GRIFFITH UNIVERSITY

National jury research published

Author

Goodman-Delahunty, Jane, Cossins, Annie, Martschuk, Natalie

Published

2016

Journal Title

Judicial Officers' Bulletin

Version

Version of Record (VoR)

Rights statement

Goodman-Delahunty, J; Cossins, A; Martschuk, N, National jury research published, Judicial Officers' Bulletin, (2016) Judicial Officers' Bulletin, pp. 45-48. © Judicial Commission of NSW. The attached file is reproduced here in accordance with the copyright policy of the publisher. Please refer to the journal's website for access to the definitive, published version.

Downloaded from

http://hdl.handle.net/10072/414710

Link to published version

https://www.judcom.nsw.gov.au/publications/bulletins-and-journals/

Griffith Research Online

https://research-repository.griffith.edu.au

Judicial Officers' Bulletin

Published by the Judicial Commission of NSW

June 2016, Volume 28 No 5

Downloaded from search.informit.org/doi/10.3316/ielapa.158825891230494. Griffith University, on 05/30/2022 11:49 AM AEST; UTC+10:00. © Judicial Officers Bulletin , 2016.

Chair of the Royal Commission, the Honourable Justice Peter McClellan AM, at the launch of the study on 23 May 2016.

National jury research published

Professor Jane Goodman-Delahunty, *Charles Sturt University*
Professor Annie Cossins, *University of NSW*
Research Associate, Natalie Martschuk, *Charles Sturt University*

The Royal Commission into Institutional Responses to Child Sexual Abuse commissioned a major study into how juries reason in child sexual abuse trials. The study, released in May, is the largest empirical study of its kind. The report forms part of the Royal Commission's research program in relation to the criminal justice system's response to child sexual abuse in institutional contexts. The following article provides an overview of the study and its main findings.

The study[1] involved 90 mock jury deliberations with more than 1,000 jury eligible community members. The study tested judicial assumptions about the risk of unfair prejudice to a defendant under statutory rules governing the admissibility and use of tendency evidence in mock trials in which evidence from three complainants had been admitted. The research compared jury reasoning about child sexual abuse offences involving three complainants versus a single complainant. The key findings were that jury decisions were logically related to the probative value of the evidence and that no verdict was based on impermissible reasoning.[2] These outcomes imply that the capacity of juries to handle cases involving a single offender with multiple victims, whether in a joint or separate trial, may be underestimated.

Aims of the study

The study aimed to:

- document the 90 juries' interpretation of cross-admissible evidence[3] in a joint child sexual abuse trial, to determine the extent to which juries engage in impermissible reasoning;
- compare the above decision-making processes with those of juries in a separate trial involving the same defendant;
- compare trial outcomes (acquittal, conviction or hung jury) in a joint versus separate trial involving the same defendant;

1 J Goodman-Delahunty et al, *Jury reasoning in separate and joint trials of institutional child sexual abuse: an empirical study*, Royal Commission into Institutional Responses to Child Sexual Abuse, Sydney, 2016. A full copy of the report is available at www.childabuseroyalcommission.gov.au/policy-and-research/our-research/published-research/jury-reasoning-in-joint-and-separate-trials, accessed 30 May 2016.

2 Defined in the report as "[r]easoning that is logically unrelated to the evidence …[including]… (a) inter-case conflation of the evidence; (b) accumulation prejudice; and (c) character prejudice": J Goodman-Delahunty, above no 1, p 17.

3 Defined in the report as evidence of a complainant which is admissible in relation to counts involving another complainant, J Goodman-Delahunty, above n 1, p 16.

Features

45 National jury researched published
Professor Jane Goodman-Delahunty, Professor Annie Cossins and Natalie Martschuk

49 Ngara Yura news: Judicial officers visit Kamay, Botany Bay National Park

Recent Decisions

51 High Court
Attwells v Jackson Lalic Lawyers Pty Ltd
Badenach v Calvert

52 Court of Criminal Appeal
Panetta v R

52 Court of Appeal
Royal Guardian Mortgage Management Pty Ltd v Nguyen

53 ACT Court of Appeal
R v Harrington

54 Education Calendar
Select Legislation
Judicial Moves
Obituaries

- examine the relationship between jurors' misconceptions about child sexual abuse, jury deliberations and decisions, and trial outcomes; and
- determine the effect of question trails on juries' reasoning and decisions.

Background and limitations of past research on joint trials

A widely acknowledged concern in trials of child sexual abuse is the low conviction rate, due in part to reliance in these cases on "word-against-word" evidence. Where the only evidence of abuse is the complainant's evidence and there is a lack of corroborating eyewitness testimony, medical evidence, or DNA evidence, it can be difficult for a jury to be satisfied beyond reasonable doubt that the alleged abuse occurred. Where there are multiple victims of a single offender, there may be evidence from other complainants or witnesses who make allegations that the offender has also abused them. An issue will be what other evidence can be admitted in the trial. A recent analysis of 137 child sexual abuse cases prosecuted in New Zealand showed that a conviction was most likely when either coincidence evidence, evidence from a witness to the offending, positive medical evidence or DNA evidence were admitted.[4] An early archival review of child sexual offences heard in 1994 in the NSW District Court also revealed that the vast majority of cases where separate trials were ordered resulted in not guilty verdicts (28 of 43).[5]

In cases of institutional child sexual abuse, where it is common for an offender to have multiple victims, evidence of the defendant's criminal conduct with other complainants is typically excluded on the grounds that its prejudicial effect outweighs its probative value. Under the *Evidence Act* 1995 (NSW), probative value of particular evidence must substantially outweigh its prejudicial effect (ss 97, 101(2)). For example, policy reasons for excluding evidence of "bad character" include that juries have been presumed to view the defendant as "a person of depraved character".[6] Whether juries are in fact prone to prejudicial reasoning in these cases has never been previously examined.

Past studies of joint trials focused exclusively on conviction rates, without taking into account that cross-admissible evidence logically increases the likelihood of conviction. This is the first study to examine jury reasoning in deliberations to discern whether conviction rates increase due to logical and permissible uses of additional inculpatory evidence or due to impermissible reasoning, logically unrelated to the evidence.

Research questions

The case law[7] disclosed three main judicial hypotheses about impermissible reasoning associated with unfair prejudice to the accused. First, that juries will confuse or conflate the evidence adduced to support different charges in a joint trial. Second, that juries will leap to the conclusion that the accused is guilty based on the cumulative impact of multiple counts or witnesses against him. Third, that juries will use evidence about the defendant's other criminal misconduct to infer a criminal disposition or character by reasoning "if he did it once, he will do it again". As a result, three research questions guided the inquiry:

1. Are juries capable of separating the counts against the defendant in reaching their verdicts in a joint trial?
2. Because of the cumulative number of complaints against a defendant in a joint trial, will juries deliver similar conviction rates for counts based on weak compared to stronger case evidence?[8]
3. Are juries in joint trials more prone than those in separate trials to convict on the basis that the defendant has a "criminal disposition"?

Methodology

The study was conducted in two stages. First, an online mock juror pilot study pre-tested the case strength of written trial summaries of weak, moderate and strong cases, using facts drawn from real cases. The cases involved historical child sexual abuse claims allegedly perpetrated by a sports coach for a boys' under-12 soccer team. A total of 300 community members who were eligible to serve on a jury served as virtual mock jurors. Each mock juror read a single case summary and returned an individual verdict. The conviction rate in the weak case with one count of indecency was 24%, and was significantly lower than the conviction rates for counts involving the other two complainants: 59% for the moderate case with one count of indecency and one count of sexual intercourse, and 74% for the strong case with two counts of indecency and one count of sexual intercourse. These differences in case strength were satisfactory to form the basis of a realistic simulated video-trial created for Stage 2 of the project.

Four full trial scripts were drafted,[9] comprised of a separate trial; trial where context evidence was led; trial where tendency evidence was led; and a joint trial. The core evidence of a single complainant with two moderately strong counts was common to all trials (the focal complainant). The trial where tendency evidence was led and the joint trial included evidence of six sexual offences involving three witnesses or co-complainants.

4 S Blackwell and F Seymour, "Prediction of jury verdicts in child sexual assault trials" (2014) 21 *Psychiatry, Psychology and Law* 567.

5 P Gallagher, J Hickey and D Ash, *Child sexual assault: an analysis of matters determined in the District Court of New South Wales during 1994*, Research Monograph No 15, Judicial Commission of NSW, 1997, p 20. The study, inter alia, examined 201 child sex offences with multiple victims, of which 158 offences were heard in joint trials and 43 in separate trials.

6 *Pfennig v The Queen* (1995) 182 CLR 461 at 513.

7 See for example *De Jesus v The Queen* (1986) 68 ALR 1, at 4–5; *Pfennig v The Queen*, ibid, at 512–513; *BRS v The Queen* (1997) 191 CLR 275 at 320–322; *KRM v The Queen* (2001) 206 CLR 221 at [97]; J Goodman-Delahunty, above n 1, p 45.

8 "Weak" evidence in the study was word-against-word evidence which included some inconsistencies, was unsupported by other witnesses or evidence, with potential motivation for secondary gain in the form of monetary compensation.

9 The four trial scripts and jury directions that formed the basis for this empirical study were written by Professor Annie Cossins. The scripts are available on the website of the Royal Commission.

Downloaded from search.informit.org/doi/10.3316/ielapa.158825891230494. Griffith University, on 05/30/2022 11:49 AM AEST; UTC+10:00. © Judicial Officers Bulletin , 2016.

The trial scripts were checked for legal accuracy by leading legal practitioners. A District Court judge presided and barristers played the roles of the legal professionals, while professional actors played the witnesses.

In all, 10 different versions of the trial were created by video-editing, keeping the evidence of the focal complainant consistent across all trial versions. The basic separate trial with three witnesses lasted 45 minutes, and the joint trial with seven witnesses lasted 110 minutes. In all trials, appropriate judicial directions were adapted from the *Criminal Trial Courts Bench Book*,[10] for example on delay in complaint; the onus on the Crown to prove the elements of each charge beyond reasonable doubt; and the requirement for a unanimous verdict. In some versions, special jury directions on context evidence and tendency evidence were added. In two versions, a fact-based question trail was added.[11]

A total of 1,029 mock jurors were recruited from the NSW Downing Centre and by a market research company. The demographic profile of the mock jurors closely matched that of real jurors who serve on trials in NSW in terms of gender, age, and employment status. The mock jurors were randomly assigned to one of 90 mock juries. They watched the video-trial and deliberated to a verdict in a group. Following deliberation, they completed a written post-trial questionnaire. All deliberations were recorded and transcribed for analysis.

Legally trained researchers[12] scrutinised the reasoning of individual jurors and the group processes of juries during deliberation for indications of unfair prejudice based on inter-case factual conflation,[13] accumulation prejudice[14] and character prejudice.[15] The data were analysed using quantitative and qualitative methods.

Results

Mock juries separated the counts against the defendant in reaching their verdicts.

The deliberations revealed that juries followed judicial directions. They distinguished different charges and complainants, and returned verdicts based on the probative value of evidence led for each count in all types of trials. Coding of all factual errors in 90 juries revealed an average of one factual error per deliberation,[16] although one third of the deliberations were entirely error free, including joint trials. In only seven of the 90 juries was the error rate higher, with 3–4 errors. Error rates were related to the overall number of witnesses (ranging from 3 to 7 witnesses depending on the type of trial), not joint trials. Two important findings emerged: first, 82% of the errors were corrected by other jurors, and no uncorrected error was an instance of inter-case conflation of the evidence. A key outcome was that not one of the 90 verdicts was premised on inter-case evidence conflation. Thus, no verdict was the result of prejudice arising from factual conflation.

Conviction rates for counts based on weak versus stronger case evidence were not uniform.

Analyses of the deliberations did not disclose a single verdict driven by the overall number of charges or Crown witnesses. The juries were not overwhelmed by the number of counts or witnesses and did not alter the standard of proof when assessing more counts or evidence from more Crown witnesses in joint trials or separate trials where tendency evidence was led. In the joint trials, the juries spent most time on the weakest claim and did not gloss over inconsistencies in the complainant's testimony, and as a result, many juries acquitted on that count. When they did convict, it was because they rated the complainant's evidence as convincing. Conviction rates for the moderately strong complaint were statistically undifferentiated whether juries assessed this evidence in the presence of 2 counts and 4 witnesses in the separate trial or 6 counts and 6 Crown witnesses in a joint trial. These results suggested that the juries were not susceptible to accumulation prejudice.

Mock juries in joint trials were not more prone than those in separate trials to convict on the basis that the defendant had a "criminal disposition"

A total of two mock jurors out of 1,029 made comments in the course of deliberation reflecting bias against the defendant. These jurors paticipated in the basic separate trial and a trial where context evidence was led, respectively, not trials where tendency evidence was led. Other jurors responded to their comments by diffusing the bias and re-focusing the discussion on the evidence. No verdict was the result of emotional arousal or unfair prejudice based on the defendant's bad character. Following deliberations, the mock jurors each recorded the main reason for their verdict. Analyses of these responses showed that decisions were based on assessments of the evidence's probative value and witness consistency, accounting for 70% of the convictions. Another fifth relied on the evidence of the

10 Judicial Commission of NSW, 1989–, at www.judcom.nsw.gov.au, accessed 30 May 2016.

11 Copies of the jury directions and question trails are appended to the full report.

12 Law school graduates who in addition received special training on the coding scheme that targeted reasoning errors and the forms of prejudice.

13 Defined in the report as "[a] type of impermissible reasoning based on substitution of the facts in evidence about one complainant for facts in evidence about another complainant, in a joint trial involving two or more complainants": J Goodman-Delahunty, above n 1, at p 17.

14 Defined in the report as "[a] type of impermissible reasoning that accords more weight to evidence than its true value, because multiple charges or multiple witnesses who give evidence against a defendant create the appearance of a stronger case against the defendant than exists in reality": J Goodman-Delahunty, above n 1, at p 13.

15 Defined in the report as "[a] type of impermissible reasoning based on the unwarranted inference of criminality in a defendant who is thus considered to deserve punishment because he or she is a bad person", J Goodman-Delahunty, above n 1, p 13.

16 For example, one mock juror in a joint trial with three complainants confused an uncharged act when the defendant put ice down the pants of the third complainant with the count of indecency alleged by the second complainant. Another mock juror corrected the error by itemising the two acts of indecency pertaining to the second complainant, and by clarifying that the ice incident had occurred with the third complainant.

Downloaded from search.informit.org/doi/10.3316/ielapa.158825891230494. Griffith University, on 05/30/2022 11:49 AM AEST; UTC+10:00. © Judicial Officers Bulletin , 2016.

grooming of the complainant(s) by the defendant. The reasons provided by mock-jurors for their verdicts were logically related to the type of evidence presented in the trial.

Credibility of the complainants was central to jury reasoning and verdicts

In assessing the oral evidence presented by the complainants about unwitnessed events of child sexual abuse, a pivotal issue confronting the juries was the credibility of the complainant (in a separate trial) or the complainants (in a joint trial). Results showed that the focal complainant's credibility varied by type of trial, with the lowest scores in the separate and context evidence trials, and increased as more inculpatory evidence against the defendant was introduced by either the complainant, but in particular, by other witnesses or co-complainants in a joint trial. Verdicts turned on the extent to which the complainant was rated as credible and the events as plausible. These assessments were further evidence of a logical response to the probative value of the evidence.

Similar comparisons of the assessments of the defendant's credibility showed the defendant was rated equally convincing in the separate trial and the trial where tendency evidence was led, but less convincing in the trial where context evidence was led.

Fairness of the trial

We further explored the perceived fairness of the various types of trials by directly asking mock jurors whether they thought the trial was fair to the defendant, whether the judge's directions were fair to the defendant, and whether they expected to be informed of other charges against the defendant. We compared responses from each of the four main trial groups. Surprisingly, the mock jurors in the separate trial rated that trial significantly less fair to the defendant than did their counterparts who rated the fairness of the joint trial. Similar results favouring the fairness of joint trials were obtained regarding the perceived fairness to the defendant of the judge's directions. Significantly more mock jurors in the separate trial expected that they would have been informed of other criminal conduct or the defendant's prior convictions. From the absence of such information, they inferred that there were none, that the allegations of the complainant in the separate trial were isolated (that this was a "one-off" offence), or that the complainant was untruthful or blameworthy. The mock juries appeared to rate the risk of miscarriage of justice against the defendant as greater in the separate than the joint trials because they assumed there were no other similar claims or offences and because the evidence of the complainant was less convincing in the absence of support from independent sources. Thus the potential for a wrongful conviction in a separate trial was perceived by the mock jurors as greater than in the context of a joint trial in which juries had a more complete picture of the history and pattern of the defendant's behaviour to evaluate.

To further assess whether a less stringent standard of proof is applied in the context of a joint or separate trial where tendency evidence is led, immediately following deliberation we asked every mock juror: "What number 0-100% represents 'beyond reasonable doubt'?" The overall average was 89%. However, the average threshold reported by mock jurors in the basic separate trial was 85.2%, which was significantly lower than the threshold reported by mock jurors in the joint trial, of 92.1%. This outcome suggested that mock jurors were unlikely to apply a lower threshold of proof in trials with tendency evidence and in joint trials. This outcome was supported by our observations of jury deliberations which showed no such trend. Only two out of 1,029 mock jurors mentioned a lower standard in jury deliberations (a balance of probabilities). Both of these individuals were in trials where context evidence was led, not trials where tendency evidence was led.

Discussion

The key outcome of this research project was that the verdicts of mock juries who deliberated in trials involving allegations of sexual abuse from multiple complainants were not swayed by unfair prejudice, whether this evidence was presented in the context of a separate trial involving a single complainant or in a joint trial involving multiple complainants. A major finding was that while using this multi-method convergent approach, we did not uncover a single verdict based on impermissible reasoning. The pattern of findings that we observed demonstrated the reverse, that the jury verdicts were logically based on the probative value of the admitted evidence.

Thus, a further important finding of the study was that in the absence of an eyewitness to the alleged crime, evidence from witnesses other than the complainant were critical in establishing the complainant's account as plausible. However, the credibility of the complainant was boosted by independent witnesses, whether their evidence was provided in a separate trial where tendency evidence was led, or as cross-admissible evidence in a joint trial.

Like all studies, this research has some limitations. Chief among them are the compressed time-frame in which the mock trials were conducted (under four hours), and the fact that the mock jurors were aware that this was a simulated and not a real trial. While the mock juries were aware that their decisions to convict would not result in a gaol sentence for the defendant, observations of the deliberations and reviews of the transcripts of the juries at work show that they undertook their task seriously and conscientiously.

Conclusion

Of course, research based on mock trials does not replicate a real trial. Jury deliberations are by no means infallible, and require careful direction. While some individual mock jurors made errors, by and large, the juries did not. The results of this study reflect the fact that the capacity of juries to return a sound verdict in joint trials and trials where tendency evidence is led may have been underestimated.

The Royal Commission will publish a consultation paper on criminal justice issues in September this year and the paper will address the tendency and coincidence issues in joint trials as well as a number of other matters. Once the consultation paper is published it is hoped that responses will be received from a wide representation of the community, including lawyers.

Downloaded from search.informit.org/doi/10.3316/ielapa.158825891230494. Griffith University, on 05/30/2022 11:49 AM AEST; UTC+10:00. © Judicial Officers Bulletin , 2016.