A re-examination of the acquittal biasing effect of offence seriousness

Samantha Lundrigan^a, Mandeep K. Dhami^b and Katrin Mueller-Johnson^c

^a*Policing Institute for the Eastern Region, Faculty of Arts, Law and Social Sciences, Anglia Ruskin University, Chelmsford, UK;* ^b*Department of Psychology, Middlesex University, London, UK;* ^c*Institute of Criminology, University of Cambridge, Cambridge, UK*

The justice system should operate free of bias, and jurors' judgements of a defendant's guilt should be based on evidential factors alone. However, research suggests that this is not always the case. The aim of this study is to investigate the biasing effect of offence seriousness – a case-related, extralegal factor – on juror decision-making. An experiment was conducted to examine the effect of this extralegal factor on 118 members of the jury-eligible public's interpretations of 'beyond reasonable doubt' (BRD), probability of commission, verdict and confidence in verdict.

It was found that defendants charged with more serious offences were judged to be less likely to have committed the crime. However, offence seriousness was not found to have a significant effect on interpretations of BRD and verdict. The present findings suggest a need to instruct jurors on the application of legal (probative) factors alone.

Key words: acquittal; beyond reasonable doubt; extralegal factors; juror decision-making; offence seriousness; probability of commission; verdict.

Introduction

Jurors are responsible for reaching verdicts in criminal cases that are based solely on the evidence admitted in court. Although some studies suggest that jurors do indeed focus on the legal factors of cases (Bornstein & Greene, 2011; Devine, Buddenbaum, Houp, Studebaker, & Stolle, 2009), others demonstrate that they can also rely on extralegal factors (for a review, see Devine, Clayton, Dunford, Seying, & Pryce, 2001). For the most part, past research on this topic focuses on the effect of person-related, extralegal factors (Devine & Caughlin, 2014), revealing factors such as defendant attractiveness that can lead to acquittal bias, and factors such as

jurors having an authoritarian attitude that can lead to conviction bias. In addition to person-related factors, research also suggests that *case-related*, extralegal factors can have a biasing effect on juror decision-making (Ruva & McEvoy, 2008; Terrance, Matheson, & Spanos, 2000).

One case-related, extra-evidential factor that has attracted research attention is the seriousness of the offence with which the defendant is charged (Kerr, 1978; Koch & Devine, 1999; Vidmar, 1972). It is suggested that jurors may be unduly biased due to thoughts about what might happen to the defendant if he or she is convicted of a serious crime that carries a severe penalty. In arriving at a judgement of guilt or innocence,

Correspondence: Samantha Lundrigan, Policing Institute for the Eastern Region, Faculty of Arts, Law and Social Sciences, Anglia Ruskin University, Chelmsford, UK. Policing Institute for the Eastern Region Faculty of Arts, Law and Social Sciences Anglia Ruskin University Bishop Hall Lane Chelmsford CM1 1SQ. Email: Samantha.lundrigan@anglia.ac.uk

however, jurors should obviously not consider the seriousness of the penalties associated with the offence with which the defendant is charged.

The present study reports the results of an experiment which examines the effects of offence seriousness on juror decision-making. The systematic study of the effects of a range of extralegal factors can help improve the theoretical understanding of juror decision-making. In addition, and following a tradition of past psycho-legal research, such research can also inform legal policies and procedures that aim to improve the process of jury decision-making, trial outcomes and perceptions of the justice system (Bornstein & Greene, 2011).

Offence seriousness

Offence seriousness can bias jurors towards acquittal or conviction. The seriousness of an offence can have a direct effect on jurors' evaluations of how likely a defendant is to have committed the crime, their interpretation of the meaning of 'beyond reasonable doubt' (BRD) and their verdict.

With one exception (see below), past research suggests that this case-related, extra-evidential factor either has no effect on juror decision-making (Freedman, Krismer, MacDonald, & Cunningham, 1994; Rind, Jaeger, & Strohmets, 1995; Simon, Snow, & Read, 2004) or leads to an *acquittal* bias (Goodman-Delahunty, Martschuk, & Ockenden, 2016; Kaplan & Simon, 1972; Kerr, 1978; Koch & Devine, 1999; Martin & Schum, 1987; McComas & Noll, 1974; Vidmar, 1972).

Only one study to date has found a conviction biasing effect of offence seriousness: de Keijser, de Lange, and van Wilsem (2014) examined the effect of offence seriousness (i.e. shoplifting, burglary and rape) on people's attitudes towards a measure of the Blackstone ratio,¹ which is a proxy for BRD. They found that the Blackstone ratio decreased as the seriousness of the offence

increased, with a lower standard of proof required in a case involving a more serious crime, thus resulting in a greater likelihood of conviction.

However, several studies – using different (and sometimes multiple) measures – report an opposite, i.e. acquittal biasing, effect of offence seriousness. Vidmar (1972) demonstrates acquittal bias in a study wherein mock jurors were provided with a combination of two, three or four offence alternatives made up of first degree murder, second degree murder, manslaughter and not guilty. He found that the jurors seldom voted not guilty when given at least one moderate offence severity option, whereas given only a serious offence severity option more than half of the jurors voted not guilty. Similarly, Goodman-Delahunty et al. (2016) found that mock jurors were significantly less likely to convict a defendant charged with terrorism than a defendant charged with attempted murder, suggesting that mock jurors may have perceived a terrorism charge as being more serious than a non-terrorism charge, biasing their judgements regarding the defendant's guilt.

Martin and Schum (1987) gave mock jurors a list of 10 criminal offences (e.g. aggravated assault, murder, petty theft, narcotics peddling) together with their possible penalties and found that the mock jurors raised their BRD threshold in response to the murder charge, thus making them more likely to acquit. Koch and Devine (1999) varied the definition of BRD and the offence seriousness, finding that juries with the option to convict on a lesser charge (i.e. voluntary manslaughter vs murder) produced more overall convictions than juries receiving only the primary charge, therefore demonstrating an acquittal bias. However, this effect was found to be limited to the condition in which BRD is undefined.

Kerr (1978) varied offence seriousness (i.e. first and second degree murder, and manslaughter) and found an acquittal biasing effect whereby the more serious the offence,

the lower the rating of the defendant's guilt and the lower the likelihood of conviction. Similarly, McComas and Noll (1974) manipulated offence seriousness (i.e. first and second degree murder, and manslaughter) and found an acquittal biasing effect of offence seriousness on jurors' probabilities of guilt and their verdicts. One study has examined the acquittal biasing effect of offence seriousness in the context of evidence evaluation: Kaplan and Simon (1972) found that, even when evidence of guilt was high or mixed, the more serious the charge, the greater the likelihood of jurors acquitting the defendant.

As mentioned, a number of studies have found no conviction or acquittal biasing effects of offence seriousness on juror decision-making. Simon et al. (2004) reported that offence seriousness (i.e. simple theft versus theft resulting in an individual's death) had no effect on jurors' conviction rates. Freedman et al. (1994) conducted a series of seven separate experiments in which they manipulated offence seriousness (e.g. robbery and murder, robbery and assault, robbery and aggravated assault, etc.), finding no effect of offence seriousness on probability of commission, verdict or verdict confidence in all scenarios. Finally, in a study examining the effect of offence seriousness on jurors' use of inadmissible evidence, Rind et al. (1995) found no direct effect of offence seriousness on juror decision-making.

In sum, the research findings on the effects of offence seriousness on juror decision-making are mixed. Specifically, while some studies demonstrate an acquittal biasing effect of offence seriousness, one has found a conviction biasing effect and others have found no effect. It is unclear to what extent this is due to differences in the measures used.

The present study

The present study aims to contribute to the above body of research in two main ways.

The first aim is to further explore the potential biasing effect of offence seriousness on interpretations of BRD, which to date has only been studied twice: Kerr (1978) found that interpretations of BRD were significantly stricter when the offence was more severe, and – compatible with this – Martin and Schum (1987) found that interpretations of BRD were significantly higher for murder than for any other offence.

The second aim is to examine a wide range of juror behaviour that may be affected by offence seriousness. To this end, probability of commission, interpretations of BRD and confidence in verdict were all measured, in the hope of contributing to a better understanding of any biasing effect of offence seriousness on the *process* of juror decision-making. For example, does the acquittal biasing effect emerge at the stage when jurors are judging probability of commission, before they decide on a verdict?

The specific objective is to investigate the effect of offence seriousness on (1) interpretations of BRD, (2) judgements of defendant guilt, (3) verdicts and (4) confidence in verdicts. Although there is some evidence to suggest an acquittal biasing effect, it is difficult to make strong a-priori predictions on the above because some past research has found no such effects and one study found the opposite (i.e. a conviction biasing) effect. In addition, a number of past studies only measure verdict and so have little to offer in terms of predictions about the effect of offence seriousness on other aspects of the juror decision-making process.

Method

Participants

A total of 118 members of the jury-eligible public were recruited from a large company in the United Kingdom (UK) that has a multi-building site and employs over 700 people in a wide range of roles from packers through clerical staff to medics. They volunteered to

participate in return for a £10 payment. Of the participants, 70% were female, the average age was 33.37 years ($SD = 11.13$; range: 18–62), 92% described themselves as Caucasian, 71% had a university degree, 53% were employed at a professional level and 5% had previously served on a jury.

Design

A between-subjects experimental design was employed using offence seriousness as the independent variable with two levels (i.e. more serious and less serious). The more serious offence is ‘attempted murder’ and the less serious is ‘common assault’.

Stimuli and measures

The participants were asked to imagine that they were serving on a jury in a criminal trial. They read a short trial summary of an adult male charged with a violent offence and were told that the defendant faced a charge of either common assault or attempted murder (see the Appendix). The case was designed to be equivocal with regard to the evidence so that it would result in some variance in verdicts. A pilot test ($n = 20$) confirmed this (i.e. 54.8% of the participants in the pilot test found the defendant guilty; none of the pilot test participants are included in the present study). In the present study, 57.1% of the participants found the defendant guilty.

After reading the case and legal arguments, the participants read the judge’s instruction on the presumption of innocence, the burden of proof and the standard of proof, i.e. ‘the defendant is presumed innocent unless the prosecution has proved guilt beyond reasonable doubt’. The participants then provided their interpretations of BRD, after which they rated the probability that the defendant had committed the crime as charged, rendered a verdict and rated their confidence in their verdict. The measurement scales for these variables are described below.

Interpretations of BRD were measured using the membership function (MF) method (Dhami, 2008; Dhami, Lundrigan, & Mueller-Johnson, 2015; Dhami & Wallsten, 2005; Lundrigan, Dhami, & Mueller-Johnson, 2016; Mueller-Johnson, Dhami, & Lundrigan, 2018) and the direct rating (DR) method (Horowitz & Kirkpatrick, 1996; Kagehiro, 1990; Kerr et al., 1976; Montgomery, 1998). The DR method requires participants to rate the minimum threshold at which they are prepared to convict the defendant from 0 to 100% (in 5% increments). This provides a point interpretation of BRD. As Figure 1 shows, the MF method involves presenting participants with the phrase ‘beyond reasonable doubt’ along with a series of 21-point scales anchored at each end with *Not at all* and *Absolutely*. Each scale corresponds to one of 11 percentage points, ordered from 0 to 100% (in 5% increments). The participants were asked to imagine each percentage point in front of the word ‘proof’ and then decide how well this percentage point substitutes for the phrase proof “beyond reasonable doubt” by marking a point along each scale from *not at all* to *absolutely*. The MF provides a ‘peak’ interpretation of BRD (i.e. the percentage value at which the participant provides a rating of *Absolutely*), as well as a minimum interpretation (i.e. the first value that has a rating higher than *Not at all*) and a range of interpretations (i.e. the difference between the minimum and maximum values). Both the DR method and the MF method have been shown to reliably capture mock jurors’ interpretations of BRD and predict their verdicts (e.g. Dhami, 2008; Dhami et al., 2015; Lundrigan et al., 2016).

Participants’ judgements of the probability that the defendant committed the crime were measured on a scale of 0 to 100% (in 5% increments). The verdict options were ‘not guilty’ or ‘guilty’ and confidence in the verdict was measured on a scale of 0 to 100% (in 10% increments).

Finally, a manipulation check question was also included, which asked participants

to rate how serious the charge was against the defendant, on an 11-point scale ranging from *not at all* to *extremely*.

Procedure

Recruitment posters were placed across the multi-building site and emails asking for volunteers were sent out to all employees. Participants were randomly assigned to one of the two experimental conditions. The data was collected individually via paper-and-pencil testing at the recruitment site. The experiment took approximately 15 to 20 minutes to complete. Demographic data was also collected on the participants, as specified above.

Results²

An independent samples *t*-test demonstrates that the manipulation of offence seriousness was successful. The attempted murder case was perceived as being significantly more serious ($M = 8.17$, $SD = 1.97$) than the common assault case ($M = 5.33$, $SD = 2.02$), $t(107) = -7.41$, $p < .001$. The effect size for this analysis, $d = 1.43$, 95% CI [1.01, 1.84], exceeds Cohen's (1988) convention for a large effect (0.80).

The participants' interpretations of BRD were measured using the MF and DR methods. For the MF method, the mean peak interpretation of BRD is 97.76% ($SD = 4.79$) in the attempted murder condition compared to 96.96% ($SD = 7.23$) in the common assault condition. For the DR method, the mean interpretation of BRD is 82.13% ($SD = 13.62$) in the attempted murder condition and 82.32% ($SD = 12.90$) in the common assault condition. The mean differences between the two conditions are non-significant for both the MF method, $p = .537$, $d = 0.12$, 95% CI [-0.26, 0.50], and the DR method, $p = .940$, $d = 0.01$, 95% CI [-0.36, 0.39]. There is no significant difference between the two conditions for either the mean minimum MF interpretation, $p = .377$, $d = 0.17$, 95% CI

[-0.21, 0.55], or the mean MF range, $p = .377$, $d = 0.17$, 95% CI [-0.21, 0.55].

Offence seriousness was found to have a significant effect on probability of commission judgements, $t(108) = 2.77$, $p = .007$, $d = 0.53$, 95% CI [0.15, 0.91]. On average, the participants in the attempted murder condition considered the defendant as being significantly less likely to have committed the crime ($M = 65.37$, $SD = 20.75$) than the participants in the common assault condition ($M = 75.26$, $SD = 16.59$).

A total of 52% of the participants in the attempted murder condition rendered a guilty verdict compared to 62% in the common assault condition. However, a Chi-square test indicates that offence seriousness is not significantly associated with verdict, $\chi^2 [1, 112] = 1.19$, $p = .275$, OR = .658.

Offence seriousness was found to have no significant effect on the participants' confidence in their verdicts, $t(109) = 1.81$, $p = .073$, $d = 0.03$, 95% CI [-0.03, 0.72]. The mean confidence ratings are 61.91 ($SD = 22.50$) in the attempted murder condition compared to 69.12 ($SD = 19.39$) in the common assault condition.

Discussion

A number of studies have demonstrated that extralegal factors may lead to an acquittal bias (e.g. Mazzella & Feingold, 1994; Schvey, Puhl, Levandoski, & Brownell, 2013). The present study investigated the possible biasing effect of one case-related, extra-evidential factor – i.e. offence seriousness – on aspects of the juror decision-making process. The verdict was measured along with interpretations of BRD, probability of commission and confidence in verdict. Past research has tended to focus on verdict alone as a measure of juror decision-making, and the present study is the first to examine the range of measures included herein. It is important to consider the multiple stages of the decision-making process in juror research in order to capture the different effects that

may occur at different stages of the process, rather than viewing juror decision-making as a single judgmental process that is similarly affected (Bellin, 2010). Furthermore, the sensitivity of different measures to experimental effects may vary (e.g., Dhami, 2008).

A significant effect of offence seriousness was found on judgements of the probability that the defendant committed the crime, consistent with McComas and Noll (1974) and Kerr (1978) but inconsistent with Freedman et al. (1994). In their study, Freedman et al. (1994) used cases wherein the amount of evidence that was needed to prove guilt was equal for all the charges under examination, whereas McComas and Noll (1974) and Kerr (1978) did not. If the probability of commission scale reflects participants' perception of the apparent weight of evidence against a defendant then this might explain the lack of effect found by Freedman et al. (1994).

However, the effect of offence seriousness did not carry over to verdict in the present study. Past research on the effect of offence seriousness on verdict is mixed. The present findings are consistent with some previous studies (Freedman et al., 1994; Rind et al., 1995; Simon et al., 2004) but inconsistent with others (Kaplan & Simon, 1972; Kerr, 1978; McComas & Noll, 1974; Vidmar, 1972). Each of the studies that found an effect of offence seriousness on verdict uses homicide as the offence, with three levels of seriousness (i.e. first degree murder, second degree murder and manslaughter). Associated with the most serious of these offences is a penalty of life imprisonment or death, which may account for the observed effect of seriousness on verdict in these studies. By contrast, the studies that found no effect of offence seriousness on verdicts, including the present one, typically consider offences with less severe penalties attached (i.e. no death penalty).

It is unclear why the effect of offence seriousness on probability of commission did not carry over to verdict in the present study, but future research may examine the extent to which interpretations of BRD diminish such

an effect. Indeed, in relation to the effects of offence seriousness on interpretations of BRD, no acquittal biasing effect was found. Two previous studies have reported such an effect, although they suffer from some methodological limitations. Martin and Schum (1987) measured interpretations of BRD across different crime types, including petty theft, tax evasion and murder, finding an effect whereby interpretations of BRD for the murder charge were significantly stricter than for all other charges. Kerr (1978) found that interpretations of BRD were significantly stricter when the potential penalty was severe compared to when it was mild, and that this effect was most pronounced for first degree murder and manslaughter. Furthermore, both studies use a very different method for measuring interpretations of BRD compared to the present study; rather than measuring interpretations of BRD directly, as is the case herein, they both use indirect measures.

The fact that no significant effect of offence seriousness was found on interpretations of BRD – measured using both the MF method and the DR method – suggests that the influence of this extra-evidential factor arises at the stage of the trial during which the jurors consider the evidence, rather than at the later stages when they are applying the standard of proof. Future research could be designed to see if other extralegal variables, such as the consequences of a potential custodial sentence for the defendant, affect the juror decision-making process in similar ways.

The mechanism most commonly put forward to account for an acquittal biasing effect of offence seriousness on juror decision-making focuses on the magnitude of the potential 'costs' associated with more serious offences (Freedman et al., 1994; Kerr, 1978). As the offence seriousness increases, jurors may be less willing to risk type I errors (i.e. convicting an innocent person) over type II errors (i.e. releasing a guilty person) because of the potential consequences of conviction. By considering the potential consequences of their

verdicts, jurors may modify their decision-making and be less likely to convict. This can be done by either reducing the probability of commission, changing the verdict or increasing the standard or proof. Herein it was found that while offence seriousness demonstrates an effect early on in the decision-making process (i.e. probability of commission), this effect does not carry over to verdict and interpretations of BRD. This may be because extralegal factors such as offence seriousness differentially impact stages of the decision-making process, or perhaps because some stages of the decision-making process are more resistant to such influence.

Limitations and implications

It can be argued that the external validity of the present findings is limited for several reasons: mock jurors were used rather than real jurors; verdicts were examined in the context of a written hypothetical case that diverges from the detailed oral and visual representations available at a real trial; and verdicts were only collected at a pre-deliberation stage. The present study uses the methodological procedures typical of research on jury decision-making, and although further research examining deliberating juries may be of value, efforts were made to minimize the limitations in a number of ways.

First, it is difficult to study real juries in real trial situations. The 1981 Contempt of Court Act in the UK currently prohibits the study of real jurors. Unlike much other juror decision-making research, the present research does not utilize a student sample. The mock jurors in the present research were all jury eligible, and 5% have been called for jury service in the past. Some studies have also shown few differences between mock and real jurors (e.g. Bornstein, 1999; MacCoun & Kerr, 1988).

Second, while the trial summary was presented in a written format, it was made as clear and plausible as possible. Additionally, Bornstein (1999) has concluded that there is

little negative impact on the validity of jury studies using written presentations of trial material.

Finally, there is evidence that the pre-deliberation distribution of verdicts predicts post-deliberation verdicts (e.g. Kalven & Zeisel, 1966; Newkirk, 1981; Sandys & Dillehay, 1995) and that interpretations of BRD differ very little from pre- to post-deliberation (Dane, 1985; Horowitz & Kirkpatrick, 1996).

From a pragmatic standpoint, the fact that jurors may be influenced by extra-evidential factors such as offence seriousness underscores the need to instruct them on the application of legal (probative) factors alone, irrespective of the potential 'cost' to the defendant. The extralegal factor studied here is legally relevant at other stages in criminal proceedings (i.e. bail and sentencing) but should not inform the decisions made by jurors. Ultimately however, the unconscious level at which extralegal influences may operate can limit the effectiveness of such instructions, which are delivered at a conscious level. This work was supported by the Arts and Humanities Research Council [grant number AH/E008607/1].

Ethical standards

Declaration of conflicts of interest

Samantha Lundrigan has declared no conflicts of interest.

Mandeep Dharmi has declared no conflicts of interest.

Katrin Mueller Johnson has declared no conflicts of interest.

Ethical approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent

Informed consent was obtained from all individual participants included in the study.

Notes

1. In this study, the Blackstone ratio was measured by the following open question: 'How many guilty people in similar cases should be acquitted in order to prevent one innocent person from wrongful conviction?'.
2. In addition to the independent samples *t*-tests reported in this section, one multivariate analysis of variance (MANOVA) was conducted that included all of the dependent measures, which replicated the results of the *t*-tests reported herein.

References

- Bellin, J. (2010). *Is punishment relevant after all? A prescription for informing juries of the consequence of conviction*. Faculty Publications. Paper 1246. Retrieved from <http://scholarship.law.wm.edu/facpubs/1246>
- Bornstein, B.H. (1999). The ecological validity of jury simulations: Is the jury still out? *Law and Human Behavior*, 26, 625–639. doi:10.1023/A:1022326807441
- Bornstein, B.H., & Greene, E. (2011). Jury decision making: Implications for and from psychology. *Current Directions in Psychological Science*, 20, 63–67. doi:10.1177/0963721410397282
- Cohen, J. (1988). The analysis of variance. In J. Cohen (Ed.), *Statistical power analysis for the behavioral sciences* (2nd ed., pp. 273–406). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Dane, F.C. (1985). In search of reasonable doubt. *Law and Human Behavior*, 9, 141–158. doi:10.1007/BF01067048
- de Keijser, J.W., de Lange, E.G.M., & van Wilsem, J.A. (2014). Wrongful convictions and the Blackstone ratio: An empirical analysis of public attitudes. *Punishment & Society*, 16, 32–49. doi:10.1177/1462474513504800
- Devine, D.J., Buddenbaum, J., Houpp, S., Studebaker, N., & Stolle, D.P. (2009). Strength of evidence, extra-evidentiary influence, and the liberation hypothesis: Data from the field. *Law and Human Behavior*, 33, 136–148. doi:10.1007/s10979-008-9144-x
- Devine, D.J., & Caughlin, D.E. (2014). Do they matter? A meta-analytic investigation of individual characteristics and guilt judgments. *Psychology, Public Policy and Law*, 20, 109–134. doi:10.1037/law0000006
- Devine, D.J., Clayton, L.D., Dunford, B.B., Seying, R., & Pryce, J. (2001). Jury decision making: 45 years of empirical research on deliberating groups. *Psychology, Public Policy and Law*, 7, 622–727. doi:10.1037//1076-8971.7.3.622
- Dhami, M.K. (2008). On measuring quantitative interpretations of reasonable doubt. *Journal of Applied Experimental Psychology*, 14, 353–363. doi:10.1037/a0013344
- Dhami, M.K., Lundrigan, S., & Mueller-Johnson, K. (2015). Instructions on reasonable doubt: Defining the standard of proof and the juror's task. *Psychology, Public Policy and Law*, 21, 169–178. doi:10.1037/law0000038
- Dhami, M.K., & Wallsten, T.S. (2005). Interpersonal comparison of subjective probabilities. *Memory & Cognition*, 33, 1057–1068. doi:10.3758/bf03193213
- Freedman, J.L., Krismer, K., MacDonald, J.E., & Cunningham, J.A. (1994). Severity of penalty, seriousness of the charge, and mock jurors' verdicts. *Law and Human Behavior*, 18, 189–202. doi:10.1007/bf01499015
- Goodman-Delahunty, J., Martschuk, N., & Ockenden, E. (2016). Effects of terrorist charges and threatening conduct on mock jurors' decisions. *Psychiatry, Psychology and Law*, 23, 5. doi:10.1080/13218719.2015.1120247
- Horowitz, I.A., & Kirkpatrick, L.C. (1996). A concept in search of a definition: The effects of reasonable doubt instructions on certainty of guilt standards and jury verdicts. *Law and Human Behavior*, 20, 655–670. doi:10.1007/BF01499236
- Kagehiro, D.K. (1990). Defining the standard of proof in jury instructions. *Psychological Science*, 1, 194–200. doi:10.1111/j.1467-9280.1990.tb00197.x
- Kalven, H., & Zeisel, H. (1966). *The American jury*. Boston, MA: Little, Brown and Company.
- Kaplan, K.L., & Simon, R.I. (1972). Latitude and severity of sentencing options, race of the victim and decisions of simulated jurors: Some issues arising from the 'Algiers Motel' trial. *Law and Society Review*, 7(1), 87–98. doi:10.2307/3052830
- Kerr, N.L. (1978). Severity of prescribed penalty and mock jurors' verdicts. *Journal of Personality and Social Psychology*, 36(12), 1431–1442. doi:10.1037//0022-3514.36.12.1431
- Kerr, N.L., Atkin, R.S., Stasser, G., Meek, D., Holt, R.W., & Davis, J.H. (1976). Guilt beyond a reasonable doubt: Effects of concept definition and assigned decision rule on the judgments of mock jurors. *Journal of Personality and Social Psychology*, 43, 282–394. doi:10.1037/0022-3514.34.2.282
- Koch, C.M., & Devine, D.J. (1999). Effects of reasonable doubt and inclusion of a lesser charge on jury verdicts. *Law and Human*

- Behavior*, 14, 409–438. doi:10.1023/A:1022389305876
- Lundrigan, S., Dhami, M.K., & Mueller-Johnson, K. (2016). Predicting verdicts using pre-trial attitudes and standard of proof. *Legal and Criminological Psychology*, 21, 95–110. doi:10.1111/lcrp.12043
- MacCoun, R.J., & Kerr, N.L. (1988). Asymmetric influence in mock jury deliberation: Jurors' bias for leniency. *Journal of Personality and Social Psychology*, 54, 21–33. doi:10.1037//0022-3514.54.1.21
- Martin, A.W., & Schum, D.A. (1987). Quantifying burdens of proof: A likelihood ratio approach. *Jurimetrics Journal*, 27, Summer, 383–402.
- Mazzella, R., & Feingold, A. (1994). The effects of physical attractiveness, race, socio-economic status, and gender of defendants and victims on judgments of mock jurors: A meta-analysis. *Journal of Applied Social Psychology*, 24, 1315–1338. doi:10.1111/j.1559-1816.1994.tb01552.x
- McComas, W.C., & Noll, M.E. (1974). Effects of seriousness of charge and punishment severity on the judgments of simulated jurors. *Psychological Record*, 24, 545–547. doi:10.1007/bf03394276
- Montgomery, J.W. (1998). The criminal standard of proof. *New Law Journal*, 148, 582–584.
- Mueller-Johnson, K., Dhami, M.K., & Lundrigan, S. (2018). Effects of Judicial Instructions and Juror Characteristics on Interpretations of Beyond Reasonable Doubt. *Psychology, Crime and Law*, 24. doi:10.1080/1068316X.2017.1394461
- Newkirk, J.M. (1981). The effect of different reasonable doubt definitions and group versus individual deliberation on the behavior of mock jurors. *Dissertation Abstracts International*, 42, 1673–B–1674–B.
- Rind, B., Jaeger, M., & Strohmets, D.B. (1995). Effect of crime seriousness on simulated jurors' use of inadmissible evidence. *Journal of Social Psychology*, 135(4), 417–424. doi:10.1080/00224545.1995.9712211
- Ruva, C.L., & McEvoy, C. (2008). Negative and positive pretrial publicity affect juror memory and decision making. *Journal of Experimental Psychology: Applied*, 14, 226–235. doi:10.1037/1076-898X.14.3.226
- Sandays, M., & Dillehay, C. (1995). First-ballot votes, pre-deliberation dispositions, and final verdicts in jury trials. *Law and Human Behavior*, 19, 175–195. doi:10.1007/BF01499324
- Schvey, N.A., Puhl, R.M., Levandoski, K.A., & Brownell, K.D. (2013). The influence of a defendant's body weight on perceptions of guilt. *International Journal of Obesity*, 37, 1275–1281. doi:10.1038/ijo.2012.211
- Simon, D., Snow, C.J., & Read, S.J. (2004). The redux of cognitive consistency theories: Evidence judgments by constraint satisfaction. *Journal of Personality and Social Psychology*, 86, 814–837. doi:10.2139/ssrn.439984
- Terrance, C.A., Matheson, K., & Spanos, N.P. (2000). Effects of judicial instructions and case characteristics in a mock jury trial of battered women who kill. *Law and Human Behavior*, 24, 207–229. doi:10.1023/A:1005411003414
- Vidmar, N. (1972). Effects of decision alternatives on the verdicts and social perceptions of simulated jurors. *Journal of Personality and Social Psychology*, 22, 211–218. doi:10.1037/h0032605

Appendix

Case materials

Charge: Paul S is charged with the **attempted murder/assault** of Roy T on Friday, 15 December 2008 at 11.30 p.m.

Background: On the evening of Friday 15 December, Paul S, aged 27, was out for Christmas drinks with his friend Chris N, aged 30, in central Cambridge. They spent the evening drinking in a number of pubs before leaving at approximately 11.30 p.m. to walk the mile back to their homes. As they walked across a park, they encountered Roy T, a man who they alleged had had an affair with both of their wives in the past. Chris had since reconciled with his wife but Paul had separated from his wife and had lost custody of their two children. The three men got into a verbal argument and a fight broke out. Roy was struck twice in the face and his nose was broken. A pair of hands then grasped his throat from behind and someone began to strangle him. The fight was interrupted by the arrival of the police, who had been called by two witnesses.

The prosecution presents a witness who claims that a week earlier he had overheard Paul and Chris in a pub discussing how they would like to get revenge on Roy. They then present two further witnesses who claim to have seen Paul strangling Roy. The defence argue that it was not Paul who strangled Roy and challenge the reliability of all three witnesses, arguing that there is some evidence that the two alleged witnesses of the crime may have been heavily intoxicated at the time.