

Note Taking and Note Reviewing Enhance Jurors' Recall of Trial Information

CRAIG THORLEY*

Department of Psychological Sciences, University of Liverpool, Liverpool, UK

Summary: Jurors forget critical trial information and this can influence their verdicts. This study assessed (1) whether or not note taking during a trial improves mock jurors' recall of trial information and (2) whether or not reviewing these notes prior to recalling the trial offers any additional enhancement. Mock jurors first watched a trial video. Three-quarters were permitted to take notes whilst watching it. One-third of these mock jurors then reviewed their notes, one-third mentally reviewed the trial only, and one-third completed a filler task to prevent any form of reviewing. The remainder took no notes during the trial and also completed a filler task afterwards. All then had their memory of the trial assessed via free recall. The principal findings were (1) note taking enhanced recall of the trial and (2) note reviewing offered an additional recall enhancement. The applied implications of these findings are discussed. Copyright © 2016 John Wiley & Sons, Ltd.

Overview

The Sixth Amendment to The United States Constitution declares that 'In all criminal prosecutions, the accused shall enjoy the right to a speedy and public trial, by an impartial jury of the State and district wherein the crime shall have been committed...' (U.S. Const., amend VI). This right results in approximately 154 000 jury trials taking place within the United States each year (U.S. Department of State, 2009). Many other countries also ask jurors to decide upon defendants' innocence or guilt during criminal trials (Hans, 2008). Memory plays a crucial role in these trials as jurors need to encode the evidence, legal arguments, and judicial instructions presented to them, store this information in mind for the duration of the trial, and recall it during deliberation in order to reach a just verdict.

Little is known about real jurors' memory of trial information as deliberations are confidential. Researchers study this issue by presenting mock jurors with a video, audio recording, or transcript of a trial and then assessing their memory of it. These memory tests can be conducted on individual mock jurors or collaborative groups of mock jurors. Irrespective of the approach taken, mock jurors forget critical trial information and what they do remember can be inaccurate (e.g., Bodenhausen, 1988; Bourgeois, Horowitz, ForsterLee, & Grahe, 1995; Fitzgerald, 2000; Kassir & Wrightsman, 1979; Pennington & Hastie, 1988; Pritchard & Keenan, 1999, 2002). Importantly, these memory failures can impact upon verdicts. For example, Costabile and Klein (2005) presented individual mock jurors with a written summary of a murder trial. Some forgot about the defendant's wiretapped confession to the murder and these same jurors were less likely to convict the defendant as a result of this. It is therefore imperative that jurors' memory of trial information is as complete and accurate as possible to ensure just verdicts are reached.

Some judicial systems have introduced reforms to try and enhance jurors' memory of trial information. For ex-

ample, jurors in The Republic of Ireland are permitted access to a full trial transcript when deliberating during fraud trials (The Criminal Justice (Theft and Fraud Offences) Act, 2001, see McGreal, 2003). One widely adopted reform is allowing jurors to take notes onto blank sheets of paper during a trial. Jurors' memory of trial information can be enhanced by note taking (e.g., Thorley, Lorek, & Baxter, 2016). In judicial systems where note taking is permitted, jurors are not given a specific period of time to review their notes prior to recalling the trial and reaching a verdict. Research from other domains suggests that being given a period of time to review one's written notes may offer an additional memory enhancement (e.g., Fisher & Harris, 1973). The present study examined whether or not the effectiveness of note taking as a memory aid can be improved by granting jurors time to review their notes prior to having their memory of the trial assessed. If so, then this is an additional reform that judicial systems could consider adopting.

Juror note taking

There are variations in practice both between and within judicial systems with regards to whether or not jurors can take notes during a trial. For example, in England and Wales, Scotland, and Ireland it is a requirement for courts to permit note taking. The United States, however, legislates note taking on a state-by-state basis with some requiring courts to allow note taking (e.g., Illinois), others permitting it at the judge's discretion (e.g., Kansas), and some only permitting it if both the prosecution and defence agree to it (e.g., Nebraska). Australia and New Zealand also permit note taking at the judge's discretion and a survey of judges in both countries shows the majority permit it (Ogloff, Clough, Goodman-Delahunty, & Young, 2006). Field research suggests jurors do not believe note taking enhances their recall of trial information (e.g., Heuer & Penrod, 1988, 1994). No field research to date has, however, examined note taking jurors' actual recall of a trial. Contrary to the belief of real jurors, mock juror research consistently shows that taking notes is beneficial with individual note takers having more complete and accurate free recall of trial information than individual non-note takers (e.g., Fitzgerald, 2000; ForsterLee & Horowitz, 1997; Rosenhan, Eisner, & Robinson, 1994).

*Correspondence to: Dr. Craig Thorley, Department of Psychological Sciences, Eleanor Rathbone Building, University of Liverpool, Liverpool, UK, L69 7ZA.
E-mail: craig.thorley@liverpool.ac.uk

Individual note takers also provide more complete and accurate responses when answering cued recall questions about trial evidence (Hope, Eales, & Mirashi, 2014). When note takers collaborate in groups of five or six (ForsterLee, Kent, & Horowitz, 2005) or groups of 12 (Horowitz & Bordens, 2002) to freely recall trial information, then they also remember more than similar sized groups of non-note takers. Similarly, when note takers collaborate in groups of five or six to complete a recognition test that requires them to discriminate between true and false trial information, then they are more accurate than equivalent sized groups of non-note takers (Horowitz & ForsterLee, 2001). Note taking therefore offers a robust general memory enhancement.

ForsterLee, Horowitz, and Bourgeois (1994) and Thorley et al. (2016) both examined why note taking enhances juror's recollection of trial information. It is possible that this enhancement is caused by the process of note taking (so an encoding benefit), having note access at retrieval (so a retrieval benefit), or both. They examined this issue by comparing the free recall of non-note takers to note takers who either did or did not have access to their notes during retrieval. Those note takers without note access at retrieval recalled more trial information than non-note takers, confirming that note taking enhances encoding. There was also no difference in the volume of trial information recalled by those with and without access to notes at retrieval, suggesting note access provides no additional recall enhancement. To explain this effect, Thorley et al. refer to findings from the educational psychology literature where it has also been demonstrated that note taking enhances the encoding, and therefore retrieval, of lectures (see Kobayashi, 2005, for a review and meta-analysis). Whilst trials and lectures are inherently different types of to-be-remembered information, they both require individuals to sit and listen to the presentation of new, sometimes complex, materials for extended periods of time with the expectation that they will remember this information. Within the educational psychology literature it has been shown that note taking enhances encoding as it encourages generative processing of the presented information (e.g., Bretzing & Kulhavy, 1979; DiVesta & Gray, 1972; Peper & Mayer, 1978, 1986). Generative processing involves actively creating connections between diverse parts of new information (or between the new information and one's own prior knowledge) so that it is stored in memory in a meaningful and organised way (Wittrock, 1992; Wittrock, Marks, & Doctorow, 1975). Generative processing techniques that note takers use include grouping related ideas together under headings, summarising sections of the presentation, and creating concept maps (see Grabowski, 2004). When these techniques are employed, they result in a more elaborate and deeper encoding of the presented information, and durable memory traces are created (Craik & Lockhart, 1972; Craik & Tulving, 1975; Kiewra, 1989; Wittrock & Carter, 1975). The benefits of generative processing, however, are not restricted to deeper encoding. When new information is stored in memory in a meaningful and organised way, it is easier to retrieve as retrieval of one piece of information cues the recall of other related pieces of information (Mayer, 1984, 1996; Tulving, 1983). Thorley et al. suggest that note access at retrieval offers no additional

enhancement as the information noted down is no greater than that already stored in memory.

Note reviewing

Reviewing refers to the process of reading over self-generated materials, such as written notes made during a presentation, or provided materials, such as a transcript of a presentation provided by a third party, in order to refresh one's memory of it. To date, the impact of reviewing self-generated notes on jurors' memory of trial information has not been examined. In the educational psychology literature, however, the benefit of reviewing self-generated notes on students' recollection of lectures has been studied. A student's own notes can be seen as analogous to a juror's own notes. An overview of this literature is therefore provided as it can offer potential insights into how note reviewing will influence jurors' memory of trial information.

A large number of studies have examined the benefits of reviewing of self-generated notes on recall of lectures.¹ In a typical study, Carter and Van Matre (1975) presented student participants with a 17-minute lecture and then assessed their memory of it. Across four conditions, participants varied with regards to whether or not they could take notes and whether or not they could review these notes prior to the memory test. There were two note taking conditions where participants either reviewed their own notes for 5 minutes prior to the memory test or had their notes confiscated and engaged in mental reviewing (i.e., thinking about the lecture) for 5 minutes. This latter condition is important as it helps establish whether or not it is the process of reviewing, irrespective of whether notes are accessible or not, that facilitates recall. There were also two non-note taking conditions where participants either listened to the lecture only and then engaged in either mental reviewing for 5 minutes or completed a letter cancellation task during the review period to prevent mental reviewing. Those participants who took notes and reviewed them recalled the most, and those who simply listened to the lecture with no form of reviewing recalled the least. There was no difference between the two mental reviewing conditions. The positive effects of note reviewing on recollection of lectures has also been demonstrated elsewhere (e.g., DiVesta & Gray, 1972; Einstein, Morris, & Smith, 1985; Fisher & Harris, 1973; Kiewra et al., 1991).

Several researchers have considered why reviewing one's own notes may benefit recall, and several mechanisms have been proposed. Kiewra (1989) suggested that note reviewing gives students an additional learning trial that helps them consolidate the noted information in memory (see also DiVesta & Gray, 1972), helps them stave off the natural forgetting process, and allows them to relearn forgotten information. Kiewra et al. (1991) also suggest that students can engage in generative processing of presented information during the reviewing period. For note takers, the degree of generative processing they engage in could be greater than

¹ Several studies have presented note taking students with a lecture, then introduced a delay of several days, and then had students review their notes immediately prior to a memory test (e.g., Kiewra, DuBois, Christensen, Kim, & Lindberg, 1989). As jurors do not typically recall trials during deliberation after such long delays, these studies are not discussed.

that during encoding. This is because students do not have to divide their attention between listening to the lecture and note taking, meaning they have more information processing resources free for generating relations among lecture ideas or between lecture ideas and prior knowledge, and this leads to a deeper encoding of the lecture. Finally, Kiewra (1989) also suggests that a student's own notes act as retrieval cues that they can use to reconstruct the presentation in their minds and recall information that was not noted down (see also Rickards & Friedman, 1978).

Aims and hypotheses

The present study examined the impact of note taking and note reviewing on individual jurors' free recall and recognition of trial information. In this study, jurors watched a trial video and were either permitted or not permitted to take notes whilst doing so. One-third of the note takers then reviewed their notes (henceforth called the N+R group), one-third engaged in mental reviewing only (henceforth called the N+MR group), and one-third of note takers and each of the non-note takers completed a filler task to prevent any form of reviewing (henceforth called the N-R and NN groups, respectively). All then completed the two memory tests. Modelling the experience of real jurors, the note takers could refer to their notes when recollecting the trial.

The decision to have jurors complete a free recall test was made as this is the type of remembering real jurors engage in. It has previously been found that note taking has no impact upon individual juror memory when it is assessed via a recognition test and it may be the case that the powerful retrieval cues within these tests cancel out any benefits of note taking (Thorley et al., 2016). Recognition tests were, however, also employed here after the free recall test had taken place in an attempt to replicate these past findings and also examine whether or not reviewing provided any additional benefit.

It was possible to formulate several hypotheses for this study. In line with past research, it was expected that note takers would recall more trial information than non-note takers (e.g., Fitzgerald, 2000; ForsterLee & Horowitz, 1997; Rosenhan et al., 1994; Thorley et al., 2016). In line with the findings from the educational psychology literature, it was also expected that note takers who reviewed their notes prior to the free recall test would remember more than note takers who did not review their notes or engaged in mental reviewing only (e.g., Cater & Van Matre, 1975; DiVesta & Gray, 1972; Einstein et al., 1985; Fisher & Harris, 1973; Kiewra et al., 1991). These note reviewing hypotheses were, however, tentative as none of the studies cited made it explicitly clear whether or not the participants, who had access to their notes during testing, were able to refer to these notes during testing or not (they could in the present study). These hypotheses were therefore based upon the assumption that participants in these studies could refer to their notes during testing. There was no past research to indicate whether or not those who took notes and engaged in mental reviewing would recall more than those who took notes and did not engage in mental reviewing so this comparison was considered exploratory. In line with past

research it was also anticipated that there would be very few free recall errors in this study (Fitzgerald, 2000; Thorley et al.), meaning that the numbers were unlikely to vary across each condition. For the recognition test, past research suggests that there would be no difference between those who did not take notes and those who took notes but did not review (Thorley et al.). As the benefits of note reviewing on the recognition of trial information have not previously been examined, analyses concerning this condition were exploratory.

METHOD

Participants

A total of 144 adults (115 females, 29 males) aged 18–62 ($M=25.29$, $SE=.83$), who were either students or employees at the author's university, acted as jurors in this study. All were required to be jury eligible in England and Wales. This meant they had to be between 18 and 70 years of age, on the electoral register, and have lived in the UK for a period of at least 5 years since the age of 13. They could not be on bail, have served a prison or youth custody sentence of more than 5 years, have been in prison or youth custody for any amount of time in the last 10 years, or have suffered from a mental health condition or mental illness at any point in their lives.

Design

This study had one between-subjects independent variable (type of reviewing for note takers) with three conditions (no reviewing; mental reviewing only; note reviewing). There was also a separate control condition where no notes were taken and, therefore, note reviewing occurred. Thirty-six jurors were assigned to each condition in a quasi-random manner whereby testing for each took place on pre-determined day of the week and jurors, unaware of the conditions being tested each day, picked a session to attend.

The primary dependent variables were the verdict reached, the number of correct and incorrect pieces of trial information noted down by note takers, the number of correct and incorrect pieces of trial information recalled during a free recall test, and the proportion of hits and false alarms generated on a true/false recognition test about the trial.

Stimuli

The jurors watched a video of a 1992 murder re-trial with the case name New Jersey vs. Daniel Bias. The trial focussed on the death of a woman, Lise Bias, who was shot in the head and killed inside the home she shared with her husband, Daniel Bias. The prosecution argued that Mrs Bias was shot by Mr Bias. The defence argued that Mrs Bias was holding a gun to her head and threatening to kill herself in front of her husband, he tried to grab the gun off her, and she accidentally shot herself. The trial footage was edited to be 30-minute long and can be segmented into three sections. Section 1 contains the opening arguments of both attorneys. Section 2 contains the cross-examination of several witnesses (including the investigating officer, two coroners, and the defendant), an audio recording of Mr Bias's phone

call to the police informing them that his wife had been shot, and a video recording of a police interview with Mr Bias. Section 3 contains both attorneys' closing arguments and the judicial instructions. The jury's verdict is not shown so that participant jurors can reach their own verdict. Past research using this footage shows non-note taking mock jurors are split on the verdict (e.g., Hope, Memon, & McGeorge, 2004; Pritchard & Keenan, 1999, 2002; Ruva & Guenther, 2015; Ruva, McEvoy, & Bryant, 2007) and that note taking does not influence the verdicts jurors return (Thorley et al., 2016).

Note takers were provided with a blank lined notepad and pen. These materials are similar to those received by real jurors. Jurors who were not eligible to review any notes were given short stories to read during the review period. The short stories chosen were *The Necklace* by Guy De Maupassant (1884) and *The Street That Got Mislaid* by Patrick Waddington (1954). These stories were chosen as the content in no way overlaps with that of the trial video, reducing the likelihood of any post-event information contaminating the jurors' subsequent memory of the trial.

A booklet with plain lined paper was provided for the free recall test. The recognition test contained 12 true and 12 false statements about the trial evidence and legal arguments. For counterbalancing purposes there were two versions of the recognition test, with the true statements in one version being turned into false statements in the other. For example, in Version 1 a true statement was '*The coroner, Dr. Mihalakis, stated the muzzle of the weapon was of a distance within the range of self-infliction from Lisa Bias' head.*', whereas in Version 2 the false statement equivalent was '*The coroner, Dr. Mihalakis, stated the muzzle of the weapon was of a distance beyond the range of self-infliction from Lisa Bias' head.*'. This recognition test was previously used by Thorley et al. (2016).

A single questionnaire from Thorley et al. (2016) was used to collect data regarding the jurors' age and gender, their verdict (guilty or not guilty), and their satisfaction with their verdict (measured from 0% to 100% with a high score indicating complete satisfaction).

Procedure

All jurors were tested within single cubicles containing a desk, chair, computer, and headphones. They first read an information sheet stating that they would be required to act as jurors by watching a 30 minute recording of a real murder trial and then reaching a verdict. It also stated that they may be asked some questions about the trial at the end but no explicit reference was made to either the free recall or recognition tests. The jurors then signed an informed consent sheet and the study commenced. Jurors in the N-R, N+MR, and N+R conditions were provided with a blank notepad and pen and told they could take notes during the trial should they wish to do so. Consistent with the experience of real jurors, no guidance was given on what they should write down, how much they should write down, or how they should structure their notes. All jurors then watched the trial footage on their computers. Once the trial had finished, the four conditions varied in what they did.

Jurors in the NN condition were given the short stories to read for 10 minutes. Jurors in the N-R condition had their notes confiscated and were then given the short stories to read for 10 minutes. These short stories were included to occupy the jurors' minds and prevent any form of reviewing taking place. Jurors in the N+MR condition had their notes confiscated and were asked to think about the trial, trying to refresh their memory of it, for 10 minutes. Jurors in the N+R condition were asked to read over their notes for 10 minutes to refresh their memory of the trial. Periods of review have not previously been examined in relation to juror memory, meaning there was no precedent for deciding how long jurors should be allowed to review their notes for. The decision to allow jurors 10 minutes was made after consulting past research on note reviewing from the educational psychology literature. Review periods within these studies vary but, in one of them, a 5-minute review period was found sufficient to enhance recall of a 17-minute presentation (Carter & Van Matre, 1975). As the presentation length here is close to double that of this earlier study it was felt appropriate to use a review period that was also double the length.

After the review period, the short stories were collected in from the NN and N-R conditions. The N-R and N+MR conditions then had their notes returned. The three note taking conditions were informed they could refer to their notes during all subsequent tasks. All jurors then completed the demographic and verdict questionnaire. They then completed the free recall test. For the free recall test they were instructed to write down as much trial information as they could remember in any order they wished, focussing upon what the trial was about, the trial evidence offered, and legal arguments offered. They were asked not to focus on case irrelevant details such as the type of clothing worn by those who featured in the trial video or what the courtroom looked like. They were also asked to refrain from writing their opinions about the case. It was emphasised that there was no time limit for this free recall. The jurors were also instructed to turn their free recall sheets over once they had finished, meaning they could not see what they had just wrote. They then completed the recognition test at their own pace. Upon completing this, the study ended and they were debriefed. The study took approximately 75 minutes.

Data coding

The jurors' notes and free recall statements were blind scored by the author to identify how many pieces of trial information were correctly and incorrectly noted down or recalled. Here, trial information of interest included descriptions of what the trial was about, the names of individuals who featured in the trial, the evidence presented, and the legal arguments offered. To ensure consistency, all notes and free recall statements were scored against the coding scheme from Thorley et al. (2016) that lists each individual piece of trial information of interest. To give an example of the scoring, consider the following statement from the trial by a counsellor who spoke to Lise Bias two years before her death: '*She told me that there was an argument tonight with her husband and she became angry and threatened to shoot herself with her husband's gun.*' Here, four pieces of

trial information are conveyed (and each is represented in the coding scheme as a discrete piece of trial information). The first is that Lise had an argument with her husband that night, the second is that she became angry, the third is that she threatened to shoot herself, and the fourth is that the threat was made with her husband's gun. Jurors received one point for each of these pieces of information if it was noted down or recalled. Blind scoring was possible as an intern provided the lead author with photocopies of the notes and free recall statements that had the conditions they belonged to removed. These conditions were only revealed once scoring was completed.

During scoring, correct trial information was identified as that present in the trial and correctly noted down or recalled (e.g., Lise was right handed), incorrect information was identified as that present in the trial but incorrectly noted down or recalled (e.g., Lise was left handed) or information that was not present in the trial at all (e.g., Lise was wearing gloves at the time of the shooting). Subjective remarks, vague remarks, and opinions were not scored. Occasionally, the same evidence or legal argument is produced by more than one individual in a trial. For example, two coroners in the trial affirm that Lise Bias's hair was rinsed prior to autopsy. When making notes or completing free recall tests, jurors often write down the correct information but fail to specify the source of this information. To ensure equivalence in the scoring across conditions, any information that was repeated by more than one person during the trial was scored only once in the note taking and free recall results. This occurred irrespective of how many times each juror wrote this information down and whether or not the information could be attributed to a specific source.

Reliability scoring was conducted on 25% of the jurors' notes and 25% of their free recall statements. This was conducted by two student interns who worked independently and were blind to the aims of the study. To ensure consistency across scorers, they also had access to the coding scheme from Thorley et al. (2016). The inter-coder agreement was high for both the notes (83%) and the free recall (86%). Disagreements between the first and second scorers were resolved by the author.

RESULTS

Verdict and verdict satisfaction

54.86% of the jurors returned a guilty verdict (NN $M=72.22$, $SE=7.54$, 95% CI=58.33–86.11; N-R $M=50.00$, $SE=8.30$, 95% CI=33.33–66.67; N+MR $M=47.22$, $SE=8.33$, 95% CI=30.56–63.89; N+R $M=50.00$, $SE=8.46$, 95% CI=33.33–66.67). Logistic regression revealed that the condition jurors were in did not predict the verdict they reached, $\chi^2(3, N=144)=6.11$, $p=.11$, with the model explaining between 0.04% and 0.06% of the variance (Cox & Snell R^2 and Nagelkerke R^2 , respectively).

The jurors were generally satisfied in their verdicts ($M=66.81\%$) and a one-way between-subjects ANOVA revealed no significant difference in the satisfaction levels across the four conditions, $F(3, 143)=.17$, $p=.92$, $\eta_p^2=.01$

(NN $M=67.22$, $SE=3.12$, 95% CI=58.34–74.99; N-R $M=64.72$, $SE=3.26$, 95% CI=58.33–70.83; N+MR $M=67.22$, $SE=3.23$, 95% CI=60.83–73.05; N+R $M=68.06$, $SE=3.29$, 95% CI=61.11–73.89).

Note taking

The mean numbers of correct and incorrect pieces of trial information noted down by jurors in the three note taking conditions were compared using two separate one-way between-subjects ANOVAs. These analyses are essential for interpreting the findings of this study. Past research shows there is a positive correlation between the completeness and accuracy of a juror's notes and the volume of trial information a juror subsequently recalls (Rosenhan et al., 1994). Establishing that there are no differences in the completeness and accuracy of the N-R, N+MR, and N+R groups' notes means that these factors can be ruled out as the cause of any differences observed between them on their subsequent memory tests. Importantly, there was no significant difference in the number of correct pieces of trial information noted down across the three note taking conditions, $F(2, 107)=1.82$, $p=.17$, $\eta_p^2=0.03$ (N-R $M=28.64$, $SE=2.18$, 95% CI=24.22–33.05; N+MR $M=29.03$, $SE=2.16$, 95% CI=24.64–33.42; N+R $M=33.83$, $SE=2.09$, 95% CI=29.59–38.08). There was also no significant difference in the number of incorrect pieces of trial information noted down across the three note taking conditions, $F(2, 107)=1.90$, $p=.15$, $\eta_p^2=0.03$ (N-R $M=0.44$, $SE=0.13$, 95% CI=0.17–0.72; N+MR $M=0.86$, $SE=0.19$, 95% CI=0.47–1.25; N+R $M=0.61$, $SE=0.12$, 95% CI=0.36–0.86).

Free recall

The mean numbers of correct and incorrect pieces of trial information recalled by jurors in each of the four conditions were compared using two separate one-way between-subjects ANOVAs. The means, standard errors, and 95% confidence intervals for each condition are in Table 1. For correct recall, there was a significant difference between all four conditions, $F(3, 143)=14.91$, $p<.001$, $\eta_p^2=0.24$. Planned comparisons comparing each of the four conditions to each other were used to explore this effect. To reduce the likelihood of Type 1 Errors, a Bonferroni correction was used to lower alpha to .008. It was found that jurors in the

Table 1. The means (M), standard errors (SE), and 95% confidence intervals (CI) for the number of correct and incorrect pieces of trial information freely recalled by jurors.

Note taking	Correct recall		Incorrect recall	
	Mean (SE)	95% CI	Mean (SE)	95% CI
NN	18.56 (1.58)	[15.35, 21.76]	1.14 (0.22)	[0.69, 1.59]
N-R	25.89 (1.74)	[22.35, 29.43]	0.72 (0.72)	[0.42, 1.02]
N+MR	26.14 (1.69)	[22.72, 29.56]	0.75 (0.15)	[0.44, 1.06]
N+R	33.58 (1.31)	[30.92, 36.25]	0.64 (0.13)	[0.37, 0.91]

Note: NN = jurors who could not take notes during the trial; N-R = jurors who could take notes but who could not engage in any form of reviewing prior to the memory tests; N+MR = jurors who could take notes but who engaged in mental reviewing only prior to the memory tests; N+R = jurors who could take notes and also review these notes prior to the memory tests.

NN condition recalled fewer correct pieces of trial information than jurors in the N-R condition, $t(140)=3.26$, $p<.008$, $d=0.73$, N+MR condition, $t(140)=3.38$, $p<.008$, $d=0.77$, and N+R condition, $t(140)=6.69$, $p<.008$, $d=1.72$. There was no significant difference in the number of correct pieces of trial information recalled by jurors in the N-R and N+MR conditions, $t(140)=.10$, $p=.92$, $d=0.02$. Fewer correct pieces of trial information were recalled by jurors in the N-R condition than in the N+R condition, $t(140)=3.53$, $p<.008$, $d=0.83$, and by jurors in the N+MR condition than those in the N+R condition, $t(140)=3.48$, $p<.008$, $d=0.82$. Finally, there was no significant difference between all four conditions for the number of incorrect pieces of trial information recalled, $F(3, 143)=1.77$, $p=.15$, $\eta_p^2=.04$. In sum, note taking enhanced correct recall of trial information, and note reviewing offered an additional correct recall enhancement.

Recognition

Signal Detection Theory (SDT) is often used to analyse recognition test data. A brief overview of SDT is now provided for readers not familiar with it (see Stanislaw & Todorov, 1999, for a detailed overview). In SDT, a hit refers to an instance where a participant correctly remembers a studied piece of information whereas a false alarm refers to an instance where a participant incorrectly claims to have studied a non-studied piece of information. Hits and false alarms are used to calculate measures called d' and C . d' is the difference between the z -transformed probabilities of hits and false alarms and indicates how accurate participants are at discriminating between studied and non-studied information. No accuracy corresponds to a d' of 0, with higher scores indicating greater accuracy. C is the average of the transformed probabilities of hits and false alarms and is a measure of response bias. No bias corresponds to a C of 0, positive values indicate a bias towards responding 'false' to test items, and negative values indicate a bias towards responding 'true' to test items. Mean recognition test hits, false alarms, d' scores, C scores, and their associated standard errors and 95% confidence intervals can be seen in Table 2. Recognition accuracy was quite high with hit rates averaging 84% and false alarm rates averaging 22%. A one-way between-subjects ANOVA revealed no significant difference in the d' scores across the four conditions, $F(3, 143)=.40$, $p=.75$, $\eta_p^2=.01$. There was also little evidence of a response bias across the conditions with the C scores

averaging $-.09$. A one-way between-subjects ANOVA revealed no significant difference in the C scores across the four conditions, $F(3, 143)=2.20$, $p=.09$, $\eta_p^2=.04$. In sum, note taking and note reviewing had no impact upon recognition test performance.

GENERAL DISCUSSION

The benefits of note taking and note reviewing on jurors recall and recognition of trial evidence were assessed. In brief, the principal findings were that note taking alone enhanced jurors correct recall of trial information and that being afforded the opportunity to review these notes for 10 minutes prior to the memory test further enhanced correct recall. Recognition memory was not influenced by either note taking or note reviewing.

Benefits of note taking and note reviewing

It is unsurprising that note takers recalled more correct trial information than non-note takers. This was predicted and is consistent with past research (e.g., Fitzgerald, 2000; ForsterLee & Horowitz, 1997; ForsterLee et al., 1994, 2005; Hope et al., 2014; Horowitz & Bordens, 2002; Rosenhan et al., 1994; Thorley et al., 2016). As note takers (irrespective of whether or not they engaged in any form of reviewing) had access to their notes at retrieval this has implications for interpreting their recall performance relative to the NN group. The decision to allow note takers access to their notes at retrieval increases the ecological validity of the study (as real note taking jurors can typically access their notes when recalling a trial during deliberation). This design decision, however, means it is impossible to determine whether or not note takers recall was enhanced by the process of note taking (so an encoding benefit), having note access at retrieval (a retrieval benefit), or both. It is beyond the scope of the present study to determine this. Past research has, however, demonstrated that note taking improves recall as a result of an encoding enhancement (ForsterLee, Horowitz, & Bourgeois, 1994; Thorley et al., 2016), and there is no reason to suspect a different cause in this study.

The observation that note takers who reviewed had greater correct recall than note takers who did not review was also predicted. This prediction was based on similar findings from the educational psychology literature (Carter & Van Matre, 1975; DiVesta & Gray, 1972; Einstein et al., 1985; Fisher & Harris, 1973; Kiewra et al., 1991). This prediction was, however, tentative as it was unclear whether

Table 2. The mean hits and false alarms (M), d' , and C scores and their associated standard errors (SE) and 95% confidence intervals (CI) for jurors on the recognition test.

Note taking condition	Hits		False alarms		d'		C	
	M (SE)	95% CI	M (SE)	95% CI	M (SE)	95% CI	M (SE)	95% CI
NN	.81 (.02)	[.77, .86]	.21 (.02)	[.16, .25]	1.87 (0.10)	[1.67, 2.07]	-0.05 (0.06)	[-0.18, 0.09]
N-R	.83 (.02)	[.79, .87]	.24 (.02)	[.19, .28]	1.80 (0.11)	[1.57, 2.03]	-0.12 (0.05)	[-0.21, -0.03]
N+MR	.85 (.02)	[.80, .90]	.20 (.01)	[.16, .22]	1.75 (0.12)	[1.51, 1.99]	-0.01 (0.04)	[-0.10, 0.08]
N+R	.86 (.01)	[.83, .89]	.24 (.02)	[.20, .29]	1.91 (0.10)	[1.70, 2.12]	-0.18 (0.03)	[-0.28, -0.08]

Note: NN = jurors who could not take notes during the trial; N-R = jurors who could take notes but who could not engage in any form of reviewing prior to the memory tests; N+MR = jurors who could take notes but who engaged in mental reviewing only prior to the memory tests; N+R = jurors who could take notes and also review these notes prior to the memory tests.

participants in these earlier studies could refer to their notes during retrieval or not. In the present study all note takers could refer to their notes at test, so reviewing offers benefits over and above this. The observed enhancement was not due to note takers being given the opportunity to mentally rehearse the trial as those who reviewed their notes recalled more trial information than those who engaged in mental reviewing. It is also not due to note reviewers having more detailed notes to refer to as there was no significant difference in the volume of correct notes made by each of the note taking groups. The cause of the note reviewing enhancement observed here is likely similar to that in the educational psychology literature. To briefly recap, Kiewra (1989) suggested that reviewing gives individuals an opportunity to consolidate the noted information in memory (see also DiVesta & Gray, 1972), it provides an opportunity to relearn forgotten information, and the notes can act as retrieval cues that help note takers to reconstruct the presentation in their minds and recall information not noted down (Rickards & Friedman, 1978). Kiewra et al. (1991) also later suggested that note takers engage in generative processing during the review period and this leads to a deeper encoding of presentations.

There were few free recall errors in this study, with jurors averaging .81 each across all four conditions. Similar levels have also been reported by others (e.g., Fitzgerald, 2000; Thorley et al., 2016). Few errors were expected as free recall requires self-guided retrieval, meaning people typically only report information they are confident about (Koriat & Goldsmith, 1994, 1996). As the number of free recall errors made was near-floor they were uninfluenced by either note taking or note reviewing. It has previously been found that non-note takers and note takers make an equivalent negligible number of free recall errors (e.g., Fitzgerald; Thorley et al.) so the present study extends this finding to also include note reviewers.

Four-fifths of questions were answered correctly on the true/false recognition test, which is comparable to other juror memory studies (Fitzgerald, 2000; Thorley et al., 2016). Whilst the number of errors made may seem high in comparison with the number made on the free recall test, this is entirely expected as recognition tests utilise externally guided retrieval whereby the researcher decides which information the respondent needs to remember. Externally guided retrieval increases recollection errors as respondents are forced to try and remember information they may be unsure of (Koriat & Goldsmith, 1994, 1996). It was also found that note taking and note reviewing had no impact upon recognition test performance. This former finding replicates Fitzgerald and Thorley et al., so the novel finding here is that note reviewing also has no effect. These null effects can be interpreted in terms of the encoding-specificity principle (Thomson & Tulving, 1970; Tulving, 1983). In brief, recognition tests provide jurors with questions/statements/response options whose content overlaps with the information presented at encoding. In the present study, the recognition test statements were quite detailed, meaning there would have been significant overlap. This overlapping information provides jurors with powerful retrieval cues, meaning their ability to recognise trial facts would be at an equivalent level irrespective of whether or not they engaged in note taking and note reviewing.

Verdicts

Overall, there was a near even split in the number of guilty verdicts and not guilty verdicts returned and jurors were quite satisfied with their verdicts. Moreover, there was no association between the condition jurors were in and the verdicts they reached² or their satisfaction with their verdicts. In previous studies using this same trial video, jurors have also been evenly split with their verdicts (e.g., Hope et al., 2004; Pritchard & Keenan, 1999, 2002; Ruva & Guenther, 2015; Ruva et al., 2007; Thorley et al., 2016). Similarly, it has previously been found that there is no association between whether or not jurors can take notes and the verdicts they reach or their satisfaction with these verdicts (Thorley et al.). Thus, the current finding replicates this past research and extends it by also showing that note reviewing has no impact upon verdicts or satisfaction with these verdicts.

The fact that note reviewing enhanced correct recall of trial information without altering verdicts can be considered a positive outcome. In most criminal trials, there is no objectively correct verdict. In this specific trial however, jurors are typically evenly split when reaching a verdict regardless of whether they have taken notes or not. If note reviewing had led jurors to favour a guilty or not guilty verdict then they would have been behaving in a highly atypical manner for this trial and that may suggest that it can have a biasing effect on verdicts.

Applied implications and future research

The findings from this study allow two suggestions to be made with regards to enhancing real jurors' recollection of trial information. The first suggestion is that all courts should consider letting jurors take notes during trials. In some judicial systems it is currently at the discretion of the judge as to whether or not jurors can take notes (e.g., in those of Australia and New Zealand), and in others it is at the discretion of both lawyers (e.g., Nebraska). The evidence here, and in a large number of other studies, suggests that permitting note taking enhances jurors' correct recall of the trial information. This is important as jurors can sometimes forget critical trial evidence and this forgetting can lead to unjust verdicts (e.g., Costabile & Klein, 2005).

The second suggestion is that courts should consider allowing jurors to review their notes prior to recalling a trial during deliberation. At present, no time is allotted to reviewing notes at the end of a trial but prior to deliberation. Doing so could further enhance note taking jurors' recall of trial information. It is acknowledged that permitting time to review a full set of notes prior to deliberation may only be practical during shorter trials. Trial lengths can vary from

² It is acknowledged that there was a trend in the data for non-note takers to return more guilty verdicts than note takers (irrespective of whether they reviewed these notes or not). This trend, however, was not significant. It is likely that this trend was an artefact of this study given that non-note taking jurors are typically evenly split between guilty and not guilty verdicts after watching this trial (e.g., Hope et al., 2004; Pritchard & Keenan, 1999, 2002; Ruva & Guenther, 2015; Ruva et al., 2007; Thorley et al., In Press).

hours to months, with the average length of a criminal trial in the U.S. being five days (U.S. Department of State, 2009). The time required to review several days, weeks, or even months' worth of notes could be off-putting for judicial systems as it could significantly extend the length of a trial. When trials do last more than a day, an alternative may be to permit jurors to review their notes on a daily basis at the start of a day's proceedings to refresh their memory of what has come before. Courts do not typically allow jurors to take their notes out of the courtroom at the end of a day's proceedings so daily reviews are not currently possible. The benefits of regularly reviewing notes throughout a trial that lasts several days were not examined here but there is no reason to expect reviewing in these circumstances to be any less beneficial to a juror's recall of trial information. This is an issue that warrants examination in future.

Limitations

This study has several limitations that affect its external validity. The trial used was considerably shorter than a real trial as it lasted 30 minutes. Note taking behaviours may differ over extended periods of time but this issue has yet to be investigated. Real trials may also place more demands on memory as they contain a larger amount of evidence, arguments, and judicial instructions and may be more complex than the trial presented here. The results from the present study may therefore be underestimating the benefits of note taking and note reviewing in real trials when the volume and complexity of information to be remembered is greater. These shortcomings, however, affect the literature as a whole (see Bornstein, 1999, and Studebaker et al., 2002, for a discussion of these issues and others). A further limitation is that memory was studied at the level of the individual juror only. In most judicial systems, jurors typically collaboratively recall trial information during deliberation and the impact of note reviewing on recall during deliberation is unknown. There are, however, some judicial systems, such as in Brazil (Novak & Elster, 2014), where jurors reach verdicts independently with no deliberation, so these findings could prove particularly informative for these.

ACKNOWLEDGEMENTS

The author thanks Jane L. Sayers, Rebecca Cosgrove, Rebecca Hutton, and Alice Marten for assistance with data collection and data scoring. The author also thanks Christine L. Ruva for supplying the trial video used in this experiment.

REFERENCES

- Bodenhausen, G. V. (1988). Stereotypic biases in social decision making and memory: Testing process models of stereotype use. *Journal of Personality and Social Psychology*, *55*, 726–737. DOI:10.1037/0022-3514.55.5.726.
- Bornstein, B. H. (1999). The ecological validity of jury simulations: Is the jury still out? *Law and Human Behavior*, *23*, 75–91. DOI:10.1023/A:1022326807441.
- Bourgeois, M. J., Horowitz, I. A., ForsterLee, L., & Grahe, J. (1995). Nominal and interactive groups: Effects of preinstruction and deliberations on decisions and evidence recall in complex trials. *Journal of Applied Psychology*, *80*, 58–67. DOI:10.1037/0021-9010.80.1.58.
- Bretzing, B. H., & Kulhavy, R. W. (1979). Notetaking and depth of processing. *Contemporary Educational Psychology*, *4*, 145–153. DOI:10.1016/0361-476X(79)90069-9.
- Carter, J. F., & Van Matre, N. H. (1975). Note taking versus note having. *Journal of Educational Psychology*, *67*, 900–904. DOI:10.1037/0022-0663.67.6.900.
- Costabile, K. A., & Klein, S. B. (2005). Finishing strong: Recency effects in juror judgments. *Basic and Applied Social Psychology*, *27*, 47–58. DOI:10.1207/s15324834basp2701_5.
- Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671–684. DOI:10.1016/S0022-5371(72)80001-X.
- Craik, F. I., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*, 268–294. DOI:10.1037/0096-3445.104.3.268.
- DiVesta, F. J., & Gray, G. S. (1972). Listening and note taking. *Journal of Educational Psychology*, *63*, 8–14. DOI:10.1037/h0032243.
- Einstein, G. O., Morris, J., & Smith, S. (1985). Note-taking, individual differences, and memory for lecture information. *Journal of Educational Psychology*, *77*, 522–532. DOI:10.1037/0022-0663.77.5.522.
- Fisher, J. L., & Harris, M. B. (1973). Effect of note taking and review on recall. *Journal of Educational Psychology*, *65*, 321–325. DOI:10.1037/h0035640.
- Fitzgerald, J. M. (2000). Younger and older jurors: The influence of environmental supports on memory performance and decision making in complex trials. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *55*, 323–331. DOI:10.1093/geronb/55.6.P323.
- ForsterLee, L., & Horowitz, I. A. (1997). Enhancing juror competence in a complex trial. *Applied Cognitive Psychology*, *11*, 305–319. DOI:10.1002/(SICI)1099-0720(199708)11:4<305::AID-ACP457>3.0.CO;2-J.
- ForsterLee, L., Horowitz, I. A., & Bourgeois, M. (1994). Effects of notetaking on verdicts and evidence processing in a civil trial. *Law and Human Behavior*, *18*, 567–578. DOI:10.1007/BF01499175.
- ForsterLee, L., Kent, L., & Horowitz, I. A. (2005). The cognitive effects of jury aids on decision-making in complex civil litigation. *Applied Cognitive Psychology*, *19*, 867–884. DOI:10.1002/acp.1124.
- Grabowski, B. L. (2004). Generative learning contributions to the design of instruction and learning. In D. H. Jonassen (Ed.), *Handbook of research on educational communications and technology* (2nd edn, pp. 719–743). Mahwah, NJ: Lawrence Erlbaum Associates.
- Hans, V. P. (2008). Jury systems around the world. *Annual Review of Law and Social Science*, *4*, 275–297. DOI:10.1146/annurev.lawsocsci.4.110707.172319.
- Heuer, L., & Penrod, S. (1988). Increasing jurors' participation in trials: A field experiment with jury notetaking and question asking. *Law and Human Behavior*, *12*, 231–261. DOI:10.1007/BF01044383.
- Heuer, L., & Penrod, S. (1994). Juror notetaking and question asking during trials: A national field experiment. *Law and Human Behavior*, *18*, 121–150. DOI:10.1007/BF01499012.
- Hope, L., Eales, N., & Mirashi, A. (2014). Assisting jurors: Promoting recall of trial information through the use of a trial-ordered notebook. *Legal and Criminological Psychology*, *9*, 316–331. DOI:10.1111/lcrp.12003.
- Hope, L., Memon, A., & McGeorge, P. (2004). Understanding pretrial publicity: Predecisional distortion of evidence by mock jurors. *Journal of Experimental Psychology: Applied*, *10*, 111–119. DOI:10.1037/1076-898X.10.2.111.
- Horowitz, I. A., & Bordens, K. S. (2002). The effects of jury size, evidence complexity, and note taking on jury process and performance in a civil trial. *Journal of Applied Psychology*, *87*, 121–130. DOI:10.1037/0021-9010.87.1.121.
- Horowitz, I. A., & ForsterLee, L. (2001). The effects of note-taking and trial transcript access on mock jury decisions in a complex civil trial. *Law and Human Behavior*, *25*, 373–391. DOI:10.1023/A:1010655602400.
- Kassin, S. M., & Wrightsman, L. S. (1979). On the requirements of proof: The timing of judicial instruction and mock juror verdicts. *Journal of*

- Personality and Social Psychology*, 37, 1877–1887. DOI:10.1037/0022-3514.37.10.1877.
- Kiewra, K. A. (1989). A review of note-taking: The encoding-storage paradigm and beyond. *Educational Psychology Review*, 1, 147–172. DOI:10.1007/BF01326640.
- Kiewra, K. A., DuBois, N. F., Christian, D., McShane, A., Meyerhoffer, M., & Roskelley, D. (1991). Note-taking functions and techniques. *Journal of Educational Psychology*, 83, 240–245. DOI:10.1037/0022-0663.83.2.240.
- Kobayashi, K. (2005). What limits the encoding effect of note-taking? A meta-analytic examination. *Contemporary Educational Psychology*, 30, 242–262. DOI:10.1016/j.cedpsych.2004.10.001.
- Mayer, R. E. (1984). Aids to text comprehension. *Educational Psychologist*, 19, 30–42. DOI:10.1080/00461528409529279.
- Mayer, R. E. (1996). Learning strategies for making sense out of expository text: The SOI model for guiding three cognitive processes in knowledge construction. *Educational Psychology Review*, 8, 357–371. DOI:10.1007/BF01463939.
- McGreal, C. (2003). *Criminal Justice (Theft and Fraud Offences) Act, 2001*. Dublin: Thomson Round Hall.
- Novak, S., & Elster, J. (Eds) (2014). *Majority decisions: Principles and practices*. Cambridge: Cambridge University Press.
- Ogloff, J. P. R., Clough, J., Goodman-Delahunty, J., & Young, W. (2006). *The jury project: Stage 1-A survey of Australian and New Zealand judges*. Melbourne, Australia: Australian Institute of Judicial Administration Incorporated.
- Pennington, N., & Hastie, R. (1988). Explanation-based decision making: Effects of memory structure on judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14, 521–533. DOI:10.1037/0278-7393.14.3.521.
- Peper, R. J., & Mayer, R. E. (1978). Note taking as a generative activity. *Journal of Educational Psychology*, 70, 514–522. DOI:10.1037/0022-0663.70.4.514.
- Peper, R. J., & Mayer, R. E. (1986). Generative effects of note-taking during science lectures. *Journal of Educational Psychology*, 78, 34–38. DOI:10.1037/0022-0663.78.1.34.
- Pritchard, M. E., & Keenan, J. M. (1999). Memory monitoring in mock jurors. *Journal of Experimental Psychology: Applied*, 5, 152–168. DOI:10.1037/1076-898X.5.2.152.
- Pritchard, M. E., & Keenan, J. M. (2002). Does jury deliberation really improve jurors' memories? *Applied Cognitive Psychology*, 16, 589–601. DOI:10.1002/acp.816.
- Rickards, J. P., & Friedman, F. (1978). The encoding versus the external storage hypothesis in note taking. *Contemporary Educational Psychology*, 3, 136–143. DOI:10.1016/0361-476X(78)90020-6.
- Rosenhan, D. L., Eisner, S. L., & Robinson, R. J. (1994). Notetaking can aid juror recall. *Law and Human Behavior*, 18, 53–61. DOI:10.1007/BF01499143.
- Ruva, C. L., & Guenther, C. C. (2015). From the shadows into the light: How pretrial publicity and deliberation affect mock jurors' decisions, impressions, and memory. *Law and Human Behavior*, 39, 294–310. DOI:10.1037/lhb0000117.
- Ruva, C., McEvoy, C., & Bryant, J. B. (2007). Effects of pre-trial publicity and jury deliberation on juror bias and source memory errors. *Applied Cognitive Psychology*, 21, 45–67. DOI:10.1002/acp.1254.
- Studebaker, C. A., Robbenolt, J. K., Penrod, S. D., Pathak-Sharma, M. K., Groscup, J. L., & Devenport, J. L. (2002). Studying pretrial publicity effects: New methods for improving ecological validity and testing external validity. *Law and Human Behavior*, 26, 19–41. DOI:10.1023/A:1013877108082.
- Thomson, D. M., & Tulving, E. (1970). Associative encoding and retrieval: Weak and strong cues. *Journal of Experimental Psychology*, 86, 255–262. DOI:10.1037/h0029997.
- Thorley, C., Baxter, R. E., & Lorek, J. (2016). The impact of note taking style and note availability at retrieval on mock jurors' recall and recognition of trial information. *Memory*, 24, 560–574. DOI:10.1080/09658211.2015.1031250.
- Tulving, E. (1983). *Elements of episodic memory*. Oxford: Clarendon Press.
- U.S. Const. amend. VI U.S. Department of State. Bureau of International Information Programs (2009). Anatomy of a Jury Trial. *eJournal USA*, 7. Retrieved from <http://photos.state.gov/libraries/amgov/30145/publications-english/EJ-jury-0709.pdf>
- Wittrock, M. C. (1992). Generative learning processes of the brain. *Educational Psychologist*, 27, 531–541. DOI:10.1207/s15326985ep2704_8.
- Wittrock, M. C., & Carter, J. F. (1975). Generative processing of hierarchically organized words. *The American Journal of Psychology*, 88, 489–501. DOI:10.2307/1421780.
- Wittrock, M. C., Marks, C., & Doctorow, M. (1975). Reading as a generative process. *Journal of Educational Psychology*, 67, 484–489. DOI:10.1037/h0077017.