JURORS’ RESPONSES TO UNUSUAL INADMISSIBLE EVIDENCE

KERRI L. PICKEL
TANYA J. KARAM
TODD C. WARNER

Ball State University

Two experiments investigated jurors’ ability to disregard unusual inadmissible evidence. Participants listened to an audio recording of a theft trial. Those in four experimental conditions heard critical testimony favoring the prosecution, which was ruled either admissible or inadmissible and which contained either neutral details or details that were unusual in terms of semantic content (Experiment 1) or form (Experiment 2). Control jurors received no critical evidence. Exposure to unusual rather than neutral evidence led jurors to see the defendant as more guilty but only if that evidence was inadmissible instead of admissible. Additionally, jurors remembered unusual evidence better than neutral evidence. The results are consistent with Wegner’s ironic-process theory and suggest that attempts at thought suppression are less successful if the forbidden information is especially memorable.

Keywords: juror decision making; psychology and law; inadmissible evidence

During a courtroom trial, witnesses sometimes inadvertently introduce inadmissible evidence while on the stand. In such cases, the judge may prefer to instruct the jurors to disregard this evidence to avoid the more expensive and less time-efficient option of granting a mistrial (Carretta & Moreland, 1983; Paglia & Schuller, 1998; Thompson, Fong, & Rosenhan, 1981). However, research indicates that jurors frequently do not disregard inadmissible evidence (Lieberman & Arndt, 2000; Lieberman, Arndt, & Vess, in press; Steblay, Hosch, Culhane, & McWethy, 2006). To explain why, researchers have considered various theoretical perspectives. However, according to Lieberman and colleagues (Lieberman & Arndt, 2000; Lieberman et al., in press), only two have shown much success in accounting for the body of data in this area: reactance theory (J. W. Brehm, 1966; S. S. Brehm & Brehm, 1981) and Wegner’s (1994) theory of ironic mental processes.

According to reactance theory (J. W. Brehm, 1966; S. S. Brehm & Brehm, 1981), individuals strongly prefer to maintain their autonomy, so when they believe their freedom is endangered, they may try to reestablish it by performing whatever behavior they think is being restricted. In the case of inadmissible evidence, jurors may sometimes perceive the judge’s order to disregard it as unfairly threatening their “freedom to consider all of the available evidence” (Wolf & Montgomery, 1977, p. 207) and interfering with their right to decide whether specific testimony is relevant (Lieberman & Arndt, 2000; Lieberman et al., in press).
in press). The jurors may therefore retaliate by weighting the inadmissible evidence more heavily than they would have if no instructions had been given. Moreover, the more strongly the judge admonishes jurors to ignore the evidence, the more his or her instructions might “backfire,” as Wolf and Montgomery (1977) found.

Reactance theory assumes that jurors may deliberately choose not to follow a judge’s orders, but in some cases, jurors may accept the judge’s declaration that certain evidence is legally off-limits and may attempt to abide by his or her instructions (Lieberman et al., in press). Wegner’s (1994) theory of ironic mental processes can explain why even well-intentioned jurors might fail to disregard inadmissible evidence. According to this theory, people can try to avoid thinking about something (such as evidence that has been ruled inadmissible) by using both (a) a conscious “operating process” that redirects the mind away from the forbidden information and toward more acceptable thoughts and (b) an unconscious, automatic “monitoring process” that searches for instances of inappropriate thoughts, bringing them into consciousness if they are detected so that the individual can deal with the failure of mental control. But, as Wegner explains, “processes that undermine the intentional control of mental states are inherent in the very exercise of such control” (p. 34). Specifically, the processes that are supposed to suppress inadmissible information may ironically increase its accessibility, triggering conscious thoughts about it, especially under conditions of high cognitive load (as is the case when jurors must comprehend and remember large amounts of evidence and legal instructions) or when jurors try exceptionally hard to ignore the inadmissible evidence (as they might when strongly admonished by the judge; Lieberman & Arndt, 2000; Lieberman et al., in press). Additionally, the off-limits information may be semantically associated with admissible testimony so that attempts to focus on acceptable evidence may prompt thoughts of inadmissible items (Lieberman & Arndt, 2000; Lieberman et al., in press).

From an ironic-process perspective, inadmissible evidence may be particularly difficult to ignore if it is very memorable, as would be the case with emotional information, for example. Edwards and Bryan (1997) asked mock jurors to read transcripts of a robbery–murder trial. The critical evidence concerned the defendant’s prior criminal record, and it was admissible or inadmissible in different conditions (and was absent in a control condition). The authors varied the emotionality of the evidence so that the critical witness described the defendant’s record either using neutral language (e.g., he was accused of “assault with a deadly weapon”) or in more lurid terms (e.g., he was accused of “hacking up a woman”). Pilot tests verified that the emotional version produced a more intense affective response in readers than did the neutral version but that the two were equivalent with respect to perceived probative value. Edwards and Bryan found that the emotional inadmissible evidence affected jurors’ judgments to a greater extent than did the neutral inadmissible evidence. Furthermore, instructions to disregard the emotional (but not neutral) evidence backfired, causing it to have even more influence than if it was ruled admissible.

The authors argued that a reactance explanation for these findings would require the assumption that jurors were more strongly motivated to restore their freedom when considering emotional evidence than when considering neutral evidence, which seems unlikely given that pilot testing indicated that the two conditions were equal in probative value. A better interpretation, according to Edwards and Bryan, is one based on ironic-process theory: Because emotional information, as compared to neutral information, attracts and maintains attention more powerfully, is more arousing, and is encoded more elaboratively,
attempts to suppress thoughts of the emotional evidence failed. Thus, jurors found it easier to disregard neutral as opposed to emotional inadmissible evidence.

As with testimony containing emotional language, evidence that includes unusual details may also be exceptionally difficult to ignore if ruled inadmissible. Previous research has demonstrated that physically unusual items, such as orthographically distinctive words (e.g., llama) are well remembered (Hunt & Elliot, 1980), as are stimuli that are conceptually unusual, such as an animal name embedded in a list of countries (Schmidt, 1985), surprising or bizarre images and sentences (Hirshman, Whelley, & Palij, 1989; McDaniel & Einstein, 1986), unusual faces (Brigham, 1990), atypical events in schematic stories (Smith & Graesser, 1981), and unexpected objects in crime scenes observed by eyewitnesses (Pickel, Ross, & Truelove, 2006). Distinctive items may receive an increased amount of attentional resources, and the resulting memory trace may be especially detailed (Schmidt, 1985, 1991). Moreover, distinctive traces share few features with other representations, which makes them less vulnerable to interference and thus more readily retrievable.

Like the examples of distinctiveness listed above, unusual details within a witness’s testimony would probably be more memorable than neutral information. From an ironic-process standpoint, these details should also be difficult to suppress (a) because their high level of accessibility should make them easily detected by the monitoring process and (b) because they would likely be related to admissible evidence items that jurors are trying to think about because of the work of the operating process (which steers the mind toward acceptable thoughts). The purpose of the present study was to investigate the possibility that unusual evidence items are more difficult to ignore than neutral items if ruled inadmissible.

In two experiments, mock jurors listened to a theft trial. In different conditions, the critical evidence included either neutral or unusual details (operationalized slightly differently in each experiment); pilot tests verified that participants perceived the two versions similarly in terms of their probative value, reliability, and several other variables. Additionally, the critical witness discussed events he allegedly either observed directly (so that the evidence was ruled admissible) or heard from another person (so that the evidence was hearsay and was inadmissible). We also included a control condition in which the critical evidence was omitted.

We predicted, on the basis of ironic-process theory, that unusual hearsay would affect guilt judgments more strongly than neutral inadmissible evidence. Moreover, we hypothesized that jurors exposed to unusual inadmissible evidence would be unable to ignore that information and would therefore make judgments that differed from those of control jurors. Finally, we also expected the critical evidence to be better remembered if it was unusual rather than neutral because of the greater accessibility of unusual information.

Previous research suggests that jurors can frequently ignore hearsay (Miene, Park, & Borgida, 1992; Paglia & Schuller, 1998; Pickel, 1995; Rakos & Landsman, 1992; although see Fein, McCloskey, & Tomlinson, 1997; Lee, Krauss, & Lieberman, 2005), whereas they are less successful with other types of inadmissible evidence, such as illegally obtained incriminating information (e.g., Carretta & Moreland, 1983; Sue, Smith, & Caldwell, 1973). In fact, jurors sometimes disregard hearsay even in the absence of instructions to do so (Paglia & Schuller, 1998), possibly because they think it would be unfair to consider it (Pickel, 1995). Therefore, by using hearsay within our study, we created a stronger test of our hypotheses than we would have with another type of inadmissible evidence.
EXPERIMENT 1

Information can be considered unusual for different reasons. In Experiment 1, the critical testimony was “unusual” if it included atypical semantic content. As noted above, past research has shown that people generally remember stimuli that are surprising or unexpected from a conceptual or schematic standpoint more accurately than items that are consistent with their expectations.

METHOD

Stimulus. An audio recording of a theft trial (approximately 24 min in length) was created by modifying the stimulus used by Pickel (1995). In this case, a fired employee is accused of stealing cash from his former supervisor’s office. The critical evidence involves an allegation that after the theft occurred, the defendant made an incriminating statement to two friends while sitting in a bar, telling them that he had suddenly acquired a large sum of money. The critical witness testifies either that he directly observed the defendant making this statement (admissible condition) or that another person told him what the defendant said (inadmissible condition). In the latter situation, the testimony is hearsay according to the U.S. Federal Rules of Evidence, because the witness has related an out-of-court statement made by a person who is not testifying with the intention of asserting that this person’s statement is true (Mauet, 1996). Although there are exceptions to the hearsay rule, in this case, it would be inadmissible.

In addition to the admissibility, we also varied the unusualness of the content of the evidence. After the defendant’s friends have expressed surprise and pleasure that he has bought drinks for everyone at the table, he describes his plans for a spending spree with his newfound wealth. Specifically, in the neutral condition, the remarks attributed to the defendant include these details:

Remember I told you I wanted to get my truck painted? Well, next week I’m getting a paint job—it’ll be blue. And I’m buying a new jacket. I’m gonna tell people to call me “Sir.”

The unusual version of the defendant’s statement is the following:

Remember I told you I wanted to get my truck painted? Well, next week I’m getting a paint job—it’ll be bright green with yellow polka dots. And I’m buying a black cape. I’m gonna tell people to call me “Batman.”

Immediately after the witness presents the critical evidence, the defense attorney objects. The judge then issues a ruling.

After the trial concludes with the attorneys’ closing arguments, the judge issues instructions to the jurors that guided them with respect to the evaluation of witnesses, the burden of proof and reasonable doubt, and the legal definition of theft. We used pattern jury instructions that are commonly employed in the state in which the participants were tested.

Design. The study’s design was 2 (admissibility) × 2 (unusualness) factorial, with an additional (control) condition in which the critical evidence was absent.

Pilot study. We conducted a pilot study to assure that the trial script successfully manipulated unusualness without changing other variables that could conceivably affect guilt
judgments. Sixty pilot participants read a brief description of the trial along with one of the two versions of the critical evidence (either neutral, \(n = 30\), or unusual, \(n = 30\)). Then they used a 10-point scale to rate the unusualness of the evidence, its reliability (i.e., the extent to which jurors believe the defendant actually made the statement), its probative value (i.e., the degree to which it incriminates the defendant), its salience (i.e., the degree to which it attracts attention), and its ability to produce an emotional response, which (following Edwards & Bryan, 1997) was measured with three items that asked about the degree to which participants were emotionally affected by the evidence, the intensity of emotional feelings elicited, and the amount of emotional content. In addition, the participants rated the defendant’s “craziness” (i.e., the degree to which he seemed to have psychological problems that interfered with his ability to separate reality from fantasy), the extent to which participants thought he intended to be humorous when he listed the items he planned to buy, and the extent to which participants actually found him to be humorous.

The results verified that the two versions differed significantly in terms of unusualness, \(t(58) = 12.75, p < .001, d = 1.43\). Also, compared to participants in the neutral condition, those in the unusual condition thought it more likely that the defendant intended to be humorous, \(t(58) = 2.86, p = .006, d = .74\). However, when they rated actual humorousness, the unusual statement received only a slightly higher mean, \(t(58) = 1.87, p = .07, d = .06\). Thus, it appears that participants interpreted the defendant’s unusual statement as an unsuccessful attempt to be funny about how he planned to spend his newly acquired funds. The two versions were similar with respect to all other tested variables (\(p s \geq .30\)).

**PARTICIPANTS AND PROCEDURE**

The participants were 166 introductory psychology students at a midwestern university. Most (93%) were White, and 54% were female. They ranged in age from 18 to 42 years (\(M = 19.05, SD = 2.17\)).

The participants were tested in groups of up to 10 individuals. Each group was assigned randomly to listen to one of the five versions of the trial recording. Afterward, they completed a questionnaire asking them to choose a verdict (given the choices of guilty of theft or not guilty) and to use a 10-point scale to rate their confidence in their choice. In addition, all jurors except those in the control condition tried to remember the critical evidence (i.e., they were instructed to write down the critical witness’s statement as specifically as possible), indicated whether they intended to use or ignore the evidence while deciding on a verdict, and rated on a 10-point scale the degree to which they thought they actually used the evidence (regardless of intention).

**RESULTS**

Evaluations of guilt. For each juror, we calculated a “guilt score” by assigning a value of +1 to a guilty verdict and −1 to a not-guilty verdict and then multiplying this value by the juror’s confidence rating (thus, a score of +10 indicates maximum confidence in a guilty verdict). This composite measure is often used in inadmissible evidence research because it is more sensitive than a dichotomous verdict (e.g., Costabile & Klein, 2005; Edwards & Bryan, 1997; Kassin & Sommers, 1997; Lee et al., 2005).

As an initial step in evaluating the results, \(2 \times 2\) analyses of variance were conducted using only data from the experimental conditions so as to reveal any main effects or interactions.
(see Table 1). There was a significant interaction, $F(1, 127) = 4.39$, $p < .04$, $\eta^2 = .03$, indicating that the impact of the inadmissibility ruling was different for neutral versus unusual evidence. Specifically, the inadmissibility ruling reduced guilt scores for neutral evidence but increased them for unusual evidence. There was also a main effect of unusualness, $F(1, 127) = 11.95$, $p = .001$, $\eta^2 = .09$, such that guilt scores were higher in the unusual than in the neutral condition. There was no main effect of admissibility ($p = .74$).

Next, simple analyses of variance that included the control condition were performed. The first revealed that guilt scores differed across conditions, $F(4, 161) = 6.06$, $p < .001$, $\eta^2 = .13$. Specific contrasts were then used to test our predictions. As hypothesized using ironic-process theory, unusual evidence affected guilt scores more strongly than neutral evidence but only if it was inadmissible, $t(161) = 4.04$, $p < .001$, $d = 1.05$, instead of admissible ($p = .31$, $d = .22$). Another contrast investigated whether a backfire effect occurred so that unusual evidence became more influential when ruled inadmissible rather than admissible. No evidence of a backfire effect was obtained, $t(161) = 1.28$, $p = .20$, $d = .28$. Finally, contrasts compared the experimental conditions to the control condition to determine whether jurors used the critical evidence as opposed to ignoring it. Jurors who heard the evidence ruled admissible had at least somewhat higher guilt scores than those who were not exposed to it: neutral, $t(161) = 1.69$, $p = .09$, $d = .43$; unusual, $t(161) = 2.70$, $p = .008$, $d = .66$. However, of the jurors who heard the evidence ruled inadmissible, only those in the unusual condition, $t(161) = 3.95$, $p < .001$, $d = 1.00$, and not those in the neutral condition ($p = .88$, $d = .05$), had guilt scores that differed from the controls.

Use of the critical evidence. Jurors who heard the critical evidence were asked whether they intended to use it in rendering a verdict. To examine this variable and the remaining ones measured in Experiment 1, $2 \times 2$ analyses of variance were used. A higher proportion of jurors in the admissible conditions (.91) than in the inadmissible conditions (.13) said that they intended to use the evidence, $F(1, 127) = 207.53$, $p < .001$, $\eta^2 = .62$. There was no main effect of unusualness and no interaction ($ps > .24$).

When asked to estimate the extent to which they actually used the evidence, jurors in the admissible conditions (overall, $M = 5.54$, $SD = 2.57$; unusual, $M = 5.48$, $SD = 2.54$; neutral, $M = 5.59$, $SD = 2.64$) gave higher ratings than did those who heard inadmissible evidence (overall, $M = 3.02$, $SD = 2.32$; unusual, $M = 3.65$, $SD = 2.50$; neutral, $M = 2.42$, $SD = 2.00$), $F(1, 127) = 34.60$, $p < .001$, $\eta^2 = .21$. We obtained no main effect of unusualness and no interaction ($ps > .12$).

---

### Table 1: Experiment 1: Mean Guilt Scores, Proportion of Correct Details Remembered, and Mean Number of Incorrect Details Remembered by Condition

<table>
<thead>
<tr>
<th>Condition</th>
<th>Guilt Score</th>
<th>Proportion of Correct Details</th>
<th>Number of Incorrect Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inadmissible</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unusual</td>
<td>1.29 (7.11)</td>
<td>.36 (.15)</td>
<td>.48 (.68)</td>
</tr>
<tr>
<td>Neutral</td>
<td>-5.09 (4.85)</td>
<td>.23 (.14)</td>
<td>.55 (.97)</td>
</tr>
<tr>
<td>Admissible</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unusual</td>
<td>-0.73 (7.28)</td>
<td>.38 (.13)</td>
<td>.33 (.60)</td>
</tr>
<tr>
<td>Neutral</td>
<td>-2.29 (6.79)</td>
<td>.26 (.11)</td>
<td>.59 (.96)</td>
</tr>
<tr>
<td>Control</td>
<td>-4.86 (5.21)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* Values in parentheses are standard deviations. Means in the same column that do not share the same subscript differ significantly ($p \leq .05$).
Memory for the critical evidence. Jurors who heard the critical evidence were asked to recall it. To create an answer key for these reports, two judges studied the transcript of each version of the critical evidence and identified the specific details contained in each. For example, in the unusual version, the defendant allegedly said that he was “buying a black cape.” Cape was one correct detail, black was another, and buying was a third. A detail reported by a juror was scored as incorrect if it did not appear in the transcript. For example, a juror could have claimed that the defendant allegedly planned to buy a black mask.

Two judges who were blind to the hypotheses and who did not participate in preparing the answer key compared jurors’ memory reports to the trial transcript. They first determined whether each juror’s response captured the gist of the critical statement, and then they counted the number of correct and incorrect details in each report using the answer key. Agreement between judges was high. They disagreed about gist only twice. Regarding scoring of the details, we used a sample of responses from 40 participants to calculate interrater reliability and found that $r = .98$. When the judges disagreed, the evaluations of the arbitrarily determined “primary judge” were entered into the data analysis.

Most (96%) of the jurors correctly reported the gist of the critical evidence. Of the five who failed to do so, three were in the neutral-inadmissible condition, one was in the unusual-inadmissible condition, and one was in the neutral-admissible condition.

Regarding the details in the reports, we recognized that the neutral and unusual versions of the critical statement contained a slightly different number of correct details as identified on the answer key (25 in the neutral version versus 28 in the unusual version). Therefore, we converted the number of correct details provided by each juror to the proportion of correct details recalled, defined as the number of correct details reported divided by the number possible for the version heard by the juror. We found that jurors who heard unusual testimony reported a higher proportion of correct details ($M = .37, SD = .14$) than did those who heard neutral testimony ($M = .25, SD = .12$), $F(1, 127) = 29.23, p < .001, \eta^2 = .19$. No other effects were significant ($ps > .27$).

An analysis of the number of incorrect details yielded no effects, probably because of a floor effect (overall $M = .49, SD = .82; ps > .27$). Approximately two thirds of the jurors reported no incorrect details.

DISCUSSION

Compared to control jurors and those who heard neutral inadmissible evidence, jurors exposed to unusual inadmissible evidence were more likely to view the defendant as guilty. Moreover, jurors remembered more correct details within the unusual evidence than within the neutral evidence. These results provide preliminary support for the hypothesis that inadmissible evidence is more difficult to disregard if it is unusual rather than neutral, as predicted by Wegner’s (1994) ironic-process theory. A more complete evaluation of the findings and their interpretation is deferred until the General Discussion.

EXPERIMENT 2

The purpose of Experiment 2 was to replicate and extend the findings of the first experiment. In contrast to the Experiment 1 conceptualization of “unusualness” as testimony including unusual semantic content, the variable was defined in the second experiment as
testimony having an unusual form. Specifically, it was distinctive because (a) it contained uncommon words and phrases and (b) one of the sentences was shouted rather than spoken at a normal, conversational level.

**METHOD**

**Stimulus.** Mock jurors listened to a recording of a theft trial that was nearly identical to the one used in Experiment 1. The only change to the script involved the testimony of the critical witness. As in Experiment 1, this witness testifies either that he directly observed the defendant making an incriminating statement (admissible condition) or that someone else told him what the defendant’s statement was (inadmissible condition). In addition to the admissibility manipulation, the evidence includes either neutral or unusual details. In Experiment 1, the unusualness of the content of the statement was varied, whereas in the second experiment, the form of the statement was varied and was either neutral or unusual. Specifically, in the neutral condition, the remarks allegedly made by the defendant (after he has bought drinks for everyone at his table) are as follows:

My boss fired me today. But afterwards I started thinking, my boss is so stupid. He doesn’t know it yet, but he’s gonna give me some money. So I sneaked into his office. Now I’ve got plenty of cash.

In the unusual condition, the defendant’s statement is the following:

My boss fired me today. But afterwards I started thinking, my boss is such a dad-gum dingle-berry. He doesn’t know it yet, but he’s gonna show me the money [shouted in a manner similar to that of the title character in the movie *Jerry McGuire*]! So I sneaked into his office, and now I’ve got plenty of cheddah.

After the critical evidence is presented, the defense attorney objects. The judge then rules on the admissibility of the testimony.

**Design.** As in Experiment 1, the study’s design was 2 (admissibility) $\times$ 2 (unusualness) factorial, with an additional (control) condition in which the critical evidence was absent.

**Pilot study.** Fifty-six pilot participants read a brief description of the trial along with either the neutral ($n = 27$) or unusual ($n = 29$) version of the critical evidence. As in Experiment 1, the participants used a 10-point scale to rate the unusualness of the evidence and its reliability, probative value, salience, and emotionality. They also rated the defendant’s craziness, intended humorousness in word choice and manner of speaking, and actual humorousness. As predicted, and consistent with the Experiment 1 pilot data, the unusual statement was considered more unusual than the neutral statement, $t(54) = 5.39, p < .001, d = 1.44$. Additionally, participants thought the defendant had a greater intention to be humorous when he spoke the unusual rather than neutral statement, $t(54) = 2.86, p = .006, d = 1.00$. However, there was no difference in the ratings of actual humorousness, $t(54) = 1.49, p = .14, d = .40$. The two versions did not differ with respect to any of the other variables ($ps \geq .29$).
The participants were 163 introductory psychology students attending the same university as those in Experiment 1; 87% were White, 63% were female, and their age range was 18 to 49 years \((M = 19.70, SD = 3.70)\). The participants were tested in groups of no more than 10 individuals. By random assignment, each group listened to one of the five versions of the trial recording. As in Experiment 1, they then completed a questionnaire requesting a verdict and a confidence rating. In addition, the jurors in the four experimental groups tried to remember the critical evidence, indicated whether they intended to use the critical evidence, and rated on a 10-point scale the degree to which they thought they actually used the evidence.

### RESULTS

**Evaluations of guilt.** We calculated a guilt score for each juror as in Experiment 1. Then we examined the results by first conducting 2 \(\times\) 2 analyses of variance using only the experimental conditions (see Table 2). We obtained a significant interaction, \(F(1, 125) = 5.69, p < .02, \eta^2 = .04\), showing that the inadmissibility ruling decreased guilt scores for neutral evidence but increased them for unusual evidence. The main effect of unusualness was also significant, \(F(1, 125) = 13.28, p = .001, \eta^2 = .10\); guilt scores were higher in the unusual than in the neutral condition. There was no main effect of admissibility \((p = .99)\).

We next conducted simple analyses of variance that included the control condition. We found that guilt scores varied by condition, \(F(4, 158) = 8.52, p < .001, \eta^2 = .18\). The first two of several contrasts demonstrated that as predicted by ironic-process theory, inadmissible evidence led to higher guilt scores if it was unusual rather than neutral, \(t(158) = 4.59, p < .001, d = 1.14\); however, there was no difference between the unusual and neutral conditions when the evidence was admissible \((p = .35, d = .21)\). Another contrast, examining whether unusual evidence exerted a more powerful effect when ruled inadmissible rather than admissible (the backfire effect), approached significance, \(t(158) = 1.81, p = .07, d = .39\). A final set of contrasts compared the experimental conditions to the control condition to establish whether jurors gave any weight to the critical evidence as they decided guilt. Jurors who heard admissible evidence had higher guilt scores compared to those who never heard it: neutral, \(t(158) = 2.22, p = .03, d = .61\); unusual, \(t(158) = 3.20, \eta^2 = .01\),

### TABLE 2: Experiment 2: Mean Guilt Scores, Proportion of Correct Details Remembered, and Mean Number of Incorrect Details Remembered by Condition

<table>
<thead>
<tr>
<th>Condition</th>
<th>Guilt Score</th>
<th>Proportion of Correct Details</th>
<th>Number of Incorrect Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inadmissible</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unusual</td>
<td>2.06c (7.09)</td>
<td>.32 (.11)</td>
<td>0.21 (0.42)</td>
</tr>
<tr>
<td>Neutral</td>
<td>−5.00c (5.18)</td>
<td>.21 (.10)</td>
<td>1.21 (1.08)</td>
</tr>
<tr>
<td>Admissible</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unusual</td>
<td>−0.75ac (7.42)</td>
<td>.33 (.13)</td>
<td>0.22 (0.42)</td>
</tr>
<tr>
<td>Neutral</td>
<td>−2.23c (6.71)</td>
<td>.20 (.09)</td>
<td>1.23 (1.54)</td>
</tr>
<tr>
<td>Control</td>
<td>−5.68c (4.44)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note.** Values in parentheses are standard deviations. Means in the same column that do not share the same subscript differ significantly \((p \leq .05)\).
\[ p = .002, \ d = .81. \] On the other hand, the inadmissible evidence increased guilt scores only if it was unusual, \( t(158) = 5.34, p < .001, \ d = 1.31, \) as opposed to neutral \( (p = .57, \ d = .14). \)

**Use of the critical evidence.** We used \( 2 \times 2 \) analyses of variance to examine the remaining variables in this experiment. Regarding jurors’ reports of their intention to use the critical evidence to determine guilt, a higher proportion of jurors in the admissible conditions (\( .70 \)) than in the inadmissible conditions (\( .14 \)) said they intended to use the evidence, \( F(1, 125) = 60.88, p < .001, \eta^2 = .33. \) We found no main effect of unusualness and no interaction \( (ps > .36). \)

Concerning ratings of the extent to which they actually used the critical evidence, jurors in the admissible conditions (overall, \( M = 5.37, SD = 2.65; \) unusual, \( M = 5.75, SD = 2.17; \) neutral, \( M = 4.95, SD = 3.05 \)) gave higher ratings than did those in the inadmissible conditions (overall, \( M = 3.53, SD = 2.27; \) unusual, \( M = 3.85, SD = 2.17; \) neutral, \( M = 3.21, SD = 2.36 \)), \( F(1, 125) = 17.90, p < .001, \eta^2 = .13. \) There was no main effect of unusualness and no interaction \( (ps > .10). \)

**Memory for the critical evidence.** The evaluation of each juror’s memory report proceeded as in Experiment 1, with two judges scoring the recall of the gist of the critical evidence and counting the number of correct and incorrect details provided. As before, we converted the number of correct details to the proportion correct because there were 20 details in the neutral version and 21 in the unusual version.

There was no disagreement among the judges in relation to gist. Regarding details, a sample of responses from 40 participants used to calculate interrater reliability yielded \( r = .95. \) When the judges disagreed about details, we used the evaluations of the primary judge.

Almost all (98%) jurors correctly reported the gist of the critical evidence. One of the two who failed to do so was in the neutral-inadmissible condition, and the other was in the neutral-admissible condition.

Jurors who heard unusual evidence reported a higher proportion of correct details \( (M = .32, SD = .12) \) than did those who heard neutral evidence \( (M = .20, SD = .09), F(1, 125) = 38.95, p < .001, \eta^2 = .24. \) No other effects were significant \( (ps > .43). \)

Similarly, unusual-condition jurors reported fewer incorrect details \( (M = .22, SD = .41) \) than did neutral-condition jurors \( (M = 1.22, SD = 1.32), F(1, 125) = 38.83, p < .001, \eta^2 = .21. \) There was no main effect of admissibility and no interaction \( (ps > .95). \)

**DISCUSSION**

The results replicated the findings of Experiment 1, even though unusualness was now operationalized in terms of form rather than semantic content. Jurors exposed to unusual inadmissible evidence considered the defendant more guilty than control jurors and those in the neutral inadmissible condition. Additionally, unusual evidence was remembered better than neutral evidence. Implications are addressed in detail in the following section.

**GENERAL DISCUSSION**

In both experiments, jurors successfully ignored the inadmissible evidence if it was neutral, as shown by the similar judgments made by neutral-inadmissible and control jurors.
This result is consistent with previous data showing that jurors can often follow instructions to disregard hearsay (Miene et al., 1992; Paglia & Schuller, 1998; Pickel, 1995; Rakos & Landsman, 1992). It was a different story, however, when the evidence was unusual: jurors were more likely to make judgments that favored conviction in the unusual-inadmissible than in the control condition, which indicated that these jurors used the critical evidence in reaching their decisions. In addition, the unusual-inadmissible judgments favored guilt more than the neutral-inadmissible judgments, showing that jurors found it harder to disregard inadmissible evidence if it was unusual rather than neutral. This discrepancy between unusual and neutral conditions was not obtained when the evidence was admissible. The unusualness effect cannot be caused by differences in probative value, reliability, salience, emotionality, or the defendant’s humorousness or craziness, because pilot data demonstrated the equivalence of the two versions with respect to these variables.

The ironic-process theory can account for the findings by postulating that unusual evidence, like emotional information (Edwards & Bryan, 1997), is more accessible than neutral evidence and is kept active by the cognitive processes that are engaged when jurors attempt to suppress unwanted thoughts of inadmissible testimony. In the two present experiments, we defined unusualness in terms of either semantic content or form, with the same results. Both versions can be seen as examples of what Schmidt (1991) called “distinctive events,” which are “incongruent with active conceptual frameworks or . . . contain salient features not present in active memory” (p. 537). Thus, for example, in Experiment 1, the defendant’s alleged statement that he plans to paint his truck bright green with yellow polka dots is inconsistent with jurors’ schemas regarding vehicle paint colors, and in Experiment 2, the shouted sentence contains auditory features not shared by the other spoken sentences in the testimony. According to Schmidt (1985, 1991), distinctive stimuli receive increased attention related to an orienting response and perhaps also increased elaboration and rehearsal. The resulting memory trace is unique, containing many features that do not overlap with other traces, which reduces susceptibility to interference and increases the probability of retrieval. Consequently, these stimuli are likely to be well remembered later.

Our jurors remembered more correct details and (in one experiment) fewer incorrect details of the evidence if it was unusual rather than neutral, regardless of whether unusualness was defined in terms of the semantic content of the testimony or its form. Thus, like emotional information (Edwards & Bryan, 1997), unusual evidence may be particularly accessible to jurors (although perhaps for different reasons, as explained by Schmidt, 1991; for example, whereas emotional stimuli may produce arousal related to a defensive response, unusual material more likely causes an orienting response). This enhanced memorability of the unusual evidence can create a problem. The data from the present experiments suggest that jurors exposed to unusual inadmissible evidence struggle to disregard it, as Wegner’s (1994) theory proposes. If the goal were simply to suppress thoughts of the inadmissible item, the conscious operating process could search for distracters that are completely unrelated to it, and jurors might more easily disregard the evidence. Unfortunately, the jurors’ primary objective is to decide on a verdict, so the operating process must seek information relevant to this task, which would include other evidence items presented during the trial as well as the attorneys’ arguments and the judge’s instructions. This collection of information is not semantically unrelated to the inadmissible evidence, so the former may serve as a retrieval cue for the latter, making suppression of it difficult (Lieberman & Arndt, 2000; Lieberman et al., in press).
While the operating process is working, the automatic monitoring process runs simultaneously. Its job is to detect failures to suppress thoughts of the inadmissible evidence. When it finds such intrusions, it increases their activation so that they will enter consciousness, and the jurors will end up explicitly thinking about the evidence. As Wegner (1994) observes, the monitor will probably easily find instances of failure because it is using a search that is specific rather than general and feature positive rather than feature negative. However, for jurors exposed to unusual rather than neutral inadmissible evidence, the situation is made much worse because the off-limits information is highly accessible and thus even more detectable. Therefore, these jurors are likely to fail to disregard this evidence, and it will probably influence their judgments.

As noted previously, reactance theory appears in the literature as an alternate explanation for jurors’ failure to ignore inadmissible evidence. This theory, however, cannot easily account for the fact that our jurors were more influenced by unusual than by neutral inadmissible evidence. First, there is no reason to suppose that jurors’ autonomy is more threatened by the former than the latter. Because both experiments’ pilot data indicated that the two versions were perceived to have the same probative value and reliability, it seems improbable that jurors would more strongly insist on using the unusual evidence to determine a verdict as a way of restoring their personal freedom. Edwards and Bryan (1997) made a similar point in arguing that the ironic-process theory explained the influence of emotional evidence better than reactance theory.

A second problem with an interpretation based on reactance is that reactance theory should apply only in cases in which jurors are motivated to use the inadmissible evidence despite the judge’s instructions (Lieberman et al., in press). For example, Wolf and Montgomery (1977) found that reactance was demonstrated by the jurors in their study who rated the importance of following the judge’s instructions as low (compared to other jurors). In the present research, however, most jurors in the inadmissible conditions reported that they intended to disregard the evidence. Another way to look at the issue is that according to an explanation based on reactance theory, those in the unusual-inadmissible condition, whose judgments were clearly influenced by the critical evidence, must have rebelled against the judge’s instructions, deliberately choosing to consider the testimony. Therefore, they should have reported that they intended to use the evidence, but the clear majority of them did not. The intention results fit more closely with the ironic-process theory, because they imply that these jurors tried to follow the judge’s order but were unsuccessful because of processes beyond their control. The results also fit with previous findings that jurors may not feel very motivated to consider hearsay when determining guilt (Miene et al., 1992; Paglia & Schuller, 1998; Pickel, 1995; Rakos & Landsman, 1992) and therefore may not display reactance (Lieberman et al., in press).

Although the discussion thus far may imply that the ironic-process theory is more useful than reactance theory when it comes to describing jurors’ reactions to inadmissible evidence in general, Lieberman and Arndt (2000) suggest that the two positions could be complementary. Reactance theory might be more applicable when jurors are motivated by crime-control concerns rather than by due-process concerns. In contrast, ironic processes may occur in other situations, particularly when jurors are cognitively overloaded or when the critical evidence is especially salient or memorable for some reason. Moreover, reactance and ironic processes might interact. Additional research is needed to address these issues.
Apart from the results discussed thus far, two others deserve additional comment. First, besides measuring jurors’ intention to use the critical evidence (as noted above), we also asked them to estimate the extent to which they actually used it. Consistent with several previous studies showing jurors’ lack of awareness that certain variables can influence their decisions (e.g., Carretta & Moreland, 1983; Kassin & Sommers, 1997; Kassin & Sukel, 1997; London & Nuñez, 2000; Pickel, 1995), our jurors did not accurately assess the impact of the critical evidence on their judgments, as can be seen by the unusual-inadmissible jurors’ failure to report ratings at least as high as those in the admissible conditions. This failure reflects the fact that across many types of tasks, individuals are not always capable of correctly describing their cognitive processes (e.g., Nisbett & Wilson, 1977).

Second, Edwards and Bryan (1997) reported a backfire effect such that emotional details had a greater impact when ruled inadmissible rather than admissible. However, we did not obtain the same result with unusual evidence, although the nonsignificant effect was in the expected direction and approached significance in one of two analyses. Edwards and Bryan’s critical evidence involved the defendant’s past criminal record instead of hearsay, so the type of evidence could matter. Because jurors often seem to consider hearsay to be unreliable, and because it typically has a relatively small influence on their decisions (Lieberman & Arndt, 2000), a backfire effect with hearsay is perhaps less likely. Additionally, a backfire effect should be less probable in the absence of a strong admonition to disregard the evidence (Lieberman & Arndt, 2000; Lieberman et al., in press), and in the current experiment, the judge did not strongly admonish the jurors.

In addition to their theoretical importance, as noted above, the present results have some practical implications as well. In actual trials, witnesses occasionally make statements that are unusual in form or content. These statements might be surprising, strange, or humorous. For example, a witness being sworn in during a murder trial responded in an odd manner by spelling rather than pronouncing her entire name and adding, “And that is not my a.k.a, also known as” when asked to identify herself to the court (Stodghill, 1999, p. 01B). Another witness, in a sex discrimination trial, shocked the jury by abruptly volunteering verbatim the profanity-laced insult he had directed toward an employee (Wolfe, 1996). Exemplifying testimony with unusual content, a murder trial witness described her past lives, which she said she had discovered by gazing into a crystal ball, noting that she thought Shakespeare wrote the part of Ophelia for her (Andrews, 1998). A witness in a different trial amused courtroom observers by relating that when she feared a neighbor’s dog was about to attack her chickens, she called them to come to her, although she admitted on cross-examination that being chickens, they typically do not come when called (Doege, 2004).

The present findings suggest that if unusual statements are also inadmissible, their influence cannot be negated by having them ruled as such. If anything, an inadmissibility ruling could increase their effect somewhat. Therefore, attorneys should weigh the costs and benefits of raising an objection after the evidence has already been introduced. If the legal advantage is not great enough, they could strategically opt not to object (Mauet, 1996). Even better would be to find a way to prevent damaging unusual testimony from emerging in the first place, and no doubt the opposing counsel would sometimes be able to raise an objection before the witness could actually articulate the critical information. On the other hand, hearsay can be difficult to avoid. Because of the complexity of the hearsay rule, it may be initially unclear whether a statement is excluded by the rule or is covered under one of many exceptions; it appears that most evidence questions that arise during trials involve
interpretations of the hearsay rule (Rothstein, 1981). Moreover, witnesses sometimes blurt out inadmissible statements before the opposing counsel can decide whether to object (Paglia & Schuller, 1998).

As with most studies of inadmissible evidence (Lieberman et al., in press), a limitation of the current experiments is that the jurors made decisions as individuals but did not deliberate. Previous studies designed to determine whether deliberation attenuates or exacerbates the potency of inadmissible evidence have yielded conflicting results (see Lieberman & Arndt, 2000). Although it is clear that further research is needed to investigate the effects of deliberation after the jurors have heard unusual inadmissible evidence, we contend that our study of individual guilt judgments represents an important starting point, as jurors’ initial positions presumably feed into the group deliberation.

Another limitation is that our participants were university students rather than members of a community sample who would be more demographically representative of actual jurors. However, some published comparisons of these two groups of participants have revealed no major differences in their responses (Bornstein, 1999; Lee et al., 2005). Therefore, although additional research using a community sample is warranted, our conclusions are probably reasonable.

Finally, we attempted to use a realistic legal case that was not overly simplistic, but actual jurors must listen to even lengthier and more complex trials that may continue for days or weeks. A longer trial might increase the cognitive load on jurors so that the influence of the inadmissible evidence would be magnified, as predicted by ironic-process theory (Lieberman et al., in press). Future research could examine this question.

REFERENCES


**Kerri L. Pickel**, PhD, is a professor in the Department of Psychological Science at Ball State University. She conducts research on eyewitness memory and juror decision making.

**Tanya J. Karam** received a BA at the University of Waterloo and an MA in cognitive and social processes at Ball State University. Currently she is pursuing her PhD in cognitive and developmental psychology at Louisiana State University. Her interests are in eyewitness memory and jury decision making. Specifically, she is currently researching source memory and memory binding for events.

**Todd C. Warner** is a research consultant at the Prevention Research Center at Arizona State University. His primary areas of interest include jury decision-making processes, interrogation practices, videotaped confessions, and retracted confessions.