Sensitivity to Eyewitness Accuracy. Although it is always difficult to differentiate between these two possible interpretations, overall the evidence presented in this thesis suggests that the latter of these inferences is more likely to be the correct one.

1. Power

Power analyses reported throughout this thesis indicate that the ability to detect effects of expert evidence which were *at least* moderate in size (w = 0. 3) ranged from 96% (in Experiment 3) to 99.8% (in Experiment 4 for the comparison between expert evidence conditions). This analysis suggests that the expert's evidence is most likely having only a small effect (w = 0. 3) on discrimination accuracy, or none at all.

2. Replication

The absence of a significant association between instruction type (including expertise level) and discrimination accuracy was observed in each of the five experiments reported in this thesis and in the study published by Wells and colleagues (Wells et al.). Moreover, these results (Experiments 1 to 5) reflect the performance of more than 1000 participants from varied sub-populations and levels of expertise. Thus, this finding appears to be very robust and replicable.

3. Internal Consistency

Although expert testimony did not have the effect on participant SEA which had been anticipated, there was clear evidence that the expert testimony resulted in predicted effects on other measures including: a) participant-juror ratings of the influence of eyewitness confidence; b) ANOVA's investigating the association between belief decision and perceived eyewitness confidence; c) regression analyses predicting belief decisions; and d) participant recall for expert instruction. Thus, there is little doubt that participants were attending to the expert evidence, or that this evidence did affect their responses. There is simply no evidence to indicate that expert testimony had an effect on the specific outcome of discrimination accuracy.

4. Design

Experiment 4 (and to some extent the combination of Experiments 2 and 3) was designed to provide an opportunity to detect a difference between the two types of expert evidence which should have resulted in the greatest possible effect (i.e., the difference between helpful and unhelpful advice). Accordingly, it is appropriate to be confident that the impact of this manipulation is substantially greater than the simple difference between the presence and absence of expert testimony which has most often been investigated by other researchers in this field. Thus, here the expert evidence has been permitted a far greater range of possible effects than ever tested before, and yet no impact was evident. Moreover, the design of Experiment 5 attempted to remove other variables which may have been inhibiting the efficacy of expert evidence (i.e., participant-jurors and abridged expert testimony), however, this refinement provided no indication of an association between expertise and improved eyewitness discrimination accuracy.

5. Validation

Each experiment presented here incorporated a comparison serving as a manipulation check in order to validate that there was a real difference between accurate and inaccurate eyewitnesses for participants to be Sensitive to and for experts to testify to. In every instance, this analysis indicated that accurate eyewitnesses were rated as significantly more confident than inaccurate eyewitnesses. Moreover, accurate and inaccurate eyewitnesses were not found to differ with respect to any other personality variable which may have served to override or dilute the extant difference in confidence. Thus, there was clear evidence that there was a real and perceptible difference between accurate and inaccurate eyewitnesses for the expert testimony to capitalise on.

6. Context

Finally, the interpretation of the null effect provided throughout this thesis complies with the methodological advice this matter:

> *while we cannot prove the null hypothesis, in many practical contexts we have to make decisions and act as though the null hypothesis were true. This is especially the case in applied research, where decisions have to be based on imperfect knowledge which only suggests that a treatment has had no detectable effect* (Cook & Campbell, 1979, p. 45).

Accordingly, even if we can't be certain that instruction type had no effect whatsoever, based on the evidence, we can be confident that the instruction type had *no detectable effect* in contexts where favourable conditions were provided (as in Experiment 4).

Taken together, all of these reasons provide a compelling rationale to suggest that there simply was no association between instruction type and discrimination accuracy, rather than a failure to detect an existing association. As a consequence, it follows that eyewitness expert evidence (or expertise) did not substantially or significantly influence participant Sensitivity to Eyewitness Accuracy.

Interpretation of Belief and Confidence Data

ANOVA analyses of belief type and eyewitness confidence ratings have been used throughout this thesis as a means by which to test the association between perceptions of confidence and belief decisions. It is important for the reader to note that this analysis provides no causal information regarding these factors, that is, it is not possible to state conclusively that perceptions of eyewitness confidence influenced belief decisions, as it is equally possible that belief decisions may have influenced perceptions of confidence. Even so, these analyses still provide useful information about the *relationship* between decision type and estimates of eyewitness confidence, suggesting that expert evidence can significantly alter this relationship when compared with control and judicial instruction conditions (see Experiments 1a, 2 & 4).

Significance & Innovation

This thesis represents a significant and original contribution to research investigating the effects of evidence or instructions regarding eyewitness testimony, as well as expert evidence and judicial instructions in general.

With regard to specific instructions on the topic of eyewitness identification issues, this thesis has added significantly to extant knowledge by: a) providing the first investigation of the pattern judicial instruction recommended for use in NSW; b) presenting the first fair comparison between judicial instruction and expert evidence by making use of real eyewitness testimony and using the SEA outcome measure; c) utilising an experimental design with the potential to detect positive and negative differences between types of eyewitness expert evidence where all previous studies have investigated only the effects of the presence or absence of such testimony; d) incorporating the first test of the ability for experts to discriminate between accurate and inaccurate identification evidence, along with the first investigation of the relationship between levels of expertise and task performance; e) updating previous surveys of the knowledge and opinions of legal professionals in the Australian forensic context; and f) increasing the generalisability of eyewitness instruction effects by incorporating representatives from various populations including community members and psychologists. Accordingly, conclusions regarding the relative efficacy of judicial instruction and expert evidence on SEA, and the role of expertise in discrimination accuracy are both novel and significant in the eyewitness domain.

More importantly, however, this thesis provides the first in-depth consideration of the experimental implications associated with the use of fictional eyewitness designs. Specifically, strong arguments have been provided to challenge the logic of inferences made from traditional indirect measures to direct measures, and consequently in favour of the adoption of real eyewitness designs. Furthermore, the research reported in this thesis *empirically* supports the arguments made regarding the equivalence (or lack of) between direct and indirect measures, as no evidence was found to support a close correspondence between Responsiveness to Expert Evidence, Sensitivity to Expert Opinions and Sensitivity to Eyewitness Accuracy. Above and beyond the other important contributions reported, this thesis provides both the rationale and the evidence necessary to prompt a reconsideration of the methodology used by researchers investigating the utility of both eyewitness expert evidence and judicial instructions.

Future Directions

Ideally, future research would treat eyewitness expert evidence as a system variable which, like any other, exerts an impact on the accuracy of an outcome, in this case juror decision-making. As such, experimenters would approach investigations of expert evidence with a discrimination accuracy criterion in mind, considering the effects of varied types of expert evidence with varying content on this important outcome, in turn modifying best practice advice to experts on the basis of the newly available empirical evidence. In keeping with this goal, future research should aim to collect and present the testimony of many real eyewitnesses to staged crimes. Importantly, the accuracy of these eyewitnesses should vary as a function of numerous different probabilistic predictors in order to begin to understand in what specific contexts, if any, expert evidence can provide assistance with eyewitness discrimination tasks. Further comparisons should also be made investigating the correspondence between direct and indirect measures with respect to various eyewitnessing factors aside from confidence. This will help to verify whether or not the logic of the arguments presented in this thesis is also borne out by empirical evidence given different samples of real eyewitnesses.

Conclusions

Overall, three key conclusions can be made on the basis of the research presented here:

1. Suggestions regarding the superiority of eyewitness expert evidence relative to judicial instruction which have been voiced by legal professionals (in Chapter 2) and by researchers (in Chapter 5) likely overstate the differences between pattern judicial instruction and expert evidence, particularly with regard to the outcome measure of discrimination accuracy. The evidence presented here provides no indication that the testimony of an eyewitness expert (regarding the confidence-accuracy issue) can reliably improve discrimination when compared with either no-instruction control or judicial instruction conditions. Consequently, the

available empirical data requires that the assumption of expert superiority ought to be re-evaluated.

2. There is compelling evidence to suggest that direct and indirect measures of discrimination accuracy provide both qualitatively and quantitatively different estimates of participant performance. Thus, in the future researchers ought to either: a) be more cautious when suggesting that significant improvements on indirect measures equate to a significant improvement in discrimination accuracy; b) utilise real eyewitness designs, thereby removing the necessity for inference altogether; or c) spend significant amounts of time constructing and evaluating the correspondence between fictional eyewitness designs which incorporate probabilistic predictors of eyewitness identification accuracy, and real eyewitness designs.
3. It is possible to explain both the inefficacy of eyewitness expert evidence and the poor performance of experts themselves. This explanation suggests that neither expert testimony, nor expert knowledge contain the necessary information to improve discrimination accuracy *where the quality of eyewitness identifications varies as a function of probabilistic predictors*. This explanation, if it withstands future research and peer review, suggests that eyewitness expert testimony will not be able to significantly improve performance accuracy until it incorporates information regarding the co-occurrence and interrelationship of various eyewitnessing factors and their impact on eyewitness accuracy. Thus, experts who choose to provide testimony in the absence of such information must be satisfied that the likely outcome of their testimony is consistent with their intentions and with the expectations of those who have requested their assistance. As at this time, based on the evidence presented in this thesis, one must entertain doubts that expert testimony will inevitably, or even frequently, lead to better outcomes for innocent accused.

REFERENCES

Aamandt, M. G., & Custer, H. (2006). Who can best catch a liar? A meta-analysis of individual differences in detecting deception. *The Forensic Examiner, 15*(1).

Ainsworth, P. B. (1998). The psychologist as expert witness. In *Psychology Law and Eyewitness Testimony*. Chichester: John Wiley & Sons.

Benton, T. R., Ross, D. F., Bradshaw, E., Thomas, W., & Bradshaw, G. S. (2006). Eyewitness Memory is Still Not Common Sense: Comparing Jurors, Judges and Law Enforcement to Eyewitness Experts. *Applied Cognitive Psychology, 20*(1), 115-129.

Blackstone, W. (1979). *Commentaries on the Laws of England* (Vol. 4). Chicago: University of Chicago Press.

Blonstein, R., & Geiselman, R. (1990). Effects of witnessing conditions and expert witness testimony on credibility of an eyewitness. *American Journal of Forensic Psychology Vol 8(4) 1990, 11-19*.

Bond, C. F., & DePaulo, B. M. (2006). Accuracy of Deception Judgments. *Personality and Social Psychology Review, 10*(3), 214-234.

Brewer, N., & Wells, G. L. (2006). The Confidence-Accuracy Relationship in Eyewitness Identification: Effects of Lineup Instructions, Foil Similarity, and Target-Absent Base Rates. *Journal of Experimental Psychology: Applied, 12*(1), 11-30.

Brigham, J. C. (1988). Is witness confidence helpful in judging eyewitness accuracy? In *Gruneberg, M* (Vol. M (Ed); Morris, pp. R).

Brigham, J. C., & Bothwell, R. K. (1983). The ability of prospective jurors to estimate the accuracy of eyewitness identifications. *Law and Human Behavior, 7*(1), 19-30.

Brigham, J. C., & Wolfskeil, M. P. (1983). Opinions of attorney and law enforcement personnel on the accuracy of eyewitness identifications. *Law and Human Behavior, 7*, 337-349.

Caputo, D. D., Dunning, D., Lindsay, R. C. L., Ross, D. F., Read, J. D., & Toglia, M. P. (2007). Distinguishing accurate identifications from erroneous ones: Post-dictive indicators of eyewitness accuracy. In *The handbook of eyewitness psychology, Vol II: Memory for people.* (pp. 427-449): Lawrence Erlbaum Associates Publishers: Mahwah.

Chaiken, S., & Maheswaran, D. (1994). Heuristic processing can bias systematic processing: Effects of source credibility, argument ambiguity, and task importance on attitude judgment. *Journal of Personality and Social Psychology, 66*(3), 460-473.

Cooper, J., & Hall, J. (2000). Reactions of mock jurors to testimony of a court appointed expert. *Behavioral Sciences & the Law, 18*(6), 719-729.

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

Cutler, B. L., Dexter, H. R., & Penrod, S. D. (1989a). Expert testimony and jury decision making: An empirical analysis. *Behavioral Sciences & the Law, 7*(2), 215-225.

Cutler, B. L., Dexter, H. R., & Penrod, S. D. (1990a). Nonadversarial methods for sensitizing jurors to eyewitness evidence. *Journal of Applied Social Psychology, 20*(14, Pt 2), 1197-1207.

Cutler, B. L., & Penrod, S. D. (1995). *Mistaken identification: The eyewitness, psychology, and the law.* Cambridge: Cambridge University Press.

Cutler, B. L., Penrod, S. D., & Dexter, H. R. (1989b). The eyewitness, the expert psychologist, and the jury. *Law and Human Behavior, 13*(3), 311-332.

Cutler, B. L., Penrod, S. D., & Dexter, H. R. (1990b). Juror sensitivity to eyewitness identification evidence. *Law and Human Behavior, 14*(2), 185-191.

Daubert v. Merrell Dow Pharmaceuticals, Inc., 113 (S.Ct 1993).

Deffenbacher, K. A. (1984). Experimental psychology actually can assist triers of fact. *American Psychologist, 39*, 1066-1068.

Deffenbacher, K. A., & Loftus, E. F. (1982). Do jurors share a common understanding concerning eyewitness behavior? *Law and Human Behavior, 6*(15-30).

Devenport, J. L., & Cutler, B. (1997). Eyewitness identification evidence: Evaluating commonsense evaluations. *Psychology, Public Policy, and Law, 3*(2/3), 338-361.

Devenport, J. L., & Cutler, B. L. (2004). Impact of defense-only and opposing eyewitness experts on juror judgments. *Law and Human Behavior, 28*(5), 569-576.

Devenport, J. L., Stinson, V., Cutler, B. L., & Kravitz, D. A. (2002). How effective are the cross-examination and expert testimony safeguards? Jurors' perceptions of the suggestiveness and fairness of biased lineup procedures. *Journal of Applied Psychology, 87*(6), 1042-1054.

Devlin, L. (1976). *Report of the Committee on Evidence of Identification in Criminal Cases*.

Dinardo, L., & Rainey, D. (1991). The effects of illumination and exposure time of facial recognition. *Psychological Record 41*(3), 329-334.

Doyle, J. M. (1998). No confidence: A step toward accuracy in eyewitness testimony. *The Champion*(January/February), 12.

Elliott, R. (1993). Expert testimony about eyewitness identification: A critique. *Law and Human Behavior, 17*(4), 423-437.

Evidence Act (Cth, 1995).

Federal Rules of Evidence (St. Paul, MN: West 1975, 1984).

Fleet, M. L., Brigham, J. C., & Bothwell, R. K. (1987). The confidence-accuracy relationship: The effects of confidence assessment and choosing. *Journal of Applied Social Psychology. Vol, 17*(2), 171-187.

Fox, S. G., & Walters, H. (1986). The impact of general versus specific expert testimony and eyewitness confidence upon mock juror judgment. *Law and Human Behavior, 10*(3), 215-228.

Frye v. United States, 293 (D.C. Cir. 1923).

Geiselman, R. (1994). *Eyewitness Expert Testimony: Handbook for the Forensic Psychiatrist, Psychologist and Attorney* (2 ed.). Balboa Island, California: ACFP Press.

Geiselman, R., & Mendez, B. A. (2005). Assistance to the fact finder: Eyewitness expert testimony versus attorneys' closing arguments. *American Journal of Forensic Psychology 23*(2), 5-14.

Geiselman, R., Putman, C., Korte, R., Shahriary, M., Jachimowicz, G., & Irzhevsky, V. (2002). Eyewitness expert testimony and juror decisions. *American Journal of Forensic Psychology Vol 20(3) 2002, 21-36.*

Goodman, J., & Loftus, E. F. (1992). Judgment and memory: The role of expert psychological testimony on eyewitness accuracy. In P. Suedfeld & P.E. Tetlock (Eds.), *Psychology and Social Policy*. NY: Hemisphere Publishing Corporation.

Greene, E. (1988). Judge's instruction on eyewitness testimony: Evaluation and revision. *Journal of Applied Social Psychology, 18*(3, Pt 1), 252-276.

Greene, E., & Loftus, E. F. (1984). Solving the eyewitness problem. *Behavioral Sciences & the Law, 2*(4), 395-406.

Gross, S. R., & O'Brien, B. (2007). *Frequency and predictors of false convictions: Why we know so little, and new data on captial cases* (No. Working Paper No. 93): University of Michigan Law School.

Hoffheimer, M. H. (1989). Effect of particularized instructions on evaluation of eyewitness identification evidence. *Law & Psychology Review, 13*, 43-58.

Hosch, H. M. (1980). A comparison of three studies of the influence of expert testimony on jurors. *Law and Human Behavior, 4*(4), 297-302.

Hosch, H. M., Beck, E., & McIntyre, P. (1980). Influence of expert testimony regarding eyewitness accuracy on jury decisions. *Law and Human Behavior, 4*(4), 287-296.

Huff, C. R. (1987). Wrongful conviction: societal tolerance of injustice. *Research in Social Problems and Public Policy*.

Hunter, J. B., Cameron, C., & Henning, T. (2005). *Litigation II: Evidence & Criminal Process* (7th ed.). Australia: Lexis Nexis.

Innocence Project. (2008) Retrieved February 6, 2008, from http://www.innocenceproject.org/

Juslin, P., Olsson, N., & Winman, A. (1996). Calibration and diagnosticity of confidence in eyewitness identification: Comments on what can be inferred from the low confidence-accuracy correlation. *Journal of Experimental Psychology: Learning, Memory, and Cognition. Vol, 22*(5), 1304-1316.

Kapardis, A. (1997). *Psychology and Law: A Critical Introduction*. Cambridge: Cambridge University Press.

Kargon, R. (1986). Expert testimony in historical perspective. *Law and Human Behavior*, *10*(1/2), 15-27.

Kassin, S. M., & Barndollar, K. A. (1992). The psychology of eyewitness testimony: A comparison of experts and prospective jurors. *Journal of Applied Social Psychology*, *22*(16), 1241-1249.

Kassin, S. M., Ellsworth, P. C., & Smith, V. L. (1989). The "general acceptance" of psychological research on eyewitness testimony: A survey of the experts. *American Psychologist*, *44*(8), 1089-1098.

Kassin, S. M., Tubb, V., Hosch, H. M., & Memon, A. (2001). On the "General Acceptance" of Eyewitness Testimony Research; A New Survey of the Experts. *American Psychologist, 56*(5), 405-416.

Katzev, R. D., & Wishart, S. S. (1985). The impact of judicial commentary concerning eyewitness identifications on jury decision making. *Journal of Criminal Law & Criminology Vol 76(3) Fal 1985, 733-745*.

Konecni, V., & Ebbesen, E. (1986). Courtroom testimony by psychologists on eyewitness identification issues: Critical notes and reflections. *Law and Human Behavior, 10*(1/2), 117-126.

Leippe, M. R. (1995). The case for expert testimony about eyewitness memory. *Psychology, Public Policy, and Law, 1*(4), 909-959.

Leippe, M. R., Eisenstadt, D., Rauch, S. M., & Seib, H. M. (2004). Timing of Eyewitness Expert Testimony, Jurors' Need for Cognition, and Case Strength as Determinants of Trial Verdicts. *Journal of Applied Psychology, 89*(3), 524-541.

Lempert, R. O. (1986). Social science in court: On "eyewitness experts" and other issues. *Law and Human Behavior, 10*(1/2).

Lindsay, D., Nilsen, E., & Read, J. (2000). Witnessing-condition heterogeneity and witnesses' versus investigators' confidence in the accuracy of witnesses' identification decisions. *Law and Human Behavior, 24*(6), 685-697.

Lindsay, R. (1994). *Expectations of eyewitness performance: Jurors' verdicts do not follow from their beliefs*. Ross, David Frank (Ed); Read, J. Don (Ed); Toglia, Michael P (Ed). (1994). Adult eyewitness testimony: Current trends and developments.

Lindsay, R., MacDonald, P., & McGarry, S. (1990). Perspectives on the Role of the Eyewitness Expert. *Behavioural Sciences and the Law, 8*, 457-464.

Lindsay, R., Wells, G. L., & O'Connor, F. (1989). Mock-Juror Belief of Accurate and Inaccurate Eyewitnesses: A Replication and Extension*. *Law and Human Behavior, 13*(3), 333-339.

Lindsay, R., Wells, G. L., & Rumpel, C. M. (1981). Can people detect eyewitness-identification accuracy within and across situations? *Journal of Applied Psychology, 66*(1), 79-89.

Loftus, E. F. (1974). Reconstructive memory: The incredible eyewitness. *Psychology Today, 8*, 116-119.

Loftus, E. F. (1979). The malleability of human memory. *American Scientist, 67*(3), 312-320.

Loftus, E. F. (1980). Impact of expert psychological testimony on the unreliability of eyewitness identification. *Journal of Applied Psychology, 65*(1), 9-15.

Loftus, E. F., & Monahan, J. (1980). Trial by data: Psychological research as legal evidence. *American Psychologist, 38*, 550-563.

Loftus, G. R. (1985). Picture perception: Effects of luminance on available information and information extraction rate. *Journal of Experimental Psychology: General, 114*, 342-356.

Maass, A., Brigham, J. C., & West, S. G. (1985). Testifying on eyewitness reliability: Expert advice is not always persuasive. *Journal of Applied Social Psychology, 15*(3), 207-229.

Maass, A., Brigham, J. C., & West, S. G. (1987). *Testifying on eyewitness reliability: Expert advice is not always persuasive*: Wrightsman, Lawrence S (Ed); Willis, Cynthia E (Ed); Kassin, Saul M (Ed).

Macmillan, N. A., & Creelman, C. D. (2005). *Detection Theory a User's Guide* (2nd ed.): Routledge.

McCloskey, M., Egeth, H., & McKenna, J. (1986). The experimental psychologist in court: The ethics of expert testimony. *Law and Human Behavior, 10*(1-2), 1-13.

McCloskey, M., & Egeth, H. E. (1983). Eyewitness identification: What can a psychologist tell a jury? *American Psychologist, 38*(5), 550-563.

McConkey, K., & Roche, S. (1989). Knowledge of eyewitness memory. *Australian Psychologist, 24*(3), 377-384.

McCullough, M. L. (2002). "Do not discount lay opinion": Comment. *American Psychologist, 57*(5), 376-377.

McKenna, J., Treadway, M., & McCloskey, M. (1992). Expert psychological testimony on eyewitness reliability. In P Suedfeld & P.E. Tetlock (Eds.), *Psychology and Social Policy*. NY: Hemisphere Publishing Corporation.

Memon, A., Hope, L., & Bull, R. (2003). Exposure duration: Effects on eyewitness accuracy and confidence. *British Journal of Psychology. Vol, 94*(3), 339-354.

Munsterberg, H. (1908). *On the Witness Stand: Essays on Psychology and Crime*: Doubleday, Page & co.

Noon, E., & Hollin, C. R. (1987). Lay knowledge of eyewitness behaviour: A British survey. *Applied Cognitive Psychology, 1*(2), 143-153.

Ogloff, J. R. P., & Rose, V. G. (2005). *The Comprehension of Judicial Instructions*. Brewer, Neil (Ed); Williams, Kipling D (Ed). (2005). Psychology and law: An empirical perspective.

Olsson, N., Juslin, P., & Winman, A. (1998). Realism of confidence in earwitness versus eyewitness identification. *Journal of Experimental Psychology: Applied, 4*, 101-118.

Park, R. C. (2003). Eyewitness identification: Expert witnesses are not the only solution. *Law, Probability and Risk, 2*, 305-308.

Penrod, S. D., & Cutler, B. (1992). Eyewitnesses, experts, and jurors: improving the quality of jury decision-making in eyewitness cases. In J. Misumi, B. Wilpert & H. Motoaki (Eds.), *Organizational and Work Psychology*. Hilldale, N.J.: Erlbaum.

Penrod, S. D., & Cutler, B. (1999). *Preventing mistaken convictions in eyewitness identification trials: The case against traditional safeguards*. Roesch, Ronald (Ed); Hart, Stephen D (Ed); Ogloff, James R. P (Ed). (1999). Psychology and law: The state of the discipline.

Pezdek, K. (2007). Expert testimony on eyewitness memory and identification. In M Costanso, D Krauss & K. Pezdek (Eds.), *Expert Psychological Testimony for the Courts*. Mahwah, NJ: Erlbaum.

Pezdek, K., Avila-Mora, E., & Sperry, K. (*in press*). Ecological validity issues in jury simulation research: Evaluating the effectiveness of eyewitness expert testimony.

Potter, R., & Brewer, N. (1999). Perceptions of witness behaviour-accuracy relationships held by police, lawyers and mock-jurors. *Psychiatry, Psychology and Law, 6*(1), 97-103.

R v D.D. (2000).

R v Sarago. (2006), NSWCC.

R v Skaf. (2006), NSWSC 394.

R v. Emery. (1993).

R v. Smith. (2000), A. Crim. 116.

R v. Turner. (1975), QB 834.

Rahaim, G. L., & Brodsky, S. L. (1982). Empirical evidence versus common sense: Juror and lawyer knowledge of eyewitness accuracy. *Law and Psychology Review, 7*, 1-15.

Ramirez, G., Zemba, D., & Geiselman, R. (1996). Judges' cautionary instructions on eyewitness testimony. *American Journal of Forensic Psychology Vol 14(1) 1996, 31-66.*

Read, D., Lindsay, D., & Nicholls. (1998). The relationship between confidence and accuracy: Is the conclusion changing? In C. P Thompson, D. J Herrman, J. D Read, D. Bruce, D. G Payne & M. P Toglia (Eds.), *Eyewitness memory: Theoretic and applied perspectives* (pp. 107-130). Mahwah, NJ: Lawrence Erlbaum.

Saunders, D. M., Vidmar, N., & Hewitt, E. C. (1983). Eyewitness testimony and the discrediting effect. In S Lloyd-Bostock & B. R Clifford (Eds.), *Evaluation Witness Evidence*. London: Wiley.

Scheck, B., Neufeld, P., & Dwyer, J. (2001). *Actual Innocence* (2 ed.). New York: Signet Printing.

Seelau, s., & Wells, G. L. (1995). Applied Eyewitness Research: The Other Mission*. *Law and Human Behavior*, *19*(3), 319-324.

Shapiro, P. N., & Penrod, S. D. (1986). Meta-analysis of facial identification studies. *Psychological Bulletin*, *100*(2), 139-156.

Shaw, J. S., Garcia, L. A., & McClure, K. A. (1999). A lay perspective on the accuracy of eyewitness testimony. *Journal of Applied Social Psychology, 29*(1), 52-71.

Sporer, S., Penrod, S., Read, D., & Cutler, B. (1995). Choosing, confidence, and accuracy: A meta-analysis of the confidence-accuracy relation in eyewitness identification studies. *Psychological Bulletin, 118*(3), 315-327.

Steinbeck, J. (1958). *TheLog From the Sea of Cortez*. London: Pan Books Ltd.

Steusser, L. (2005). Experts on eyewitness identification: I just don't see it [Electronic Version]. *International Commentary on Evidence*, 3.

Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences* (4th ed.): Lawrence Erlbaum Associates.

U.S. v. Greene. (1979), 591 F.2d 471 8th Cir.

U.S. v Telfaire. (1972), D.C. Cir. 469 F.2d 552.

U.S. v. Zeiler. (1970). 427 F.2d 1305 3d Cir.

United States v. Amaral. (1973), 9th Cir. 488 F.2d 1148.

Van Wallendael, L. R., Cutler, B., Devenport, J. L., & Penrod, S. D. (2007). Mistaken Identification = Erroneous Conviction? Assessing and improving legal safeguards. In R. Lindsay, David F. Ross, Don Read & M. P Toglia (Eds.), *Handbook of Eyewitness Psychology: Memory for People* (Vol. 2, pp. 557-572). Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.

Vidmar, N., & Schuller, R. A. (1989). Juries and expert evidence: Social framework testimony. *Law and Contemporary Problems, 52*(4), 133-176.

Volokh, A. (1997). *n* Guilty Men [Electronic Version]. *University of Pennsylvania Law Review*, 173, 146 from http://www.law.ucla.edu/volokh/guilty.htm#2

Wagenaar, W. A., & Van Der Schreir, J. H. (1996). Face recognition as a function of distance and illumination: a practical tool for use in the courtroom. *Psychology, Crime & Law, 2,* 321-332.

Weber, N., & Brewer, N. (2003). The effect of judgment type and confidence scale on confidence-accuracy calibration in face recognition. *Journal of Applied Psychology, 88*(3), 490-499.

Weber, N., & Brewer, N. (2004). Confidence-Accuracy Calibration in Absolute and Relative Face Recognition Judgments. *Journal of Experimental Psychology: Applied, 10*(3), 156-172.

Wells, G. L. (1978). Applied eyewitness-testimony research: System variables and estimator variables. *Journal of Personality and Social Psychology, 36*(12), 1546-1557.

Wells, G. L. (1986). Expert psychological testimony: Empirical and conceptual analyses of effects. *Law and Human Behavior, 10*(1-2), 83-95.

Wells, G. L., & Leippe, M. R. (1981). How do triers of fact infer the accuracy of eyewitness identification? Using memory for peripheral detail can be misleading. *Journal of Applied Psychology, 66*, 682-687.

Wells, G. L., Lindsay, R., & Ferguson, T. J. (1979). Accuracy, confidence, and juror perceptions in eyewitness identification. *Journal of Applied Psychology, 64*(4), 440-448.

Wells, G. L., Lindsay, R., & Tousignant, J. (1980). Effects of expert psychological advice on human performance in judging the validity of eyewitness testimony. *Law and Human Behavior*, *4*(4), 275-285.

Wells, G. L., & Murray, D. M. (1983). What can psychology say about the Neil v. Biggers criteria for judging eyewitness accuracy? *Journal of Applied Psychology*, *68*(3), 347-362.

Wells, G. L., & Olson, E. A. (2003). Eyewitness testimony. *Annual Review of Psychology Vol 54 2003, 277-295 http://psych annualreviews org/.*

Wigmore, J. (1909). Professor Munsterberg and the psychology of testimony. *Illinois Law Review, 3*, 399-445.

Wise, R., Pawlenko, N., Meyer, D., & Safer, M. (*under review*). A survey of defense attorneys knowledge and beliefs about eyewitness testimony. *The Champion*.

Wise, R., & Safer, M. (2004). What US judges know and believe about eyewitness testimony. *Applied Cognitive Psychology, 18*, 427-443.

Wogalter, M., Malpass, R., & McQuiston, D. (2004). A national survey of US police on preparation and conduct of identification lineups. *Psychology, Crime & Law, 10*(1), 69-82.

Woller, S. (2004). Rethinking the role of expert testimony regarding the reliability of eyewitness identifications in New York. *New York Law School Law Review, 48*, 323-352.

Yarmey, A. (2001). Expert testimony: Does eyewitness memory research have probative value for the courts? *Canadian Psychology Vol 42(2) May 2001, 92-100*.

Yarmey, A., & 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 witness evidence* (pp. 13-40). Chichester, England: Wiley.

Yarmey, D. (1986). Verbal, Visual, and Voice Identification of a Rape Suspect Under Different Levels of Illumination. *Journal of Applied Psychology, 71*(3), 363-370.

APPENDICIES

Appendix A

Expert Testimony Questionnaire (ETQ) – Please see enclosed disc.

Appendix B

Telfaire Instruction (1972) – Please see enclosed disc

Appendix C

Eyewitness Evidence Protocol I – Stage 1.

Stage 1. of a real eyewitness design involves the completion of the following three sub-stages; a) crime video construction; b) eyewitnessing and identification; and c) eyewitness interviews. Details regarding the latter two of these sub-stages are reported here, as this study uses a pre-existing crime video.

Participant-witnesses in for this study were recruited via a colleague’s experiment, which will not be fully reported here. That study involved undergraduate participants watching a crime video before making an identification from a target-present or target-absent simultaneous lineup. Any witnesses who selected the target from a target-present lineup, or who made any selection from a target-absent lineup were asked if they would be willing to participate in this study. The first four witnesses to accept this offer (two accurate and two inaccurate identifications) then completed interviews in the style of an examination in chief and a cross-examination.

Stage 1b : Eyewitnessing & Identification

Materials

Videotaped Crime

Participant-witnesses watched a video of a crime where a young man is seen attempting to enter locked offices along a corridor. He finds one door unlocked and the camera follows him into the office where he searches for valuables. Upon finding a wallet the thief takes it and leaves the office. The video of the incident lasted one minute thirty seconds during which time thief’s face was clearly visible from many different angles, distances and levels of illumination (as he moved from the corridor to the office).

Lineups

Each participant-witness viewed one of two lineups, either a simultaneous target-present lineup or a simultaneous target-absent lineup. Allocation to one of the lineup conditions was randomly determined. The lineups consisted of nine high-resolution colour “mug shot” style photographs. The images were positioned in two rows, the top with five photos, and the bottom with four. Target-present lineups included eight foils and the target, while target-absent lineups were composed of the same eight foils plus one additional foil.

Procedure

Experimenters informed participant-witnesses that they would be shown a video and were asked to “watch carefully” before being presented the videotaped crime. At the end of the video, participant-witnesses spent approximately five minutes answering filler questions about their knowledge of Australian trivia before being presented with either the target-absent or target-present version of the lineup. Participants were then asked to try to identify the perpetrator from a lineup and were instructed that the perpetrator “may or may not be present”. Subsequent to making their identification, participants rated their confidence in the accuracy of their decision on a scale from 1-100. If the participant-witness made a correct identification from a target-present lineup or any identification from a target-absent lineup, the experimenter asked if they would be willing to participate in another study. The first four witnesses to consent were then scheduled for phase two. The nature of their identification was withheld from the experimenters involved in the interviews.

Stage 1c : Eyewitness Interviews

Participant-Witness Testimony

Two assistants (one acting as council for the prosecution the other acting as council for the defence) recorded interviews with the four participant-witnesses. These assistants were blind to the accuracy of the identification made by each participant-witness.

The first interview, conducted by council for the prosecution, was in the style of an examination-in-chief. In this interview the witness was asked to describe what they saw, and to outline the details of the identification process and their resulting decision. The cross-examination style interview, conducted by council for the defence, focused on the witness’ estimate of the duration of the incident and other peripheral details as well as information regarding the perpetrators appearance (see Appendix C on the enclosed disc for the complete interview schedules).

Procedure

Participant-witnesses returned for their interviews approximately one week after completing the witnessing experiment. Two assistants (one acting as counsel for the prosecution the other acting as counsel for the defence), who were trained to administer a standard interview schedule, recorded interviews with each participant-witness. These assistants did not know whether the participant-witness had made an accurate or inaccurate identification. Participant-witnesses were interviewed in random order.

The eyewitness was greeted by the confederate interviewer acting as council for the prosecution. Council for the defence was not present until recording commenced. It was explained to the eyewitness that they would be interviewed by two people, and that the first interview would permit them to describe what they saw, while the second interview would aim to challenge their account of events. Participant-witnesses were told that the examination-in-chief would be conducted twice, once in the presence of council for the prosecution (pre- interview), and once in the presence of both interviewers. Only the second of these interviews would be recorded. It was explained that it was of *utmost importance* for the witness to be aware that: a) the interviewer did not know what the eyewitness had seen and therefore could not aid the witness in giving their testimony; and b) that the eyewitness should not fabricate answers to questions which they did not know the answers to; rather they should indicate when and if they “did not know”. All questions in the pre-record interview were open-ended, thereby allowing the eyewitness to relate details of the crime scenario according to their own recollections. In the recorded version of the interview questions were at times necessarily leading in order to facilitate the communication of all the information already provided by the eyewitness in the pre-recorded interview.

Upon completion of the pre-recorded interview the counsel for the prosecution invited the counsel for the defence into the interview room and recording began. All questions presented to the eyewitness in the cross-examination were video recorded at their first presentation. After both interviews were recorded, the eyewitness was debriefed about the experiment and thanked for their participation in the study.

Appendix D

Pre-trial Instructions

All Conditions

Participants-jurors were read the following instructions:

What you are about to see is the testimony of an actual witness to a theft. First, this witness saw a crime being committed and was then asked to identify the person they saw from a lineup. This witness was able to identify the suspect from that lineup, and as a result was asked to testify at the trial of the accused.

What follows is a recording of the testimony given by the witness. The footage takes the form of an examination in chief and a cross-examination.

The examination in chief is conducted by the prosecution and allows the witness to provide their account of the theft and the identification they made. The cross-examination of the witness is conducted by those defending the accused and seeks to challenge the witness' version of the events.

I would like you to watch these interviews as though you were a juror in the trial of the man accused by this witness. As such you are to watch this footage keeping in mind that the role of a juror in a trial such as this is to determine the guilt or innocence of the accused based on the evidence presented.

In addition the following judicial warning (taken from JCNSW, 2006; s3-61) was issued:

s3-610

It is important to note then, that the evidence given by the witness you are about to see is the only evidence relating to the guilt or innocence of the accused in this trial. Accordingly it is your concern as a juror to assess not only what the witness says, but also the witness' honesty and reliability. You should examine and scrutinize the testimony of this witness with great care before you accept their evidence.

Appendix E

Expert Evidence Pre-trial Instructions – Please see enclosed disc

Appendix F

Judicial Instructions to Jury Pre-Deliberation – Please see enclosed disc

Appendix G

Jury Study Case Facts & Legal Instructions – Please see enclosed disc

Appendix H

Jury Study Case Questionnaire (Expert) – Please see enclosed disc

Appendix I

Eyewitness Evidence Protocol II – Stage 1

As a result of the limitations observed in Experiment 1 (the jury study), another protocol was constructed to collect eyewitness identification testimony. In Protocol II the intention was to manipulate the quality of eyewitnessing conditions by systematically varying factors known to affect eyewitness identification accuracy rates. As a result of this, it is hoped that both accurate and inaccurate identifications will be accompanied by a range of cues that participant-jurors can use to estimate the likely accuracy of the identification, therefore introducing real world predictors into the experiment. This is an improvement on Protocol I.

Stage 1 of this protocol is composed of three sub-stages: a) crime video construction; b) eyewitnessing and identification; and c) eyewitness interview.

Stage 1a : Crime Video Construction

The aim of this procedure was to develop two movies which differed significantly with respect to a range of the witnessing and identification variables about which an eyewitness expert would later testify in court. This was necessary so that the participant-juror has the opportunity to demonstrate Sensitivity to the witnessing conditions and the expert's opinions regarding them.

Design

Four witnessing variables were manipulated in order to create a range of high- and low-quality witnessing scenarios. Sixteen different versions of the same general crime scenario (see Table I.1 below) could be constructed by systematically varying lighting (from day to night as in Dinardo & Rainey, 1991; Fox & Walters, 1986; Geiselman et al., 2002; Wagenaar & Van Der Schreir, 1996; Yarmey, 1986); angle of view (profile or no profile, Lindsay, D. et al., 2000; Shapiro & Penrod, 1986); disguise (in this case hat or no hat as in, Cutler et al., 1989a; Cutler et al., 1990a; Cutler et al., 1989b; Geiselman et al., 2002) and viewing distance (near or far, Blonstein & Geiselman, 1990; Fox & Walters, 1986; Geiselman et al., 2002; Wagenaar & Van Der Schreir, 1996; Wells & Wright, 1983 cited in Wells, 1986). Of these combinations, eight were filmed as shown by the shaded cells in Table I.1 below. The crime itself was a bag-snatching, where the perpetrator approached the victim from behind while she was talking on a mobile phone. The perpetrator, grabbed the victims bag and struggled over it for a few minutes before escaping.

Consistent with the principle of maximising processing at encoding time and maximizing information available at recognition time outlined in Shapiro and Penrod (1986), it was hypothesised that: 1) performance in darkness would be worse than daylight; 2) performance when the suspect was disguised would be worse than when undisguised; 3) performance at a shorter distance would be better than at a farther distance; and 4) that the provision of more viewing angles would better facilitate accurate identification than when fewer angles were provided.

Table I.1 : Factorial Structure of Witnessing Factors in Recorded Crime Videos

	Near				**Far**			
	No Disguise		**Disguise**		**No Disguise**		**Disguise**	
	Profile	**No Profile**	**Profile**	**No Profile**	**Profile**	**No Profile**	**Profile**	**No Profile**
Day	1	2	3	4	5	6	7	8
Night	9	10	11	12	13	14	15	16

Stage 1b : Eyewitnessing & Identification (Pilot 1)

Two of the eight available videos, numbers 1 and 10, were selected for use during the eyewitnessing phase of the protocol as upon inspection of the quality of the crime videos, it was suspected that identifications made from videos 3, 4, 5, 14, 15 and 16 might be so poor as to reduce performance to floor levels.

In movie 1 – or the “good” version of the crime – the thief was visible for 16 seconds, the crime took place in daylight and during the film the thief was clearly seen in both his left and right profiles and in full-face view. In movie 10 – the “poor” version of the crime – the thief was visible for two seconds less (14 seconds), the crime took place at night under chromatic overhead lighting and the video did not show a clear view of either profile. In all other respects, as far as possible, the videos were the same.

Method

Participants

The participant-witnesses were 56 staff and students from the University of New South Wales (16 males and 40 females) aged between 18 and 56 years ($\overline{x}$ = 27.0 yrs, $\overline{\sigma}$ = 8.6 yrs) who responded to advertised requests for experimental participation in return for charitable donations made in their name.

Design

This study employed a 2 (witnessing conditions: good vs. poor) x 2 (lineup type: simultaneous target-present lineup vs. simultaneous target-absent lineup) between subjects factorial design. The experimenter was blind to the random allocation of participant-witnesses to these conditions.

Materials

Lineups

Each participant-witness viewed one of two lineups, either a simultaneous target-present lineup or a simultaneous target-absent lineup. Allocation to one of the lineup conditions was randomly determined by a computer program. The lineups consisted of nine high-resolution colour "mug shot"–style photographs. The images were randomly positioned in two rows, with five photographs on the top and four in the row below. Target-present lineups included eight foils and the target, while target-absent lineups were composed of the same eight foils plus one additional foil. For the purposes of this study, any identification from this lineup was considered a "suspect" identification. In addition, both lineups also included a "Not Present" option which always appeared in the bottom right corner of the lineup array allowing the participant to indicate that they did not think the perpetrator appeared in the lineup.

Witnessing Program

A computer program presented participant-witnesses with one of the four possible combinations of the witness video (good or poor) and lineup types (target-absent or target-present). The experimenter was blind to the allocation of participants to these conditions. The witnessing program recorded the type of identification made, the amount of time taken to make the identification and the participant-witness's rating of their own confidence in their decision.

Procedure

The participant-witnesses completed the study on their own to ensure that the experimenter was blind to witnessing and lineup condition. The witnessing program instructed participant-witnesses to "watch carefully" and then presented a version of the videotaped crime. At the end of the video the participant-witness spent approximately five minutes answering 15 yes-no filler questions about their cognitive style. The program then asked participant-witnesses to try to identify (select with the computer mouse) the perpetrator from a lineup. They were instructed that the perpetrator "may or may not be present" and that they may choose the "Not Present" option if they believed that to be the case. Participants then rated their confidence in the accuracy of their decision on a scale from 1-100 before asking the experimenter back into the room.

Results

Effect of Witnessing Conditions

There was no significant association between witnessing conditions (i.e. good vs. poor) and the accuracy of participant-witnesses identification ($\chi^2_{(1)} = .286$, $p = .593$) (see Table I.2. for the breakdown of identification type by witnessing condition).

Table I.2 : Identification Type by Version of Crime

	Identification Type					
	Target-Absent		Target-Present			Total
	Suspect	Correct Rejection	Foil	Perp.	Incorrect Rejection	
Good	4	10	1	5	8	28
Poor	5	9	3	4	7	28
Total	9	19	4	9	15	56

However, differences between the good and poor witnessing conditions did emerge when the association between confidence and accuracy was considered. Table I.3 shows the observed confidence-accuracy relation for all witnesses, for those who chose any individual from either lineup (“choosers”) and for the subset of these witnesses who either correctly selected the target from a target-present lineup or made a selection from a target-absent lineup. It is this sub group who are most likely to be asked to give evidence in court (henceforth referred to as “court-choosers”).

Table I.3 : Confidence-Accuracy Correlations for Total Sample, Choosers and Court-Choosers in Good and Poor Witnessing Conditions

	Eyewitness Sample								
	Total Sample			**Choosers (Perp. + Susp. + Foil)**			**Court-Choosers (Perp. + Susp.)**		
	N	r_{pbi}	***p***	***n***	r_{pbi}	***p***	***n***	r_{pbi}	***p***
Good	28	.000	.999	10	.700	.024	9	.827	.006
Poor	28	-.161	.413	12	.294	.354	9	.367	.331
Overall	56	-.072	.599	22	.444	.038	18	.568	.014

Overall, the correlations between confidence and accuracy for all participant-witnesses was not significant (r_{pbi} = -0.072, p = 0.599), although confidence was found to be a significant predictor of accuracy for positive identifications made under good witnessing conditions (choosers: r_{pbi} = 0.7, p = .024; court-choosers: r_{pbi} = .827, p = .006). No significant confidence-accuracy relationship was found for identifications made for the poor video. (r_{pbi} = -.161, p = .413). The patterns observed above for the good witnessing conditions were replicated overall witnessing conditions (choosers: r_{pbi} = 0.444, p = .038; court-choosers: r_{pbi} = .568, p = .014). These positive confidence-accuracy correlations were significantly stronger for choosers (z = 2.05, p = .02) and for court-choosers (z = 2.45, p = .007) than for all participant-witnesses.

Eyewitness Accuracy

By using participant-juror lineup selections to determine whether the 10 faces in the lineup were plausible alternatives for the perpetrator (see Table I.4 below for the pattern of decisions made in each lineup type) it was established that the probability of making an accurate decision by chance is at most 0.2. Compared to this, participant-jurors were observed to make accurate decisions significantly more often than would be expected by chance alone both overall ($\chi^2_{(1)} = 31.5$, $p = .000$), and within witnessing conditions (good: $\chi^2_{(1)} = 19.72$, $p = .000$; poor: $\chi^2_{(1)} = 15.78$, $p = .000$). No significant differences were observed between the average amount of time taken by participant-eyewitnesses to make an accurate decision ($\bar{x} = 19.39$ sec, $\bar{\sigma} = 9.53$) compared with an inaccurate decision ($\bar{x} = 23.07$ sec, $\bar{\sigma} = 21.72$, $t_{(54)} = 0.821$, $p = .415$). This was also true for both chooser ($t_{(20)} = 1.08$, $p = .292$) and court-chooser ($t_{(16)} = 1.4$, $p = .181$) groups of eyewitnesses.

Table I.4 : Lineup Member Selected by Lineup Type

	Lineup Member Selected										
	Foil 1	Foil 2	Foil 3	Foil 4	Foil 5	Foil 6	Foil 7	Foil 8	Perp.	Reject	**Total**
Target-Absent	1	4	2	2	0	0	0	0	0	19	28
Target-Present	1	1	0	1	1	0	0	0	9	15	28
Total	2	5	2	3	1	0	0	0	9	34	56

Pilot 1 Discussion

Overall, there was no significant between the quality of view (good or poor) and identification accuracy. This failure to influence witness accuracy appears to reflect the overall similarity of the two tested versions of the film, indicating that changes in lighting and angle of view alone were not sufficient to significantly influence identification accuracy. The apparent similarity of the witnessing scenarios appears to be further compounded by the frequency with which participant-witnesses chose to reject the lineup (64% rejection in the good condition and 57% in the poor).

Accordingly, it does not appear that the failure to identify differences across conditions can be attributed to either floor or ceiling effects.

In spite of this failure, however, one significant difference between conditions was observed. Analysis of the association between confidence and identification accuracy revealed significant positive correlations for choosers and court-choosers. This was true both across conditions for only identifications made under good witnessing conditions. This finding adds support to existing evidence which suggests the presence of a moderating effect of identification type on the confidence-accuracy relationship (Brewer & Wells, 2006; Juslin, Olsson, & Winman, 1996; Lindsay, Nilsen, & Read, 2000; Leippe, 1995; Olsson, Juslin, & Winman, 1998; Read, Lindsay, & Nicholls, 1998; Sporer, Penrod, Read, & Cutler, 1995; Weber & Brewer, 2003, 2004). This pattern of association was not identified among those witnesses from the poor condition and therefore suggests that although the original manipulation attempted was not powerful enough to significantly influence identification accuracy, it was sufficient to affect the predictive power of the eyewitness's confidence estimates.

Stage 1b : Eyewitnessing & Identification (Pilot 2)

Due to the failure to identify significant differences in participant-witness identification accuracy across witnessing conditions, a third version of the crime video was piloted. The sample in this study was composed of University of New South Wales staff and students who participated either voluntarily in return for book donations or for partial course credit. It was hoped that this sample of participants may have a higher choosing rate than community members alone, and that increased variability in witnessing conditions may help to significantly differentiate the good version of the crime scenario from its poorer counterpart.

Method

Participants

Pilot 2 included the 28 participant-witnesses who viewed the good version of the crime from Pilot 1, plus an additional 49 staff and students from the University of New South Wales (32 male and 45 female) aged between 18 and 48 years ($\bar{x}$ = 23.1

yrs, $\overline{\sigma}$ = 6.8 yrs) who either responded to advertised requests for experimental participation in return for charitable donations made in their name, or received course credit. In total, 38 of these witnesses viewed the very poor version of the crime and 39 viewed the good version of the crime.

Design

This study employed a 2 (witnessing conditions: good vs. very poor) x 2 (lineup type: simultaneous target-present lineup vs. simultaneous target-absent lineup) between subjects factorial design.

Materials

Crime Video

Movie 15 – or the "very poor" version of the crime scenario – was used as it offered the most challenging scenario for viewing the thief out of the eight filmed versions of the crime. In this version of the crime the thief was visible for 16 seconds, the crime occurred at night under chromatic lighting, the thief wore a baseball-style cap shading his eyes throughout which concealed his hair style and colour. Both his right and left profile were shown clearly and as far as possible, this video was the same as videos 1 and 10 in all other respects.

Lineups

The lineups used in Pilot 2 were identical in all respects to those used in Pilot 1.

Witnessing Program

In this study, the experimenter was required to allocate participant-witnesses to witnessing condition in order to ensure a larger sample in the very poor witnessing condition. The experimenter did, however, remain blind to both lineup and identification type. The witnessing program recorded the type of identification made, the amount of time taken to make the identification and the participant-witness's rating of their own confidence in their decision.

Procedure

The procedure was identical to that described for Pilot 1 above.

Results

Effect of Witnessing Condition

Again, there was no significant association between witnessing conditions (i.e. good vs. very poor) and the accuracy (vs. inaccuracy) of participant-witnesses identification ($\chi^2_{(1)}$ =1.628, p = .202). See Table I.5. for a description of identification type by witnessing condition.

Table I.5 : Identification Type by Version of Crime

	Identification Type					
	Suspect	**Correct Rejection**	**Foil**	**Perp.**	**Incorrect Rejection**	**Total**
Good (Pilot 1)	4	10	1	5	8	28
Good (Pilot 2)	2	5	0	0	4	11
Very Poor	9	10	6	4	9	38
Total	9	19	4	9	15	77

Consideration of the confidence-accuracy relationship (reported in Table I.6 below) again revealed differences between the good and very poor witnessing conditions. For this sample, the overall correlations between confidence and accuracy was significant for all participant-witnesses (r_{pbi} = 0.293, p = 0.01), and for court-choosers (r_{pbi} = 0.410, p = 0.047). As in Pilot 1, confidence was also found to be a significant predictor of accuracy for positive identifications made under good witnessing conditions (choosers: r_{pbi} = 0.649, p = .023; court-choosers: r_{pbi} = .736, p = .01). These positive confidence-accuracy associations were significantly stronger for both "choosers" (z = 1.75, p = .04) and "court-choosers" (z = 2.1, p = .01) relative to the correlation observed for all participant-witnesses. While as before, no significant confidence-accuracy relationship existed for positive identifications made in the very poor condition (choosers: r_{pbi} = 0.014, p = .955; court-choosers: r_{pbi} = .061, p = .843), a significant association between confidence and accuracy was observed when all identification types were considered (r_{pbi} = 0.385, p = 0.017).

Table I.6 : Confidence-Accuracy Correlations for Total Sample, Choosers and Court-Choosers in Good and Very Poor Witnessing Conditions

	Eyewitness Sample								
	Total Sample			**Choosers**			**Court-Choosers**		
	N	r_{pbi}	***p***	***n***	r_{pbi}	***p***	***n***	r_{pbi}	***p***
Good	39	.12	.469	12	.649	.023	11	.736	.01
Very Poor	38	.385	.017	19	.014	.955	13	.061	.843
Overall	77	.293	.01	31	.312	.087	24	.410	.047

Eyewitness Accuracy

Of the 10 possible decisions participant-witnesses could make from each lineup type only seven were ever selected (See Table I.7 below). If we take this as an indication of the "functional size" of the lineup, the probability of making an accurate decision by chance alone becomes one in seven. Compared to this, participant-jurors were observed to make accurate decisions significantly more often than would be expected by chance alone, both overall ($\chi^2_{(1)} = 56.07$, $p = .000$), and within witnessing conditions (good: $\chi^2_{(1)} = 43.57$, $p = .000$; very poor: $\chi^2_{(1)} = 15.78$, $p = .000$). No significant differences were observed between the average amount of time taken by participant-eyewitnesses to make an accurate decision ($\bar{x}$ = 21.41 sec, $\bar{\sigma}$ = 13.41) compared with an inaccurate decision ($\bar{x}$ = 23.14 sec, $\bar{\sigma}$ = 19.40, $t_{(75)} = 0.442$, p = .660). There was no significant difference between the mean decision time under good and poor witnessing conditions, either for choosers ($t_{(29)} = .477$, $p = .637$), or court-choosers ($t_{(22)} = .619$, $p = .542$).

Table I.7 : Lineup Member Selected by Lineup Type

	Lineup Member Selected											
	Foil 1	Foil 2	Foil 3	Foil 4	Foil 5	Foil 6	Foil 7	Foil 8	Foil 9	Perp.	Reject	**Total**
Target-Absent	1	5	1	5	1	1	0	0	0	0	25	40
Target-Present	2	0	0	2	1	1	1	0	0	9	21	37
Total	3	5	1	7	2	2	1	0	0	9	46	77

Pilot 2 Discussion

As with the good and poor versions of the crime scenario, no significant association between witnessing condition and identification accuracy was observed under good and very poor conditions. In this case, however, this finding represents a more surprising result as these films differed with respect to three witnessing variables rather than two. In the good version of the crime, the bag-snatching took place in bright daylight, the perpetrator was undisguised, and spent the duration of the film in close proximity to the camera, indeed exiting the scene by running toward the camera itself. By contrast, in the very poor version of the crime, the incident took place at night under chromatic lighting, the thief wore a baseball cap concealing his hair and casting shadows over his eyes throughout, and finally the perpetrator was largely viewed from a further viewing distance, reflected by the fact that he both entered and exited the scene from further away from the camera than in the good condition. In spite of this, however, both the overall accuracy rates and the frequency with which different types of identification decisions were made, suggests that the two versions of the crime presented an equivalent challenge to participant-witnesses. This does not appear to be consistent with Shapiro and Penrod's (1986, p. 152) principle that "performance will be improved to the extent that viewing conditions…make more information available at identification". However, this is an interpretation which should be made somewhat tentatively in light of the high rates of lineup rejections observed across both witnessing conditions (69% rejection in the good condition and 50% rejection in the very poor condition).

The pattern of associations observed between eyewitness confidence and identification accuracy in Pilot 2 was broadly similar to those observed in Pilot 1. As in Pilot 1, a significant association was noted for both choosers and court-choosers making identifications from the good version of the crime. The same relationship was not observed among identifications made under very poor witnessing conditions. However, in contrast with earlier results, significant associations between confidence and accuracy were found for all eyewitnesses both overall and within the very poor witnessing conditions. This suggests that those eyewitnesses from the very poor condition who rejected the lineup (either correctly or incorrectly) were acting consistently with their confidence ratings. Conversely, expressions of confidence made by eyewitnesses rejecting the lineup under good conditions were sufficiently poorly calibrated to undermine the diagnosticity of the confidence estimates provided by those who made positive identifications from the lineup.

General Discussion – Stages 1a & b

When it comes to predicting and indeed manipulating eyewitness recognition performance, both Loftus (1985) and Shapiro and Penrod (1986) suggest that the more information available at the time of encoding, the better the performance at retrieval will be. In accordance with this principle, Pilot 1 and 2 sought to vary the quality of witnessing conditions through the manipulation of four estimator variables (Wells, 1978), each of which has been shown in isolation to effect identification accuracy: illumination, angle of view, distance of view and disguise. Neither the effects of illumination and angle of view (in Pilot 1), nor a combination of illumination, distance of view and disguise (in Pilot 2) were sufficient to significantly influence eyewitness identification performance. This is surprising as such a result is in stark contrast with previous research which found the manipulation of illumination alone reliably affected eyewitness identification accuracy (Dinardo & Rainey, 1991; Wagenaar & Van Der Schreir, 1996; Yarmey, 1986). At this time it is unclear how to account for the discrepancy between previous research and the results of the current pilot studies. However, further exploration of results would require the testimony of additional participants. Despite this, it is important to note that these tests were only conducted to generate real eyewitnesses, and not to investigate the combined influence of witnessing and identification factors on likely identification accuracy. It may be that these studies lack sufficient power to establish a significant difference as Pilot 1 had only a 61% chance of detecting an effect of moderate size ($1-\beta = 0.61$, $\alpha = .05$, $w = 0.3$), while Pilot 2 had a 75% chance of detecting an effect of the same size ($1-\beta = 0.75$, $\alpha = .05$, $w = 0.3$). In addition, further investigation of the common element between all three levels of witnessing conditions tested – the lineup – may provide an alternate account for the absence of the significant effects sought.

The significant confidence-accuracy correlations observed for choosers, including the group we have called “court-choosers”, from the good witnessing condition is consistent with the association perceived by participant-jurors in Experiment 1 and with previous research on the mediating effect of decision type on the confidence-accuracy relation (Brewer & Wells, 2006; Sporer et al., 1995). The same association between confidence and accuracy was not evident for either choosers or court-choosers when identifications were made in poor or very poor witnessing conditions. This pattern of results suggests that although the manipulation of witnessing condition did not significantly affect eyewitness identification accuracy rates, it did systematically influence the predictive power of the subjective judgements made by eyewitnesses. This finding is in keeping with the results obtained from previous research investigating the effect of illumination on both identification performance and subjective estimates of performance. Although not significant, the evidence reported by Yarmey (1986) suggests that subjects are more confident about correct decisions made in daylight than those made at twilight, leading to a significant confidence-accuracy correlation for daylight identifications only. Yarmey also found that confidence expressed for inaccurate positive identifications was higher for observations made in reduced light than in daylight. Together these results suggest that the diagnosticity of the confidence-accuracy relationship is greater in daylight than in twilight conditions. The results reported from both Pilot 1 and Pilot 2 are consistent with this interpretation.

Given the manifest differences in the confidence-accuracy correlation across witnessing conditions, it was concluded that the aims of Stages 1a and 1b of the eyewitness evidence protocol had been achieved as the confidence-accuracy relationship is an issue that eyewitness experts are highly likely to testify to in court, (Kassin et al., 1989; Kassin et al., 2001) and this relationship varies for these groups of participant-witnesses. Therefore the eyewitnesses from the pilot studies were considered suitable for use in subsequent stages of the investigation.

Stage 1c : Eyewitness Interviews

The participant-witnesses who had viewed either the good or poor crime scenario were interviewed in the form of an examination-in-chief, followed by a cross-examination. These interviews were video recorded to serve as the stimulus materials from which participant-jurors would endeavour to evaluate the eyewitness’s evidence.

Method

Participants

Of the 56 witnesses from Pilot 1 (good vs poor witnessing conditions), only the 18 court-choosers were asked to participate in the Phase 3 interviews. Fourteen of these witnesses (5 male and 9 female) aged between 19 and 40 years ($\overline{x}$ = 25.9 yrs, $\overline{\sigma}$ = 6.2 yrs) actually completed the interview, while four failed to attend (one accurate eyewitness from each of the good and poor conditions and two inaccurate eyewitnesses from the poor condition). In return for their extra time, additional charitable donations were made on behalf of all participant-witnesses who attended their interview.

Materials

Participant-Witness Testimony

Two assistants (one acting as counsel for the prosecution the other acting as council for the defence) recorded interviews with each of the 15 participant-witnesses. These assistants did not know whether the participant-witness had made an accurate or inaccurate identification, or which version of the crime they had witnessed.

The first interview, conducted by council for the prosecution, was in the style of an examination-in-chief (or direct examination). In this interview the witness was asked to describe what they saw, and to outline the details of the identification process and their resulting decision. The cross-examination style interview, conducted by council for the defence, focused on the witness's estimate of the duration of the incident and other details regarding the perpetrator's appearance. During this interview the cross-examiner introduced into evidence the confidence estimate provided by the eyewitness at the time of their identification by stating, "*[i]sn't it true that at the time of your identification you were **X%** confident?*" This question was asked irrespective of whether there was any difference between the original and subsequent (i.e., made during direct examination) confidence estimates provided by the eyewitness. The same interview schedule was adopted as far as was possible for each of the 15 witnesses interviewed (See Appendix I on the enclosed disc for the complete interview schedules).

Procedure

At the time the eyewitness made their identification the computer program informed the experimenter if a particular eyewitness was required for a follow-up interview. At this time these witnesses were asked to come back for an interview. Participants who consented to do so made an appointment for an interview the following week. Aside from this, the interview procedure used in this protocol was identical to the one adopted in Protocol I.

Results

On average the examination-in-chief lasted just under four minutes (range: 3 min 37 sec to 6 min 40 sec). On average the cross-examination lasted between three and four minutes (range: 3 min 21 sec to 5 min 49 sec).

The eyewitness's responses to those questions which ought to have differentiated "good" eyewitnesses from "poor" eyewitnesses (i.e., describing the quality of lighting, the quality of the footage and the presence or absence of profile view) were independently evaluated by two raters blind to eyewitness accuracy.

Eyewitness Memory for Viewing Conditions

In order to check that the participant-witnesses had provided information which accurately reflected the conditions under which they had seen the perpetrator, raters coded the testimony of each eyewitness. After viewing the good and poor crime scenarios, and co-rating two eyewitness interviews, blind-raters were required to indicate whether the responses were consistent with the good or poor witnessing conditions. The raters were found to be in 100% agreement in their classifications of eyewitness descriptions of the lighting conditions and the quality of the footage (which was significantly more "grainy" in the poor condition as an artefact of the dim lighting), with both raters providing the same classification for each eyewitness (either rating the description as consistent with the poor or good condition). In all instances, these classifications were consistent with the crime scenario that the eyewitness actually viewed. That is, all witnesses who viewed the good crime scenario correctly reported the quality of lighting in the footage that they saw when asked about it in direct- or cross-examination.

Eyewitness Recall for Profile Presentation

The raters also independently coded the accuracy of the eyewitnesses' descriptions regarding the presence or absence of a profile view of the offender. Those participants in the good condition would be expected to indicate that they had seen the offender's profile, while those participants in the poor condition should indicate that they had not. Independent ratings of the interviews found that five eyewitnesses provided insufficient information to determine the accuracy of their descriptions (e.g., witness MC stated "His face was not clear, I think I could not see the front of his face quite clearly", revealing nothing about whether they saw a profile or not). The raters agreed that three eyewitnesses were inaccurate in their descriptions (two from the good condition and one from the poor), and five eyewitnesses were accurate in their descriptions (three from the good condition and two from the poor).

Eyewitness Recall for Exposure Duration

Finally all eyewitnesses were required to estimate the amount of time they had to view the perpetrator. Those participant-eyewitnesses from the good condition had 16 seconds to view the perpetrator and those in the poor condition had 14 seconds to view the perpetrator. On average all participants considerably overestimated the amount of time they had to view the perpetrator, with eyewitnesses in the "good" condition stating that they had 30.4 seconds on average to view the perpetrator (range: 19 to 40 sec), while those from the "poor" condition provided an average estimate of 54.5 seconds of exposure (range: 7 to 210 sec). Although the difference between these estimates was not significant ($t_{(5.08)} = 0.75$, $p = .484$), significantly more variance was present in the estimates provided by participants in the poor condition, than those in the good ($F = 5.82$, $p = <.05$).

Discussion

Participant-witnesses were clearly able to recall enough information to allow a third party to determine whether they viewed the crime during the day or at night. These participant-witnesses, however, were not as accurate in their recall or reporting of the angles of view they were able to see the perpetrator from. Specifically, participant-eyewitnesses were as likely to accurately report the angles of view provided to them, as they were to provide inconclusive statements (36% in both cases). Moreover, consistent with previous research, participant-witnesses in this sample showed a tendency to considerably overestimate the duration of their exposure to the perpetrator (Brigham & Bothwell, 1983; Caputo et al., 2007; Wells & Murray, 1983).

Appendix J

Expertise Study Questionnaire (Electronic) – Please see enclosed disc

Appendix K

Experiment 4 Cued Recall Questions – Please see enclosed disc

Appendix L

Eyewitness Evidence Protocol III

As in the first experiment, participant-witnesses in this study viewed a pre-existing crime video, thus only Stages 1b and 1c are reported here.

Stage 1b : Eyewitnessing & Identification

Materials

Videotaped Crime

Participant-witnesses watched a video of a crime where a young man is seen emerging from the fire stairs on the roof of a multistorey car park. The camera follows the man as he moves between the parked cars testing to see if any have been left unlocked. After attempting to gain entry to several vehicles, the man is successful and is followed into the car by the camera. Whilst being viewed from the front passenger's seat, the man searches that car before discovering a mobile phone and taking it. The perpetrator then leaves the car and is seen to return to the fire stairs he entered from.

Lineups

Each participant-witness viewed one of two lineups, either a simultaneous target-present lineup or a simultaneous target-absent lineup. The lineups were presented on a computer, and allocation to lineup condition was randomly determined. The lineups consisted of 9 high-resolution colour "mug shot" style photographs. The images were positioned in two rows, with three rows of three.

Procedure

Experimenters informed participant-witnesses that they would be shown a video and were asked to "watch carefully" before being presented the videotaped crime on the computer. At the end of the video, participant-witnesses spent approximately five minutes answering filler questions about their cognitive style before being presented with either the target-absent or target-present version of the lineup. The participants who viewed this lineup were not given clear instructions that the "perpetrator may or may not be present", instead they read the following instruction on the computer screen before viewing the mug-shots; “please identify the perpetrator from the following set of photographs”. Subsequent to making their identification, participants rated their confidence in the accuracy of their decision on a scale from 1-100. If the participant-witness made a correct identification from a target-present lineup or any identification from a target-absent lineup, the experimenter asked if they would be willing to participate in the second phase of the study. The first accurate and inaccurate eyewitnesses to consent to the follow-up interview were then scheduled to complete the next phase of the protocol. As before, the nature of their identification was withheld from the experimenters running the interview phase until all interviews were completed.

Stage 1c : Eyewitness Interview

Participant-Witness Testimony

Two assistants (one acting as council for the prosecution the other acting as council for the defence) recorded interviews with the four participant-witnesses. These assistants were blind to the accuracy of the identification made by each participant-witness.

The first interview, conducted by council for the prosecution, was in the style of an examination-in-chief. In this interview the witness was asked to describe what they saw, and to outline the details of the identification process and their resulting decision. The cross-examination style interview, conducted by council for the defence, focused on the witness’ estimate of the duration of the incident and other peripheral details as well as information regarding the perpetrators appearance (see Appendix L on the enclosed disc for the complete interview schedules).

Procedure

The interview procedure followed in this study was identical to that reported in the earlier protocols.

Appendix M

Experiment 5 : Comparative Analysis – Trial One Only

This analysis includes the data relevant only to the first judgement made by each participant in Experiment 5. This analysis is arguably the most ecologically sound as it replicates real life circumstances where jurors will be asked to evaluate the accuracy of an eyewitness's testimony without prior practice at the task and generally in the absence of relative comparisons. In addition, this is the more statistically conservative approach, as it avoids the problem of multiple, non-independent observations. However, the drawback with this approach relative to that reported in Chapter 10 is that it lacks statistical power due to being based on one third the number of observations. The experiment was designed such that the first eyewitness seen by participants from the community and postgraduate and expert samples was always from Protocol II. This was done because Protocol II had the largest number of witnesses, and thus the analysis of the response to the first trial only would be based on the largest number of different eyewitnesses.

Qualitative Evaluations of Eyewitnesses

Jurors evaluated eyewitnesses on four personality and performance dimensions using a seven-point Likert scale. There were no differences in ratings of accurate and inaccurate eyewitnesses for credibility ($t_{(143)}$ = 1.75, p = .082), attractiveness ($t_{(143)}$ = -0.43, p = .671) or trustworthiness ($t_{(143)}$ = -0.02, p = .983). Ratings of eyewitness confidence will be discussed in subsequent sections.

Effect of Expertise on Decisions

A two-way between groups analysis was conducted to explore the effect of expertise level and eyewitness accuracy on respondent estimates of eyewitness confidence. A significant main effect was observed for eyewitness accuracy ($F_{(1,144)}$ = 11.03, p < .005, η_p^2 = .08), such that accurate eyewitnesses ($\bar{x}$ = 3.80, $\bar{\sigma}$ = 1.17) were rated as being significantly more confident than inaccurate eyewitnesses ($\bar{x}$ = 3.16, $\bar{\sigma}$ = 1.24). There was also a significant main effect of expertise ($F_{(3,144)}$ = 4.61, p < .005, η_p^2 = .09), with post-hoc comparisons revealing that those respondents in the undergraduate group ($\bar{x}$ = 3.93, $\bar{\sigma}$ = 1.04) attributed significantly higher confidence to eyewitnesses than postgraduate or expert participants (postgraduate $\bar{x}$ = 3.29, $\bar{\sigma}$ = 1.44; expert $\bar{x}$ = 3.05, $\bar{\sigma}$ = 1.19). No significant interaction effect was observed ($F_{(3,144)}$ = 0.82, p = .487, η_p^2 = .02).

A subsequent two-way between groups analysis was conducted to investigate the impact of expertise and belief on respondent estimates of eyewitness confidence. In this analysis, both belief and expertise were found to significantly impact upon estimates of confidence. The main effect of belief was such that more confident eyewitnesses were more likely to be believed ($\bar{x}$ = 3.74, $\bar{\sigma}$ = 1.08) than less confident eyewitnesses ($\bar{x}$ = 3.02, $\bar{\sigma}$ = 1.42; $F_{(1,143)}$ = 7.88, p = <.01, η_p^2 = .06). Post-hoc analyses of the significant main effect of expertise ($F_{(3,143)}$ = 7.95, p = <.0005, η_p^2 = .15) again revealed that the confidence ratings attributed by members of the undergraduate group ($\bar{x}$ = 3.93, $\bar{\sigma}$ = 1.04) were significantly higher than those from either the postgraduate or expert groups (postgraduate = 3.29, $\bar{\sigma}$ = 1.44; expert $\bar{x}$ = 3.05, $\bar{\sigma}$ = 1.19). Furthermore, a significant interaction was observed ($F_{(3,143)}$ = 4.14, p = <.01, η_p^2 = .01). The community, undergraduate and expert groups all rated the confidence of eyewitnesses they did not believe as lower than those they did believe, however, postgraduates showed the reverse pattern; attributing more confidence to disbelieved than believed eyewitnesses (believed $\bar{x}$ = 3.10, $\bar{\sigma}$ = 1.45; disbelieved $\bar{x}$ = 3.75, $\bar{\sigma}$ = 1.50). See Figure M.1 below.

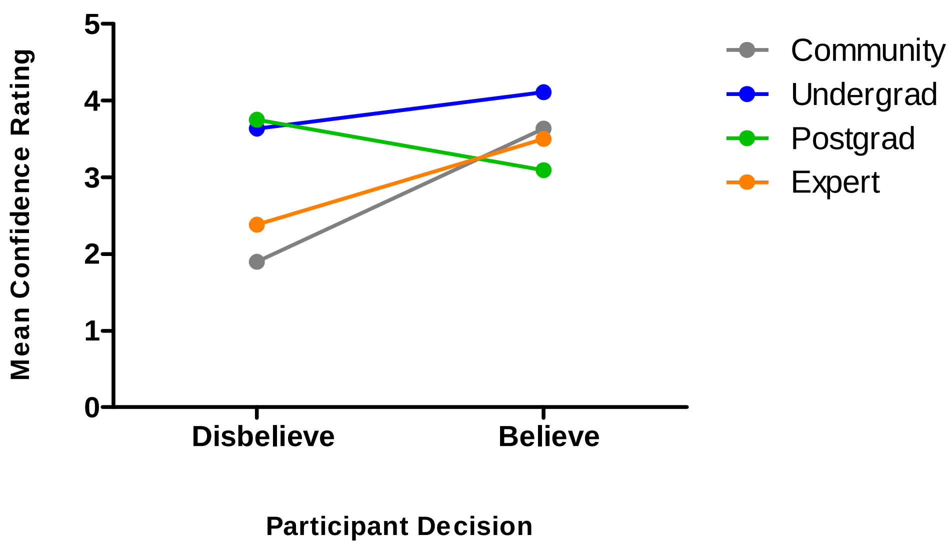

Figure M.1 Mean confidence ratings by juror belief decisions

Decision Criterion

A chi-squared goodness-of-fit test indicated all participants believed significantly more than 50% of the eyewitnesses ($\chi^2_{(1)} = 21.78$, $p = <.0005$). A two-way chi-squared test-of-independence revealed no significant association between expertise level and belief decision ($\chi^2_{(3)} = 5.50$, $p = .139$). However, there appears to be some evidence of trend towards lower levels of belief with increasing experience; undergraduates(62.1% belief, $\chi^2_{(1)} = 3.38$, $p = .066$), postgraduates (71.4% belief, $\chi^2_{(1)} = 2.57$, $p = .109$) and experts (60.0% belief, $\chi^2_{(1)} = 0.80$, $p = .377$) were not significantly more likely to believe an eyewitness, while those in the community sample did show a significant bias towards belief (80.8% belief, $\chi^2_{(1)} = 19.69$, $p = <.0005$).

Considerations Affecting Belief Decisions

A series of (2x4) ANOVA’s were conducted, each investigating the effect of belief and expertise level on participant-jurors’ impression of the extent to which each of the four variables (eyewitness confidence, eyewitness manner, eyewitnessing conditions) influenced their judgements. A (2x3) ANOVA was also conducted investigating the effect of belief type and expertise level on ratings of the impact of identification conditions on judgements. The undergraduate sample was not included in this analysis as they were not asked to evaluate the role of identification conditions in the formulation of their judgements.

Post-hoc analyses revealed that community members rated eyewitness as significantly more influential ($\overline{x} = 2.81$, $\overline{\sigma} = 0.74$) than did postgraduates ($\overline{x} = 1.71$, $\overline{\sigma} = 1.27$) or experts ($\overline{x} = 1.95$, $\overline{\sigma} = 1.13$, $F_{(3,142)} = 6.36$, $p < .0005$, $\eta_p^2 = .124$). Moreover, confidence was also rated to be significantly more influential for the undergraduates ($\overline{x} = 2.47$, $\overline{\sigma} = 0.96$) than the experts. No significant main effect for belief type ($F_{(1,142)} = 1.23$, $p = .269$) or interaction effect ($F_{(3,142)} = 1.14$, $p = .334$) was reported. Reliance on eyewitness manner also differed significantly with expertise ($F_{(3,141)} = 3.41$, $p = <.05$, $\eta_p^2 = .071$). Here post-hoc analyses revealed that experts ($\overline{x} = 1.78$, $\overline{\sigma} = 0.88$) reported being influenced by eyewitness manner significantly less than the community ($\overline{x} = 2.62$, $\overline{\sigma} = 0.80$) and undergraduate respondents ($\overline{x} = 2.50$, $\overline{\sigma} = 0.92$).

No main effect for belief type ($F_{(1,141)}$ = 0.10, p = .751) or interaction ($F_{(3,141)}$ = 1.26, p = .291) was identified. Analysis of the estimates of the influence of witnessing conditions on eyewitness evaluations revealed no significant main or interaction effects ($F_{(3,143)}$ = 0.91, p = .439), with all participants rating witnessing conditions as equally influential ($F_{(3,143)}$ = 2.28, p = .083), irrespective of belief type ($F_{(1,143)}$ = 1.40, p = .239). The final (2x3) ANOVA indicated that identification conditions were rated to be equally influential when eyewitnesses were believed as disbelieved ($F_{(1,83)}$ = 0.12, p = .727), and this rating did not vary as a function of levels of expertise ($F_{(2,83}$ = 0.29, p = .747). Moreover, there was no significant interaction effect ($F_{(2,83)}$ = 0.82, p = .446).

Qualitative Predictors of Belief Decisions

A binary logistic regression was conducted in order to investigate whether any of the participants' ratings of the eyewitness's characteristics predicted the participants' decision to believe or disbelieve the eyewitnesses they saw. This analysis was conducted separately for each group of participants. The predictor variables used included; credibility, confidence, attractiveness, and trustworthiness as the predictors in the model. Each of these variables was rated by participants on a 7-point scale

For the community group the overall model was significant ($\chi^2_{(4)}$ = 18.26, p < .005) correctly classifying 88.2% of respondents decisions, however, no single predictor in the model was shown to be significant. For undergraduates the model was also found to be significant ($\chi^2_{(4)}$ = 22.13, p = <.0005), accurately classifying participant decisions 69% of the time. In this case participant estimates of eyewitness credibility (β = -2.15, p < .005) and trustworthiness (β = 0.89, p < .05) were both significant predictors in the model. Analysis of responses in the postgraduate group also revealed a significant model ($\chi^2_{(4)}$ = 16.75, p = <.005), which accurately classified respondents 76.2% of the time. Here credibility alone was found to be a significant predictor of belief judgements (β = 1.13, p < .0005). The model for the expert participants was also significant ($\chi^2_{(4)}$ = 15.09, p < .01) correctly classifying 85% of judgements, however, no individual predictor in the model was shown to be significant.

Sensitivity to Eyewitness Accuracy

Overall, respondents accurately classified eyewitnesses 59.3% of the time, a chi-squared goodness-of-fit test indicated that this was significantly better than could be expected by chance alone (i.e., greater than 50%; $\chi^2_{(1)} = 5.03$, $p = <.05$). Respondents in the community, undergraduate, postgraduate and expert conditions attained 64.2%, 51.7%, 64.3% and 65% accuracy respectively, however, a chi-squared test-of-independence indicated that there was no significant association between expertise and accuracy ($\chi^2_{(3)} = 2.31$, $p = .511$). Even so performance accuracy was significantly better than chance in the control condition ($\chi^2_{(1)} = 4.25$, $p = <.05$), but not at the higher levels of expertise (undergraduate: $\chi^2_{(1)} = 0.07$, $p = .793$; postgraduate: $\chi^2_{(1)} = 1.14$, $p = .285$; expert: $\chi^2_{(1)} = 1.80$, $p = .180$).

Chapter 10 includes a discussion of the major differences between this analysis and the one presented in “All Trials” analysis.

Appendix A

Expert Testimony Questionnaire (ETQ)

Please note the accuracy of each of the following statements by selecting the appropriate box e.g. ☒

Definitely True	Probably True	Probably False	Definitely False	I hadn't considered this an issue	
☐	☐	☐	☐	☐	1. Very high levels of stress impair the accuracy of eyewitness testimony.
☐	☐	☐	☐	☐	2. The presence of a weapon does not impair an eyewitness's ability to accurately identify the perpetrator's face.
☐	☐	☐	☐	☐	3. The use of a one-person showup instead of a full lineup increases the risk of misidentification
☐	☐	☐	☐	☐	4. The more that members of a lineup resemble the suspect, the higher the likelihood that the identification of the suspect is accurate.
☐	☐	☐	☐	☐	5. Police instructions can influence whether or not an eyewitness makes a selection from a lineup.
☐	☐	☐	☐	☐	6. The less time an eyewitness has to observe an event, the less well he or she will remember it.
☐	☐	☐	☐	☐	7. An eyewitness's confidence is a good predictor of his or her identification accuracy.
☐	☐	☐	☐	☐	8. Memory for an event declines most rapidly immediately after its occurrence and more slowly thereafter.
☐	☐	☐	☐	☐	9. An Eyewitness' testimony about an event often reflects not only what they actually saw but information they obtained later on.
☐	☐	☐	☐	☐	10. Judgments of colour made under monochromatic light (e.g. an orange streetlight) are highly unreliable.
☐	☐	☐	☐	☐	11. An eyewitness's testimony about an event remains invariant no matter what the wording of the questions asked.
☐	☐	☐	☐	☐	12. Eyewitnesses sometimes identify as a culprit someone they have seen in another situation or context.
☐	☐	☐	☐	☐	13. Police officers and other trained observers are more accurate as eyewitnesses than the average person.

☐	☐	☐	☐	☐	14. Hypnosis increases the accuracy of an eyewitness's reported memory.
☐	☐	☐	☐	☐	15. Hypnosis decreases suggestibility to leading and misleading questions.
☐	☐	☐	☐	☐	16. An eyewitness's perception and memory for an event may be affected by his or her attitudes and expectations.
☐	☐	☐	☐	☐	17. Eyewitnesses have more difficulty remembering violent than non-violent events.
☐	☐	☐	☐	☐	18. Eyewitnesses are more accurate when identifying members of their own race than members of other races.
☐	☐	☐	☐	☐	19. An eyewitness's confidence can be influenced by factors that are unrelated to identification accuracy
☐	☐	☐	☐	☐	20. Alcoholic intoxication impairs an eyewitness's later ability to recall persons and events.
☐	☐	☐	☐	☐	21. Exposure to mug shots of a suspect increases the likelihood that the witness will later choose that suspect in a lineup.
☐	☐	☐	☐	☐	22. Traumatic experiences can be repressed for many years and then recovered.
☐	☐	☐	☐	☐	23. Memories people recover from their childhood are usually highly accurate.
☐	☐	☐	☐	☐	24.It is possible to reliably differentiate between true and false memories.
☐	☐	☐	☐	☐	25. Young children are more accurate as witnesses than are adults.
☐	☐	☐	☐	☐	26. Young children are more vulnerable than adults to interviewer suggestion, peer pressures, and other social influences
☐	☐	☐	☐	☐	27. The more that members of a lineup resemble a witness's description of the culprit, the more accurate an identification of the suspect is likely to be.
☐	☐	☐	☐	☐	28. in a lineup, the way in which photographs are presented to witnesses (e.g. simultaneously or sequentially) affects the accuracy of identifications
☐	☐	☐	☐	☐	29. Elderly eyewitnesses are less accurate than younger adults.
☐	☐	☐	☐	☐	30. The more quickly a witness makes an identification upon seeing the lineup, the more accurate he or she is likely to be.

Background Information

1. How long have you been a public defender? ________yrs

2. Approximately what percentage of cases you have been involved included disputed eyewitness identification evidence? _______%

3. Have you ever commissioned an expert to testify to the limitations of eyewitness testimony?

☐ Y ☐ N

If Yes, approximately how many times? ________

Please continue over the page

4. Of those cases where you have attempted to introduce expert evidence, what percentage have been excluded from testifying? ______%

5. Briefly describe what issues your excluded experts were planning to testify to, and the reasons for their exclusion.

6. Do you believe that jurors understand the instructions given by the judge regarding the limitations of eyewitness testimony? (please tick the most appropriate statement)

☐ Definitely Yes ☐ Probably Yes ☐ Probably No ☐ Definitely No

7. Do you believe that jury decision-making is influenced by the instructions given by the judge regarding the limitations of eyewitness testimony? (please tick the most appropriate statement)

☐ No ☐ Yes – favouring prosecution ☐ Yes – without bias ☐ Yes – favouring defence

8. Do you believe that direction from the judge can replace the testimony of a relevant expert in the case of eyewitness testimony?

☐ Y ☐ N

If No, why?

THANK YOU FOR YOUR TIME AND PARTICIPATION

If you would like to know the current status of the psychological evidence on any of these issues please contact Dr Richard Kemp (Richard.kemp@unsw.edu.au) or Kristy Martire (kmartire@psy.unsw.edu.au).

Additionally if you would like to be involved in future research in this area please fill in the space below or email Kristy Martire (as above) with your contact details.

NAME

PHONE

E-MAIL

Appendix B

Telfaire Instruction

One of the most important issues in this case is the identification of the defendant as the perpetrator of the crime. The Government has the burden of proving identity, beyond a reasonable doubt. It is not essential that the witness himself be free from doubt as to the correctness of the witness' statement. However, you, the jury, must be satisfied beyond a reasonable doubt of the accuracy of the identification of the defendant before you may convict. If you are not convinced beyond a reasonable doubt that the defendant was the person who committed the crime, you must find the defendant not guilty.

Identification testimony is an expression of belief or impression by the witness. Its value depends on the opportunity the witness had to observe the offender at the time of the offense and to make a reliable identification later.

In appraising the identification testimony of a witness, you should consider the following:

1) Are you convinced that the witness had the capacity and an adequate opportunity to observe the defender?
Whether the witness had an adequate opportunity to observe the offender at the time of the offense will be affected by such matters as how long or short a time was available, how far or close the witness was, how good were lighting conditions, whether the witness had had occasion to see or know the person in the past.

2) Are you satisfied that the identification made by the witness subsequent to the offense was the product of his own recollection? You may take into account both the strength of the identification and the circumstances under which the identification was made.
If the identification by the witness may have been influenced by the circumstances under which the defendant was presented to him for identification, you should scrutinize the identification with great care. You may also consider the length of time that lapsed

between the occurrence of the crime and the next opportunity of the witness to see the defendant, as a factor bearing on the reliability of the identification.

3) Finally, you must consider the credibility of each identification witness in the same way as any other witness, consider whether the witness is truthful, and consider whether the witness had the capacity and opportunity to make a reliable observation of the matter covered in his testimony.

I again emphasize that the burden of proof on the prosecutor extends to every element of the crime charged, and this specifically includes the burden of proving beyond a reasonable doubt the identity of the defendant as the perpetrator of the crime with which the defendant stands charged. If after examining the testimony, you have a reasonable doubt as to the accuracy of the identification, you must find the defendant no guilt

Appendix E

Pre-trial Instruction

Expert Evidence Conditions - JCNSW, 2006; s2-1110

In addition to the testimony of the eyewitness, you will also see the testimony of [*expert's name*] who has been called as an expert witness in this case.

An expert is a person who has specialized knowledge based on their training, study or experience. Unlike other witnesses, a witness with such specialized knowledge may express an opinion on relevant matters within his or her particular area of expertise. Other witnesses may speak only as to facts, that is what they saw or hear and are not permitted to express their opinions. This expert evidence is admitted to provide you with scientific opinion which is within the witness' expertise, but which is likely to be outside the experience and knowledge of the average lay person.

The expert evidence is before you to assist with an assessment of the accuracy of the eyewitness' account of the crime. You should bear in mind that, if having given the matter careful thought, you do not accept the evidence of the expert, you do not have to act upon it.

You will see two people testify. The screen will go black for a short period between these videos, please do not stop, fast forward or rewind at any point."

Appendix F

Judicial Instructions to Jury Pre-Deliberation

The jury must reach a unanimous verdict.

In his summary given at the end of the trial, the judge indicated that if you as jurors **believe the witness's** account of the incident, and the accuracy of their identification, then the legal requirements for the crime of larceny (direct students to the list on the overhead) have been met and you are required to find the defendant guilty.

Appendix G

Jury Study Case Facts & Legal Instructions

Case Facts

1) Three hundred dollars was stolen from the victim in the theft.

2) The accused was arrested with $328 in possession for which he did not account.

3) The accused was arrested in the vicinity of the theft.

4) The accused did not have an alibi for the time of the theft.

5) At the time of his arrest the accused was wearing clothes similar to those described by the witness.

Legal Instructions

As jurors, you must be satisfied beyond a reasonable doubt that the actions of the accused meet the following criteria –

That the ACCUSED person
took and carried away
the property of another
with the intent to permanently deprive the owner of the property
and the taking is without the owner's consent

You must be satisfied of each of these criteria before you can render a verdict of GUILTY for the crime of larceny. If you can not satisfy all of these criteria beyond reasonable doubt, you must find the accused NOT GUILTY.

Appendix H

Experiment 1 – Pre- and Post-Deliberation Questionnaires (Expert)

COMPLETE THESE QUESTIONS BEFORE VIEWING FILM

A) **Student number:** **z __ __ __ __ __ __ __**

B) **Age:** __ __

C) **Gender:** Male ☐ Female ☐

D) **Are you a native English speaker?** Yes ☐ No ☐

If no, how many years have you been speaking English? _______

E) **Are you eligible to vote in Australia?** Yes ☐ No ☐

F) **What nationality are you?** ____________________________

DO NOT TURN THIS PAGE UNTIL INSTRUCTED

---

1. In *your opinion as a juror,* do you believe the witness identified the perpetrator?

Yes ☐ No ☐

2. How much *confidence* do you have in the choice you made in Q1? *(please tick one)*

☐	☐	☐	☐	☐	☐	☐
Not at all	Very little	A little	Moderate	Highly	Very high	Extremely

3. In *your opinion as a juror,* do you believe <u>beyond a reasonable doubt</u> that the person the witness
identified is guilty of theft?

Yes ☐ No ☐

4. How much *confidence* do you have in the choice you made in Q3? *(please tick one)*

☐ ☐ ☐ ☐ ☐ ☐ ☐

5. Rate the *credibility* of the witness *(please tick one)*

☐	☐	☐	☐	☐	☐	☐
Not at all	Very little	A little	Moderate	Highly	Very high	Extremely

6. Rate the *confidence* of the witness *(please tick one)*

☐	☐	☐	☐	☐	☐	☐
Not at all	Very little	A little	Moderate	Highly	Very high	Extremely

7. Rate the *attractiveness* of the witness *(please tick one)*

☐	☐	☐	☐	☐	☐	☐
Not at all	Very little	A little	Moderate	Highly	Very high	Extremely

8. How *trustworthy* was the witness? *(please tick one)*

☐ Not at all ☐ Very little ☐ A little ☐ Moderate ☐ Highly ☐ Very high ☐ Extremely

You also saw an expert give evidence. Answer the following questions in regards to the *expert*.

9. How *credible* was the expert? *(please tick one)*

☐ Not at all ☐ Very little ☐ A little ☐ Moderate ☐ Highly ☐ Very high ☐ Extremely

10. How *useful* was the testimony of the expert in evaluating the witness? *(please tick one)*

☐ Not at all ☐ Very little ☐ A little ☐ Moderate ☐ Highly ☐ Very high ☐ Extremely

11. Was the testimony of the expert *understandable/clear*? *(please tick one)*

☐ Not at all ☐ Very little ☐ A little ☐ Moderate ☐ Highly ☐ Very high ☐ Extremely

12. When deciding whether or not the *eyewitness* correctly identified the man from the lineup, how influential was;

a) The confidence of the witness?

☐ Not at all ☐ A little ☐ Moderately ☐ Substantially ☐ Greatly

b) The manner of the witness? e.g. facial expressions, tone of voice, body language, eye contact etc.

☐ Not at all ☐ A little ☐ Moderately ☐ Substantially ☐ Greatly

c) The conditions under which the identification was made? e.g. duration of view, quality of view, lighting etc.

☐ Not at all ☐ A little ☐ Moderately ☐ Substantially ☐ Greatly

13. Did you recognize this witness as a person *familiar* to you? i.e. have you ever met this person?

Yes ☐ No ☐

Student number: z __ __ __ __ __ __ __

2

1. **Did your *jury* reach an anonymous verdict?** *(please tick one)*

☐ Yes- Guilty ☐ Yes – Not Guilty ☐ No

2. **How much *confidence* do you have in the verdict of the jury?** *(please tick one)*

☐ Not at all ☐ Very little ☐ A little ☐ Moderate ☐ Highly ☐ Very high ☐ Extremely

3. **After having discussed the case, do *<u>you</u>* now believe that the witness identified the right person?**

Yes ☐ No ☐

4. **How much *confidence* do you have in the choice you made in Q3?** *(please tick one)*

☐ Not at all ☐ Very little ☐ A little ☐ Moderate ☐ Highly ☐ Very high ☐ Extremely

5. **After having discussed the case, do *<u>you</u>* now believe <u>beyond a reasonable doubt</u> that the person identified**

by the witness is guilty of theft?

Yes ☐ No ☐

6. **How much *confidence* do you have in the choice you made in Q5?** *(please tick one)*

☐ Not at all ☐ Very little ☐ A little ☐ Moderate ☐ Highly ☐ Very high ☐ Extremely

7. **Rate the *credibility* of the witness** *(please tick one)*

☐	☐	☐	☐	☐	☐	☐
Not at all	Very little	A little	Moderate	Highly	Very high	Extremely

8. Rate the *confidence* of the witness *(please tick one)*

☐	☐	☐	☐	☐	☐	☐
Not at all	Very little	A little	Moderate	Highly	Very high	Extremely

9. Rate the *attractiveness* of the witness *(please tick one)*

☐	☐	☐	☐	☐	☐	☐
Not at all	Very little	A little	Moderate	Highly	Very high	Extremely

10. How *trustworthy* was the witness? *(please tick one)*

☐	☐	☐	☐	☐	☐	☐
Not at all	Very little	A little	Moderate	Highly	Very high	Extremely

You also saw an expert give evidence. Answer the following questions in regards to the *expert*.

11. How *credible* was the expert? *(please tick one)*

☐	☐	☐	☐	☐	☐	☐
Not at all	Very little	A little	Moderate	Highly	Very high	Extremely

12. How *useful* was the testimony of the expert in evaluating the witness? *(please tick one)*

☐	☐	☐	☐	☐	☐	☐
Not at all	Very little	A little	Moderate	Highly	Very high	Extremely

13. Was the testimony of the expert *understandable/clear*? *(please tick one)*

☐ Not at all ☐ Very little ☐ A little ☐ Moderate ☐ Highly ☐ Very high ☐ Extremely

14. What were the *main points* covered in the experts testimony? *(describe briefly below)*

1.

2.

3.

4.

5.

15. When deciding whether or not the *eyewitness* correctly identified the man from the lineup, how influential was;

a) The confidence of the witness?

☐ Not at all ☐ A little ☐ Moderately ☐ Substantially ☐ Greatly

b) The manner of the witness? e.g. facial expressions, tone of voice, body language, eye contact etc.

☐ Not at all ☐ A little ☐ Moderately ☐ Substantially ☐ Greatly

c) The conditions under which the identification was made? e.g. duration of view, quality of view, lighting etc.

☐ Not at all ☐ A little ☐ Moderately ☐ Substantially ☐ Greatly

16. Complete the following four statements so they *accurately* reflect what the expert said

Select option 1, 2, 3 or 4 below to most accurately complete the statement:

a) regarding eyewitness memory

1. Eyewitness' "perform at a chance level of accuracy".
2. It is "more error prone than people would normally believe".
3. It is "a fallible subjective process"
4. It is "less error prone than people would normally believe".

Select option 1, 2, 3 or 4 below to most accurately complete the statement:

b) regarding the confidence accuracy relationship

1. "confidence is a good predictor of accuracy".
2. "nervous uncertain witness' can be right".
3. "confident witness' are generally right".
4. "confidence isn't necessarily a predictor of accuracy".

Select option 1, 2, 3 or 4 below to most accurately complete the statement:

c) regarding human memory

1. “we remember our own interpretation which is likely to include errors and omissions”.
2. “it is highly reliable”.
3. “it is like a video recorder”.
4. “is an interactive and reconstructive process”.

Select option 1, 2, 3 or 4 below to most accurately complete the statement:

d) regarding the role of an expert in court

1. “I can tell you about things that are generally true of eyewitness’”.
2. “I cant tell you whether this witness was accurate”.
3. “I can tell you whether this witness was accurate”.
4. “My research makes clear which witnesses are accurate and which witnesses are not”

THIS IS THE END OF THE STUDY – PLEASE TELL THE EXPERIMENTER THAT YOU HAVE FINISHED.

THANKYOU

Appendix J

Experiments 2 and 3 Questionnaire (Expert)

COMPLETE THESE QUESTIONS BEFORE VIEWING FILM

A) **Do you agree to your data from today's tasks being used anonymously in educational research?**

Yes ☐ No ☐

B) **Video Code** *(please write the code/file name shown the start of the video here)*: - __ __ - 2 __

C) **Tutorial Day/Time:** ________________

D) **Student number:** z __ __ __ __ __ __ __

E) **Age:** __ __

F) **Gender:** Male ☐ Female ☐

G) **Are you a native English speaker?** Yes ☐ No ☐

If no, how many years have you been speaking English? _______

H) **Are you eligible to vote in Australia?** Yes ☐ No ☐

I) **What nationality are you?** ____________________________

1. In *your opinion as a juror,* was the person the witness chose from the lineup the person who committed the robbery? *(please tick one)*

Yes ☐ No ☐

2. How much *confidence* do you have in the choice you made in Q1? *(please tick one)*

☐	☐	☐	☐	☐	☐	☐
Not at all	Very little	A little	Moderate	Highly	Very high	Extremely

3. Did you recognize this witness as a person *familiar* to you? i.e. have you ever met this person *(please tick one)* Yes ☐ No ☐

4. Rate the *credibility* of the witness *(please tick one)*

☐	☐	☐	☐	☐	☐	☐
Not at all	Very little	A little	Moderate	Highly	Very high	Extremely

5. Rate the *accuracy* of account given by the witness *(please tick one)*

☐	☐	☐	☐	☐	☐	☐
Not at all	Very little	A little	Moderate	Highly	Very high	Extremely

6. Rate the *confidence* of the witness *(please tick one)*

☐	☐	☐	☐	☐	☐	☐
Not at all	Very little	A little	Moderate	Highly	Very high	Extremely

7. Rate the *attractiveness* of the witness *(please tick one)*

☐	☐	☐	☐	☐	☐	☐
Not at all	Very little	A little	Moderate	Highly	Very high	Extremely

8. How *likeable* was the witness? *(please tick one)*

☐	☐	☐	☐	☐	☐	☐
Not at all	Very little	A little	Moderate	Highly	Very high	Extremely

9. How *trustworthy* was the witness? *(please tick one)*

☐	☐	☐	☐	☐	☐	☐
Not at all	Very little	A little	Moderate	Highly	Very high	Extremely

10. How *angry* was the witness? *(please tick one)*

☐	☐	☐	☐	☐	☐	☐
Not at all	Very little	A little	Moderate	Highly	Very high	Extremely

You also saw an expert give evidence. Answer the following questions in regards to the *expert*.

11. How *credible* was the expert? *(please tick one)*

☐	☐	☐	☐	☐	☐	☐
Not at all	Very little	A little	Moderate	Highly	Very high	Extremely

12. How *useful* was the testimony of the expert in evaluating the witness? *(please tick one)*

☐	☐	☐	☐	☐	☐	☐
Not at all	Very little	A little	Moderate	Highly	Very high	Extremely

13. Was the testimony of the expert *understandable/clear*? *(please tick one)*

☐	☐	☐	☐	☐	☐	☐
Not at all	Very little	A little	Moderate	Highly	Very high	Extremely

14. What were the *main points* covered in the experts testimony? *(describe briefly below)*

1.

2.

3.

4.

5.

If you need more space please continue on a separate page.

15. When deciding whether or not the *eyewitness* correctly identified the man from the lineup, how influential was;

a) The confidence of the witness?

☐ Not at all ☐ A little ☐ Moderately ☐ Substantially ☐ Greatly

b) The manner of the witness? e.g. facial expressions, tone of voice, body language, eye contact etc.

☐ Not at all ☐ A little ☐ Moderately ☐ Substantially ☐ Greatly

c) The conditions under which the identification was made? e.g. duration of view, quality of view, lighting etc.

☐ Not at all ☐ A little ☐ Moderately ☐ Substantially ☐ Greatly

16. Which of the following statements (one only) which most accurately reflects what the expert said

(circle one)

a) regarding eyewitness memory.

1. Eyewitness' "perform at a chance level of accuracy".
2. It is "more error prone than people would normally believe".
3. It is "a fallible subjective process"
4. It is "less error prone than people would normally believe".

b) regarding the confidence accuracy relationship

1. "confidence is a good predictor of accuracy".

2. "nervous uncertain witness' can be right".

3. "confident witness' are generally right".

4. "confidence isn't necessarily a predictor of accuracy".

c) regarding human memory

1. "we remember our own interpretation which is likely to include errors and omissions".

2. "it is highly reliable".

3. "it is like a video recorder".

4. "is an interactive and reconstructive process".

d) regarding the role of an expert in court

1. "I can tell you about things that are generally true of eyewitness'".

2. "I cant tell you whether this witness was accurate".

3. "I can tell you whether this witness was accurate".

4. "My research makes clear which witnesses are accurate and which witnesses are not"

THERE ARE NO FURTHER QUESTIONS – THANK YOU

Appendix K

Experiment 4 : Cued Recall Questions (Expert)

For each statement in *italics*, which of the options below it most accurately reflects what the expert said

(circle one option(1-4) for each statement in italics (A-D))

A) regarding eyewitness identification

1. It will be "both accurate and inaccurate".
2. It may be "accurate or inaccurate".
3. It is "always accurate"
4. It is "always inaccurate".

B) regarding the confidence accuracy relationship

1. "confidence is a good predictor of accuracy".
2. "confidence is a poor predictor of accuracy".
3. "confident witness' are generally right".
4. "confident witness' are generally mistaken".

C) regarding human memory

1. "contains accurate information, inaccurate information and a mixture of those two things".
2. "it is highly reliable".
3. "it is like a video recorder".
4. "is an interactive and reconstructive process".

D) regarding the role of an expert in court

1. "I can comment on the processes of memory".
2. "I can't tell you whether this witness was accurate".
3. "I can tell you whether this witness was accurate".
4. "My research makes clear which witnesses are accurate and which witnesses are not"

Thank you for your interest in this study.

As a participant in this study you will be asked to watch three videos and answer some questions about yourself and what you have seen. The entire process will take approximately 40 minutes of your time.

I greatly appreciate your willingness to help with my research and would be happy to send you a report of the results if you are interested in the outcome.

PLEASE NOTE This form is designed to be emailed back to me once it is completed. If you have any concerns about your email address being associated with your responses, please follow the postal instructions at the bottom of this form to preserve your anonymity.

Please continue to the next page to view the statement of research ethics and begin the eyewitness evaluation task.

9 June 2006

Kristy Martire
PhD Candidate
School of Psychology, UNSW
Sydney, NSW, 2052
Australia

UNSW
School of Psychology

STATEMENT OF RESEARCH ETHICS

You are invited to participate in a study of juror evaluations of eyewitness testimony. We hope to learn more about the cues used by eyewitness evaluators to determine eyewitness accuracy. You were selected to participate in this study because you were identified to have expertise in the field of eyewitness identification issues.

Any information that is obtained in connection with this study and that can be identified with you will remain confidential and will be disclosed only with your permission except as required by law. We plan to publish the results of this research in academic journals. In any publication, information will be provided in such a way that you cannot be identified.

Your decision whether or not to participate will not prejudice your future relations with The University of New South Wales. If you decide to participate, you are free to withdraw your consent and to discontinue participation at any time without prejudice.

If you have any questions regarding this research, please contact the researcher, Kristy Martire, School of Psychology, University of New South Wales, Sydney 2052 AUSTRALIA (ph: +61 2 9385 3049, email kmartire@psy.unsw.edu.au).

Any complaints about this research may be directed to the Ethics Secretariat, The University of New South Wales, SYDNEY 2052 AUSTRALIA (ph: +61 2 9385 4234, fax +61 2 9385 6648, email ethics.sec@unsw.edu.au). Approval Number: 534.

INSTRUCTIONS

As a participant in this study you will be directed to watch some short video clips ("Witness.wmv" files on this CD) and then answer some questions about what you have seen. You will be asked to watch the "Witness.wmv" files in a specific order, answering the questions relating to each before you view the next.

You will require a version of Windows Media Player (preferably version 9) to be able to view these files. If you do not have WMP on your computer the installation files for both MAC and PC operating systems are also enclosed on this CD. Please double click on the appropriate icon on the CD to install, or go directly to http://www.microsoft.com/windows/windowsmedia/download/alldownloads.aspx **(copy and paste the address into your internet browser) to download directly from the microsoft website.**

After you have watched all three "Witness.wmv's" you will be asked some questions about yourself. PLEASE NOTE; It is expected that you will answer all of the questions that apply to you, but we do not anticipate that every participant will have an answer for all questions in the "About me" section.

Once you have completed the questionnaire you will be directed to email or post a copy of your responses. PLEASE NOTE; a completed version of this questionnaire CAN NOT BE SAVED so please put aside approximately 40mins to complete this study in one session.

PLEASE CONTINUE ON TO THE NEXT PAGE WHEN YOU ARE READY TO BEGIN

Please read this when you are ready to begin

What you are about to see is the testimony of three actual witnesses.

First, these witnesses saw a crime being committed and were then asked to identify the person they saw from a lineup. Each of these witnesses identified the suspect from the lineup, and as a result were asked to testify at the trial of the accused. Each of these witnesses viewed a different crime committed by a different perpetrator.

What follows is a recording of the testimony given by the witness. The footage takes the form of an examination in chief and a cross examination.

The examination in chief is conducted by the prosecution and allows the witness to provide their account of the theft and the identification they made. The cross examination of the witness is conducted by those defending the accused and seeks to challenge the witness's version of the events.

I would like you to watch these interviews as though you were a juror in the trial of the man accused by each witness. As such, you are to watch this footage keeping in mind that the role of a juror in a trial such as this is to determine the guilt or innocence of the accused based on the evidence presented.

It is important to note then, that the evidence given by the witness you are about to see is the only evidence relating to the guilt or innocence of the accused at this trial. Accordingly, it is your concern as a juror, **knowing what you know about eyewitness identification issues**, to assess not only what the witness says, but also the witness's honesty and reliability. You should examine and scrutinize the testimony of this witness with great care before you decide whether to accept their evidence.

To watch the first video, double click the CD file: "Witness1.wmv" NOW

Answer these questions only AFTER watching "Witness1.wmv"

WITNESS 1

1. In *your opinion as a juror*, do you believe the witness selected the thief from the lineup? Yes No

2. How much *confidence* do you have in the decision you made in Question 1?

Not at all | Very little | A little | Moderate | Highly | Very high | Extremely

3. Rate the *credibility* of the witness

None | Very little | A little | Moderate | High | Very high | Extremely high

4. Rate the *confidence* of the witness

None | Very little | A little | Moderate | High | Very high | Extremely high

5. Rate the *attractiveness* of the witness

None | Very little | A little | Moderate | High | Very high | Extremely high

6. Rate the *trustworthiness* of the witness

Not at all | Very little | A little | Moderate | Highly | Very high | Extremely

7. When deciding whether or not the eyewitness correctly identified the thief in the lineup how influential was each of the following factors?

a. Witness' confidence?

Not at all | A little | Moderately | Substantially | Greatly

b. Witness' manner?

Not at all | A little | Moderately | Substantially | Greatly

c. Witnessing conditions?

Not at all | A little | Moderately | Substantially | Greatly

d. Identification conditions?

Not at all | A little | Moderately | Substantially | Greatly

To watch the next video, double click the CD file: "Witness2.wmv" NOW

Answer these questions only AFTER watching "Witness2.wmv"

WITNESS 2

8. In *your opinion as a juror*, do you believe the witness selected the thief from the lineup? ◯ Yes ◯ No

9. How much *confidence* do you have in the decision you made in Question 8?

◯ Not at all ◯ Very little ◯ A little ◯ Moderate ◯ Highly ◯ Very high ◯ Extremely

10. Rate the *credibility* of the witness

◯ None ◯ Very little ◯ A little ◯ Moderate ◯ High ◯ Very high ◯ Extremely high

11. Rate the *confidence* of the witness

◯ None ◯ Very little ◯ A little ◯ Moderate ◯ High ◯ Very high ◯ Extremely high

12. Rate the *attractiveness* of the witness

◯ None ◯ Very little ◯ A little ◯ Moderate ◯ High ◯ Very high ◯ Extremely high

13. Rate the *trustworthiness* of the witness

◯ Not at all ◯ Very little ◯ A little ◯ Moderate ◯ Highly ◯ Very high ◯ Extremely

14. When deciding whether or not the eyewitness correctly identified the thief in the lineup how influential was each of the following factors;

a. Witness' confidence?

◯ Not at all ◯ A little ◯ Moderately ◯ Substantially ◯ Greatly

b. Witness' manner?

◯ Not at all ◯ A little ◯ Moderately ◯ Substantially ◯ Greatly

c. Witnessing conditions?

◯ Not at all ◯ A little ◯ Moderately ◯ Substantially ◯ Greatly

d. Identification conditions?

◯ Not at all ◯ A little ◯ Moderately ◯ Substantially ◯ Greatly

To watch the next video, double click the CD file: "Witness3.wmv" NOW

Answer these questions only AFTER watching "Witness3.wmv"

WITNESS 3

15. In *your opinion as a juror*, do you believe the witness selected the thief from the lineup? ◯ Yes ◯ No

---

16. How much *confidence* do you have in the decision you made in Question 15?

◯ Not at all ◯ Very little ◯ A little ◯ Moderate ◯ Highly ◯ Very high ◯ Extremely

---

17. Rate the *credibility* of the witness

◯ None ◯ Very little ◯ A little ◯ Moderate ◯ High ◯ Very high ◯ Extremely high

---

18. Rate the *confidence* of the witness

◯ None ◯ Very little ◯ A little ◯ Moderate ◯ High ◯ Very high ◯ Extremely high

---

19. Rate the *attractiveness* of the witness

◯ None ◯ Very little ◯ A little ◯ Moderate ◯ High ◯ Very high ◯ Extremely high

---

20. Rate the *trustworthiness* of the witness

◯ Not at all ◯ Very little ◯ A little ◯ Moderate ◯ Highly ◯ Very high ◯ Extremely

---

21. When deciding whether or not the eyewitness correctly identified the thief in the lineup how influential was each of the following factors;

a. Witness' confidence?

◯ Not at all ◯ A little ◯ Moderately ◯ Substantially ◯ Greatly

b. Witness' manner?

◯ Not at all ◯ A little ◯ Moderately ◯ Substantially ◯ Greatly

c. Witnessing conditions?

◯ Not at all ◯ A little ◯ Moderately ◯ Substantially ◯ Greatly

d. Identification conditions?

◯ Not at all ◯ A little ◯ Moderately ◯ Substantially ◯ Greatly

Please continue to the next page

ABOUT ME

22a. Age: Years [ ]

b. Gender: ◯ M ◯ F

c. What is your country of residence? type here: [ ]

---

23. How would you describe yourself?

◯ Practicioner ◯ Practicioner/Academic ◯ Academic

Other [ ]

---

24. Which of the following best describes your employer?

◯ University or other research institution ◯ Government forensic institution ◯ Private company ◯ Self employed

Other [ ]

---

25. How would you describe your training, qualifications and experience?

Please describe here :

eg Bachelor of Science degree in subject ***x***, 5 years experience undertaking ***y***

---

26. Have you ever been asked to prepare a report for court relating to eyewitness identification issues?

◯ Yes ◯ No

---

27. Have you ever been asked to testify in court as an expert on matters relating to eyewitness identification issues?

◯ Yes ◯ No

---

28. Within which of the following geographical areas have you been involved (i.e. were consulted, asked to prepare a report or testify) in eyewitness identification cases?

☐ U.K. & Ireland ☐ Canada ☐ Australasia

☐ Mainland Europe ☐ USA ☐ None

Other [ ]

29. Considering the 5 year period from Jan 2001 to Dec 2005, please indicate the number of times you have been:

a. Contacted/asked to provide **advice** relating to eyewitness identification issues for a case ☐

b. Instructed to prepare a **report** addressing eyewitness identification issues for a court or similar body ☐

c. Required to **give evidence** in a court or other body regarding eywitness identification issues ☐

If you answered **0** to **all three parts** of Question 29, please go straight to **Question 31.**

---

30. Considering all those occasions in the period 2001-2005 in which you have either prepared a written report and/or sttended court (or similar body) to give evidence, on what percentage of occasions has the court rejected or refused to consider your evidence? **%** ☐

a. Where a court has rejected your evidence, please select the grounds for exclusion:

☐ **Relevancy** - Scientific findings ruled to be not directly pertinent
(Daubert v. Merrell Dow Pharmaceuticals, Inc., 1993)

☐ **Legal sufficiency -** Prejudicial value said to outweigh the probative value
(Daubert)

☐ **Reliability** - the scientific research was deemed untestable, the error rate too high/unknown, absence of peer review or an absence of general accptance.
(Daubert)

☐ Failure of the field to develop generally accepted scientific techniques
(Frye v. U.S. 1923)

Other

Please describe: ☐

---

31. If asked, would you testfy to support the valiidity of the following statements?

	Would testify	Would not testify
The potential for the wording of questions put to a witness to affect their testimony	○	○
The confidence of an eyewitness can be influenced by factors that are unrelated to identification accuracy	○	○
Instructions can influence an eyewitness's willingness to make an identification	○	○
An eyewitness's confidence is not a good predictor of their identification accuracy	○	○
The presence of a weapon negatively impacts eyewitness identification accuracy	○	○

(Based on Kassin, Hosch & Memon, 2001)

32. Please list your primary areas of research below:

Area 1

Area 2

Area 3

---

33. If any, how many times have you published on matters relating to eyewitness identification issues?

Number of published peer reviewed articles (approx)

Number of **other** publications (approx)

---

34. If any, how many times have you published on the issue of eyewitness expert testimony?

Number of published peer reviewed articles (approx)

Number of **other** publications (approx)

If you know of anyone else who might be interested in completing this study, please enter their postal address below so we can send them a copy of this CD.

Name & Address:

Do you have any questions for, or comments/feedback you would like to pass on to the experimenter?

Please enter here:

Would you like to be sent information regarding the results of this study ?

If **YES**, please enter your email address here:

THANK YOU FOR PARTICIPATING IN THIS STUDY

PLEASE CLICK THE "SUBMIT BY EMAIL" BUTTON BELOW TO SUBMIT YOUR COMPLETED QUESTIONNAIRE ELECTRONICALLY

Submit by Email

PLEASE CLICK THE "PRINT FORM" BUTTON BELOW TO POST A COPY OF YOUR COMPLETED QUESTIONNAIRE

Print Form

POST TO: Kristy Martire
School of Psychology, UNSW
Sydney, NSW, 2052 AUSTRALIA