Oñati International Institute for the Sociology of Law
Instituto Internacional de Sociología Jurídica de Oñati
Oñati Socio-Legal Series
ISSN: 2079-5971

Oñati Socio-legal Series, v. 6, n. 2 (2016) – Juries and Mixed Tribunals across the Globe: New Developments, Common Challenges and Future Directions
ISSN: 2079-5971

Programmatic Pretest-posttest Research to Reduce Jury Bias in Child Sexual Abuse Cases

JANE GOODMAN-DELAHUNTY*
NATALIE MARTSCHUK*
ANNE COSSINS*

Goodman-Delahunty, J., Martschuk, N., Cossins, A., 2016. Programmatic Pretest-posttest Research to Reduce Jury Bias in Child Sexual Abuse Cases. *Oñati Socio-legal Series* [online], 6 (2), 283-314. Available from: http://ssrn.com/abstract=2786072

Abstract

When lay jurors are unfamiliar with key evidentiary issues, expert evidence, judicial instructions and group deliberation may enhance their understanding of this evidence. Systematic steps to assess the relationship between juror biases in cases of child sexual abuse are offered as an example to illustrate a programmatic research approach. Using pretest-posttest research designs, the effectiveness of three traditional legal procedural safeguards to reduce common jury misconceptions in the context of simulated trials were tested and compared. By measuring mock-juror knowledge before and after each intervention, knowledge gains attributable to these interventions were distinguished from practice effects. Unexpected increases in acquittals following deliberation underscored the importance of adding adequate control groups and of testing deliberation effects in jury simulation studies. Benefits of this research paradigm to assist courts, legal counsel and policy makers in

Support for this research was provided by Australian Research Council Discovery Project DP110103706 awarded to Jane Goodman-Delahunty and Anne Cossins.

* Trained in law and experimental psychology, Jane Goodman-Delahunty, JD, PhD, is a Research Professor at Charles Sturt University. She served as editor of *Psychology, Public Policy, and Law*, and president of both the American Psychology-Law Society and the Australia and New Zealand Association of Psychiatry, Psychology and Law. Her research promotes evidence-based policies to enhance justice, informed by her experience as a litigator, administrative judge, mediator, Commissioner with the New South Wales Law Reform Commission and Member of the New South Wales Civil and Administrative Tribunal. Charles Sturt University Manly Campus. P.O. Box 168. Manly NSW 1655, Australia. jdelahunty@csu.edu.au

* Natalie Martschuk, Dipl-Psych, is a Research Associate at Charles Sturt University. Her research interests are jury decision making, expert evidence, investigative interviewing strategies, witness reliability and advanced statistical methods. She is working towards her PhD at the University of Giessen on the reliability of elderly eyewitnesses. Previously, she practiced at the State Hospital for Forensic Psychiatry in Stendal, where she interviewed suspects and offenders, and provided expert evidence in court. Her postgraduate major was Legal Psychology. Charles Sturt University Manly Campus. P. O. Box 168. Manly NSW 1655, Australia. nmartschuk@csu.edu.au

* Anne Cossins, PhD, is a Professor of Law and Criminology in the Faculty of Law, University of New South Wales. She is the Co-convenor of the Gendered Violence Research Network at UNSW and has been involved in several law reform initiatives with government bodies over the past decade. She is a member of the NSW Victims Advisory Board and the Advisory Board of the Global Foundation on Justice for Child Witnesses. She has several external grants in relation to jury studies and the sentencing of Indigenous offenders for sexual assault offences. Her research interests include: the treatment, recidivism and classification of sex offenders; sexual assault law reform, jury decision making, theoretical criminology and expert forensic evidence. a.cossins@unsw.edu.au

devising effective methods to enhance jury decisions in complex criminal cases are discussed.

Key words

Child sexual abuse; deliberation; expert evidence; jury bias; pretest-posttest research design; specialized knowledge; wrongful acquittal

Resumen

Cuando los miembros del jurado popular no están familiarizados con los temas probatorios clave, las pruebas periciales, las instrucciones judiciales y la deliberación de grupo pueden mejorar su comprensión de las evidencias. Se ofrecen pasos sistemáticos para evaluar la relación entre sesgos del jurado en casos de abusos a menores como un ejemplo para ilustrar un enfoque de investigación programática. Se probó y comparó la efectividad de tres garantías procesales jurídicas tradicionales para reducir los malentendidos comunes de los jurados en el contexto de juicios simulados, usando patrones de investigación pre y post análisis. Al medir el conocimiento de jurados en juicios simulados antes y después de cada intervención, se distinguían los conocimientos adquiridos atribuibles a estas intervenciones. Un aumento inesperado de absoluciones después de la deliberación ponía de manifiesto la importancia de añadir grupos de control adecuados y analizar los efectos de la deliberación en los estudios de jurados en juicios simulados. Se analizan los beneficios de este paradigma de investigación para ayudar a tribunales, abogados y legisladores en el desarrollo de métodos eficaces para mejorar las decisiones del jurado en casos penales complejos.

Palabras clave

Abusos a menores; deliberación; pruebas periciales; sesgo del jurado; diseño de investigación pre- y post-prueba; conocimiento especializado; absoluciones erróneas

Table of contents

1. Introduction

Scholars conducting research on legal decision making express divergent views about the trustworthiness of a group of lay jurors untrained in the law, such as those who serve on juries in Australia, Canada, Spain, the United Kingdom, and the United States of America (Kaplan and Martin 2006), compared to professionally trained judges alone, as is the practice in the Netherlands (van Koppen 2009) or mixed tribunals comprised of both lay jurors and professional judges, as are used in France and Japan (Marder and Hans 2015). In Australia, the site of the present jury research, a small number of indictable or serious crimes, estimated at less than 1-9% of all criminal cases, are eligible for trial by jury (Goodman-Delahunty 2015, Goodman-Delahunty and Tait 2006). A high proportion of those trials are child sexual assault cases (Cossins and Goodman-Delahunty 2013). In most of the nine Australian jurisdictions, the defendant can elect or consent to a trial by a judge alone (Goodman-Delahunty 2015). The Australian jury system was implemented in 1833 along the lines of the jury system in England and Wales (Chesterman 1999). A panel of twelve citizens is selected from citizens on the roll of registered voters to hear the evidence and determine whether the accused is guilty beyond reasonable doubt (Goodman-Delahunty *et al*. 2008). At the close of the trial, the judge orally summarizes the relevant evidence, outlines the pertinent law, and directs the jury on the application of the legal principles to the facts (Ogloff *et al*. 2007). In most Australian states, after a minimum period of deliberation (eight hours in NSW; six hours in Victoria), a majority verdict of all but one of the jurors is accepted, although in some states, such as Queensland, the jury verdict must be unanimous even after protracted deliberation (Nolan and Goodman-Delahunty 2015).

In general, social science research has established that juries perform their duties conscientiously and reach decisions substantially similar to those of judges (Eisenberg *et al*. 2005, Bornstein and Greene 2011, Kim *et al*. 2013), but add value to the criminal justice system by bringing community values to bear on the legal process (Delli Carpini *et al*. 2004, McClellan 2008). Community participation through jury service can legitimize justice outcomes, as was demonstrated by research on procedural justice and deliberative democracy (Gastil *et al*. 2010, Gastil and Xenos 2010). By contrast, judge-only trials and mixed tribunals are criticised as more susceptible to political pressure to reach a particular verdict in high-profile cases. Compared to a decision by a judge alone, the benefits of a lay jury include the diversity of jurors' backgrounds, experiences and perspectives, the ability to recollect facts as a group rather than rely on one individual's memory, and the opportunity in deliberation to challenge each others' views and correct errors and misconceptions to render a group verdict that reflects 'common sense judgment' (Marder and Hans 2015, p. 799).

The verdict reached by a jury after deliberation has been described as a reflection of the story developed by the group from the evidence presented at trial (MacKillop and Vidmar 2015) although it may include inaccuracies, especially when the evidence presented is unreliable or incomplete (Lempert 2015). In cases of child sexual abuse (CSA), jurors must frequently make a decision based on the credibility of the child complainant, a form of "soft" evidence ("word against word" oral evidence alone) without supporting eyewitnesses or "hard" evidence (e.g., DNA profiling evidence, medical or photographic evidence). Thus, in CSA cases, jurors are likely to rely more extensively on their common knowledge to interpret the alleged events (Cossins and Goodman-Delahunty 2013).

Overall, general attitudinal biases in jurors do not pose a substantial threat to justice outcomes, as these predispositions are expected to balance out in a jury composed of twelve community members, and to exert a less powerful influence on verdicts than the evidence itself. Experimental research has confirmed that the evidence in a case predicts about 80% of trial simulation verdicts, whereas "extralegal" factors such as jury demographic attributes, personality traits, and

Oñati Socio-legal Series, v. 6, n. 2 (2016), 283-314
ISSN: 2079-5971

general attitudes are typically associated weakly and unreliably with jurors' verdicts (MacCoun 1989). When key issues in the case lie outside of the scope of lay jurors' common experience and expertise, the potential for inaccuracies and misinterpretation of evidence increases (Freckelton *et al*. 2016).

Cases that result in a wrongful conviction attract the most public and research attention; cases that result in wrongful acquittal are also legally problematic. In CSA cases, the acquittal rate at trial far exceeds that of other crimes (Fitzgerald 2006, Oliver 2006). The risk of wrongful acquittal may increase in the absence of corroborative evidence and judicial or expert guidance on how to discern which party is the more credible or reliable source (Gabora *et al*. 1993, Australian Institute of Judicial Administration Committee (AIJA) 2015). As a consequence, concerns about specific jury predispositions or biases associated with common misperceptions about children's responses to sexual abuse have become a focus of research exploring the causes of high attrition and acquittal rates in CSA trials (Cossins *et al*. 2009, Cossins and Goodman-Delahunty 2013, Tabak and Klettke 2014, AIJA Committee 2015). Interest has similarly increased in cases susceptible to "failures of justice" (Ransley *et al*. 2012) such as those in which expert evidence is presented on issues beyond the general knowledge of most lay jurors (Edmond *et al*. 2013).

1.1. A research approach to assess jury bias in cases of child sexual abuse

This paper describes empirical research strategies devised to explore potential jury bias in CSA cases and legal safeguards to minimize this impact in jury trials. Rather than report results of a single study, it offers a detailed example of a series of separate experimental studies conducted in the course of a systematic research program. The aim of the research program was to discern (a) the scope of juror error or bias that can pose a risk to a fair and just verdict, and (b) the effectiveness of traditional legal procedures in reducing identified jury bias, particularly interventions addressing the credibility of the child complainant found in past jury studies (Donnelly 2007, Donnelly and Huggett personal communication, 15 May 20121) devised to redress findings showing the low perceived credibility of child complainants (Gabora *et al*. 1993).

Theoretically, this research program draws on social cognitive Dual Process models of decision making (Chaiken and Trope 1999, Smith and DeCoster 2000), namely the Elaboration Likelihood Model (Pennington and Hastie 1986) and the Heuristic-Systematic Model (Petty and Cacioppo 1986, Chaiken *et al*. 1989), that have been applied to jury decisions (Freckelton *et al*. 2016). These models are compatible with the cognitive Story Model and well-suited to examine the influence of extralegal information on jury decisions, as juries are exposed to a series of persuasive messages from legal counsel, witnesses and the trial judge during the course of a trial, and from other jurors during the course of deliberations. Dual Process Models distinguish two ways in which decisions about witness credibility and verdicts are made by jurors: (a) the central route involves more effortful thinking, and is used when jurors have the cognitive capacity and time to think about an issue that they regard as important; and (b) the peripheral route involves less thinking and is used more frequently by jurors who are distracted, or who have limited time to think about the message content, or about a topic viewed as less significant (Freckelton *et al*. 2016).

In this paper, a worked example of a pretest-posttest trial simulation paradigm outlines one research approach that can be applied to address the complex problem of jury bias and jury decision making in CSA trials. "A worked example is a step-by-

[1] H. Donnelly and S. Huggett. *Specialized knowledge of the impact of sexual abuse.* Presentation at the National Symposium on Truth, Testimony and Relevance: Improving the Quality of Evidence in Sexual Offence Cases, Australian Institute of Criminology and Australian Centre Study of Sexual Assault, 15-16 May 2012, MCG Melbourne, Victoria, Australia.

step demonstration of how to perform a task or how to solve a problem" (Clark *et al*. 2006, p. 190). By using a pretest-posttest trial simulation paradigm across a series of separate CSA studies, changes in juror knowledge about key trial issues were uncovered. This new knowledge had an impact on jury appraisals of the credibility of the child complainant and on jury verdicts.

First, this research paradigm is described. Next, the sequential steps or stages of the research program are illustrated by excerpts interspersed with figures drawn from prior studies conducted over a period spanning six years, to explicate the worked example. The research examples offered in this paper describe factors that can increase or decrease conviction rates and the respective risk of wrongful acquittal in CSA trials. More broadly, the worked example in this paper outlines a series of considerations in implementing an evidence-based approach to assist jurors in rendering fair and impartial verdicts in CSA and other types of cases.

2. Pretest-posttest research method

Several different experimental designs can be applied to test intervention effects. Some include only a posttest. Weaknesses of a posttest only include threats to both internal validity (the extent to which treatment causes change) and external validity (the extent to which the treatment effect is generalizable). To avoid validity threats, pretest-posttest designs are recommended so that the equivalence of two groups at the outset can be assured, based on comparisons of the group pretest scores (Bonate 2000). Moreover, in the absence of pretest scores, researchers cannot control for any preexisting jury bias, or interaction effects due to a possible selection bias (Dimitrov and Rumrill 2003). Thus, designs that include both a pretest and a posttest are superior.

Two major types of pretest-posttest designs are distinguished depending on whether two or four groups of participants are included in a study: (a) two-group randomized or nonrandomized control group designs (Groups 1 and 2 in Table 1); and (b) Solomon four-group pretest-posttest designs, as shown in Table 1. The Solomon four-group pretest-posttest design combines features of the randomized posttest design and the two-group design, and includes two control groups, as shown in Table 1.

Table 1. Solomon four-group pretest-posttest research design

Experimental group	Pretest	Treatment	Posttest
1	✓	✓	✓
2	✓	-	✓
3	-	✓	✓
4	-	-	✓

The Solomon four-group pretest-posttest research design is the most reliable research approach to assess the presence of juror or jury bias as the treatment or intervention group (Group 3 in Table 1) is compared with a no treatment or no intervention group (Group 2 in Table 1), and both groups complete the same pre- and posttest (dependent measures). To fulfill requirements for an experiment (as opposed to a quasi-experiment), participants are randomly assigned to experimental groups, comprising a “randomized control group posttest design.” The presence or absence of bias is inferred from statistical comparisons of the group posttest scores.

Using this design, researchers can test the effectiveness of interventions by comparing experimental and control groups on mean “gain scores” by subtracting the difference between each participant’s posttest and pretest scores, while controlling for pretest differences. Further, the interaction of pretest and treatment effects can be statistically controlled. This approach is rigorous as it reduces error variance, and generally yields gain scores that are reliable.

Oñati Socio-legal Series, v. 6, n. 2 (2016), 283-314
ISSN: 2079-5971

Table 2. Mean juror age, gender and pretrial CSA misconception scores by study sample and type

Study	Sample	*N*	Mean age	Gender (% women)	Aim of study	Trial mode	No. of items (scale)	Mean score (raw)	Proportion score
Cossins *et al.* 2009	Community & students	230 429	48.0 20.0	66.5 59.1	Survey CSA knowledge	--	20 (1-6)	53.6	44.7
Goodman-Delahunty *et al.* 2010	Community members	130	50.6	65.0	Expert evidence and judicial direction change to CSA knowledge	written	22 (1-7)	81.9	45.0
Goodman-Delahunty *et al.* 2011	Students	118	19.4	72.0	Expert evidence and judicial direction change in CSA knowledge	written	26 (1-7)	87.0	47.8
Goodman-Delahunty *et al.* 2013	NSW jurors	443	43.3	43.8	Evidence strength and judicial direction change in CSA knowledge	videotape	28 (1-5)	79.1	43.5
Goodman-Delahunty *et al.* 2014	NSW jurors	876	43.4	42.0	Expert evidence and judicial direction impact on verdict	videotape	20 (1-5)	49.1	49.1
Goodman-Delahunty *et al.* 2016 (Study 1)	NSW jurors	843	43.3	41.8	Validation of CSA knowledge	--	9 (1-5)	25.4	56.4
Goodman-Delahunty *et al.* 2016 (Study 2)	NSW jurors	813	43.3	40.9	Cross-validation, predictive validity in weak and strong CSA cases	videotape	9 (1-5)	25.4	56.4

Note. Proportion scores reflect endorsement of CSA misconceptions. The higher scores reported in Goodman-Delahunty *et al.* (2016) are the result of refinements to the questionnaire, excluding items which 45% or more of the participants answered accurately.

2.1. Overview of pretest-posttest research findings

Table 2 summarizes the studies conducted in the course of programmatic research undertaken to assess and reduce jury bias in CSA trials. Excerpts of key findings from these studies are presented below interspersed with figures to illustrate the paradigmatic pretest-posttest methodology.

2.1.1. Participants

Participants in the research program included undergraduate students (Cossins *et al.* 2009, Goodman-Delahunty *et al.* 2011), community volunteers (Cossins *et al.* 2009, Goodman-Delahunty *et al.* 2010) and NSW nonempanelled jurors (Goodman-Delahunty *et al*. 2013, 2014, 2016). Community members and students participated individually online without remuneration for their participation; undergraduate students received course credit for participation. Nonempanelled jurors participated in-person on court premises and either completed the survey independently (nondeliberating jurors) or after deliberating in jury groups of 8-12 for up to ninety minutes (deliberating jurors). Deliberating jurors were provided with a light lunch. At the outset of deliberation they selected a foreperson. In Australian trials, the jury chooses a foreperson who serves as the spokesperson for the jury and who delivers the verdict on behalf of the jury.

2.1.2. Trial simulation materials

Two mock-trials written by lawyer Annie Cossins with input from the first author were based on actual CSA cases. A defendant, who was known to a female child complainant aged 8 or 11 years of age, was alleged to have committed a single penetrative offense without use of force. All trial materials included prosecution and defence opening and closing addresses, evidence-in-chief and cross-examination of the child complainant and an expert witness (if present), and a judicial summation. The accused did not testify at trial.

In each study, participants were challenged by case facts that included different types of counter-intuitive behaviour of sexually abused children, such as reporting delay, the absence of emotional displays, continuing to spend time with the abuser after the alleged abuse event, inconsistencies in the account of the alleged abuse, a retracted statement, or manipulation by adults. In certain trials, the evidence strength was varied by adding a corroborating witness for the Crown who observed ambiguous but potentially inculpatory behaviour of the defendant, and the disclosure of abuse was immediate rather than delayed. Two different modes of trial presentation were tested: (a) a written transcript, and (b) a professionally role-played videotrial (see Table 2).

2.1.3. Educative interventions to increase jurors' knowledge about child sexual abuse

Educative information to counteract common CSA misconceptions was provided to mock-jurors in the course of a trial simulation either by an expert witness or by the trial judge in a judicial instruction, as follows (a) specialized information presented by a clinical psychological expert witness; (b) specialized information presented by an expert witness identified as a research psychologist; (c) a judicial direction presented before the testimony of the child complainant; (d) a judicial direction presented during the summing up; and (e) no specialized information (Control Group). The specialized knowledge incorporated several empirical findings on counter-intuitive behaviour of child sexual abuse victims, children's suggestibility in response to questioning by adults, and the reliability of their accounts. The experts further reported that the complainant's behaviour was consistent with that of a sexually abused child, but did not state that the complainant had been abused. The clinical psychologist additionally stated that he had interviewed the complainant. The judge presented the same information as the experts without referring directly

to the behaviour of the child complainant. An example of the educative direction presented orally by the judge in the judicial summation at the end of the videotrial is attached as Appendix A (954 words). In the written mock-trial transcripts when the educative information was provided by an expert witness as evidence-in-chief and on cross-examination, it was transmitted in question and answer format in approximately 1980 words, but was conveyed more succinctly (about 574 words) when presented by the judge in a judicial direction (Goodman-Delahunty *et al.* 2010, 2011). Similarly, in the videotaped mock-trial lasting 40-55 minutes, depending on the experimental condition, educative information was provided orally by an expert witness (about 2,200 words) or orally by the judge (about 950 words) (Goodman-Delahunty *et al.* 2013, 2014, 2016).

3. Results of studies investigating juror child sexual abuse bias

3.1. Identifying juror child sexual abuse bias

Prior CSA studies of juror responses revealed that jurors were prone to acquit due to common misunderstandings and uncertainties about how to assess a child complainant's credibility (Cossins 2008). One common objective method to identify unwanted juror predispositions or bias is by means of a pretrial or posttrial juror questionnaire. Drawing on a review of the research literature, we designed a written questionnaire to test the extent to which community members endorsed common CSA misconceptions (Cossins *et al*. 2009, Goodman-Delahunty *et al*. 2010). The items covered several topics including the counter-intuitive behaviour of sexually abused children, children's suggestibility to questioning by adults, common offence characteristics, and the reliability of children's abuse reports, as shown in Table 3. Items were worded so that stronger agreement reflected a greater number of misconceptions. A subset of the items was negatively worded to avoid biased response styles, i.e., higher agreement indicated no misconception. Participants indicated their agreement on a rating scale (strongly disagree-strongly agree) without an option to indicate "do not know." Scores in the scale midpoint indicated uncertainty about a particular item. The questionnaire format in all studies was similar but the particular set of questions varied across studies in length (9–28 items) and type of rating-scale (5–7 points), as specified in Table 1. An 18-item questionnaire, shown in Table 3, was administered to several hundred jurors who presented for jury duty in New South Wales (NSW), Australia.

Table 3. CSA Knowledge Questionnaire (Goodman-Delahunty *et al.* 2016).

Reliability of children's testimony	
1	Memories of children aged 7-8 years about traumatic events are not as accurate as those of adults.
2	Children sometimes make false claims of sexual abuse to get back at an adult.*
3	Children who change their reports of sexual abuse were probably lying in the first place.
4	Children who cry when testifying are more likely to be truthful than children who do not.
5	A child who is very confident is more likely to be truthful than a child who is reticent or halting.
6	Inconsistencies in a child's report of sexual abuse indicate that the report is false.
7	A child's description of sexual abuse disclosed over time, with more details reported each time the child is interviewed, indicates a false description.
Susceptibility to suggestion by adults	
8	Children aged 7-8 years are easily manipulated to give false reports of sexual abuse.*
9	Children are easily coached to make false claims of sexual abuse.*
10	Repeating questions such as: "What happened? What else happened?" leads children to make false abuse claims.*
Children's reactions to sexual abuse	
11	A sexually abused child typically cries out for help and tries to escape.*
12	Children who are sexually abused display strong emotional reactions afterwards.*
13	A child victim of sexual abuse will avoid his or her abuser.*
14	Child victims of sexual abuse respond in a similar way to the abuse.*
15	A child who continues to spend time with the alleged abuser is unlikely to have been sexually abused.
16	A child who has been sexually abused will tell someone soon afterwards.
Offence characteristic	
17	A friend or family member is less likely to sexually abuse a child than a stranger.
Forensic evidence	
18	A medical examination almost always shows whether or not a child was sexually abused.*

Note. Correct answer is disagreement**.** *Included in the final 9-item CSA Knowledge Questionnaire.

3.2. Diagnostic item analyses and scale construction

Diagnostic item analyses were conducted to eliminate items that were less reliable in assessing common CSA misconceptions, errors, and uncertainties (Goodman-Delahunty *et al.* 2016). Where a substantial proportion of eligible jurors (over 45%) endorsed errors or expressed uncertainty about an item, the item was retained in the questionnaire. Examples of items in this category are displayed in Table 4.

Oñati Socio-legal Series, v. 6, n. 2 (2016), 283-314
ISSN: 2079-5971

Table 4. Pretest item analyses (Goodman-Delahunty *et al.* 2016).

Misconception	Uncertain	Error	Sum
Children who are sexually abused display strong emotional reactions afterwards.	33.5	44.8	78.3
A child victim of sexual abuse will avoid his or her abuser.	30.9	42.0	72.9
Children aged 7-8 years are easily manipulated to give false reports of sexual abuse.	40.0	25.4	65.4
A medical examination almost always shows whether or not a child was sexually abused.	26.1	27.2	53.3
Repeating questions such as: "What happened? What else happened?" leads children to make false abuse claims.	40.9	15.0	45.9

Note. Agreement denotes greater susceptibility to CSA misconceptions. Percent; *N* = 1656.

These item analyses revealed that positively worded items were more robust than negatively worded counterpart items, due to the instability of the latter in conjunction with the positive and negative poles of the rating scale. Negatively worded items appeared to confuse participants in using the scale, resulting in within-participant response inconsistencies and more extreme variation in responses (Goodman-Delahunty *et al.* 2016).

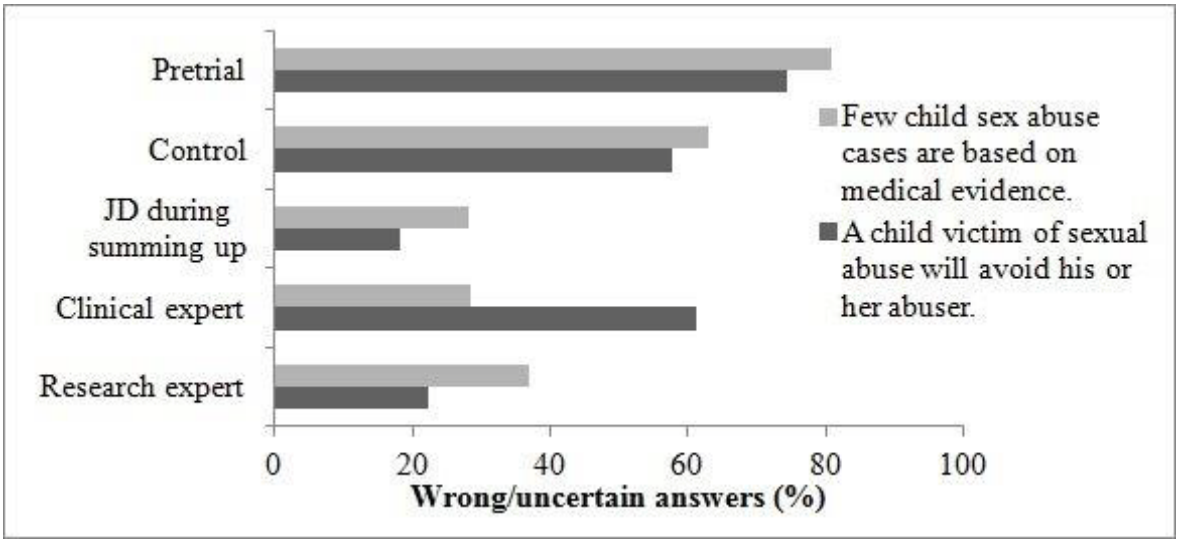
Next, factor analysis was undertaken (Goodman-Delahunty *et al.* 2016) to further refine and validate the robust questions. From the eighteen items displayed in Table 3, the nine-item Child Sexual Abuse Knowledge Questionnaire (CSA-KQ) was psychometrically derived. Correlational analyses yielded an absence of low inter-item correlations ($r < .25$) and multicollinearity ($r > .80$). Exploratory factor analysis yielded a two-factor structure comprised of nine items (Goodman-Delahunty *et al.* 2016, Study 1). Factor 1 represented the Impact of Sexual Assault on Children, including the psychological and physical consequences or lack thereof (five items). Factor 2 represented the Contextual Influences on Report (four items). Confirmatory factor analysis consolidated the factor structure and the predictive validity of the CSA Knowledge Questionnaire (Goodman-Delahunty *et al.* 2016, Study 2). The final nine-item questionnaire was used in a series of experimentally controlled interventions aimed at increasing jurors' CSA knowledge.

3.3. Pretest-posttest research designs

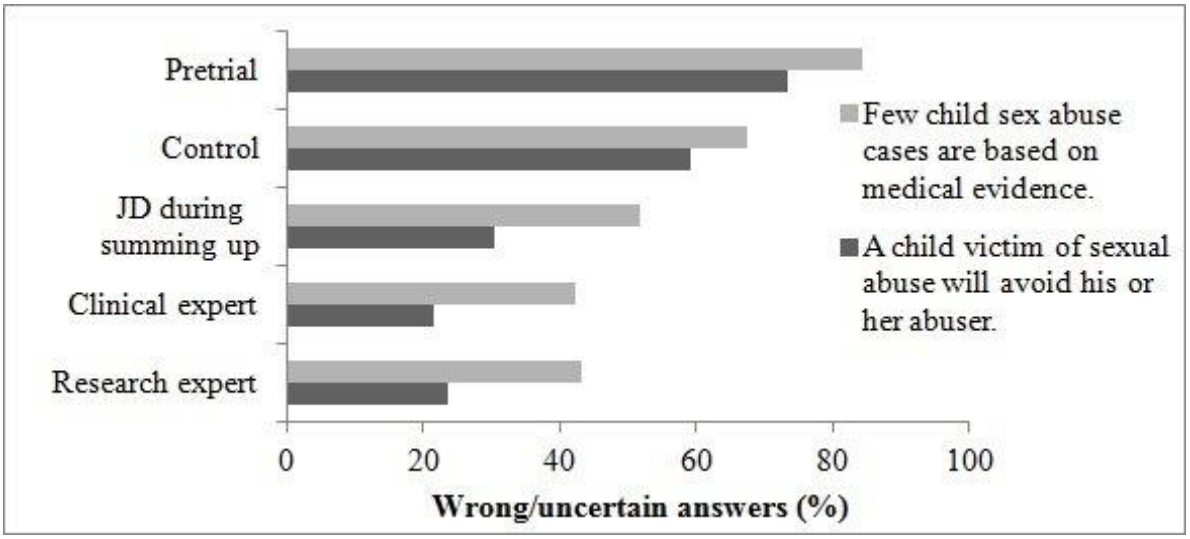
Participants completed the CSA Knowledge Questionnaire prior to viewing the trial materials and it was re-administered following exposure to one of a series of experimental interventions presented in the context of a simulated CSA trial to assess which intervention was most effective in educating NSW jurors on that topic. These interventions were tested in separate experimental studies yielding results cited here for illustrative purposes. Figure 1 shows pretest and posttest scores on two items extracted from the questionnaire to demonstrate changes in juror responses after exposure to either no educative intervention (Control) or one of three different interventions, with or without group deliberation: (a) specialized information presented by a clinical psychologist expert witness; (b) specialized information presented by an expert witness who was a research psychologist; or (c) a judicial direction presented during the judge's summing up (Goodman-Delahunty *et al*. 2014). All three educative interventions provided the jury the same information on these two topics.

Figure 1. Errors/uncertainties on two CSA knowledge items before and after exposure to educative trial interventions by an expert witness or the judge (JD) (%).

(a) Individual jurors



(b) Deliberating groups

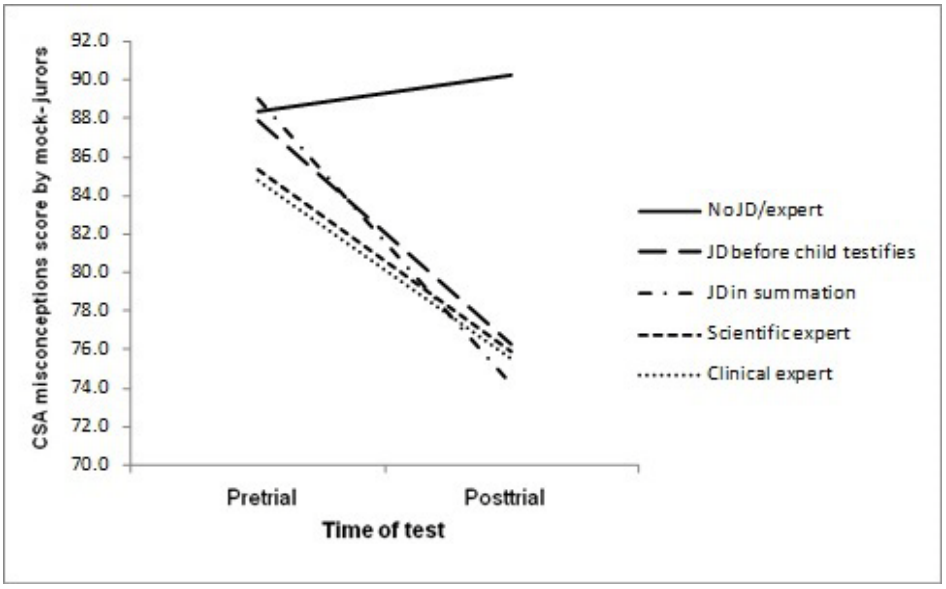


Aggregated questionnaire scores for jurors assigned to the experimental and control groups were compared to assess the effectiveness of the interventions in increasing jurors' CSA knowledge. Figure 2 depicts results of a comparison between responses from mock-jurors exposed to one of three intervention groups versus a control group that was purposefully not exposed to any educative CSA information.

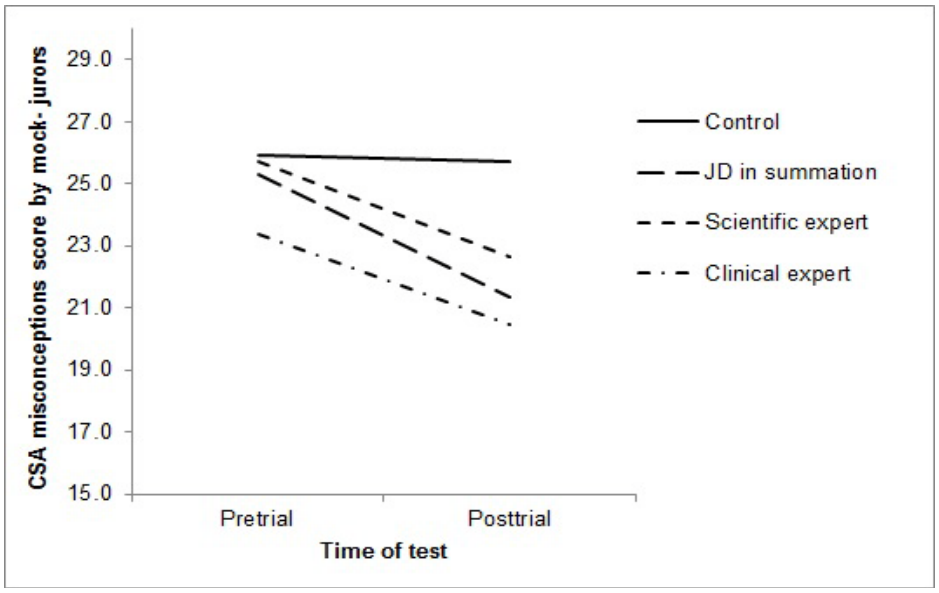
Oñati Socio-legal Series, v. 6, n. 2 (2016), 283-314
ISSN: 2079-5971

Figure 2. Comparison of overall mean pretest vs. posttest CSA misconception scores by intervention and type of simulated trial.

(a) Written trial simulation (26 items; Goodman-Delahunty *et al.* 2011)



(b) Videotrial simulation (9 items)



In both studies, the posttrial CSA knowledge of jurors in the control group remained unchanged whereas exposure to the educative CSA information significantly decreased jurors' CSA misconceptions (Goodman-Delahunty *et al.* 2011). The intervention in the form of a clinical psychological expert witness was the most successful in terms of average pretest versus posttest knowledge gains. In both the written and videotrial studies, the CSA misconception scores were lowest following exposure to the specialised information presented by the clinical expert. However, inspection of the jurors' average pretrial CSA misconception scores revealed that jurors who watched the videotrial with the clinical expert had started with significantly lower scores than those of other groups, as is shown in Figure 2b. Use of this experimental design allows the researchers to take those pretrial scores into account and control for any preexisting group differences so the effects of the interventions can be tested more precisely.

3.3.1. Interventions with written trial simulations

Using a pretest-posttest measure of mock-juror CSA knowledge, a study conducted with online written trial simulation materials demonstrated that all experimental groups were substantially equivalent at the outset, with mean pretest CSA knowledge scores in a similar range (Goodman-Delahunty *et al.* 2011). After the educative interventions, only the mock-jurors in the control group (no exposure to specialized CSA information) retained scores in that range. In all other groups, the CSA knowledge of mock-jurors increased significantly following exposure to the trial interventions (Goodman-Delahunty *et al.* 2011).

3.3.2. Interventions in the videotrial simulations

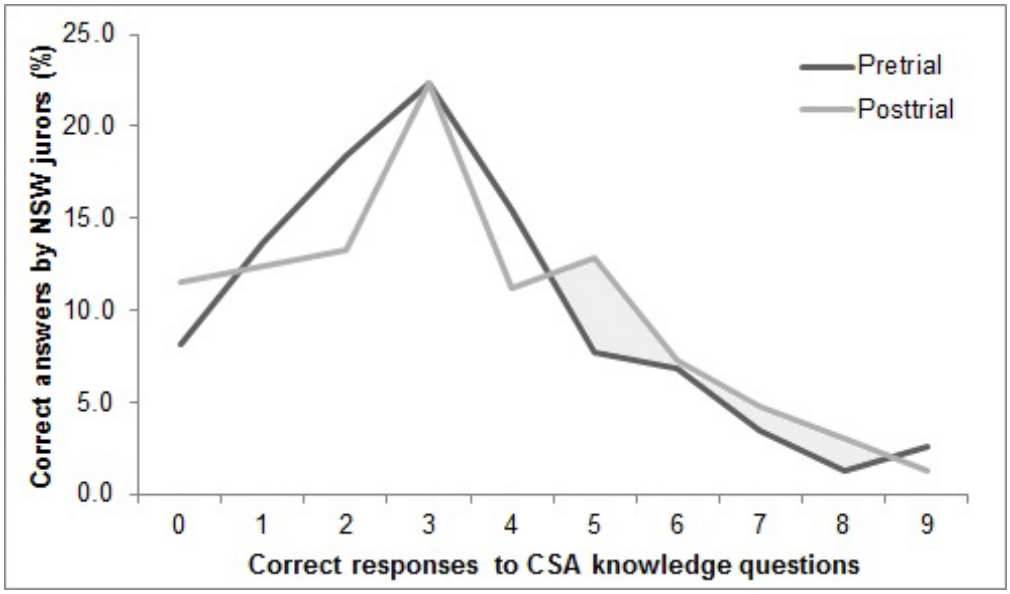
To test the effectiveness of the educative interventions in a more realistic trial simulation, using a videotrial, a study was designed in which the trial evidence was varied to create either weak or strong evidence (Goodman-Delahunty *et al.* 2013). In the weak version, inculpatory evidence was provided by the testimony of a child complainant aged 13 years (11 years old at the time of the offence). In the strong version, the same evidence was accompanied by testimony from a corroborating witness, the child's grandmother, who observed ambiguous conduct that was plausibly supportive of the child's allegations of sexual abuse by her grandfather.

Pretest and posttest CSA knowledge of jurors who were exposed to the specialized judicial direction or not after viewing either strong or weak cases, were compared to assess the interactions between evidence strength and the educative trial interventions. Responses to the CSA knowledge questions were recoded so that disagreement was a correct response (range: 0-9 correct responses, after the final 9-item questionnaire was applied to the data). No significant interaction emerged between case strength and time (pre- v. posttrial). However, there was a significant main effect for time showing higher CSA knowledge posttrial[2] than pretrial.[3] This effect was, however, moderated by the intervention. Without the intervention, the mean pretest and posttest correct responses to the CSA knowledge questions were not significantly different,[4] as shown in Figure 3a. By comparison, the number of correct posttrial responses of mock-jurors exposed to the trial intervention in the form of a special judicial direction increased significantly.[5] These results are depicted in Figure 3b. In sum, these results indicated that CSA knowledge increased in both weak and strong cases following exposure to the educative information. Conversely, CSA misconceptions persisted following exposure to both weak and strong evidence in the control groups (Goodman-Delahunty *et al.* 2013, see Figure 3a).

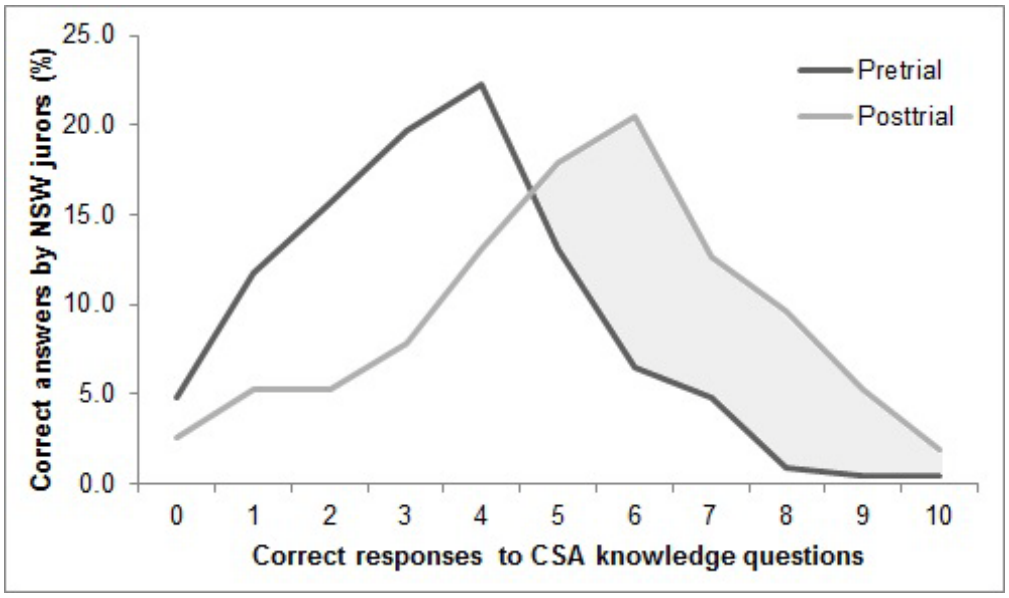
[2] $M = 4.21$, $SD = 2.39$
[3] $M = 3.29$, $SD = 1.97$, $p < .001$
[4] Pretrial: $M = 3.17$, $SD = 2.08$; posttrial: $M = 3.27$, $SD = 2.21$
[5] Pretrial: $M = 3.41$, $SD = 1.84$; posttrial: $M = 5.17$, $SD = 2.19$

Figure 3. Mean correct CSA item responses (percent) before and after trial simulation

(a) in the absence of any educative intervention



(b) with an educative judicial direction in summation



3.4. Applying the four-group pretest-posttest research design

A more complex randomized control group design was applied to test the impact of educative trial interventions on jurors' CSA knowledge independently of exposure to the pretest questionnaire, to ensure that observed posttrial improvements were not due to practice effects as a result of repeated exposure to the CSA knowledge questions.

3.4.1. Testing juror decisions as the unit of analysis

In particular, 661 NSW nonempanelled excused jurors were randomly assigned to one of the three experimental groups in the four-group pretest-posttest design as follows: (a) jurors completed the CSA knowledge questionnaire before and after viewing the professionally-acted video of a CSA trial in which specialized CSA knowledge was provided by an experimental psychologist (Experimental group 4 in

Table 5); (b) jurors watched the same video and completed the CSA knowledge questionnaire only after rendering a verdict (Experimental group 3 in Table 5); and (c) jurors completed the CSA Knowledge Questionnaire before and after viewing a videotrial containing no specialized CSA intervention (Experimental group 1 in Table 5).

Although a fully crossed four-group pretest-posttest design includes a control group without any intervention and a posttest measure only (Group 4 in Table 2), we found no practice effect of the questionnaire in the intervention groups 3 and 4 (see Table 5). The specialized CSA information presented by an expert witness, in this case, an experimental psychologist, significantly improved jurors' CSA knowledge independently of exposure to the pretrial CSA questionnaire. Due to logistical difficulties (winter court recess without jury trials) and high costs of extending data collection to add a further 100 participants from a limited pool of nonempanelled jurors, we omitted the above-mentioned control condition.

Table 5. Pretest-posttest mean (*SD*) CSA knowledge scores by (1) specialized CSA information presented by an expert experimental psychologist and (2) jury deliberation.

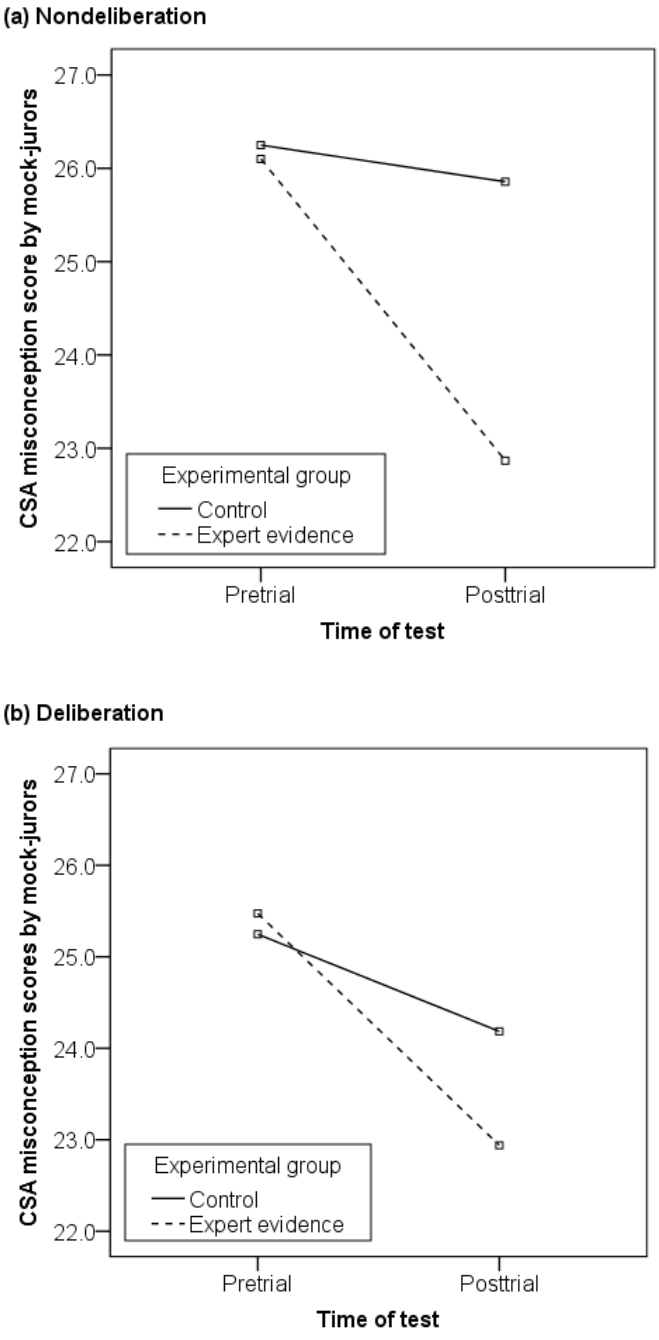
Experimental group	***N***	$M_{Pretest}$ (*SD*)	$M_{Posttest}$ (*SD*)
1. No intervention, no deliberation	214	26.29 (5.21)	25.97 (5.13)
2. No intervention, deliberation	214	25.26 (4.94)	24.26 (5.31)
3. Expert evidence, no deliberation, posttest only	227	--	22.87 (5.21)
4. Expert evidence, no deliberation	220	26.10 (5.10)	22.91 (5.11)
5. Expert evidence, deliberation	118	25.44 (5.30)	22.92 (5.24)

Results revealed a strong positive correlation between mock-jurors' pretrial and posttrial CSA knowledge,[6] and a significant posttrial decrease in juror CSA misconceptions.[7] However, the extent of jurors' CSA knowledge change (or gain) was dependent on the intervention.[8] CSA misconceptions did not change when jurors watched the videotrial without expert evidence presenting specialised CSA information,[9] and jurors' CSA misconceptions decreased significantly after exposure to the expert information,[10] as shown in Table 5. The effect of time of test and deliberation (Intervention 2) was not significant.[11] This was, however, moderated by exposure to the educative information from the expert witness (Intervention 1).[12] On the posttest, mock-jurors who deliberated with and without exposure to the expert, and individual mock-jurors who were exposed to the expert evidence without an opportunity for deliberation endorsed significantly fewer CSA misconceptions than mock-jurors in the control group without expert evidence or group deliberation, as is shown in Figure 4.

[6] $r = .69$, $p < .001$
[7] A mixed between-within participants analysis of variance, Wilks' Lambda = .84, $F(1, 751) = 148.63$, $p < .001$, $\eta^2_{partial} = .17$
[8] Wilks' Lambda = .93, $F(1, 751) = 53.03$, $p < .001$, $\eta^2_{partial} = .07$
[9] $M_{diff} = 0.73$, $SD = 3.75$
[10] $M_{diff} = 2.99$, $SD = 4.17$
[11] Wilks' Lambda = 1.00, $F(1, 751) = 0.00$, $p = .961$, $\eta^2_{partial} = .00$
[12] Wilks' Lambda = .99, $F(1, 751) = 5.33$, $p = .021$, $\eta^2_{partial} = .01$

Oñati Socio-legal Series, v. 6, n. 2 (2016), 283-314
ISSN: 2079-5971

Figure 4. Mean pretest vs. posttest CSA misconception scores by experimental group.



Notably, the issues associated with measuring knowledge change following exposure to an educative intervention are complex. Reliance on pretest-posttest changes in raw scores alone can be misleading. For example, observed changes may be due to test items that are (too) easy, and learning may take place simply from exposure to the questionnaire rather than the intervention. This source of "noise" was reduced by refinements to the questionnaire by omitting items that were answered correctly by a high proportion of NSW jurors, as described above. Similarly, when group deliberation is included, pretest-posttest changes can be dependent on social factors such as cohesion and group conformity, norms within a group deliberation, the status of the speakers and impact of one or more dominant persons in the group (Gordon 2015).

Aside from the two examples of questionnaire items included in Figure 1b, the foregoing findings examined the impact of educative interventions on individual jurors. Jury deliberation about the evidence is thought to provide an opportunity to correct juror errors and misconceptions before a group verdict is rendered (Marder and Hans 2015). To test this premise, and the impact of the interventions on the jury as unit of analysis (Devine *et al.* 2001, Devine 2012), group deliberation was added to the four-group pretest-posttest design. As is shown in Table 5, in two experimental conditions (Groups 2 and 5), jurors completed the CSA questionnaire before viewing the videotrial and after group deliberation. In one group, specialized CSA knowledge was provided by an experimental research psychologist; and in another experimental group, jurors viewed a videotrial without any specialized educative CSA intervention (Goodman-Delahunty *et al.* 2014).

Descriptive analyses revealed a similar pattern to that observed in nondeliberating individual NSW jurors. In other words, the CSA knowledge of deliberating jurors increased after exposure to a videotrial in which specialized CSA knowledge was presented by a psychologist, and CSA misconceptions increased somewhat after deliberations about a trial containing no educative specialized CSA information.

The nonindependence of the deliberation data violated prerequisites for analysis of variance or regression analyses of the individual jurors' posttest responses, both of which require independence of the data. Binomial tests in which differences of multiple observations (e.g., questionnaire items) of each of the groups are compared only allow pairwise comparisons (McGarthy and Smithson 2005), and are therefore unsuitable for analysing results of a more complex research design, as interaction effects would not be detected, and between- and within-group differences ignored. For this reason, statistical techniques such as multi-level modelling are necessary to separate treatment effects from trend effects. These models, based on linear regression for continuous dependent measures, or logistic regression for binary measures (such as verdict), allow analysis of nested designs and consideration of inter-individual and intergroup differences. To ascertain the cumulative impact of both educative trial interventions presenting specialized CSA knowledge plus group deliberation, changes in juror knowledge must be considered within juries as the unit of analysis by applying statistical techniques that take into account both individual differences (within-group level) and group differences (between-group level), using hierarchical or multilevel modelling (Raudenbush and Bryk 2002, Wright *et al.* 2011).

Using multilevel modelling, an analysis was conducted in which demographic covariates, pretrial CSA knowledge, intervention type (educative information provided by the trial judge in summation, an experimental psychologist, a clinical psychologist or no intervention) and decision type (individual juror v. jury) were taken into account within-groups as predictors of posttrial CSA knowledge (Goodman-Delahunty *et al.* 2014). The mean posttrial CSA knowledge score was entered into the model between-groups. Results revealed that demographic covariates were significant for gender[13] reflecting higher CSA knowledge scores following the mock-trial by female rather than male jurors; age[14] and age squared[15] demonstrating a monotonic decrease in posttrial CSA knowledge as juror age increased. Juror education level and type of decision were not associated with posttrial CSA knowledge. The levels of pretrial and posttrial CSA knowledge were significantly associated.[16] After controlling for juror gender, age, education, and pretrial CSA knowledge score, compared to the control group, on the posttrial tests, jurors and juries in all three intervention groups were significantly less likely to endorse CSA misconceptions. Interestingly, the clinical psychologist was most

[13] $b = -1.41$, $p < .01$
[14] $b = -0.03$, $p < .05$
[15] $b = 0.002$, $p < .05$
[16] $b = 0.75$, $p < .01$

Oñati Socio-legal Series, v. 6, n. 2 (2016), 283-314
ISSN: 2079-5971

effective in communicating CSA knowledge and reducing CSA misconceptions,[17] followed by a judicial direction,[18] and the experimental psychologist.[19]

3.4.2. Assessing the impact of juror bias on verdict

In the research program described here, a main effect of evidence strength emerged as anticipated. The mean conviction rate in response to trials including a corroborating witness was higher than that in the trial in which only the child complainant testified about abuse (see Figure 5). The educative intervention presented by the trial judge increased the conviction rate by approximately 5%, but this increase was not statistically significant (Goodman-Delahunty *et al.* 2013). Analysis of the mean conviction rate by proportion of correct items on the posttrial CSA knowledge questionnaire (high versus low score groups) revealed that jurors with fewer misconceptions in both the control and intervention groups voted to convict at about the same rate, whereas jurors who endorsed more misconceptions were significantly less likely to vote to convict, as shown in Figure 6. These findings confirmed the relationship between jurors' posttrial CSA knowledge and verdict.

Figure 5. Mean conviction rate as a function of evidence strength and judicial intervention in summation.

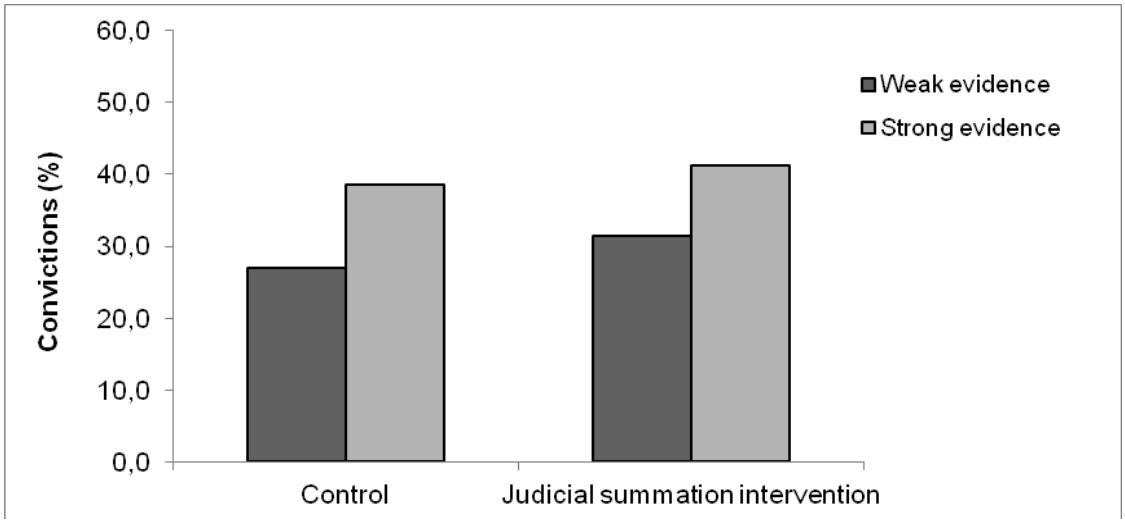
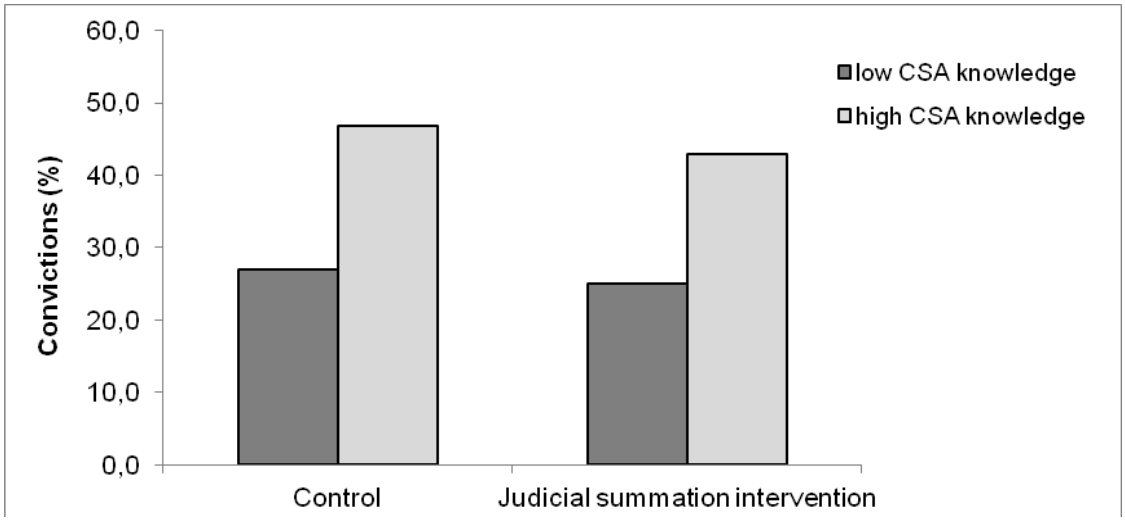


Figure 6. Mean conviction rate as a function of CSA knowledge.



[17] b = -3.871, p < .01
[18] b = -3.784, p < .01
[19] b = -2.855, p < .01

However, the results displayed in Figures 5 and 6 did not permit determinations regarding the most effective intervention to increase jurors' CSA knowledge and consequently, the conviction rate in cases of child sexual abuse. Notably, the results displayed in Figures 5 and 6 reflect individual juror decisions, without taking group deliberation into account.

3.4.3. Testing jury decisions as the unit of analysis

When group deliberation was considered in the forty-three juries included in these analyses, the mean conviction rates decreased in all intervention groups, but not the control group (Goodman-Delahunty *et al.* 2014), as is shown in Table 6.

Table 6. Mean individual juror conviction rate by intervention type and deliberation (percent)

Intervention	No deliberation	Deliberation
None (control)	32.6	31.7
Judicial direction in summation	44.3	24.8
Research expert	44.0	24.0
Clinical expert	54.7	32.8

The foregoing conviction rates are individual verdicts reported in the posttrial questionnaires completed by each juror, and did not correspond precisely with the jury verdicts returned on verdict forms at the conclusion of the group deliberation. (Dependent on intervention type, guilty verdicts reported by individual jurors varied between 0.2% and 9.1% from those of the deliberating juries.)

The observed conviction rates indicate that of the three interventions presenting educative CSA information, the clinical expert witness was more effective than both the research expert witness and a specially drafted judicial direction in the summation at increasing the conviction rate. However, when individual verdicts were gathered from jurors who deliberated to a verdict after exposure to any of these three educative interventions, the mean conviction rate was lower by approximately 20% in comparison with the mean verdicts reported by individual jurors who viewed the same videotrial interventions in the absence of group deliberation. In other words, the deliberation process itself reduced the beneficial effects of the educative interventions, irrespective of the type of intervention.

Further analyses were conducted with deliberating jurors using the mean CSA knowledge score, perceived complainant credibility (scored on the Witness Credibility Scale, Brodsky *et al*. 2010) and deliberation time for each jury (N = 443 jurors, 43 juries). A significant correlation emerged between juries' post-trial CSA knowledge scores and the perceived credibility of the complainant.[20] Accordingly, the higher the mean CSA knowledge of the jury as a group, the more likely the jury was to rate the complainant as credible. Perceived complainant credibility was significantly associated with deliberation time[21] such that juries perceiving the complainant as more credible spent more time deliberating about the case. No differences emerged in mean jury deliberation time between the different experimental groups.

3.4.4. Exploring verdict in intervention and control groups

A serial mediation analysis conducted to assess the impact of specific juror bias (posttrial CSA knowledge scores, reverse coded) on verdict, with the perceived

[20] r = .30, p = .050
[21] r = .53, p < .001

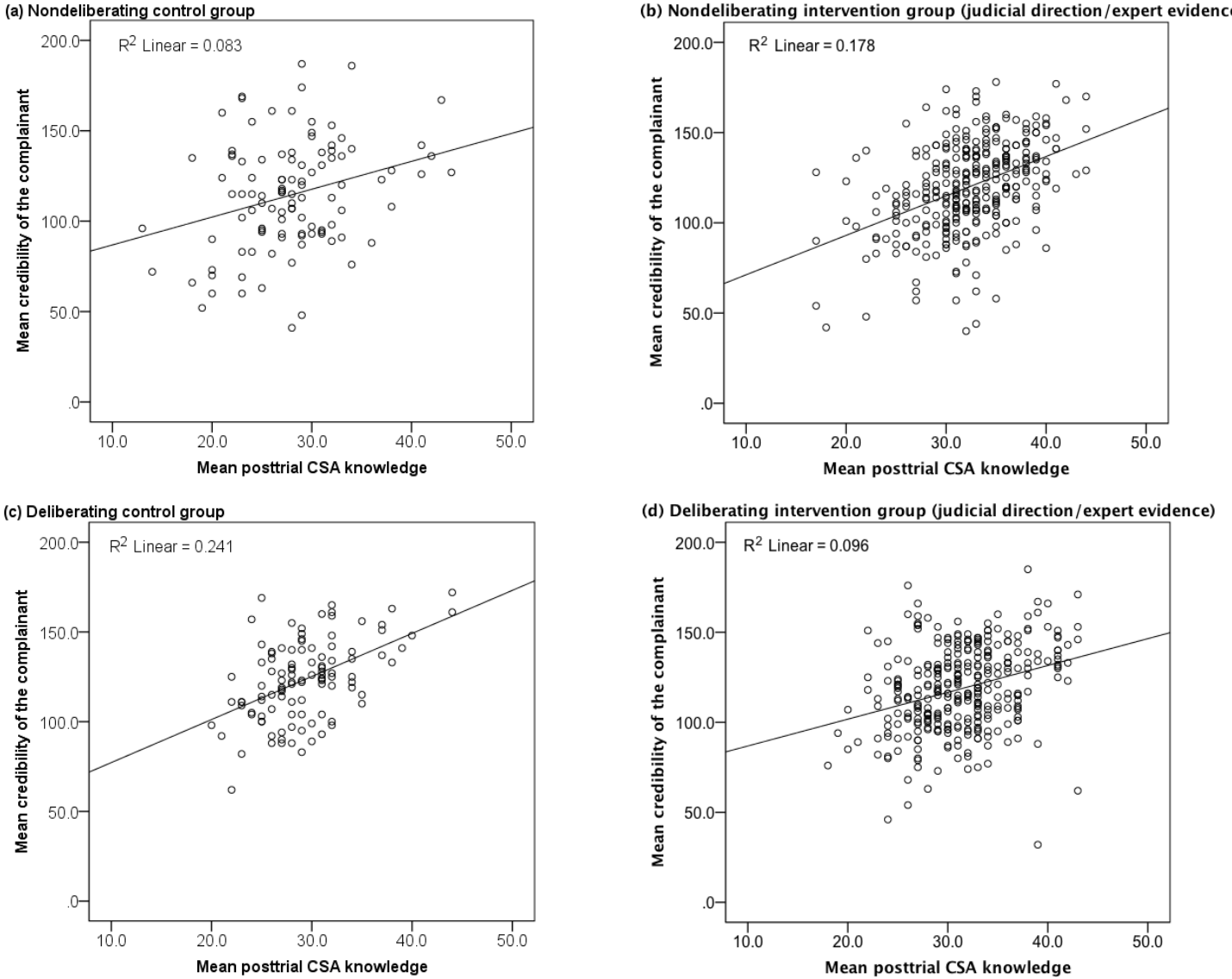
credibility of the child complainant and the corroborative witness reconfirmed that juror CSA knowledge at the close of the trial, after exposure to the educative intervention, was a significant predictor of a juror's vote to convict or acquit. Further, juror CSA knowledge increased perceived credibility of the child and corroborative witness. The assessed credibility of both witnesses, in turn, fully mediated the association between specific juror bias and verdict among the nondeliberating jurors and among the deliberating jurors. In other words, perceived witness credibility of the child complainant and the corroborative witness explained the association between posttrial CSA knowledge and verdict completely. These analyses replicated and extended prior mediation analyses based on individual juror decisions (in response to a different CSA trial) showing that jurors with more accurate CSA knowledge perceived the child complainant as more credible, and that the perceived credibility of the complainant mediated the effect between specific juror bias and verdict (Goodman-Delahunty *et al.* 2011).

Inspection of the jury verdicts (not guilty, guilty, and hung) revealed that juries voting not guilty had significantly lower mean CSA knowledge scores[22] than juries that voted guilty.[23] Further, juries voting to acquit rated the complainant as significantly less credible[24] than juries voting to convict[25] or hung juries,[26] and they also spent less time deliberating about the case.[27] These results indicate that consensus on the jury verdict was reached more swiftly by juries that discredited the complainant. In contrast, juries that perceived the complainant as more credible spent more time discussing the case, possibly trying to reach consensus about the credibility of the complainant before voting to convict or reporting that they were unable to reach a unanimous verdict in the available deliberation time.

To assist in understanding the effect of deliberation on verdict following exposure to the experimental interventions, scatter plots were examined to explore the association between mean posttrial CSA knowledge scores and the perceived credibility of the child complainant as a function of decision type (nondeliberation vs. deliberation) and intervention (control group vs. interventions in the form of expert evidence or a judicial direction). Figure 7 displays scatter plots with a regression line in each of four experimental groups. In the control groups, these analyses revealed a stronger relationship between posttrial CSA knowledge and the perceived credibility of the complainant among deliberating (Figure 7c) rather than nondeliberating jurors (Figure 7a).[28] This effect, however, was reversed when jurors received specialized educative information about CSA (Figures 7b and 7d) such that the correlation tended to be stronger among nondeliberating than deliberating jurors.[29]

[22] M = 29.78, SD = 1.68, p < .001
[23] M = 32.66, SD = 2.32
[24] M = 112.27, SD = 9.59
[25] M = 121.38, SD = 7.58, p < .001
[26] M = 124.74, SD = 12.25, p = .004
[27] Not guilty: M = 30.51, SD = 18.08, Guilty: M = 48.66, SD =18.36, Hung jury: M = 63.97, SD = 17.71
[28] Fisher's z = -2.190, p = .029.
[29] Fisher's z = 1.770, p = .077.

Figure 7. Scatter plots for the association between mean posttrial CSA knowledge and mean complainant credibility



Oñati Socio-legal Series, v. 6, n. 2 (2016), 283-314
ISSN: 2079-5971

4. Discussion

The foregoing programmatic studies illustrate some of the intricacies in identifying and testing the impact of potential jury biases in CSA cases, and the effectiveness of traditional legal safeguards to manage and minimize the effects of these biases at the level of the juror and at the level of the jury. The study outcomes disclose that many jurors endorsed CSA biases controverted by empirical findings, and that jurors' CSA knowledge predicted their credibility assessments of the complainants and their verdicts.

4.1. Impact of interventions to reduce juror CSA biases

In the context of jury simulation studies using a pretest-posttest research paradigm, administered via written and videotaped modes of presentation, jurors with high CSA misconceptions at the pretest stage benefited more from the educative interventions presented in the course of a simulated trial. Conversely, jurors who achieved high knowledge scores at the pretest stage learned less from the educative trial interventions than jurors with low pretest knowledge, as those in the former group were already familiar with the information presented by the experts or the judge in the educative intervention. Although legislation enabling trial innovations such as the expert evidence presented in these studies was implemented in Australia (Cossins 2013), the uptake of expert evidence in real trials has been negligible. The more persuasive force of an educative intervention presented by an expert witness identified as a clinical rather than a research psychologist replicated findings in prior jury simulation research conducted in the United States of America on other topics (Krauss *et al.* 2004, 2011), and in child sexual abuse cases (Gabora *et al.* 1993). This effect may be attributable to the greater familiarity of many jurors with medical rather than research experts, or the perception that the opinion of an expert who met with the child complainant in the case at hand was more reliable than one who did not.

4.2. The importance of control groups and deliberation in jury research

The most important lesson from the implementation of the pretest-posttest research design applied in this research program was the value of including (a) control groups that might seem unnecessary and costly, especially if they do not reflect legal practicalities; and (b) deliberation and analysis of jury in addition to juror decisions (Waters and Hans 2009). Without the inclusion of the deliberating control group that was not exposed to any of the educative interventions, the researchers might erroneously have accounted for the reduced rate of convictions following deliberation on grounds that deliberation (a) provided an opportunity for misconceptions stimulated by the cross-examination of the child complainant to be aired and reinforced; (b) prompted jurors to apply a higher threshold of proof due to their focus on the judicial direction that the prosecution must establish all elements of the offence beyond reasonable doubt (Wright and Hall 2007); or (c) produced a "leniency asymmetry" effect (Kerr and MacCoun 2012). However, none of these explanations was viable as all were inconsistent with the finding that that the acquittal rate following deliberation by jurors in the control group did not increase. Although social interactions and group dynamics within a deliberation group, jurors' personal motivation and experience, and the impact of a dominant juror on the decision process (Gordon 2015) are some additional factors that may have contributed to the observed disparity in deliberation effects, the pattern of observed verdicts in the control group tended to indicate that exposure to the interventions distinguished decision making in those juries from that of their counterparts in the control group juries.

In light of the Heuristic-Systematic Dual Process Model, one possible explanation for the increase in acquittals by juries exposed to educative interventions following deliberation is the added burden and complexity of the specialized educative

informational content from the expert witnesses or the judge, increasing their cognitive load and prompting more peripheral and heuristic assessments of the credibility of the complainant in deliberation. In the limited time available to reach a unanimous verdict, these jurors may have resorted to less thorough and less elaborated decision strategies. By comparison, the cognitive load on the deliberating jurors in the control group juries was lower, as they had fewer witnesses to consider and less information to integrate when evaluating the child's credibility, and were not faced with the further task of modifying their preexisting beliefs and attitudes to take on new information that controverted common jury CSA misconceptions.

A content analysis of group deliberations that will provide more insight into the group processes in these juries is beyond the scope of this paper. Detailed content analyses of deliberations will examine the extent of central versus heuristic and peripheral processing in the sixty-five intervention and control group juries to assess this potential explanation for differences in the observed effects of deliberation. Application of the story model of jury decision making (Hastie *et al.* 1983, Pennington and Hastie 1993, Penrod and Hastie 1979) may reveal that the narratives developed by deliberating juries placed more emphasis on their pretrial knowledge and their expectations and experiences thus, interfering with their reliance on the educative interventions (Smith 1993, MacKillop and Vidmar 2015). Further research is required to account for the effect of deliberation following the interventions and to discern ways to maintain the individual jurors' knowledge gains in the deliberation process (Ellsworth 1989).

4.3. Limitations of the studies

The pretest-posttest trial simulation research program described above is subject to a number of limitations. Chief among these is the truncated nature of trial simulations in comparison with the length of actual CSA trials. Some of the verdicts may have changed if jurors were afforded additional deliberation time. Moreover, these jurors deliberated as a group without an extensive period together in advance of the deliberations to get better acquainted, as would be the case of jurors serving alongside each other in a more protracted real trial.

Although both the written and videotrial versions presented the trial information to the mock-jurors in the same manner and sequence as is used in the course of an actual trial, the presentation was condensed to make it feasible to secure participants for 60-90 minutes in the versions administered online, and for up to four hours in-person. Even though online studies have many advantages, saving time, space and costs, the generalizability of the findings is limited by the lack of realism in the trial simulations, lack of control over the circumstances of participation (e.g., environment, third-party influence, time spent reading/answering), and possible self-selection bias of the participants. The time commitment by jurors who participated in the videotrial simulations was on average three to four hours, taking into account completion of the pre-trial questionnaire, the trial simulation, group deliberation and completion of a posttrial questionnaire. This time commitment and the use of a videotaped trial presentation conveying the information orally to jurors made the jury task more realistic than that in many experimental trial simulations. Nonetheless, the relatively brief deliberation period coupled with awareness by participants in these trial simulations that their verdicts had no real consequences for the accused may have affected their decisions. Although experimental simulations may reflect ecological verisimilitude in numerous respects, they will inevitably fail to replicate the trial experience.

4.4. Implications of the findings

Interventions that aim to sensitize jurors to the trial issues and thereby avoid or reduce jury error or bias deserve more consideration. The least controversial

Oñati Socio-legal Series, v. 6, n. 2 (2016), 283-314
ISSN: 2079-5971

approach is to apply interventions compatible with existing legal procedures, such as *voir dire*, cross-examination, expert witness testimony, judicial instructions, and deliberation guidance. The comparative effectiveness of such interventions has rarely been systematically evaluated, in part because intervention procedures suited to one legal system may not be available in another, or because the same available legal safeguard may not be implemented in the same way across jurisdictions. For example, a sizable literature documents the gap between lay and expert knowledge on factors surrounding potential inaccuracies in eyewitness memory (Benton *et al.* 2006). In the United States, to increase juror sensitivity to these factors and thus avoid miscarriages of justice in cases that rely on incriminating eyewitness testimony against the accused, expert evidence on eyewitness memory is often admitted (Pezdek 2007), and in some jurisdictions, its exclusion constitutes an abuse of judicial discretion.[30] Presenting jurors with specialized knowledge via expert evidence on law enforcement procedures known to reduce eyewitness errors has been shown to reduce mock-jury perceptions of defendant culpability and conviction rates (Lampinen *et al.* 2005). However, admissibility criteria for expert evidence in some common-law jurisdictions, such as those in Australia (Queensland, South Australia, Western Australia) and New Zealand, may preclude this type of intervention on the basis that these matters come within the "common knowledge" of the jury members (Freckelton 2014, Freckelton *et al.* 2016).

Detailed analyses of preexisting juror knowledge and juror responses to different types of educative interventions such as those outlined in this paper can be used to construct a profile of what the average jurors know and do not know about a specialized topic that is a key issue in a criminal trial. Results of these analyses can be used by judges, lawyers and experts to target and refine the trial presentations to jurors in the context of a criminal trial on this topic. For example, evidence can be aimed at jurors with the lower levels of this knowledge, to bring their familiarity with the critical subject matter to the same level as that of their more sophisticated counterparts who achieve higher knowledge scores at the outset. By conducting analyses of this nature to assess knowledge deficits in a group of jury-eligible citizens about a specific topic, adjustments can be made to content to target specialized information presented by an educative intervention (the "dose") as well as the level of detail of the educative information (the intensity of the intervention). In the example presented in this paper, an educative intervention in the form of in-person expert testimony from a clinical psychologist was most effective. Jury research on effective methods to convey expert evidence to benefit tribunals of fact in cases requiring specialized knowledge is compatible with efforts by courts and legal counsel to ensure that only high quality, reliable expert evidence is admitted at trial (Jurs 2012, Edmond 2015).

4.5. Conclusion

When lay jurors are susceptible to common misconceptions about a key issue involving specialized knowledge in a criminal trial, such as child sexual abuse, this may lead to a failure of justice in the form of wrongful acquittal in the absence of an intervention by the parties or the court to educate jurors on that issue. These outcomes can be minimised by applying findings derived from a systematic empirical research program to assess the scope of relevant jury bias, by devising an effective educative intervention, and testing the impact of that intervention on jury knowledge and case outcomes. The foregoing worked example of jury CSA knowledge illustrated how insights gained by applying a pretest-posttest methodology in a programmatic research program can be used to develop recommendations to judges, legal advocates and policy-makers by identifying

[30] E.g., State v. Chapple, 660 P.2d 1208 (AZ. 1983); Ex parte Williams, 594 So.2d 1225 (AL. 1992); People v. Campbell, 847 P.2d 228 (CO. App. 1992); State v. Whaley, 406 S.E.2d 369 (SC 1991); Echavarria v. State, 839 P.2d 589 (NV. 1992); State v. Copeland, 226 S.W.3d 287 (TN 2007).

sources of the risks and the most effective measures to reduce the risk of an unfair trial.

References

Australian Institute of Judicial Administration Committee, 2015. *Benchbook for children giving evidence in Australian courts* [online]. Melbourne, Victoria: Australian Institute of Judicial Administration. Available from: https://www.aija.org.au/Child%20Witness%20Bch%20Bk/Child%20Witness%20BB%20Update%202015.pdf [Accessed 10 May 2016].

Bonate, P.L., 2000. *Analysis of pretest-posttest designs.* Boca Raton: CRC Press.

Benton, T.R., *et al.,* 2006. Eyewitness memory is still not common sense: comparing jurors, judges and law enforcement to eyewitness experts. *Applied Cognitive Psychology,* 20 (1), 115-129.

Bornstein, B.H., and Greene, E., 2011. Jury decision making: implications for and from psychology. *Current Directions in Psychological Science,* 20 (1), 63-67.

Brodsky, S.L., Griffin, M.P., and Cramer, R.J., 2010. The witness credibility scale: an outcome measure for expert witness research. *Behavioral Sciences and the Law,* 28 (6), 892-907.

Chaiken, S., Liberman, A., and Eagly, A.H., 1989. Heuristic and systematic information within and beyond the persuasion context. *In*: J.S. Uleman and J.A. Bargh, eds. *Unintended thought*. New York: Guilford Press, 212-252.

Chaiken, S., and Trope, Y., 1999. *Dual process theories in social psychology.* New York: Guilford Press.

Chesterman, M., 1999. Criminal trial juries in Australia: from penal colonies to a federal democracy. *Law and Contemporary Problems* [online], 62 (2), 69-102. Available from: http://scholarship.law.duke.edu/lcp/vol62/iss2/4 [Accessed 10 May 2016].

Clark, R.C., Nguyen, F., and Sweller, J.K., 2006. *Efficiency in learning: evidence-based guidelines to manage cognitive load*. San Francisco: Pfeiffer.

Cossins, A., 2008. Children, sexual abuse and suggestibility: what laypeople think they know and what the literature tells us. *Psychiatry, Psychology and Law,* 15 (1), 153-170.

Cossins, A., 2013. Expert witness evidence in sexual assault trials: questions, answers and law reform in Australia and England. *International Journal of Evidence and Proof,* 17 (1), 74-113.

Cossins, A., and Goodman-Delahunty, J., 2013. Misconceptions or expert evidence in child sexual assault trials: enhancing justice and jurors' 'common sense'. *Journal of Judicial Administration,* 22 (4), 171-190.

Cossins, A., Goodman-Delahunty, J, and O'Brien, K., 2009. Uncertainty and misconceptions about child sexual abuse: implications for the criminal justice system. *Psychiatry, Psychology and Law,* 16 (3), 435-452.

Delli Carpini, M.X., Cook, F.L., and Jacobs, L.R., 2004. Public deliberations, discursive participation and citizen engagement: a review of the empirical literature. *Annual Review of Political Science,* 7 (1), 315-344.

Devine, D.J., 2012. *Jury decision making.* New York University Press.

Devine, D.J., *et al.,* 2001. Jury decision making: 45 years of empirical research on deliberating groups. *Psychology, Public Policy, and Law,* 7 (3), 622-727.

Dimitrov, D.M., and Rumrill, P.D., 2003. Pretest-posttest designs and measurement of change. Work, 20 (2), 159-165.

Donnelly, H., 2007. Delay and the credibility of complainants in sexual assault proceedings. *Judicial Officers' Bulletin,* 19 (3), 17.

Edmond, G., 2015. Forensic science evidence and the conditions for rational (jury) evaluation. *Melbourne University Law Review* [online], 39 (1), 77-127. Available from: https://www.researchgate.net/publication/283802530_Forensic_science_evidence_and_the_conditions_for_rational_Jury_evaluation [Accessed 10 May 2016].

Edmond, G., *et al*., 2013. Admissibility compared: the reception of incriminating expert evidence (i.e., forensic science) in four adversarial jurisdictions. *University of Denver Criminal Law Review* [online], 3 (1), 31-109. Available from: http://www.law.unsw.edu.au/sites/law.unsw.edu.au/files/docs/posts/admissibility_compared_2013_reduced.pdf [Accessed 10 May 2016].

Eisenberg, T., *et al.,* 2005. Judge-jury agreement in criminal cases: a partial replication of Kalven and Zeisel's 'The American Jury.' *Journal of Empirical Legal Studies,* 2 (1), 171-207.

Ellsworth, P.C., 1989. Are twelve heads better than one? *Law and Contemporary Problems,* 52 (4), 205-224.

Fitzgerald, J., 2006. The attrition of sexual offences from the New South Wales criminal justice system. *Contemporary Issues in Crime and Justice* [online], 92. Available from: http://www.austlii.edu.au/au/journals/NSWCrimJustB/2006/3.pdf [Accessed 10 May 2016].

Freckelton, I., 2014. Admissibility of expert opinions on eyewitness evidence: international perspectives. *Psychiatry, Psychology and Law,* 21 (6), 821-836.

Freckelton, I., *et al.*, 2016. *Expert evidence and criminal jury trials*. Oxford University Press.

Gabora, N.J., Spanos, N.P., and Joab, A., 1993. The effects of complainant age and expert psychological testimony in a simulated child sexual abuse trial. *Law and Human Behavior,* 17 (1), 103-119.

Gastil, J., *et al.,* 2010. *The jury and democracy: how jury deliberation promotes civic engagement and political participation*. New York: Oxford University Press.

Gastil, J., and Xenos, M., 2010. Of attitudes and engagement: clarifying the reciprocal relationship between civic attitudes and political participation. *Journal of Communication* [online], 60 (2), 318-343. Available from: http://onlinelibrary.wiley.com/doi/10.1111/j.1460-2466.2010.01484.x/pdf [Accessed 10 May 2016].

Goodman-Delahunty, J., 2015. The jury box and the urn: Containing our expectations. *Pandora's Box: Crime, Justice and the People* [online], 9-16. Available from: http://static1.squarespace.com/static/53d069cde4b05f100c854faf/t/561f8ff9e4b00e50405dfd7b/1444909049765/Pandoras+Box+2015.pdf [Accessed 10 May 2016].

Goodman-Delahunty, J., *et al.*, 2008. *Practices, policies and procedures that influence juror satisfaction in Australia* [online]. Canberra, ACT: Australian Institute of Criminology. Available from: http://www.aic.gov.au/media_library/publications/rpp/87/rpp087.pdf [Accessed 10 May 2016].

Goodman-Delahunty, J., Cossins, A., and Martschuk, N., 2013. *Countering misconceptions in child sexual abuse cases with judicial directions.* Poster presentation at the Annual Meeting of the American Psychology-Law Society, 7-9 March 2013 Portland, USA (on file with authors).

Goodman-Delahunty, J., Cossins, A., and Martschuk, N., 2014. Persistent misconceptions about child sexual abuse: the impact of specialized educative information and deliberation in a simulated trial. *In*: K.C. Haung, ed. *Proceedings of the 3rd International Conference on Empirical Studies of Judicial Systems: Citizen participation around the world*, 5-6 September 2014 Taipei, Taiwan. Taipei: Academia Sinica.

Goodman-Delahunty, J., Cossins, A., and O'Brien, K., 2010. Enhancing the credibility of complainants in child sexual assault trials: the effect of expert evidence and judicial directions on. *Behavioural Sciences and the Law,* 28 (6), 769-783.

Goodman-Delahunty, J., Cossins, A., and O'Brien, K., 2011. A comparison of expert evidence and judicial directions to counter misconceptions in child sexual abuse trials. *Australian and New Zealand Journal of Criminology,* 44 (2), 196-217.

Goodman-Delahunty, J., Martschuk, N., and Cossins, A., forthcoming 2016. Validation of the Child Sexual Abuse Knowledge Questionnaire. Manuscript under revision for *Psychology, Crime & Law*.

Goodman-Delahunty, J., and Tait, D., 2006. Lay participation in legal decision making in Australia and New Zealand: jury trials and administrative tribunals. *In*: M.F. Kaplan and A.M. Martin Rodriguez, eds. *Understanding world juries through psychological research.* New York: Psychology Press, 47-71.

Gordon, S., 2015. All together now: using principles of group dynamics to train better jurors. *Indiana Law Review* [online], 48 (2), 415-459. Available from: http://mckinneylaw.iu.edu/ilr/pdf/vol48p415.pdf [Accessed 10 May 2016].

Hastie, R., Penrod, S.D., and Pennington, N., 1983. *Inside the jury*. Cambridge, UK: Harvard University Press.

Jurs, A.W., 2012. Balancing legal process with scientific expertise: a comparative assessment of expert witness methodology in five nations, and suggestions for reform of post-*Daubert* U.S. reliability determinations. *Marquette Law Review,* 95 (4), 1329-1415. Available from: http://scholarship.law.marquette.edu/mulr/vol95/iss4/12/ [Accessed 10 May 2016].

Kaplan, M.F., and Martin, A.M., eds., 2006. *Understanding world juries through psychological research*. New York: Psychology Press.

Kerr, N.L., and MacCoun, R.J., 2012. Is the leniency asymmetry really dead? Misinterpreting asymmetry effects in criminal jury deliberation. *Group Processes and Intergroup Relations,* 15 (5), 585-602.

Kim, S., *et al.,* 2013. Judge-jury agreement in criminal cases: the first three years of the Korean jury system. *Journal of Empirical Legal Studies,* 10 (1), 35-53.

Krauss, D.A., McCabe, J.G., and Lieberman, J.D., 2011. Dangerously misunderstood: Representative jurors' reactions to expert testimony on future dangerousness in a sexually violent predator trial. *Psychology, Public Policy, and Law,* 18 (1), 18-48.

Krauss, D.A., Lieberman, J.D., and Olsen, J., 2004. The effects of rational and experiential information processing of expert testimony in death penalty cases. *Behavioral Sciences and the Law,* 22 (6), 801-822.

Lampinen, J.M., *et al.,* 2005. The reactions of mock jurors to the Department of Justice Guidelines for the Collection and Preservation of Eyewitness Evidence. *Basic and Applied Social Psychology,* 27 (2), 155-162.

Lempert, R., 2015. The American jury system: a synthetic overview. *Chicago-Kent Law Review* [online], 90 (3), 825-859. Available from:

http://scholarship.kentlaw.iit.edu/cklawreview/vol90/iss3/4/ [Accessed 10 May 2016].

MacCoun, R.J., 1989. Experimental research on jury decision making. *Science,* 244 (4908), 1046-1050.

MacKillop, K., and Vidmar, N., 2015. Decision-making in the dark: how pre-trial errors change the narrative in criminal justice trials. *Chicago-Kent Law Review* [online], 90 (3), 957-980. Available from: http://scholarship.kentlaw.iit.edu/cklawreview/vol90/iss3/8/ [Accessed 10 May 2016].

Marder, N.S., and Hans, V.P., 2015. Introduction to juries and lay participation: American perspectives and global trends. *Chicago-Kent Law Review* [online], 90 (3), 789-824. Available from: http://scholarship.kentlaw.iit.edu/cklawreview/vol90/iss3/3/ [Accessed 10 May 2016].

McClellan, P., 2008. *The Australian justice system in 2020*. Presentation at the National Judicial College of Australia, Australian National University, 25 October 2008 Canberra, Australia. Available from: http://www.austlii.edu.au/au/journals/NSWJSchol/2008/23.pdf [Accessed 10 May 2016].

McGarthy, C., and Smithson, M., 2005. Independence and nonindependence: a simple method for comparing groups using multiple measures and the binomial test. *European Journal of Social Psychology,* 35 (2), 171-180.

Myers, D.G., and Lamm, H., 1976. The group polarization phenomenon, *Psychological Bulletin,* 83 (4), 602-627.

Nolan, M., and Goodman-Delahunty, J., 2015. *Law and psychology in Australia*. Pyrmont: Lawbook, Thompson Reuters.

Ogloff, J.P., *et al.,* 2007. Enhancing communication with Australian and New Zealand juries: a survey of judges. *Journal of Judicial Administration,* 16 (4), 235-255.

Oliver, J., 2006. The legislation changed, what about the reality? *Queensland University of Technology Law and Justice Journal,* 6 (1), 55–65.

Pennington, N., and Hastie, R., 1986. Evidence evaluation in complex decision making. *Journal of Personality and Social Psychology,* 51 (2), 242-258.

Pennington, N., and Hastie, R., 1993. The story model for juror decision-making. *In:* R. Hastie, ed. *Inside the juror: the psychology of juror decision making.* New York: Cambridge University Press, 192-221.

Penrod, S., and Hastie, R., 1979. Models of jury decision making: a critical review. *Psychological Bulleting,* 86 (3), 462-492.

Petty, R.E., and Cacioppo, J.T., 1986. The Elaboration Likelihood Model of persuasion. *In*: L. Berkowitz, ed. *Advances in experimental social psychology*. New York: Academic Press, 123-205.

Pezdek, K., 2007. Expert testimony on eyewitness memory and identification. *In*: M. Costanzo, D. Krauss, and K. Pezdek, eds. *Expert psychological testimony for the courts.* Mahwah, New Jersey: Lawrence Erlbaum Associates, 99-117.

R v Tait, [1992] 2 NZLR 666.

Ransley, J., Dioso-Villa, R., and Goldsmith, A.J., 2012. Failures of justice. *In:* M. Marmo, W. de Lint and D. Palmer, eds. *Crime and justice a guide to criminology.* 4th ed. Pyrmont, NSW: Lawbook, 609-630.

Raudenbush, S.W., and Bryk, A.S., 2002. *Hierarchical linear models: applications and data analysis methods*. 2nd ed. Newbury Park, CA: Sage.

Smith, V.L., 1993. When prior knowledge and law collide: helping jurors use the law. *Law and Human Behavior,* 17 (5), 507-536.

Smith, E.R., and DeCoster, J., 2000. Dual process models in social and cognitive psychology: conceptual integration and underlying memory systems. *Personality and Social Psychology Review,* 4 (2), 108-131.

Tabak, S.J., and Klettke, B., 2014. Mock jury attitudes towards credibility, age, and guilt in a fictional child sexual abuse scenario. *Australian Journal of Psychology,* 66 (1), 47-55.

van Koppen, P.J., 2009. The diversity of nations and legal systems: contrasting the Dutch and the Americans. *In*: M.E. Oswald, S. Bieneck, and J. Hupfeld-Hienemann, eds. *The social psychology of punishment of crime*. Chichester, UK: John Wiley & Sons, Ltd, 3-17.

Waters, N.L., and Hans, V.P., 2009. A jury of one: opinion formation, conformity, and dissent on juries. *Journal of Empirical Legal Studies,* 6 (3), 513-540.

Wright, D.B., and Hall, M., 2007. How a "reasonable doubt" instruction affects decisions of guilt. *Basic and Applied Social Psychology,* 29 (1), 91-98.

Wright, D.B., Strubler, K.A., and Vallano, J.P., 2011. Statistical techniques for juror and jury research. *Legal and Criminological Psychology,* 16 (1), 90-125.

Appendix A. Sample judicial direction conveying specialised educative knowledge to the jury in the oral judicial summation

Because the complainant in this trial is a child under the age of 16 years of age and because she testified about a sexual offence, I am required by law to give you the following additional instructions. These instructions are about children's responses to sexual abuse.

It is my duty to tell you that a child will not react to sexual abuse in the ways that an adult reacts in such a situation. Children who have been sexually abused commonly delay reporting the abuse. Most children do not resist or try to escape from the abuser. It is common for the child to continue to feel and show affection for the abuser after the abuse has occurred, especially where there is an established relationship of love and trust between them. These behaviors may appear to be counter-intuitive to the adult layperson because of their expectations about the way a victim of sexual abuse should react.

Not all children who have been sexually abused will display the same symptoms or behaviours. There is no one set of symptoms or behaviours that all children who have been sexually abused will exhibit. A child who has not been sexually abused may also have some of the symptoms and behaviours present in an abused child.

It is common for child victims of sexual abuse not to tell anyone about the fact that they have been abused. In fact, a majority of children do not report the abuse immediately, within one month, or within the first year of the abuse. A lapse of time between the abuse and the report is a common occurrence.

The reasons for delayed reporting are varied and can depend on the age of the child and the relationship between the abuser and the child. In the majority of cases the abuser is known to the child since strangers only account for about ten per cent of child sexual abuse cases. Thus, most abuse happens when there is a preexisting relationship between the child and the abuser, particularly one involving a degree of authority over the child. Often the abuser is a parent, family member or someone else who has a close relationship with the child. Children abused by family members are likely to delay reporting for longer than one month and many never report their abuse at all. On the other hand, children abused by strangers are more likely to disclose their abuse within one month.

Children who experience multiple incidents or types of abuse are also less likely to report. The younger the child, the less likely it is that she or he will report it to anyone. The child may have been threatened by the abuser or told to keep the abuse secret. Sometimes an abuser may threaten violence or harm to the child or the child's loved ones. Children often remain silent for fear that they will not be believed or will be blamed. They may feel responsible for allowing the abuse to happen, or may not report because they feel ashamed or embarrassed.

If there is a close relationship between the child and abuser, the victim may experience conflict because the abuser is someone they love and trust, and they may want to protect this person. Because of this relationship, the child may fail to recognize that the abuse is wrong or the abuser may mislead the child into believing that the abusive acts are normal. Younger children may not have the language skills to explain what has happened to them.

There is generally no physical evidence of the abuse such as DNA or other medical evidence because of delays in reporting. In other cases, DNA evidence may not be helpful because of innocent reasons for the presence of an abuser's DNA on the child. The vast majority of children who are assessed for suspected sexual abuse will have normal genital examinations. The absence of physical evidence to support an allegation of abuse is not necessarily an indication that the abuse did not occur.

I am also required to tell you that children can be reliable witnesses if questioned in a neutral, non-suggestive manner. When children are questioned about their experiences they are more likely to make errors when suggestive or leading questions are used rather than free-recall, and open-ended prompts, such as “What else can you tell me?” Older children are not more susceptible to misleading information about an event than adults.

Children’s memories of abuse have been shown to be generally reliable. Because a child’s brain and memory continue to develop with age, some details will be captured with precision, such as the acts of abuse. Other details, such as the times, dates or the sequence of events, are harder for some age groups to remember. It is common for children to give incomplete or inconsistent accounts of events that have happened.

Children as young as six years of age have been shown to be as accurate as adults when asked to distinguish between their own memories and what someone else has said to them.

Older children are more likely to give correct answers to questions about events that have happened, despite inaccurate suggestions being put to them compared to pre-schoolers who are more vulnerable to suggestions. This may be due to the greater obedience to adult authority figures by younger children.

Younger children are more likely to endorse incorrect details about actual events, but usually in relation to peripheral matters rather than agreeing to participate in an event that did not occur.

With those final directions, I have now completed my summing-up. With a final reminder that any verdict you reach must be unanimous, I ask you to retire to the jury room to consider your verdict.

Oñati Socio-legal Series, v. 6, n. 2 (2016), 283-314
ISSN: 2079-5971