

Adversarial Forensic Science Experts: An Empirical Study of Jury Deliberation

Jane Goodman-Delahunty and Kosuke Wakabayashi*

Abstract

Surprisingly little is known about ways that juries resolve differences of opinion between competing scientific forensic experts. Concerns have been raised that juries defer unduly to scientific experts and are susceptible to the ‘white coat effect’. The study reported in this article examined jury decision-making in the context of a live, simulated homicide trial that incorporated traditional legal procedural safeguards against jury error: cross-examination, the use of rebuttal expert witnesses, judicial directions and group deliberation. Following the presentation of opposing expert opinions on analyses of trace evidence, 12 juries deliberated to a verdict. Deliberation transcripts were systematically analysed to discern prominent topics discussed in the jury room. Using a text-mining method, shifts in deliberation focus were explored in response to two interventions: (a) a judicial caution about the non-binding nature of expert illustrative visual aids; and (b) uneven versus evenly balanced rebuttal expert evidence. Results indicated that juries were not deferential to the experts, but did not ignore them either. They found the experts’ sophisticated interactive visual aids useful, but appropriately discerned that the scientific evidence did not resolve the issues before them. No evidence emerged for ‘the white coat effect’. Excerpts from deliberations substantiated the quantitative outcomes.

Introduction

Typically, in a criminal case involving forensic scientific evidence, an expert appears for the prosecution, and the defence relies on cross-examination to challenge that evidence. Although scientific experts often disagree on what constitutes scientifically reliable and valid knowledge (Römkens 2000), or on inferences drawn from data produced by robust methods (Goodman-Delahunty et al 2012:132), only infrequently does the defence call a rebuttal expert. In single expert cases that featured more hostile and confrontational disputes between the prosecution and defence about DNA profiling evidence, juror comprehension appeared to be reduced (Findlay 2008), raising questions about the extent to which exposure to a battle of the experts facilitates or impedes jury comprehension of forensic scientific evidence. When a rebuttal expert testifies, commentators have compared the resulting duel

* Professor Jane Goodman-Delahunty, Charles Sturt University (Manly Campus), Australian Graduate School of Policing and Security and School of Psychology, email: jdelahunty@csu.edu.au; Kosuke Wakabayashi, PhD candidate, Ritsumeikan University, Ritsumeikan Global Innovation Research Organization, Department of Psychology. With thanks to Thea Gumbert, who assisted with thematic coding.

of the experts to a medieval joust (Wuffle 1985). Innovations to manage opposing experts, such as the presentation of concurrent expert evidence, nonetheless leave the task of resolving differences between experts to the trier of fact (Edmond 2009).

Numerous previous studies have examined jury responses to a single expert. However, questions about the impact of opposing expert witnesses have rarely been the focus of empirical research, and surprisingly few studies have investigated ways that triers of fact resolve differences of opinion between well-credentialed, opposing expert witnesses. The collaborative, interdisciplinary study discussed in this article addressed this topic by analysing jury deliberations following exposure to live, controverted expert witness evidence in a criminal trial. Using an innovative text-mining methodology,¹ the aim of the study was to analyse the content of jury deliberations in response to well-qualified, adversarial forensic scientific experts. If the expert evidence presented at trial is influential in determining a verdict, we expect jurors will discuss this evidence in the course of their deliberations. In line with the Elaboration Likelihood Model of persuasion, which asserts that engagement with the central content of a message entails elaborative cognitive processing of that message (Petty and Cacioppo 1984), in the present study, more extensive cognitive engagement with the trial content was expected to elicit a higher frequency of words associated with that material during jury deliberations.

Prior research on juror responses to a single forensic scientific expert

Anecdotally, some judges have reported that juries faced with a forensic expert ‘become bored or confused, inevitably leading them to defer mechanistically to the opinion of an expert or even relinquish their decision-making power’ (McClellan 2011:17). The theory that jurors suspend disbelief in response to expert credentials suggests that jurors faced with complex evidence tend to rely on cognitive shortcuts or ‘peripheral cues’ in lieu of ‘central processing’ of the content (Evans 2008). Reliable quantitative methods have revealed that after hearing forensic scientific evidence in real and in simulated trials, individual ‘jurors comprehend and engage deeply with the expert evidence’ (Diamond and Casper 1992:558; Hans 2008). Whether the group deliberation process increases or diminishes jury reliance on central processing of scientific expert evidence is less clear.

Research on jury exposure to adversarial experts

Concerns have been raised that opposing experts will neutralise each other, leading jurors to ignore the content of their evidence, and that aspects unrelated to the reliability of the specialised knowledge will influence jury decisions: ‘When there are competent experts on both sides, and they offer contradictory or confusing opinions, jurors may resolve the differences by relying on general impressions of character and veracity’ (Devine et al 2001:624). Others contend that status differences between institutionalised and less-established scientific disciplines, and gender bias, will have more influence (Römkens 2000). Most research on adversarial experts has addressed jury responses to experts on eyewitness memory. These studies yielded two types of effects: in some studies, the rebuttal

¹ Text mining was first developed to cluster similar data items from different linguistic sources according to their meaning to apply data mining methods. Text-mining methods have been widely and successfully applied in the field of marketing to assess consumer preferences. Within computer sciences using digital archiving methods, visualisation techniques were developed to allow users to see, explore, and understand large amounts of information at once.

expert impacted the credibility of other witnesses; and in others, the rebuttal expert impacted the conviction rate.

In one simulated criminal trial, videotaped evidence presented to mock-jurors ($N = 497$) contained one of three experimental manipulations: (a) no expert testimony; (b) a single defence expert; or (c) that expert plus a rebuttal prosecution expert (Devenport and Cutler 2004). Exposure to an opposing expert diminished the credibility of the defence expert in the eyes of the mock-jurors, but no significant differences emerged in the conviction rate. A subsequent study tested whether the presence of an opposing expert prompted mock-jurors to scrutinise the scientific validity of the expert evidence more thoroughly. Using a written trial transcript, the methodological quality of testimony by an exculpatory defence expert on eyewitness memory was systematically varied, as was the type of opposing prosecution expert evidence: (a) critical of the first expert's method; or (b) critical of the first expert's conclusions (Levett and Kovera 2008). The presence of the rebuttal expert did not sensitise jurors to flaws in the defence expert's testimony, but generated scepticism about all expert evidence, resulting in unanticipated increases in the conviction rates: significantly more guilty verdicts followed exposure to two adversarial experts than to a single expert, regardless whether the opposing testimony questioned the methodology applied by the original expert or the validity of that expert's opinion. Similarly, in a trial simulation study in which mock-jurors were exposed to (a) no expert; (b) a single defence eyewitness expert; or (c) defence and prosecution rebuttal expert (Pezdek, Avila-Mora and Sperry 2010), participants were influenced by the presentation of expert testimony, but not differentially between the presentation of one, compared with two, experts. Specifically, participants were less certain of the culpability of the defendant after the defence expert testimony was presented, but there was no significant difference in culpability decisions between participants exposed to a single defence expert and participants exposed to both a defence and prosecution rebuttal expert. These results replicated those of Devenport and Cutler (2004).

The foregoing trial simulation studies yielded significant insights into individual jurors' responses to expert witnesses, but were uninformative about the decision-making processes of the jury as a group. Whether the same findings persist after jury deliberation is largely untested, as few experimental studies have explored the influence of group deliberation on assessments of adversarial forensic experts. A rare exception comparing pre-deliberation and post-deliberation individual juror decisions revealed that following deliberation, mock-juries exposed to a rebuttal defence expert in a civil case viewed the plaintiff as less credible and returned significantly fewer pro-plaintiff verdicts (Buck and Warren 2010). In line with this finding, a comparison of individual jurors' verdicts before and after deliberation following exposure to the rebuttal expert in a criminal trial yielded an increase in the rate of acquittals of six percentage points (Tait 2011). The current study examined the decision-making processes of the jury as a group in response to adversarial forensic scientific experts.

Significance of research on deliberation about expert evidence

Jury deliberation is a key legal safeguard intended to protect the verdict against error (Edmond 2011; Goodman-Delahunty et al 2012). To assess the influence of expert evidence on jury decision-making, ideally one should observe real deliberating juries. However, investigations into jury deliberations in real trials are constrained by law. In New South Wales (NSW), soliciting information about deliberations from a juror or former juror is prohibited unless the research is conducted with the permission of the Attorney-General

(*Jury Act 1977* (NSW), s 68(1) and (3)). Parallel legislation exists in other Australian states and territories, in Canada, New Zealand and the United Kingdom (Vidmar 2005).

Some Australian jurisdictions have permitted studies of forensic expert evidence by means of post-deliberation interviews with jurors (Findlay 2008; Fordham 2009; Verrender and Goodman-Delahunty 2011; Wheate 2006). An acknowledged weakness of reliance on post-trial surveys and interviews to study jury decisions about forensic scientific experts is that what jurors report is not necessarily a reliable indicator of their performance (Goodman-Delahunty and Hewson 2010). Realistic trial simulations followed by videotaped mock jury deliberations can provide more precise measures and insights into jury performance.

For example, confirmation that jury deliberation reduced errors about scientific evidence was gathered in a realistic jury simulation by administering objective, multiple choice tests to non-empanelled jurors who watched a video-trial at court and deliberated to a verdict. Comparisons of their pre- and post-deliberation scores revealed that comprehension of forensic scientific evidence presented by a single expert increased significantly following group deliberation (Hans et al 2011). A study in which a single court-appointed or partisan expert testified disclosed that deliberating jurors were more resistant to the content of testimony by a non-adversarial than a partisan expert, and that deliberation served to increase juror scrutiny of the court-appointed expert evidence (Brekke et al 1991). In sum, studies of deliberation following jury exposure to a single expert indicated that jurors were attentive to the content and source of the expert evidence. However, these studies did not address jury responses to adversarial scientific experts, leaving a gap in the literature on that issue. The current study employed a realistic simulation methodology to analyse the content of jury deliberations about adversarial expert testimony more directly.

In the current study, we hypothesised that following the presentation of prosecution expert visual evidence, in the absence of judicial warnings or countervailing visual evidence, circumstances favouring the prosecution would exist, yielding more peripheral processing and less deliberation content about the expert testimony. Judicial warnings prior to the prosecution's expert visual evidence, however, were expected to prompt central processing by jurors to scrutinise this evidence, leading in turn to increased deliberation content about the expert evidence. The presentation of more balanced opposing expert testimony in conjunction with repeated judicial warnings, on the other hand, was expected to result in significantly diminished deliberation content pertaining to the experts and their computer simulations and more elaboration reflecting more central processing of other evidence.

The current study

Using a text-mining approach, the current study examined jury deliberation content following a live simulated criminal trial in which 12 juries responded to a battle between two forensic scientific experts. This study complemented a research report summarising the influence on verdict of individual juror predispositions and demographic characteristics assessed by means of individual pre-deliberation and post-deliberation questionnaires (Tait 2011) in which the unit of analysis was the juror. In the current study, the unit of analysis was the jury.

Theoretically, the more centrally a topic is processed, the more extensively that topic will be discussed, yielding a higher frequency of words related to this topic in transcripts of jury

deliberations. Thus, examination of deliberation transcripts and the frequency of words related to expert and other evidence can provide insight into the extent to which jurors scrutinise or ignore the expert evidence, or defer to the scientifically trained experts, a phenomenon referred to as ‘the white coat effect’ (Vidmar 2005).

Numerous jury simulations have focused primarily on jury deliberation effects (Devine et al 2001), but few studies have analysed the content of group deliberations using a text-mining approach (Hotta and Fujita 2007), and very few studies have assessed deliberations following a live presentation of evidence (cf American Bar Association 1989). Researchers rarely conduct live simulated trials because they are logistically difficult, time-consuming and expensive (Bornstein and McCabe 2005). Funding to conduct a live simulated criminal trial involving adversarial forensic experts was obtained from the Australian Research Council (LP0667764; Goodman-Delahunty, Rossner and Tait 2011).

Method

Materials

A trial script was prepared incorporating opening statements, witness examinations, closing statements, summation and directions to the jury (Goodman-Delahunty, Rossner and Tait 2011). The accused faced five counts of murder for knowingly placing an explosive device on a suburban train. Bomb-making instructions and 500 grams of plastic explosive were found in his home. The accused admitted leaving his sports bag on the train, but testified that he saw a cardboard box under the seat adjacent to his sports bag.

Forensic scientific evidence about the source of the bomb explosion was presented by two opposing experts. Dr Green, the prosecution witness (in real life, a scientist employed by the Australian Federal Police), explained that traces of bomb residue on the defendant’s sports bag linked the defendant to the bomb. The expert’s oral evidence was supplemented by a computer animation showing a man entering a train carriage, leaving a sports bag under the seat from which an explosion originated, scattering debris, explosives residue, and injuring or killing the passengers. The defence expert, Dr Lange, identified as a forensic scientist with explosives expertise working in private practice (in real life a Professor of Forensic Science), posited that the source of the explosion was a bomb planted in a cardboard box adjacent to the defendant’s sports bag. He testified that the explosive residue and damage originating from a bomb located in a sports bag or the box would be virtually identical. The defence expert disputed the conclusions, but not the methodology of the prosecution expert.

The verbal content of the evidence by the two forensic scientists was invariant in all three versions of the enacted trial. In the first, only the prosecution expert used illustrative visual evidence (‘PIVE’) in the form of a computer animated reconstruction. In a second enactment, to balance the technological duel between the experts, the defence scientist used similar illustrative visual evidence (‘DIVE’). The study also tested the effectiveness of a judicial caution about the weight of the expert computer simulation (Appendix A). In trial PIVE, the judge administered no caution. In trial ‘JI’, a jury instruction in the form of a judicial caution about the weight of the interactive visual evidence (‘IVE’) was presented prior to the testimony of the prosecution expert and was repeated in the judicial summation. In the DIVE trial, three judicial instructions about the weight of the IVE were provided: one prior to the prosecution IVE, the second prior to the defence IVE, and the third caution in the summation.

Participants

Mock-jurors were 144 Australian jury-eligible volunteers who responded to advertisements placed in local newspapers in Sydney, achieving a mix of age ranges, occupational interests, and educational backgrounds similar to that of actual juries (Tait 2011).

Procedures

The trial was performed live three times in succession in the King Street Court in Sydney, NSW. Each enactment was presided over by a real judge and attended by four juries who watched the trial simultaneously, yielding a total of 12 juries who deliberated to a verdict. Deliberations were conducted in real jury rooms and recorded. To ensure that each jury commented on the expert evidence within the available time, deliberation facilitators (members of the research team) provided four key questions to the jury at the outset of deliberation:

- 1) What do you see as the main issues in the case?
- 2) How did you respond to the expert witnesses?
- 3) How did you respond to the visual and verbal evidence?
- 4) How did you respond to the instructions from the judge?

The facilitators did not participate in the discussion and expressed no views on the evidence. Deliberations lasted an average of 92 minutes. Transcripts of the 12 jury deliberations comprised the data reviewed in this study.

Hypotheses

In the PIVE trial, in the absence of any judicial warnings or rebuttal illustrative visual evidence from the defence expert, jurors had few cues to scrutinise and critique the prosecutorial IVE, thus conditions were optimal to elicit a 'white coat effect' favouring the prosecution. The prosecution evidence was, therefore, expected to be more salient and uncritically accepted by those four juries. PIVE deliberations were expected to yield a lower frequency of words relating to the computer simulation (expert visual evidence) compared to deliberations by juries exposed to the JI and DIVE trials (Hypothesis One).

In the JI trial in which the prosecution IVE was preceded and followed by a judicial warning and the defence expert did not present countervailing IVE, juries were expected to scrutinise the prosecution IVE more critically and discuss the prosecution IVE more extensively, resulting in a higher frequency of words in JI deliberation transcripts about the judicial directions and the expert evidence, compared to the frequency of these words in the PIVE and DIVE deliberations (Hypothesis Two).

In the DIVE trial in which four juries were exposed to balanced prosecution and defence IVE and three judicial warnings about its weight, we expected the prominence or salience of both experts' visual evidence to diminish in deliberations, resulting in a lower frequency of words related to the forensic experts themselves and their visual presentations, and conversely, a higher frequency of words related to the other evidence (Hypothesis Three).

Preparation for data analysis

Transcripts of the 12 deliberation sessions were analysed using Tiny Text Miner (Matsumura and Miura 2012), a free text-mining software application for English and Japanese languages. The transcripts were prepared for analysis as follows: first, synonyms within each deliberation were identified and replaced with a single word to reduce the number of word categories and enable more accurate results. Plural nouns were replaced by singular nouns so they would be counted as same word by the software, for example, 'expert' and 'experts' were replaced by 'expert'. After this preliminary work, the software counted word frequencies generated by jury members within each deliberation.

Second, criteria were applied for the inclusion of morphemes or meaningful semantic units. Morphemes belonging to any of the following groups were included in the analysis: adjective-main, adjective-suffix, adjective-auxiliary, noun-verbal, noun-common, noun-adjective, or noun-proper-organisation. Using words that met the aforementioned criteria, we conducted correspondence analysis to generate a two-dimensional map of the content of statements made throughout the deliberation. This analysis of categorical data provides a means of displaying or summarising a set of words in two-dimensional graphical form based on correlation and co-occurrence between high frequency words, similar to principal component analysis (Benzécri 1992). The outcome of correspondence analysis is the clustering of high frequency indices close to the centre of the figure. The magnitude of the distance between the speaker and the topic indicates the frequency of the use of that concept in deliberation by the speaker. As the speaker and the topic appear more frequently, the distances diminish. Each figure depicts the relationships of correlation and co-occurrence between each jury and the respective deliberation topics derived from the correspondence analysis as shown in Figure 2. These statistics permit inferences about the contribution of each jury member to the major topics discussed in deliberation.

Results

Preliminarily, the structure of communication in each deliberation was examined as a manipulation check. This confirmed that the facilitator was positioned separately from the discussion between jurors, not centrally, and that jury deliberations were independent of the facilitator (Goodman-Delahunty and Wakabayashi 2012).

Next, the 50 words used most frequently in deliberations by the four juries exposed to each trial presentation (PIVE, JI and DIVE) were identified. As displayed in Table 1, almost all high frequency words (70%) were prevalent across all group deliberations, reflecting the common focus on these topics by all 12 juries exposed to the same evidence. Notably, only three of these words ('expert', 'issue' and 'judge') had appeared in the facilitators' questions.

Table 1: High frequency common and unique words in deliberation by trial presentation group

PIVE			JI			DIVE					
evidence	334	<u>visual</u>	62	bag	359	number	56	evidence	412	fact	55
bag	262	point	57	evidence	335	important	55	bag	291	information	54
box	177	number	56	box	172	<u>simulation</u>	55	box	252	explosive	53
guilty	142	fact	56	people	155	explosion	53	bomb	175	group	47
people	133	<u>trial</u>	55	bomb	140	residue	50	people	161	word	46
train	128	information	54	question	126	room	48	juror	157	judge	44
question	119	court	51	guilty	121	reality	46	question	151	comments	43
expert	113	tennis	50	phone	114	Mr Wheeler	45	phone	138	discussion	41
phone	108	reasonable	48	train	112	information	44	train	125	explosion	40
bomb	99	forensic	45	way	95	person	44	expert	123	seat	37
weigh	96	Mr Wheeler	42	sort	92	judge	44	guilty	119	jury	37
case	94	important	41	issue	91	cardboard	43	time	110	Mr Wheeler	35
issue	89	reality	41	expert	85	 motive	41	issue	112	tennis	35
defence	85	person	39	doubt	83	witness	41	defence	91	residue	33
doubt	82	residue	35	mobile	80	mind	40	point	89	motive	32
sort	79	guy	35	time	78	tennis	38	witness	83	guy	32
jury	79	cardboard	35	visual	75	forensic	37	way	79	right	32
witness	79	stuff	34	point	72	guy	36	person	77	deliberation	31
room	78	convincing	34	defence	71	discussion	35	prosecution	77	decision	30
time	73	house	33	case	70	word	33	case	75	trial	29
judge	73	instruction	33	different	63	computer	32	number	69	item	26
prosecution	65	decision	33	explosive	63	brother	31	cardboard	69	carriage	26
mobile	64	<u>simulation</u>	32	reasonable	61	group	30	room	59	cent	25
explosive	63	explosion	32	fact	61	decision	30	doubt	58	instruction	24
word	62	possible	31	prosecution	59	Mr Green	30	sort	57	fragments	23
										situation	23

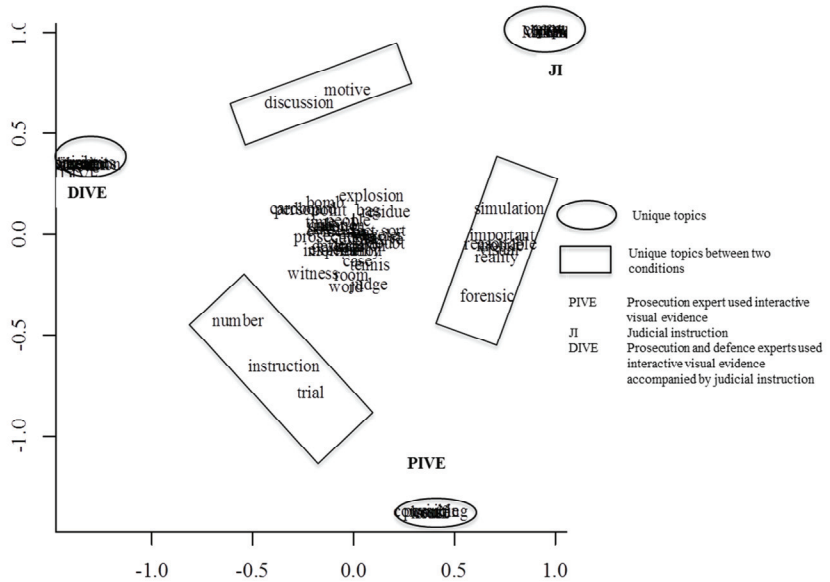
Note: Masked words are unique to deliberations in each trial presentation group. Words without underline, boldface or mask are common to all trial group deliberations. Underlined words are common to the PIVE and JI deliberations. Boldface words are common to all PIVE and DIVE deliberations. Italicised words are common to all JU and DIVE deliberations.

We then conducted a correspondence analysis to visually depict frequently repeated words by trial presentation. Figure 1 displays data converted from Table 1 into a two-dimensional graphical form. Circles contain all unique high frequency words used in deliberation following each trial presentation, and rectangles contain all common high frequency words between two trial presentations. Since the unique and common words that overlap in the graphic display are difficult to discern, they are listed below.

- Unique high frequency words that characterised the PIVE deliberations were: ‘count’, ‘stuff’, ‘convincing’, ‘house’ and ‘possible’.
- Unique topics discussed in the JI deliberations were: ‘different’, ‘mind’, ‘computer’, ‘brother’, ‘group’ and ‘Mr Green’ (prosecution expert).
- Unique topics discussed in the DIVE deliberations were: ‘comments’, ‘seat’, ‘right’, ‘deliberation’, ‘item’, ‘carriage’, ‘cent’, ‘fragments’, and ‘situation’.
- Deliberation topics shared by PIVE and JI juries (enclosed by rectangles) were ‘mobile’, ‘visual’, ‘trial’, ‘reasonable’, ‘forensic’, ‘important’, ‘reality’ and ‘simulation’.
- Deliberation topics shared by PIVE and DIVE juries were ‘jury’ and ‘instruction’.
- Deliberation topics ‘motive’ and ‘discussion’ were the only topics in common and shared by deliberating JI and DIVE juries.

These analyses by two dimensions yielded an accumulated contribution ratio of 56.6%. The magnitude of the distance between trial presentation and the topic of discussion designates the frequency of occurrence in deliberation of that topic. Distances diminish as speakers and the topic of discussion co-occur more frequently.

Figure 1: The correlation and co-occurrence of high frequency words by trial presentation group



Next, we conducted a correspondence analysis between high frequency words common to all three trial conditions to reveal common deliberation topics (Figure 2). Using the Ward sampling method and Euclidian square distance, a hierarchical cluster analysis was performed on the component score of the 35 words in the two dimensions identified by means of the foregoing correspondence analysis. Although the correspondence analysis positioned words by their proximity to PIVE, DIVE and JI, it did not provide a straightforward partitioning of the data. For this purpose, cluster analysis was conducted to obtain an efficient and statistical account of distinctive groupings within the common topics of deliberation.

Figure 2: Correspondence and cluster analysis of the correlation and co-occurrence between words common to deliberations following all trial presentations.

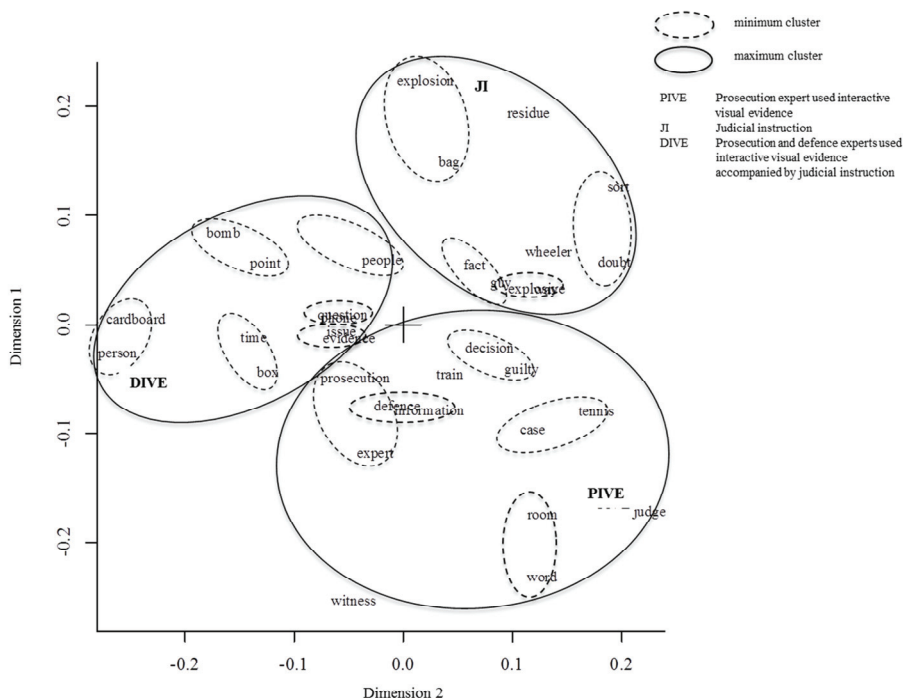


Figure 2 shows the arrangement of each cluster and the relations of correlation and co-occurrence between the common high frequency deliberation topics overall and within the three conditions, derived from the correspondence analysis. Each ellipsis in solid and dashed lines indicated the range of groupings in each cluster. The dashed-line ellipse depicted 15 clusters by minimum distance and the solid-line ellipse depicted three clusters by the maximum distance. The proximity of words and trial labels (PIVE, JI and DIVE) depicted the frequency of use of those concepts in deliberation by speakers in those respective conditions: smaller distances between the names of trial presentation groups and topical words reflected more frequent use of those terms by jurors in those deliberation groups.

Finally, we conducted chi-square tests using the Ryan method on the frequency of common words between three conditions. Results of these analyses are shown in Table 2.

Table 2: Differences in the frequency of common deliberation topics by trial presentation group.

Word	Trial presentation			Chi χ^2	<i>P</i> <i>p</i>	Differences
	PIVE	JI	DIVE			
evidence	334	335	412	11.3	<.01	PIVE = JI < DIVE
box	177	172	252	20.1	<.01	
time	73	78	110	9.3	<.01	
person	39	44	77	16.0	<.01	
cardboard	35	43	69	12.9	<.01	
bag	262	359	291	16.3	<.01	PI = DIVE < JI
witness	79	41	83	15.9	<.01	PI = DIVE > JI
bomb	99	140	175	21.0	<.01	PIVE < JI = DIVE
judge	73	44	44	10.4	<.01	PIVE > JI = DIVE
room	78	48	59	7.5	<.05	PIVE > JI
word	62	33	46	9.0	<.05	
point	57	72	89	7.1	<.05	PIVE < DIVE
expert	113	85	123	7.3	<.05	JI < DIVE
sort	79	92	57	8.2	<.05	JI > DIVE

Chi-square tests revealed statistically significant differences in the frequency of deliberations about the 15 most common topics by trial presentation condition.

Hypothesis One, that PIVE juries would devote less discussion to the prosecution expert's evidence than other juries, was unsupported. In all trials, the frequency of the word 'defence' outstripped the frequency of the word 'prosecution, but only in PIVE juries did the cluster analysis establish that 'prosecution' was associated with the 'expert', disconfirming the hypothesis that salience of PIVE in the absence of judicial warnings would diminish deliberation about the prosecution expert. However, this difference was statistically significant only in comparison with DIVE and not JI deliberations. Although deliberations about the word 'bomb' were significantly fewer following the PIVE trial than the other two trial presentations, significantly more mention was made by PIVE juries of the defendant's room in which bomb-related literature and materials were found.

Jurors exposed to the judicial caution following the prosecution computer simulation (JI trial) spoke significantly more about the defendant's sports bag, but made fewer explicitly references to the expert/witness than did jurors in other trial groups. Their increased scrutiny of the prosecution expert theory that the bomb was in the sports bag and the high mention of 'doubt' by JI juries confirmed Hypothesis Two. As anticipated, following the DIVE presentation where the battle between the experts was most even, deliberation centred more on the defence theory that the source of the bomb was the cardboard box and not the bag belonging to the accused, shown by significantly more mentions of the concepts 'evidence', 'box', 'time', 'person', and 'cardboard', confirming Hypothesis Three.

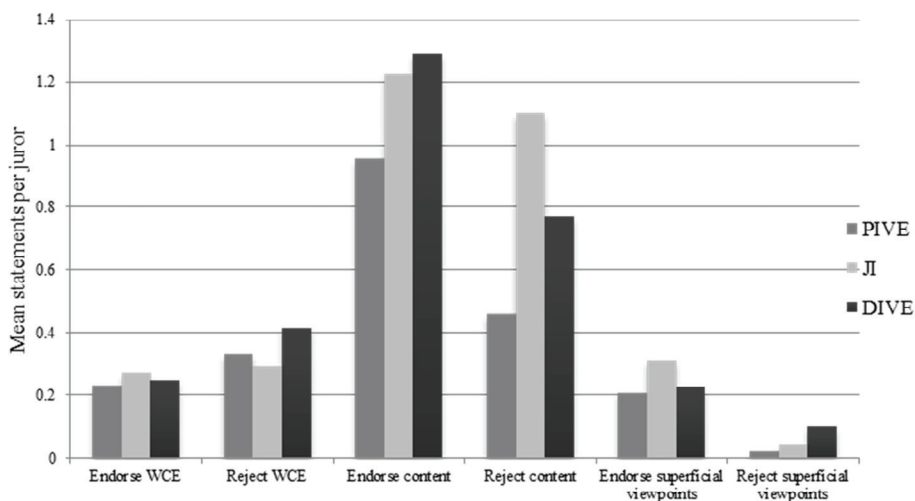
Verdicts

In response to the circumstantial evidence, overall 37 % of the jurors voted to convict, 63% voted to acquit, and two abstained. Four jurors changed their votes from conviction to acquittal as a result of deliberation. In the time available for deliberation none of the juries reached a unanimous verdict, and more changes may have occurred had deliberation continued. Conviction rates in JI were lowest (32%); higher in PIVE (36%) and highest in DIVE (44%).

Deliberations about expert evidence

To cross-validate the findings on jury uses of common words in deliberation, the content of juror comments about expert evidence in the deliberation transcripts following each trial presentation was reviewed and thematically coded for statements: (a) endorsing; or (b) refuting the ‘white coat’ effect; (c) assessing the content of the expert testimony; or (d) commenting on superficial indicia of expert credibility (appearance, qualifications etc). Descriptive results of this analysis, displayed in Figure 3, revealed that in all juries, discussion of the content of the expert evidence dominated; considerably fewer comments about superficial features of the experts were offered. Jurors who observed the PIVE trial made fewer statements that were critical of the experts than jurors who were exposed to the judicial instruction advising them that they could reject the expert’s interpretation and opinions. Jurors who observed JI and DIVE trials elaborated more on the content of the expert evidence than did PIVE jurors. In all juries, statements endorsing a ‘white coat’ effect were counterbalanced by statements rejecting this viewpoint; with respect to statements about the appearance or credentials of the experts, fewer challenges were made.

Figure 3: Mean number of juror statements per expert theme, by trial presentation



Excerpts from deliberations following observations of the experts in the three trial groups are presented below to further illustrate the nature of the jury discussions.

PIVE jury discussions about the prosecution expert’s un rebutted computer animation reflected some individual juror deference to that expert. One juror stated:

‘I found the most *convincing* was the forensic evidence, especially the DVD simulation. What I found least *convincing* was the accused when he testified’. (PIVE-3)

Through the group discussions, the weight of the PIVE was challenged and ultimately dismissed:

‘Both sides agreed that that’s where the explosion started from, and that’s where it all happened, and that the defendant was sitting there. So I felt that, for example, all of that visual that the first expert witness brought in was redundant’. (PIVE-1)

‘I didn’t think the computer generated simulation — it didn’t really add to it at all, because they just said, “this is where the location was.” We could have got that from the verbal. I didn’t think it really added to saying, “this man is guilty”’. (PIVE-4, Juror A)

‘I think that the [prosecution expert], despite his qualifications, was very, very *unconvincing*’. (PIVE-4, Juror B)

Contrary to Hypothesis One, PIVE juries did not uncritically accept the prosecution IVE. One juror eloquently elaborated this point:

‘A digital image does not qualify someone to be an expert. All that does is qualify someone to be a presenter and point — the red dots mark this, they yellow dots mark that — any of us could have done that’. (PIVE-1)

These excerpts demonstrate that decisions to convict were not the result of a ‘white coat’ effect.

As predicted in Hypothesis Two, jurors in the JI trial debated the content of the prosecution expert IVE extensively. JI juries were conscious of the disparity in presentation caused by the one-sided use of a computer simulation by the prosecution. Jurors in JI-1 expressed this concern:

‘I felt [the defence expert’s] testimony had sufficient impact to balance out what I heard from Dr Green. So, I found it quite convincing.’ (Juror A)

‘So Green’s testimony followed up by his audiovisual display was totally equaled by Dr Lange saying a box was a possible under the bench?’ (Juror B)

‘Yes, it balanced out’ (Juror A)

‘I don’t think that’s balanced at all.’ (Juror B)

‘In my view, I just found it completely unbalanced’. (Juror C)

Once again, when individual jurors expressed views deferential to the prosecution expert, other jurors interjected their critique:

‘Actually, when [the prosecution expert] got up and gave his presentation and he charted where all the debris had landed, that was very convincing. It’s sort of like hard evidence — the only sort of hard evidence that you had there. Whereas, [the defence expert] I consider just to be a hired mouth that could say anything at all’. (Juror C)

‘They’re both hired, and they’re both paid a lot of money to say what they’re supposed to say’. (Juror D)

The JI juries expressed scepticism about the value of the computer simulation presented by the expert for the prosecution:

‘I found that their TV — the computer simulation — had virtually no bearing on whether I felt the guy was guilty or not.’ (Juror A)

‘I agree.’ (Juror B)

‘Yes, I agree with that.’ (Juror C)

‘I felt [defence expert’s] testimony had sufficient impact to balance out what I heard from [prosecution expert]. So I found it quite convincing. So the prosecution expert testimony followed up by his audio visual display was totally equaled by the defence expert saying a box was possible under the bench? Yes, it balanced out’. (Juror B)

Deliberations by other JI juries reflected similar themes, dismissive of the prosecution’s simulation. One juror stated:

‘You can make computer graphics do anything. One of my — or my youngest child is a bit of a whiz bang — or he thinks he is, anyway, and he can make any scene on his computer screen do anything he wants it’. (JI-2)

Jurors in the third JI jury commented:

‘I thought the visuals were really good, but they had no — little — bearing for me on the case.’ (Juror A)

‘No it didn’t change my views, but it helped me to understand in my head where everything went.’ (Juror B)

‘No, that’s just CSI stuff on TV, where, you know, like all the stuff on TV, and you know ... the truth comes out ... and you know, it looks nice. I mean ... it does — but, like I said, CSI visuals’. (Juror C)

The hypothesis that JI deliberations about the IVE would exceed those in other trial groups was supported by the quantitative results and the foregoing transcript excerpts. The juries appeared motivated to compensate for the disparity between expert presentations. As hypothesised, the frequency of explicit references to the judicial caution did not increase, but was lowest in deliberations by JI juries. The effect of the judicial caution in evoking elaborated scrutiny of the expert evidence was apparent in the following exchange by a JI jury (JI-3):

‘Well, the expert is really — his expert opinion is not necessarily right. That’s what we have to remember. Quite often we get expert advice and it is not right.’ (Juror D)

‘And when it comes around to the experts, you know, they equaled one another out... They just equaled one another out in as much as saying do we really know if there was a box there. Well, we don’t know if there is, but we don’t know if there wasn’t, either. And both of them can’t prove that there wasn’t a box and they both can’t prove that there is one. So who, then, do we believe?’ (Juror B)

‘So expert witnesses are not credible, to the extent that they earn their living by being expert witnesses. They are there to be invited by the prosecution or the defence to present a particular view.’ (Juror E)

‘A biased view’. (Juror F)

‘But one that is supported by evidence and knowledge, but it is blinkered view, lacking balance. I would suggest that they like to think if themselves as being unblinkered, but I think every human being eventually becomes blinkered.’ ... So the police and the forensic scientists are blinkered because they are within that culture. The independents are blinkered because they earn a living by offering expert evidence. So, a juror has to look at both of them and say why they maybe saying what they were saying’. (Juror E; JI-3)

As predicted in Hypothesis Three, where both experts relied upon IVE to illustrate and support their opinions (DIVE trial), jurors did not place much emphasis in deliberation on topics ‘visual’, ‘simulation’, ‘mobile’, ‘reasonable’, ‘forensic’, ‘important’, and ‘reality’, suggesting that the equipoise between the experts relieved them of the task of reviewing the realism of the computer simulated visual evidence, and allowed them instead focus on the content of the other evidence. Jurors realised that the visual illustrative evidence was not dispositive, but found it useful:

‘The visual was handy to give you an idea of where it happened and what happened, but I found the verbal was much more convincing for me’. (DIVE-4, Juror A)

‘Well, I think the visual kind of set the scene, but also, something you just mentioned, in the visual you are looking at the number, but the signs and actions of witnesses’. (DIVE-4, Juror B)

These results confirmed our hypothesis that the balanced presentation of computer simulations in DIVE would diminish the salience of those simulations and decrease deliberation on that topic. Concerns that the use by both experts of the sophisticated IVE technology would be unduly persuasive and would elicit a ‘white coat’ effect were unfounded.

The balanced use of visual evidence allowed the juries to better understand and debate the meaning of the trace evidence. One juror stated: ‘From what I recall, they didn’t say the *fragments* were found inside the bag. They said the residue of the plastic explosive was found on the inner surface of the bag lining’ (DIVE-4). This showed the jurors were considering whether the fragments depicted in the IVE proved how the explosion occurred, and perceived the IVE as an illustration, not as substantive evidence. Another juror confirmed this: ‘That didn’t prove anything to me, or didn’t start to prove anything to me, but just seeing the visual representation helped me, I suppose, get a picture, really, and of the distribution of the *fragments* of the bag’ (DIVE-1). The presence of the countervailing expert visual presentation enabled jurors to focus productively on topics pertinent to the core issues and evidence, and to avoid the white coat effect.

Discussion

The hypothesis that juries in PIVE would be more susceptible to the white coat effect was unsupported. Although comments reflective of this deference were at times made by individual jurors, this view did not survive deliberation, as it was rejected by other jurors. Indeed, none of the juries exposed to the well-qualified forensic scientists were unduly deferential; the presence of an opposing defence expert appeared to have rendered the prosecution expert less persuasive. Thematic analyses showed that the lower rate of comments rejecting the expert evidence observed in PIVE than JI and DIVE juries appeared to be related to the absence of the judicial caution about the expert and the uneven expert evidence, rather than the white coat effect per se. Overall, these findings are consistent with

research discussed earlier affirming that jurors are not intimidated by expert testimony and instead 'engage deeply' with the material (Diamond and Casper 2001:558; Hans 2008). When given the opportunity to ask questions about expert evidence in civil trials, for instance, 90% of jurors did so, demonstrating pro-activity and engagement in understanding the evidence (Vidmar and Diamond 2001). This outcome also appeared to be related to efforts by the JI juries to compensate for the lack of visual evidence by the defence. Their reaction to the disparity, intensified by the judicial caution, appeared to motivate them to reject the prosecution evidence to a greater degree than occurred when the same judicial instruction was presented in the DIVE trial, but both experts employed a computer simulation.

As might be anticipated, given the identical trial scripts, the topics discussed most frequently in deliberation were similar and common across all three trial presentations. This congruence, reflected in Tables 1 and 2, indicated that notwithstanding the diversity of individual jurors, the different dynamics of each of the 12 juries, administration of different technologies and judicial directions, a remarkably similar set of core issues was addressed and debated. This finding demonstrated that what occurs in the secrecy of the jury room is considerably less random and unpredictable than some jury sceptics contend. Furthermore, it indicated that jurors were not distracted by the expert evidence, but were able to focus on the core issues and reduce evidential errors, as shown in prior studies (Hans et al 2011; Brekke et al 1991).

Given that the trial content across all conditions was invariant, significant differences that emerged in deliberation suggested that procedural variations in the trial presentation can shift the emphasis in deliberation. Where the defence's oral presentation was 'outgunned' by the prosecution's sophisticated visual simulation of the explosion, the debate about that technology and its realism was more salient, as reflected both in the text-mining analysis and deliberations excerpts of PIVE and JI deliberations. This scrutiny of the prosecution IVE was consistent with earlier research demonstrating that jurors are not deferential to expert testimony, however sophisticated the presentation (Vidmar 2005).

Identification of unique deliberation topics in the different trial conditions confirmed subtle changes in deliberation responsive to procedural variations. High frequency words in DIVE deliberations (Figure 1) suggested that these deliberations were more strongly focused on the non-expert evidence. Again, this finding is consistent with earlier research demonstrating that adversarial experts can diminish the salience of both experts (Devenport and Cutler 2004; Pezdek, Avila-Mora and Sperry 2010) and increase the conviction rate (Levett and Kovera 2008).

Conclusion

Juries were not daunted or overwhelmed by a battle of forensic science experts. The forensic science did not eclipse other evidence in the case; quantitative analyses confirmed that other deliberation topics were more prominent following all trial presentations. Overall, the 12 juries were not blinded by science, the experts' qualifications, or their sophisticated computer simulations. Even under conditions most optimal to induce deference to the prosecution expert, no white coat effect persisted through deliberation. Deliberation about two adversarial experts enhanced the juries' awareness of the limits of the expert evidence in resolving the issues in the case.

Nonetheless, these findings are limited in several respects. The battle of experts was not compared with a single expert, and the logistics of the live simulation did not permit a fully crossed experimental design (Tait 2011). Future research using these experimental controls in a fully crossed design is advised.

The innovative text-mining analysis of frequencies of deliberation topics was compatible with the Dual-Processing Theory and the Elaboration Likelihood Model as a means to add rigour in evaluating subtle differences evoked in 12 different juries to the persuasive impact of forensic expert evidence that was evenly or unevenly balanced or accompanied by a judicial caution. The countervailing expert enhanced jury scrutiny of the central content of forensic scientific testimony, confirming earlier research on the impact of adversarial experts on other topics (Brekke and Borgida 1988; Cutler, Dexter and Penrod 1989). These findings extended previous research by identifying factors that discouraged juries from uncritically endorsing the scientific expert evidence (Vidmar and Diamond 2001). Because deliberations were truncated and no unanimous jury verdicts rendered, caution is advised in drawing inferences from differences in the juror conviction rates across groups. The presence of the judicial direction (JI) indicated that this intervention by the judge exerted a powerful influence on jury assessments, diminishing the weight of the prosecution expert and reducing the conviction rate. However, when the expert use of computer simulated visual aids was balanced (DIVE), despite the presence of the same judicial warnings, juries devoted more time to a discussion of the core issues in the case and whether the charges were proved beyond reasonable doubt, and the conviction rate increased. In sum, the results of this study provided some empirical support for the relative effectiveness of three distinct legal procedural safeguards available to minimise jury errors when forensic scientific evidence is used: the use of a rebuttal expert, a judicial direction about the non-binding nature of expert evidence, and group deliberation.

Legislation

Jury Act 1977 (NSW)

References

- American Bar Association (1989) 'Jury Comprehension in Complex Cases: Report of the Special Committee on the ABA Section of Litigation', The American Bar Association, Chicago
- Benzécri J P (1992) *Correspondence Analysis Handbook* Marcel Dekker, New York
- Bornstein B and McCabe S (2005) 'Jurors of the Absurd—The Role of Consequentiality in Jury Simulation Research', *Florida State University Law Review* 32, 443–67
- Brekke N and Borgida E (1988) 'Expert Psychological Testimony in Rape Trials: A Social-Cognitive Analysis', *Journal of Personality and Social Psychology*, 55(3), 372–86
- Brekke N, Enko P, Clavet G and Seelau E (1991) 'Of Juries and Court-Appointed Experts: The Impact of Nonadversarial Versus Adversarial Expert Testimony', *Law and Human Behavior* 15(5), 451–75
- Buck JA and Warren AR (2010) 'Expert Testimony in Recovered Memory Trials: Effects on Mock Jurors' Opinions, Deliberations and Verdicts', *Applied Cognitive Psychology* 24(4), 495–512

- Cutler BL, Dexter HR and Penrod SD (1989) 'Expert Testimony and Jury Decision Making: An Empirical Analysis', *Behavioral Sciences and the Law* 7, 215–25
- Devenport JL and Cutler BL (2004) 'Impact of Defense-Only and Opposing Eyewitness Experts on Juror Judgments', *Law and Human Behavior* 28(5), 569–76
- Devine DJ, Clayton LD, Dunford BB, Seying R and Pryce J (2001) 'Jury Decision Making: 45 Years of Empirical Research on Deliberating Groups', *Psychology, Public Policy, and Law* 7(3), 622–727
- Diamond SS and Casper JD (1992) 'Blindfolding the Jury to Verdict Consequences: Damages, Experts, and The Civil Jury', *Law & Society Review* 26, 513–63
- Edmond G (2009) 'Merton and The Hot Tub: Scientific Conventions and Expert Evidence in Australian Civil Procedure', *Law and Contemporary Problems* 72, 159–89
- Edmond G (2011) 'Actual Innocents? Legal Limitations and Their Implications for Forensic Science and Medicine', *Australian Journal of Forensic Science* 43, 177–212
- Evans JSBT (2008) 'Dual-processing Accounts of Reasoning, Judgement, and Social Cognition', *Annual Review of Psychology* 59(1), 255–78
- Findlay M (2008) 'Juror Comprehension and the Hard Case: Making Forensic Evidence Simpler', *International Journal of Law, Crime and Justice* 36(1), 15–53
- Fordham J (2009) 'Bad Press: Does the Jury Deserve It?', Paper presented at the 36th Australian Legal Convention, Perth, Western Australia, 17–19 September
- Goodman-Delahunty J, Rossner M and Tait DT (2011) 'Simulation and Dissimulation in Jury Research: Credibility in a Live Mock Trial' in Bartels L and Richardson K (eds) *Qualitative criminology: stories from the field*, Federation Press, Sydney
- Goodman-Delahunty J, Saunders P, Dhimi M and Tinsley Y (2012) *Strengthening Forensic Science in Korea*, Korean Institute of Criminology, Seoul
- Goodman-Delahunty J and Wakabayashi K (2012) 'Professional Facilitators and Judges in Traditional and Mixed juries: Text Mining to Analyse Deliberative Democracy', Paper presented at the European Association of Psychology and Law Conference, Nicosia, Cyprus, 10–13 April 2012
- Hans VP (2008) 'Judges, Juries, and Scientific Evidence', *Journal of Law and Policy* 16, 19–46
- Hans VP, Kaye DH, Dann BM, Farley EJ and Albertson S (2011) 'Science in the Jury Box: Jurors' Comprehension of Mitochondrial DNA Evidence', *Law and Human Behaviour* 35(1), 60–71
- Hotta S and Fujita M (2007) 'The Analysis of Mock Trials (part 2): From the Viewpoint of Quantitative Linguistics', *Quarterly Keiji-Bengo* 52, 58–61
- Levett L and Kovera M (2008) 'The Effectiveness of Opposing Expert Witnesses for Educating Jurors about Unreliable Expert Evidence', *Law and Human Behavior* 32(4), 363–74
- Matsumura M and Miura A (2012), *TTM Tiny Text Miner β version* <<http://mtmr.jp/ttm/>>
- McClellan P (2011) 'Looking Inside the Jury Room', *Bar News*, The Journal of the NSW Bar Association, 64–72

Petty R and Cacioppo J (1984) 'Source Factors and the Elaboration Likelihood Model of Persuasion', *Advances in Consumer Research* 11(1), 668–72

Pezdek K, Avila-Mora E and Sperry K (2010) 'Does Trial Presentation Medium Matter in Jury Simulation Research? Evaluating the Effectiveness of Eyewitness Expert Testimony', *Applied Cognitive Psychology* 24, 673–90

Römken R (2000) 'Ambiguous Responsibilities: Law and Conflicting Expert Testimony on the Abused Woman Who Shot Her Sleeping Husband', *Law and Social Inquiry* 25(2), 355–91

Tait D (2011) 'Deliberating About Terrorism: Prejudice and Jury Verdicts in a Mock Terrorism Trial', *Australian & New Zealand Journal of Criminology* 44(3), 387–403

Verrender P and Goodman-Delahunty J (2011) 'Jurors' Evaluations of Expert Evidence: What's More Important--The Message or the Messenger?', Paper presented at the Annual Conference of the Australian and New Zealand Association of Psychiatry, Psychology, and Law, Wellington, New Zealand, 16–20 November 2011

Vidmar N (2005) 'Expert Evidence, the Adversary System, and the Jury', *American Journal of Public Health* 95, S137–143

Vidmar N and Diamond SS (2001) 'Juries and Expert Evidence', *Brooklyn Law Review* 66(4), 1121–80

Wheate R (2006) 'Australian Juries and Scientific Evidence', *Australian Journal of Forensic Sciences* 38, 75–84

Wuffle A (1985) 'Expert vs Expert: Lessons from *Badham v. Eu*', *PS* 18(3), 576–81

Appendix A

Judicial direction about illustrative expert evidence

'Dr Green (expert for the Crown), is about to present evidence visually using an interactive computer simulation. Dr Green was not present in person to observe the events in question, but will testify based on information collected after the incident, and a reconstruction of what might have happened. What really happened is for you to decide based on all the evidence you hear and see in this case. In a simulation of this type, you will experience a three-dimensional virtual environment, where you can move around, look at the scene from different perspectives, and interact with what you see, for example, by moving things around in the scene. In the simulation, the witness may show you the scene from different angles, focus on particular points, and demonstrate different possible events. You may choose to accept some parts of the evidence and not others. You may accept some evidence, but disagree with the interpretation the witness places on it. The decision is yours. It is for you, the jury, to decide how much value to give any evidence provided in the interactive display. The simulation doesn't have any special status. It is evidence like all other evidence. Give as much weight to evidence presented in visual form as you think it deserves. How each of you hears or sees particular aspects of the evidence may be different from the juror sitting next to you. When you come to review the evidence as a group, you will have the opportunity to consider the responses of different jurors to various aspects of the evidence, including that presented visually.'