
On the Many Impacts of Inadmissible Testimony: Selective Compliance, Need for Cognition, and the Overcorrection Bias

Samuel R. Sommers
University of Michigan

Saul M. Kassin
Williams College

A mock juror study tested three hypotheses: (a) Jurors comply selectively with instructions to disregard inadmissible testimony, (b) this effect is greater among jurors who are high rather than low in the need for cognition (NC), and (c) high-NC decision makers sometimes overcorrect against the perceived biasing agent of inadmissible testimony. Participants read a trial summary in which the admissibility of an incriminating wiretap and the basis for this ruling were manipulated. Consistent with the prediction that jurors would be motivated to reach a "just" verdict, participants disregarded an inadmissible wiretap when it was deemed unreliable but not when it violated due process. This pattern of selective compliance was found only among high-NC jurors. High-NC participants also exhibited signs of bias overcorrection, estimating the probability of the defendant's guilt to be lower in the inadmissible-unreliable condition than in the no-wiretap control. The theoretical and practical implications of these results are discussed.

Before jurors begin deliberating, they are instructed to render a verdict based solely on the evidence presented in court and the application of the law as formalized in the judge's instructions. Over the years, however, researchers have found that jurors are also influenced by nonevidentiary factors such as pretrial publicity (Kramer, Kerr, & Carroll, 1990; Studebaker & Penrod, 1997); the defendant's race (Sommers & Ellsworth, 2000); presumptuous cross-examination questions (Kassin, Williams, & Saunders, 1990); exposure to a death qualification voir dire (Haney, 1984); hearsay as communicated by an expert witness (Schuller, 1995); and inadmissible testimony that reveals a defendant's prior record, financial resources, illegally wiretapped

conversations, or possibly coerced confessions (for a review, see Kassin & Studebaker, 1998).

According to the rules of evidence, information is generally admissible at trial if it has logical relevance to the case and probative value. There are exceptions, however, such as when information is excluded because it was illegally obtained or because it is deemed to be too prejudicial, redundant, or confusing to the jury. When an attorney or witness discloses an inadmissible item of evidence and opposing counsel objects to this disclosure, the judge, as a matter of procedure, will typically strike the information from the trial record and admonish the jury to disregard it.

On the question of whether jurors are willing and able to comply with such judicial directives, psychological theory and research have been mixed. On one hand, studies have shown that people can at times ignore information on command (Johnson, 1994). This effect has been observed in a forensic context, as when, for example, mock jurors discount eyewitness testimony that is later discredited (Elliot, Farrington, & Manheimer, 1988; Schul & Manzury, 1990; Weinberg & Baron, 1982). On the other hand, belief perseverance studies have shown that people often cling to newly created beliefs

Authors' Note: The authors wish to thank Duane Wegener and three anonymous reviewers for their comments on an earlier version of this article. Address correspondence concerning this article to Samuel R. Sommers, 3225 East Hall, Social Psychology, University of Michigan, Ann Arbor, MI 48109-1109, or Saul M. Kassin, Department of Psychology, Williams College, Bronfman Science Center, Williamstown, MA 01267; e-mail: ssommers@umich.edu or skassin@williams.edu.

PSPB, Vol. 27 No. 10, October 2001 1368-1377
© 2001 by the Society for Personality and Social Psychology, Inc.

even after the evidentiary basis for these beliefs is discredited (Anderson, Lepper, & Ross, 1980; Johnson & Seifert, 1994; Schul & Burnstein, 1985), as when, for example, the conviction rate among mock jurors increases significantly after incriminating disclosures ruled inadmissible by the judge (Carretta & Moreland, 1983; Greene & Dodge, 1995; Sue, Smith, & Caldwell, 1973; Wissler & Saks, 1985). Some researchers have found that a judge's admonishment can even backfire by further increasing the impact of the forbidden information on jurors (Broeder, 1959; Edwards & Bryan, 1997; Pickel, 1995; Wolf & Montgomery, 1977).

In an effort to reconcile the inconsistencies evident in past research, Kassir and Sommers (1997) proposed that jurors comply on a selective basis with instructions to disregard, depending on the stated or presumed reason for that ruling. Specifically, we argued that jurors are motivated in large part by a desire to render a "just" verdict, regardless of whether this decision conforms to the rules of evidence. Achieving a just verdict is a function of both being accurate about whether the defendant committed the acts in question and believing that the acts in question, if committed, merit punishment. This latter determination is related to jury nullification and is most likely to become an issue in cases involving controversial laws or application of laws (see Horowitz & Willging, 1991). Other researchers have used different terminology to refer to this motivation of jurors to render a verdict that they perceive to be just. Fleming, Wegener, and Petty (1999), for example, use the phrase "accuracy motivation" to relate this effect to the more general psychological literature on attitudes and bias correction. In the jury domain, our terminology of "just verdict motivation" has the advantage of being consistent with the connotation of similar legal phrases such as "just desserts."

To test this just verdict hypothesis, Kassir and Sommers (1997) presented mock jurors with a murder trial summary in which the judge ruled an incriminating wiretap admissible, inadmissible because it was illegally obtained, or inadmissible because it was not reliable (there was also a no-wiretap control group). As measured by conviction rates, participants were influenced by the incriminating evidence when it was ruled admissible and when it was ruled inadmissible because it was obtained illegally (a so-called legal technicality). This influence also was present in online interpretations of the trial evidence, as the disclosure of both the admissible and technically inadmissible wiretaps affected mock jurors' ratings of subsequent items of evidence. Participants did comply with the instruction to disregard an unreliable inadmissible wiretap, however, presumably because this evidence was not relevant to the pursuit of a just, accurate verdict. Other studies have yielded the comparable finding that the biasing impact of inadmissi-

ble testimony and pretrial publicity can be erased when mock jurors are led to suspect that the source of the information has an ulterior motive or lacks credibility (Fein, McCloskey, & Tomlinson, 1997; Fein, Morgan, Norton, & Sommers, 1997).

Recent research suggests that there also might be individual differences as well in jurors' motivation and ability to correct for the biasing influence of inadmissible evidence. The present study was designed to investigate one such difference: the need for cognition (NC). Cacioppo and Petty (1982) defined NC as a tendency to engage in and enjoy effortful cognitive activities. Because jurors are charged with the responsibilities of attending to information, integrating facts, and making judgments, we believed that individual differences in NC might have a bearing on juror motivation and performance. As a result, the present study was conceived with two goals in mind: (a) to examine the relationship between the influence of inadmissible evidence and jurors' NC and (b) to test the hypothesis that high-NC jurors would exhibit an overcorrection bias against the alleged biasing agent of inadmissible testimony.

Using Cacioppo and Petty's scale (1982; Cacioppo, Petty, & Kao, 1984), researchers over the years have found that people who are high in NC devote close attention to ongoing tasks, actively seek out relevant information, scrutinize arguments carefully, and base their decisions primarily on the strength of the available evidence. In contrast, those who are low in NC engage in relatively little message-relevant critical thinking and, as such, are influenced more by superficial cues such as a speaker's physical appearance or the reactions of others in the audience (for a review, see Cacioppo, Petty, Feinstein, & Jarvis, 1996). Previous studies of bias correction have revealed that when potential biases are obvious or salient, high-NC individuals are more likely than lows to exert the cognitive effort required for correction (for a review, see Wegener, Kerr, Fleming, & Petty, 2000). To our knowledge, however, such an investigation has never been conducted using a mock juror paradigm, despite the clear relevance of NC to the study of juror decision making.

In light of previous findings, we expected that high-NC mock jurors would be more likely than lows to engage in active correction for the perceived biasing influence of inadmissible evidence. After all, differences in processing styles should make highs more responsive than lows to variations in evidence during the course of a trial, as well as more likely to detect the possibility of bias. But beyond cognitive differences, we expected that NC also would be associated with different motivational levels among jurors. We predicted that high-NC jurors, who are accustomed to viewing mental tasks as challenges to be solved accurately and who tend to actively monitor

the need to adjust for potential biases in decision making, would exhibit a greater motivation to render a just verdict. And because recent theories of bias correction suggest that motivation guides correction (Wegener & Petty, 1997), we expected that selective compliance with judicial instructions to disregard would be more pronounced among high-NC mock jurors than among lows.

Not only does Wegener and Petty's (1997) Flexible Correction Model hold that motivation and NC can influence bias correction, but it also postulates that personal and situational factors can influence whether perceivers overestimate the potential influence of biasing information and adjust their perceptions too far in the opposite direction. The model suggests that people who are highly motivated to identify, evaluate, and correct for potential sources of bias in their judgments are most likely to demonstrate this overcorrection. Researchers have examined this phenomenon in conjunction with NC and have found that when biases are easily detected, high-NC individuals are more likely than lows to overcorrect (e.g., DeSteno, Petty, Wegener, & Rucker, 2000; Martin, Seta, & Crelia, 1990). But in the legal setting, whereas numerous psychologists have identified situations in which jurors do not fully comply with instructions to disregard, little attention has been paid to the possibility that jurors sometimes go too far in their bias correction.

In the trial context, an overcorrection hypothesis would suggest that jurors with a strong motivation for accuracy—or a sense of personal involvement with the case—might bend over backward to assess the defendant as innocent when they view an incriminating item of testimony as a possible contaminant. Indeed, this phenomenon may have occurred in the infamous criminal trial of O. J. Simpson. In that case, Detective Mark Fuhrman, a key witness for the prosecution, had his credibility impeached when taped conversations disclosed that he had used racist language. After Simpson's controversial acquittal, many of the jurors said they suspected that Fuhrman was part of a conspiracy to frame Simpson, a belief that led them to view the rest of the state's case through skeptical eyes and to make adjustments in their decision making accordingly (Dunne, 1995). By their own admission, jurors in that case felt highly invested in its outcome, making them more likely to overcorrect. Jurors high in NC might share this tendency for overcorrection, especially because the potential biasing influence of inadmissible evidence becomes obvious when a judge issues an instruction to disregard.

In the present study, participants who were pretested on the Need for Cognition Scale (Cacioppo, et al., 1984) read a murder trial summary in which the presence and admissibility ruling of an incriminating wiretap were varied. Three hypotheses were tested. The first hypothesis

was that participants would comply with an instruction to discount the wiretap device if it was excluded for a lack of reliability but not if it was excluded because it violated due process; this result would replicate the findings of Kassin and Sommers (1997). The second hypothesis was that mock jurors high in NC would be more likely than lows to exhibit this pattern of selective compliance (which takes into account the causal basis of the judge's ruling). The third hypothesis was that their motivation to arrive at a just verdict also would lead high-NC jurors to overcorrect against the perceived biasing influence of the unreliable inadmissible evidence. No overcorrection was expected for reliable inadmissible evidence, which, given a just verdict motivation, would still be considered relevant to a final determination of guilt.

METHOD

Participants

Participants included 105 introductory psychology students who assumed the role of mock jurors in exchange for either course credit or financial compensation. All participants were randomly assigned to one of the four trial conditions (no-wiretap control, admissible wiretap, inadmissible—due process, inadmissible-unreliable), which yielded a 2 (high vs. low NC) \times 4 (trial condition) factorial design.¹ One month prior to the study, participants completed the short version of the Need for Cognition Scale in a mass testing situation (those who were not present at this session completed the scale after their posttrial questionnaires). The scale consisted of 18 items measured on a 9-point scale ranging from -4 to +4. Via a median split, participants scoring higher than 26 were classified as high in NC, whereas those scoring less than 26 were classified as low in NC (2 participants with scores of 26 were assigned at random to the high- and low-NC groups).

Materials

Participants received one of seven versions of a criminal trial summary titled *NY v. Givens*: three different experimental trial versions crossed with early versus late wiretap disclosure (see Note 1) in addition to one no-wiretap control group. The case involved a man charged with murdering his ex-wife and male neighbor. The prosecution alleged that the defendant had killed the two victims in a fit of jealous rage and presented as evidence testimony from two police officers, a private investigator hired by the defendant, the coroner, and an eyewitness. The defendant claimed that he had returned to his former home to retrieve some personal papers when he stumbled upon the bodies, and a friend of the defendant testified as well. The trial presentation consisted of 26 paragraphs that summarized the opening statements,

the direct and cross-examinations of seven witnesses, closing arguments, and the judge's charge to the jury. Various objections and rulings on admissibility were included in these paragraphs. The summary was presented to each participant individually on a Power Macintosh 6100/60 using SuperLab 1.4.

In the no-wiretap control version, the state's case against the defendant was circumstantial, incomplete, and ambiguous. In three experimental versions, however, a police officer revealed that a wiretap from an unrelated case had picked up a tape-recorded phone conversation in which the defendant confessed to a friend moments after fleeing the scene ("I cut him up good . . . this stud in a red ski cap. The blade, yeah, I threw it in the bay. No, the bay."). In all conditions, the defense immediately objected to this disclosure. In the admissible condition, the judge overruled the objection, admitted the tape into evidence, and explained that it was a proper form of evidence. In the inadmissible—due process condition, the judge sustained the objection and admonished the jury to disregard the tape because it was obtained without a proper warrant; he then explained that to ensure a fair trial, a jury should not consider evidence that was illegally obtained. In the inadmissible-unreliable condition, the judge also sustained the objection but then admonished the jury to disregard the tape because it was barely audible due to background noise and static. He explained that to ensure a fair trial, a jury should not consider evidence that is unreliable.

Procedure

After arriving, participants were told that they would be taking part in a study of juror decision making. They were told that to simplify and condense the trial presentation, we had prepared a written summary of an actual case and that the facts would be presented in paragraphs and displayed on a computer. To examine the online impact of the wiretap evidence and judge's ruling, as well as their possible effects on perceptions of subsequent evidence, each participant was asked to rate each item of evidence during the trial presentation. The evidence ratings were made on a computer keyboard using a scale of 1 to 9, with 5 defined as the neutral point. Participants were told that if the information within a paragraph led them to see the defendant as guilty, they should enter a number greater than 5; if it led them to see the defendant as innocent, they should enter a number less than 5 ("So a rating of 9 means the paragraph makes the defendant look *very guilty*, a rating of 1 means the paragraph makes him look *very innocent*, and a rating of 5 means the paragraph is *not informative or influential one way or the other*"). Participants were asked to enter a response as soon as they were ready to do so and were told that each response would activate the presentation of the next

paragraph. For each item of evidence, SuperLab 1.4 stored both the online rating and the response (reading and decision making) time at an accuracy of 16 milliseconds. The trial presentation took an average of 10 minutes.

Afterward, participants completed a written questionnaire. First, they rendered a verdict (guilty or not guilty) and rated their confidence in that judgment on a 10-point scale (1 = *not at all confident*, 10 = *very confident*). With verdicts a dual function of a juror's subjective probability that the defendant committed the crime and the standard of proof deemed necessary for conviction, both of these variables also were assessed. Participants thus estimated the likelihood that the defendant committed the crime by circling a percentage from 0 to 100, and they then completed the sentence, "The defendant should be found guilty if there is at least a ___% chance that he committed the crime." They were then asked to list, in order of importance, the factors that led to their verdicts—open-ended responses that were later coded for mention of the wiretap.

Finally, to ensure that participants had correctly recalled the information presented within their respective conditions, they were asked the following questions: "According to the prosecutor, what did Givens say to his friend in the taped telephone call? As best you can, write down what he said word for word" and "Did the judge allow the taped telephone statement into evidence (yes or no)—why or why not?" Those in the experimental groups were also asked directly whether they were influenced in their decision making by what the defendant had said in the taped conversation (yes or no) and to indicate why or why not.

RESULTS

Participants in all experimental conditions were asked to recall the judge's admissibility ruling as well as the basis for that ruling. Across experimental conditions, their accuracy rates were uniformly high on both manipulation checks—100% recalled the judge's admissibility ruling and 98.9% correctly recalled his reason. There were no between-group differences on either measure, and groups did not differ in their explicit memory for what was said on the wiretap recording.

Overall, participants estimated that there should be at least an 89.2% probability that the defendant committed the crime before voting guilty, a quantification of "beyond a reasonable doubt" that is virtually identical to estimates obtained in prior studies (e.g., Kagehiro, 1990; Kassir & Sommers, 1997). There were no between-group differences as a function of trial condition, $F(3, 97) = 1.05$, *ns*; NC, $F(1, 97) < 1$; or their interaction, $F(3, 97) < 1$.

Verdict-Related Measures

Collapsing across conditions, 42 mock jurors voted guilty and 63 voted not guilty, yielding a conviction rate of 40%. As in Kassin and Sommers (1997), the pattern of verdicts across conditions indicated the selective compliance that is predicted by a just verdict motivation in jurors. Specifically, the conviction rate was 60% in the admissible wiretap condition, compared with only 27% in the inadmissible-unreliable and no-wiretap control groups, $\chi^2(2, n = 105) = 5.3, p < .05$. There was a 40% conviction rate in the inadmissible-due process condition, which was not significantly different than rates in the other conditions.

The pattern of results differed for jurors who were high versus low in NC (see Figure 1). For high-NC jurors, the conviction rate was 80% in the admissible wiretap condition, significantly higher than the no-wiretap control (27%), inadmissible-due process (44%), and inadmissible-unreliable (13%) conditions, $\chi^2(3, n = 53) = 14.3, p < .005$. In contrast, there were no significant between-group differences among low-NC participants, who demonstrated conviction rates of 25% in the no-wiretap control group, 40% in the admissible and inadmissible-unreliable conditions, and 36% in the unreliable-due process condition, $\chi^2(3, n = 52) < 1.2$. For both high- and low-NC mock jurors, the conviction rate in the inadmissible-unreliable condition was not significantly lower than in the control group, failing to provide evidence of overcorrection.

To obtain a more sensitive measure of verdict preferences, a scalar variable was created by multiplying each participant's confidence rating by either a -1 for a not guilty verdict or a $+1$ for a guilty verdict (scores thus ranged from -10 to $+10$). A 2×4 ANOVA on these scaled verdict confidence scores revealed a significant main effect for trial condition, $F(3, 97) = 3.95, p < .02$. To test for selective compliance, we performed a planned comparison of scaled verdict confidence scores in the admissible condition with scores in the control and inadmissible-unreliable conditions.³ As predicted, compared with the admissible wiretap condition ($M = 1.73$), verdict scores were significantly lower in the no-wiretap control and inadmissible-unreliable conditions ($M_s = -3.07$ and -3.33 , respectively), $t(97) = 3.27, p < .01$. The difference between verdict confidence scores in the admissible condition and inadmissible-due process condition ($M = -0.37$) was not significant, $t(97) = 1.31, ns$, suggesting that the reliable yet technically inadmissible wiretap was not fully discounted by mock jurors.

There was no main effect on this measure for NC, $F(1, 97) < 1$. As predicted, however, there was a significant two-way interaction, $F(3, 97) = 3.30, p < .05$. Figure 2 shows that the pattern of selective compliance was exhibited only by high-NC jurors, among whom the defendant was seen as less guilty in the no-wiretap control and

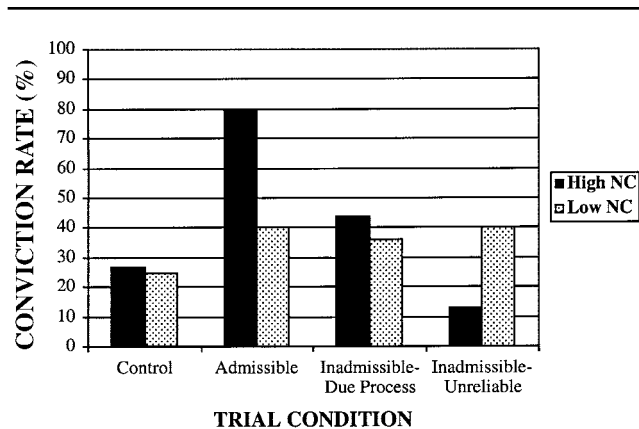


Figure 1 Conviction rates by trial condition and juror need for cognition (NC).

inadmissible-unreliable conditions ($M_s = -2.71$ and -5.47 , respectively) than in the admissible condition ($M = 4.60$), $t(97) = 4.05, p < .001$. Among these same high-NC participants, verdict scores were not significantly lower in the inadmissible-due process condition ($M = 0.56$) than in the admissible condition, $t(97) = 1.81, ns$, suggesting that compliance was selective. As shown in Figure 2, low-NC mock jurors did not exhibit this pattern, both planned comparison $t_s(97) < 1$.

The continuous nature of the scaled verdict-confidence measure also allowed us to perform a more direct test of our prediction that high- and low-NC jurors would differ in the extent to which they demonstrated selective compliance. We used a planned comparison to test for an interaction between NC and evidence reliability in the four inadmissible evidence cells in the 2×4 design.⁴ This interaction contrast was significant, $t(97) = 2.01, p < .05$, providing the most direct support for our hypothesis that selective compliance would be more pronounced among high- rather than low-NC jurors. Finally, comparing scaled verdict scores in the control and inadmissible-unreliable conditions did not reveal evidence of overcorrection among either high- or low-NC jurors, $t_s(97) < 1$.

After rendering a verdict and rating their confidence in it, participants were asked to estimate the likelihood that the defendant had committed the crime on a 0-to-100-point scale. Closely paralleling the verdict-confidence data, a 2×4 ANOVA revealed a significant main effect for trial condition, $F(3, 97) = 5.00, p < .005$. Specifically, a planned comparison revealed that probability-of-commission estimates were higher in the admissible condition ($M = 74.7$) than in the control ($M = 68.0$) and inadmissible-unreliable conditions ($M = 58.8$), $t(97) = 2.51, p < .02$; this result was driven by the difference between the admissible and inadmissible-unreliable conditions, $t(97) = 3.31, p < .005$. Inadmissible-due pro-

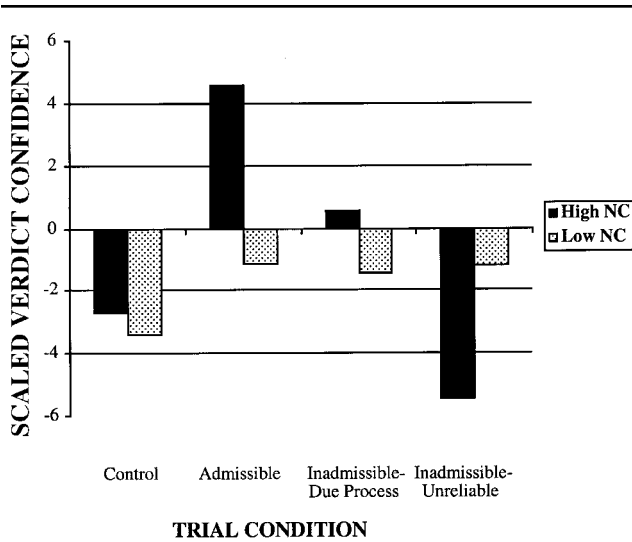


Figure 2 Scaled verdict confidence scores by trial condition and juror need for cognition (NC).

NOTE: Scaled verdict confidence was computed by multiplying participants' confidence in their verdict (on a scale of 1 to 10) by -1 for a not guilty verdict and +1 for a guilty verdict. The computed scores therefore range from -10 (*extreme confidence in a not guilty verdict*) to +10 (*extreme confidence in a guilty verdict*).

cess condition estimates ($M = 75.5$) were not significantly lower than those in the admissible condition, $t(97) < 1$.

As with verdict scores, high- and low-NC participants did not differ in their probability estimates overall, $F(1, 97) = 2.20, ns$. There was, however, a significant interaction, $F(3, 97) = 2.90, p < .05$, as depicted in Figure 3. Among high-NC mock jurors, probability ratings of the defendant's guilt were higher in the admissible condition ($M = 80.3$) than in the control ($M = 76.4$) and inadmissible-unreliable conditions ($M = 52.7$), $t(97) = 2.45, p < .02$; again, this significant result was driven by the difference between the admissible and inadmissible-unreliable conditions, $t(97) = 4.06, p < .001$. There were no significant differences between highs in the admissible and inadmissible-due process ($M = 79.1$) conditions, $t(97) < 1$. As with the verdict score data, there was no evidence of selective compliance among low-NC jurors, all planned $t_s(97) < 1$.

Once again, we analyzed the interaction between NC and evidence reliability in the four inadmissible evidence cells to obtain a direct test of the hypothesis that selective compliance would differ by NC (see Note 4). This planned contrast was significant, $t(97) = 2.08, p < .05$, suggesting that the basis for the judge's ruling of admissibility did have different effects on high- versus low-NC mock jurors. Finally, consistent with the hypothesis that high-NC jurors would exhibit bias overcorrection, their probability estimates were significantly lower in the inadmissible-unreliable condition than in

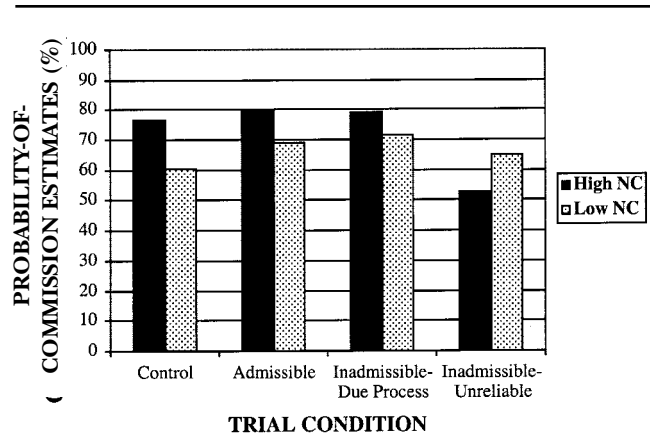


Figure 3 Probability-of-commission estimates by trial condition and juror need for cognition (NC).

NOTE: Participants were asked to estimate the likelihood that the defendant committed the crime in question by circling one number on a scale that ranged from 0% to 100% in increments of 5 percentage points.

the no-wiretap control group, $t(97) = 2.91, p < .01$. No such overcorrection was demonstrated by low-NC jurors, $t(97) < 1$.

To assess the self-reported influence of the wiretap evidence, we asked participants to list all factors that contributed to their verdict. We then coded these open-ended responses for whether the wiretap was on the list. Among those in the experimental conditions, the wiretap was cited by 57% in the admissible condition and 60% in the inadmissible-due process condition but only by 17% in the inadmissible-unreliable condition, $\chi^2(2, n = 90) = 14.1, p < .001$. This same pattern emerged in the analysis of only the high-NC participants' responses, as they listed the wiretap more often in the admissible (73%) and inadmissible-due process (69%) conditions than in the unreliable condition (7%), $\chi^2(2, n = 46) = 16.8, p < .001$. Among lows, however, the differences were not significant, as the frequencies were 40%, 50%, and 27% in the admissible, due process, and unreliable conditions, respectively, $\chi^2(2, n = 44) = 1.7, ns$. Across trial conditions, the difference between high- and low-NC participants' reported use of the wiretap evidence (50% vs. 38%) was not significant, $\chi^2(1, n = 90) = 1.2, ns$.

Non-control-group participants were also more pointedly asked whether they used the wiretap in reaching their verdict. Overall, more said they used the wiretap in the admissible condition (93%) than in the inadmissible-due process (50%) and unreliable (37%) conditions, $\chi^2(2, n = 90) = 8.1, p < .02$. High-NC jurors—100% of whom said they used the admissible wiretap—were more likely to admit having used the forbidden wiretap when it was excluded for reasons of due process than for a lack of

reliability (63% vs. 20%, respectively), $\chi^2(2, n = 46) = 20.2, p < .001$. This latter result indicates that among high-NC participants, noncompliance with the instruction to disregard was overt and motivated. For low-NC jurors—87% of whom reported using the admissible wiretap—the difference between self-reported wiretap use in the inadmissible–due process and unreliable conditions (53% vs. 36%) was not significant.

Online Evidence Ratings

During the trial presentation, participants rated each paragraph on a scale of 1 to 9 for the extent to which it led them to view the defendant as innocent or guilty. As expected, some items of evidence were rated as more incriminating than others, $F(25, 2425) = 62.3, p < .00001$. To test the hypothesis that disclosure of an inadmissible wiretap would influence the way in which jurors perceived subsequent evidence, a $2 \times 3 \times 2$ mixed-model ANOVA with two between-subject variables (trial condition and NC) and one within-subject variable (average pre- vs. postwiretap item ratings) was performed on both the early- and late-condition participants (see Note 1).

Unlike in Kassin and Sommers (1997), there was no indication that the disclosure of the wiretap influenced mock jurors' ratings of subsequent evidence. Among participants in the early-disclosure conditions, the ANOVA revealed no main effect for trial condition, $F(3, 52) < 1$, or NC, $F(1, 52) = 1.61, ns$, and no interaction, $F(3, 52) = 1.53, ns$. The repeated-measures analysis failed to reveal a significant main effect, $F(1, 52) < 1$, or interactions with trial condition, $F(3, 52) = 2.23, ns$, or NC, $F(1, 52) < 1$. In addition, no three-way interaction was demonstrated, $F(3, 52) = 1.49, ns$. The effects for the late-condition participants were similarly nonsignificant.

Response times to each item were measured in milliseconds during the computer presentation of the trial summary. A 2×4 ANOVA revealed no significant between-group differences in the total amount of time participants took to read and respond to the trial evidence, either as a function of trial condition, $F(3, 97) = 1.39, ns$, or NC, $F(1, 97) < 1$; the interaction term also failed to approach significance, $F(3, 97) < 1$. These results indicate that NC was not associated with differences in reading or decision time during the trial summary presentation. Thus, NC did not affect decision making merely by influencing the amount of time participants spent on the experimental task.

DISCUSSION

The present results supported the first hypothesis that mock jurors would demonstrate selective compliance with an instruction to disregard. Individual differences in NC aside, the present data replicated the findings of Kassin and Sommers (1997), as conviction rates, scaled

verdict-confidence scores, and probability-of-commission estimates indicated that jurors were somewhat influenced by a wiretap ruled inadmissible because it violated due process but not by one ruled inadmissible because it was unreliable. This hypothesis was also supported by participants' self-reports of how influential the wiretap was in their decision making. Indeed, most mock jurors admitted that the inadmissible yet reliable wiretap influenced their verdict, but few reported being influenced by the unreliable wiretap.

These findings further bolster the hypothesis articulated by Kassin and Sommers (1997) that the motivation to reach a just verdict is a crucial factor in determining whether jurors comply with instructions to disregard. Other researchers have suggested that alternative motivations, such as a concern with procedural justice or a desire to uphold the law, also influence juror compliance (Fleming et al., 1999). The present results do not rule out the possibility that other such motivations exist. To the contrary, in the present study and in Kassin and Sommers's study, conviction rates and guilt judgments in the inadmissible–due process condition were not as high as in the admissible condition, although this difference was not statistically significant. This result implies that concerns about procedure did lead to some discounting of a reliable inadmissible wiretap, even if this effect was smaller than the discounting in the inadmissible-unreliable condition. At the same time, these findings do seem to suggest that reaching a just verdict is often the strongest motivator of jurors, a conclusion that is consistent with the vast literature documenting examples of jury nullification. Moreover, the strength of jurors' resolve to render just verdicts is also demonstrated by the finding that it seems to require unusual trial circumstances to overwhelm this motivation with procedural concerns (e.g., an egregious violation of due process in which police officers break into every house within a 10-mile area in the middle of the night without a search warrant, as in Fleming et al., 1999). Additional studies that identify the circumstances under which each of these different motivations drive juror performance are clearly needed.

The present study failed to replicate the findings of Kassin and Sommers (1997) in one regard. Among participants in the two inadmissible conditions, there were no differences in online ratings of the postwiretap items of evidence. One possible reason for this absence of integrative processing is that the instructions given to participants in the present study placed a greater emphasis on the requirement that they consider each item of evidence separately. Another possibility is that limiting online ratings in the present study to a scale of 1 to 9—in contrast to the 0 to 100 scale used in Kassin and Sommers—reduced variability and masked the influence of the wiretap on subsequent items of evidence.

Further research will be necessary to retest the online influence that inadmissible evidence can have on juror decision processes. For now, however, it is safe to conclude that in actual trials, when evidence is not disclosed in discrete, isolated chunks as it is in the laboratory, it is more difficult for jurors to interpret information independently from the evidence that has preceded it.

We were also able to shed light on the motivation underlying selective compliance by considering individual differences in NC. The results from the scaled verdict scores and probability-of-commission estimates supported the hypothesis that evidence reliability and NC would interact to influence juror performance. The conviction rate data also supported the prediction that selective compliance would be more pronounced among high-NC jurors, but the dichotomous nature of the variable prevented the difference between the two inadmissible conditions from reaching significance. Motivational differences between high- and low-NC mock jurors provide a plausible explanation for these results, and additional data collection supported the conclusion that high-NC jurors have a stronger just verdict motivation than do low-NC jurors: Analysis of 60 participants' responses to a brief questionnaire indicated that NC scores were positively correlated with a self-reported just verdict motivation among mock jurors, $r(n=60) = .29$, $p < .03$.⁵

Evidence for processing differences between high- and low-NC jurors was less clear-cut, however. Consistent with the observations of Cacioppo et al. (1996), high-NC mock jurors seemed better able to process information methodically and base judgments on rational and empirical information. Highs made logical decisions in their pursuit of a just verdict: When the wiretap was admissible and apparently reliable, the vast majority (80%) voted guilty. They did not fully disregard a reliable wiretap that was inadmissible because of a due process violation, revealing their motivation for accuracy even when such a verdict was inconsistent with the letter of the law. High-NC jurors also reported that the reliable inadmissible wiretap was as influential as an admissible one and was far more influential than one that was unreliable.

In contrast, low-NC participants did not exhibit a significant increase in convictions when an admissible wiretap was disclosed, even though this item of evidence was intended to be highly incriminating. This is consistent with the conclusion of Cacioppo et al. (1996) that low-NC individuals give more weight to peripheral considerations and heuristics than to rational arguments and probative facts in decision making. Our written trial summary was devoid of the peripheral cues often present at trial (e.g., the physical appearance of individuals, the confidence and credentials of witnesses, nonverbal reactions of others, etc.), and this could have contributed to

the reduced sensitivity of low-NC jurors to the differences between versions of the case. But overall, there was no direct evidence of processing differences by NC. Lows spent just as much time reading and deliberating on the trial as did highs, and they did not differ in their explicit memory for the facts of the case or the experimental manipulations. In fact, the only observable differences between highs and lows were found in their final decisions about the case and their self-reported use of the inadmissible evidence; unlike high-NC jurors, lows did not differentiate between the two inadmissible wiretaps.

The absence of direct evidence of processing differences between high- and low-NC individuals in this study highlights the need for additional investigations of NC and juror performance. Perhaps a trial presentation that also included peripheral cues of persuasion would magnify processing differences between high- and low-NC jurors. Manipulation of the strength and quality of trial evidence also could allow for the same sort of comparisons that are often made between highs and lows in the broader persuasion and attitudes literature. Of course, it is important as well to point out the correlational nature of the present study when it comes to NC. Third variables such as intelligence, moral reasoning, or faith in the legal system cannot be ruled out as influences on juror decision making using the present design. Earlier in this article, we suggested that jurors' sense of involvement in the outcome of the trial might influence bias correction. Therefore, manipulating personal relevance in a case might be one way to clarify these issues surrounding causality. In sum, the results reported here support our predictions in this initial investigation of NC and juror performance, but several issues remain in need of empirical attention.

The present results also provided some support for our third hypothesis that high-NC mock jurors would overcorrect for the perceived bias of an unreliable wiretap. Among high-NC participants, conviction rates and verdict-confidence scores in the inadmissible-unreliable condition were somewhat, although not significantly, lower than those in the no-wiretap control group. Significant support for a "Mark Fuhrman effect" among highs was provided by their probability-of-commission estimates in the inadmissible-unreliable condition, which were lower than those in the baseline control group. This significant overcorrection provides additional evidence that high-NC jurors in particular are motivated to reach a just verdict. When these participants first read about the wiretap, they were influenced in the direction of a guilty verdict (as seen in the high conviction rate in the admissible group). But those in the inadmissible-unreliable condition were informed immediately thereafter that the wiretap was barely audible and was not reliable enough to be admitted into evidence. As a result,

they actively engaged in bias correction. As the Flexible Correction Model would predict, given the obvious biasing potential of this information, high-NC mock jurors sometimes overcorrected.

It is not terribly surprising that overcorrection among highs was significant for probability-of-commission estimates but not for verdicts (although the latter results were in the predicted direction). Wegener and Petty (1995) found that overcorrection is most likely when the biasing information is only mildly influential, not when the actual impact of the information is great. In the present study, the wiretap evidence had a much stronger impact on verdicts and verdict scores than it did on probability estimates, rendering the latter measure more susceptible to overcorrection. Furthermore, as in past mock jury studies, the present participants quantified the standard of proof "beyond a reasonable doubt" to mean that there should be at least an 89% probability that the defendant committed the crime to support a conviction. Among high-NC participants, the average probability-of-commission estimate was only 76% in the control group and 53% in the inadmissible-unreliable group. Thus, most mock jurors were leaning toward acquittal even before the inadmissible wiretap was introduced. For overcorrection to have influenced verdicts, a high conviction rate would have been required in the baseline control group. In other words, overcorrection can affect verdict rates only when the rest of the evidence is highly incriminating.

Taken as a whole, the present findings demonstrate that NC, a variable rarely considered by psychologists who study juror decision making, can play a key role in the influence of inadmissible evidence at trial. Preliminary indications suggest that high-NC jurors are strongly motivated by the desire to reach a just verdict and that their bias correction in the face of inadmissible evidence is guided by this motivation. Low-NC jurors, on the other hand, do not appear to be as sensitive to variations in the probative value of evidence, and they seem less willing and/or able to exert the cognitive resources necessary to correct for the biasing agent of inadmissible evidence. Furthermore, the present study is, to our knowledge, the first demonstration of overcorrection in the jury context. Given the premium placed by the legal system on avoiding the biasing influence of extralegal information, it should be of interest to discover that bias in the opposite direction can occur as well under certain situational and personal circumstances.

It is also worth considering how serving on an actual jury, with the fate of a real defendant at stake, would alter the decision-making strategies and processing styles of jurors.⁶ Schul and Manzury (1990) have argued that a court setting serves to "activate a schema that leads people to correct for potential biases on the relevant guilt

judgments" (p. 337), suggesting that actual trials are so personally involving for jurors that even low-NC individuals would expend the effort necessary to reach a just verdict. Kassin, Reddy, and Tulloch (1990) were even more explicit in their view that trials are highly engaging events that elicit close consideration and attention from all jurors, regardless of their normal motivational and cognitive tendencies. If actual trials do indeed compel low-NC jurors to strive for a just verdict and process information in a high-NC manner, then selective compliance and overcorrection may be even more pervasive in actual trials than the present findings suggest.

NOTES

1. The timing of the wiretap evidence and ruling within the case (early vs. late in the proceedings; i.e., in paragraphs 9 and 10 out of 26 vs. 22 and 23 out of 26) was also varied to test whether this disclosure would influence jurors' interpretations of subsequently presented evidence (thus rendering the wiretap that was disclosed early on in the trial more influential). There were no significant main effects or interactions involving this factor of timing, however, so the data were collapsed across this variable for all analyses unless otherwise stated.

2. The overall conviction rates were 43% and 37% among high- and low-need for cognition (NC) mock jurors, respectively—a difference that was not statistically significant, $\chi^2(1, n = 105) < 1$.

3. In Kassin and Sommers (1997), selective compliance was demonstrated when jurors' perceptions of guilt in the admissible condition were significantly different from perceptions in the no-wiretap control and inadmissible-unreliable conditions but not significantly different from perceptions in the inadmissible—due process condition. It is this pattern of results, whereby jurors completely discount an unreliable but not a reliable wiretap, that guided the present analyses. Conducting the analyses as separate comparisons of the admissible condition with both the inadmissible—due process and inadmissible-unreliable conditions yielded results comparable with those presently reported.

4. The contrast codes for the eight cells (four trial conditions \times two NC levels) in this comparison were (0 0 1 -1 0 0 -1 1).

5. Just verdict motivation was measured by participants' responses to the following three statements: (a) "It is more important for a criminal jury to arrive at what it perceives to be an accurate verdict than it is for the jury to follow the letter of the law exactly," (b) "Even if evidence in a criminal trial seems reliable and important, a juror should disregard it if instructed to do so by the judge," and (c) "A jury's main role is to apply the law to a set of facts, not to determine the 'truth.'" Responses were made on a scale ranging from 1 (*strongly disagree*) to 9 (*strongly agree*), and responses to the second and third item were reverse scored and added to responses to the first item to create a positive index of just verdict motivation.

6. In addition to judicial admonishments to disregard, the legal system has two other safeguards intended to prevent inadmissible evidence from influencing jurors, neither of which was included in the present study: voir dire and deliberation. However, Dexter, Cutler, and Moran (1992) found that a voir dire with explicit warnings against inadmissible information failed to prevent jurors from rendering verdicts biased by pretrial publicity. Carretta and Moreland (1983) demonstrated that inadmissible evidence influenced jurors' pre- as well as postdeliberation decisions (cf. Kerwin & Shaffer, 1994).

REFERENCES

- Anderson, C. A., Lepper, M. R., & Ross, L. (1980). Perseverance of social theories: The role of explanation in the persistence of discredited information. *Journal of Personality and Social Psychology, 39*, 1037-1049.
- Broeder, D. (1959). The University of Chicago Jury Project. *Nebraska Law Review, 38*, 744-760.

- Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42, 116-131.
- Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W.B.G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin*, 119, 197-253.
- Cacioppo, J. T., Petty, R. E., & Kao, C. F. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment*, 48, 306-307.
- Carretta, T. R., & Moreland, R. L. (1983). The direct and indirect effects of inadmissible evidence. *Journal of Applied Social Psychology*, 13, 291-309.
- DeSteno, D., Petty, R. E., Wegener, D. T., & Rucker, D. D. (2000). Beyond valence in the perception of likelihood: The role of emotion specificity. *Journal of Personality and Social Psychology*, 78, 397-416.
- Dexter, H. R., Cutler, B. L., & Moran, G. (1992). A test of voir dire as a remedy for the prejudicial effects of pretrial publicity. *Journal of Applied Social Psychology*, 22, 819-832.
- Dunne, D. (1995). The "N" word: Not guilty. *Vanity Fair*, pp. 134-148.
- Edwards, K., & Bryan, T. S. (1997). Judgmental biases produced by instructions to disregard: The (paradoxical) case of emotional information. *Personality and Social Psychology Bulletin*, 23, 849-864.
- Elliot, R., Farrington, B., & Manheimer, H. (1988). Eyewitnesses credible and discreditable. *Journal of Applied Social Psychology*, 18, 1411-1422.
- Fein, S., McCloskey, A. L., & Tomlinson, T. M. (1997). Can the jury disregard that information? The use of suspicion to reduce the prejudicial effects of pretrial publicity and inadmissible testimony. *Personality and Social Psychology Bulletin*, 23, 1215-1226.
- Fein, S., Morgan, S. J., Norton, M. I., & Sommers, S. R. (1997). Hype and suspicion: The effects of pretrial publicity, race, and suspicion on jurors' verdicts. *Journal of Social Issues*, 53, 487-502.
- Fleming, M. A., Wegener, D. T., & Petty, R. E. (1999). Procedural and legal motivations to correct for perceived judicial biases. *Journal of Experimental Social Psychology*, 35, 186-203.
- Greene, E., & Dodge, M. (1995). The influence of prior record evidence on juror decision-making. *Law and Human Behavior*, 19, 67-78.
- Haney, C. (1984). On the selection of capital juries: The biasing effects of the death qualification process. *Law and Human Behavior*, 8, 121-132.
- Horowitz, I. A., & Willging, T. E. (1991). Changing views of jury power: The nullification debate, 1787-1988. *Law and Human Behavior*, 15, 165-182.
- Johnson, H. M. (1994). Processes of successful intentional forgetting. *Psychological Bulletin*, 116, 274-292.
- Johnson, H. M., & Seifert, C. M. (1994). Sources of the continued influence effect: When misinformation in memory affects later inferences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 1420-1436.
- Kagehiro, D. K. (1990). Defining the standards of proof in jury instructions. *Psychological Science*, 1, 194-200.
- Kassin, S. M., Reddy, M. E., & Tulloch, W. F. (1990). Juror interpretations of ambiguous evidence: The need for cognition, presentation order, and persuasion. *Law and Human Behavior*, 14, 43-52.
- Kassin, S. M., & Sommers, S. R. (1997). Inadmissible testimony, instructions to disregard, and the jury: Substantive versus procedural considerations. *Personality and Social Psychology Bulletin*, 23, 1046-1054.
- Kassin, S. M., & Studebaker, C. A. (1998). Instructions to disregard and the jury: Curative and paradoxical effects. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches*. Hillsdale, NJ: Lawrence Erlbaum.
- Kassin, S. M., Williams, L. N., & Saunders, C. L. (1990). Dirty tricks of the cross-examination: The influence of conjunctural evidence on the jury. *Law and Human Behavior*, 14, 373-384.
- Kerwin, J., & Shaffer, D. R. (1994). Mock jurors versus mock juries: The role of deliberations in reactions to inadmissible testimony. *Personality and Social Psychology Bulletin*, 20, 153-162.
- Kramer, G. P., Kerr, N. L., & Carroll, J. S. (1990). Pretrial publicity, judicial remedies, and jury bias. *Law and Human Behavior*, 14, 409-438.
- Martin, L. L., Seta, J. J., & Crelia, R. A. (1990). Assimilation and contrast as a function of people's willingness and ability to expend effort in forming an impression. *Journal of Personality and Social Psychology*, 59, 27-37.
- Pickel, K. L. (1995). Inducing jurors to disregard inadmissible evidence: A legal explanation does not help. *Law and Human Behavior*, 19, 407-424.
- Schul, Y., & Burnstein, E. (1985). When discounting fails: Conditions under which individuals use discredited information in making a judgment. *Journal of Personality and Social Psychology*, 49, 894-903.
- Schul, Y., & Manzur, F. (1990). The effects of type of encoding and strength of discounting appeal on the success of ignoring invalid testimony. *European Journal of Social Psychology*, 20, 337-349.
- Schuller, R. A. (1995). Expert evidence and hearsay: The influence of "secondhand" information on jurors' decisions. *Law and Human Behavior*, 19, 345-362.
- Sommers, S. R., & Ellsworth, P. C. (2000). Race in the courtroom: Perceptions of guilt and dispositional attributions. *Personality and Social Psychology Bulletin*, 26, 1367-1379.
- Studebaker, C. A., & Penrod, S. D. (1997). Pretrial publicity: The media, the law, and common sense. *Psychology, Public Policy, and Law*, 3, 428-460.
- Sue, S., Smith, S. E., & Caldwell, C. (1973). Effects of inadmissible evidence on the decisions of simulated jurors: A moral dilemma. *Journal of Applied Social Psychology*, 3, 345-353.
- Wegener, D. T., Kerr, N. L., Fleming, M. A., & Petty, R. E. (2000). Flexible corrections of juror judgments: Implications for jury instructions. *Psychology, Public Policy, and Law*, 6, 629-654.
- Wegener, D. T., & Petty, R. E. (1995). Flexible correction processes in social judgment: The role of naive theories in corrections for perceived bias. *Journal of Personality and Social Psychology*, 68, 36-51.
- Wegener, D. T., & Petty, R. E. (1997). The flexible correction model: The role of naive theories of bias in bias correction. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 29). New York: Academic Press.
- Weinberg, H. I., & Baron, R. S. (1982). The discreditable eyewitness. *Personality and Social Psychology Bulletin*, 8, 60-67.
- Wissler, R. L., & Saks, M. J. (1985). On the inefficacy of limiting instructions: When jurors use prior conviction evidence to decide on guilt. *Law and Human Behavior*, 9, 37-48.
- Wolf, S., & Montgomery, D. A. (1977). Effects of inadmissible evidence and judicial admonishment to disregard on the judgments of mock jurors. *Journal of Applied Social Psychology*, 7, 205-219.

Received July 17, 2000

Revision accepted September 21, 2000