Jurors’ Use of Standards of Proof in Decisions about Punitive Damages

William Douglas Woody, Ph.D.* and Edie Greene, Ph.D.†

Standards of proof define the degree to which jurors must be satisfied that a fact is true, and plaintiffs in civil lawsuits assume the burden of proving their claims to the requisite standard of proof. Three standards—preponderance of evidence, clear and convincing evidence, and beyond a reasonable doubt—are used by different jurisdictions in trials involving liability for punitive damages. We investigated whether individual mock jurors apply these standards appropriately by instructing them to read two personal injury trial summaries and to use one of three standards in either qualitative or quantitative format when deciding punitive liability. Results showed that jurors tended not to incorporate the standard into their judgments: defendants were just as likely to be found liable when the plaintiff’s burden was high (“beyond a reasonable doubt”) as when the burden was low (“preponderance of evidence”). The format of the instruction also had a negligible effect. We suggest that nonuse of the standard of proof is related to jurors’ preferences for less effortful or experiential processing in situations involving complicated or ambiguous material. Copyright © 2012 John Wiley & Sons, Ltd.

INTRODUCTION

Compensatory and punitive damage awards serve distinct legal functions in civil cases. Compensatory damages compensate the plaintiff for past, present, and future losses that result from the defendant’s behavior (Wissler, Evans, Hart, Morry, & Saks, 1997). If the defendant’s behavior has been particularly egregious, the plaintiff may also seek punitive damages, intended to punish a willfully malicious or reckless defendant and deter that defendant and others from similar actions in the future (Vidmar & Wolfe, 2009).

In trials involving punitive damages, jurors must decide whether the defendant is liable for punitive damages, and, if so, in what amount (Greene & Bornstein, 2003). But they receive “precious little guidance” (Greene & Bornstein, 2000, p. 743) regarding decisions about punitive damage awards. Their instructions are generally limited and vague, and significant concerns exist that unclear instructions increase the possibility that bias will inform decisions about punitive awards. The U.S. Supreme Court has expressed concerns regarding the lack of guidance in punitive damages instructions (see, e.g., Honda Motor Co. Ltd. et al. v. Karl L. Oberg, 1994; TXO Production Corp. v. Alliance Resources Corp. et al., 1993), and some commentators have argued that, without clearly defined directives to guide them, juries impose substantial and unwarranted punitive damage awards (Kuhlik & Kingham, 1990).
We focus on one directive intended to guide decisions about punitive damages: the standard of proof. The standard of proof in a civil trial varies depending on the nature of the decision (i.e., decisions regarding liability, compensatory damages, and punitive damages may entail different standards of proof) as well as the type of case (e.g., cases involving civil commitment and termination of parental rights typically require a higher standard of proof than those involving personal injury; see Addington v. Texas, 1979; Stantosky v. Kramer, 1982).

A plaintiff who makes a claim for punitive damages carries the burden of proof. This means that, in order to prevail, the plaintiff must demonstrate that the defendant behaved in ways that showed willful or malicious disregard of others and must convince the judge or jury to the degree specified by the standard of proof (see, e.g., In re Winship, 1970; Addington v. Texas, 1979, for judicial discussions of standards of proof). Standards of proof for punitive liability vary by jurisdiction. The present study investigated whether mock jurors are attentive to differences in standards of proof regarding punitive liability and, in particular, whether their judgments reflect consideration of the relevant standard.

**Why Examine Standards of Proof for Liability for Punitive Damages?**

Relatively few civil cases actually go to trial (Wissler et al., 1997), but decisions regarding whether to settle or proceed to trial involve many complex factors and the relevant standard of proof may influence these choices, particularly if attorneys’ beliefs reflect legal expectations that standards of proof shape jurors’ decisions. Arguably, plaintiffs seeking punitive damages may be less likely to settle in jurisdictions in which their burden of proof for punitive damages is low and their chances of prevailing at trial are high. On the other hand, they may be more likely to settle when their burden of proof is high and their chances of prevailing are low. However, this line of thinking presupposes that jurors understand and correctly apply the standard of proof. Whether they do and whether variations in instructions influence their abilities to do so are empirical questions we address in this study.

In addition to influencing settlement choices, the standard of proof can have other effects on the course of litigation. Plaintiffs who seek punitive damages may file their cases (“forum shop”) in sympathetic venues, particularly those that impose relatively lax standards of proof for punitive liability (Bassett, 2006). Another indication of the importance of the standard of proof regarding punitive damages is that legal reformers, including the American Tort Reform Association (ATRA), have sought to legislate more stringent standards of proof for liability for punitive damage awards (ATRA, 2007). The U.S. Supreme Court provided support for these reforms with their dicta in Pacific Mutual Life Insurance Company v. Haslip et al. (1991), in which the majority favored raising the standard of proof for liability for punitive damages. Additionally, in their discussion of the adverse effects of punitive damage awards on pharmaceutical research, Kuhlik and Kingham (1990) advocate for a heightened standard of proof for punitive damages. Finally, Hastie, Schkade, and Payne (1998) proposed that, if juries are biased in favor of injured plaintiffs and against unsympathetic defendants, increasing the required standard of proof should reduce the frequency of improper verdicts. These recommendations all presuppose that jurors attend to and abide by the requisite levels of proof in assessing liability for punitive damages. Do they?
Standards of Proof Regarding Punitive Damages

Neither a judge nor a jury can ever be absolutely certain of facts in dispute. At best, they can attempt to determine what probably happened; sometimes they err. Recognizing the probabilistic nature of trial-related decisions and attempting to allocate fairly the risks of error, the law imposes rules—standards of proof—that define the degree to which decision makers (i.e., judge and jury) must be satisfied that a fact is true (In re Winship, 1970). The standard of proof serves as a criterion for decision making, a benchmark jurors and judges should use to determine whether each of the requisite elements has been established to the specified level of proof and hence whether the plaintiff should prevail. Jurisdictions employ one of three standards of proof in decisions concerning punitive damages: (1) beyond a reasonable doubt, (2) by clear and convincing evidence, and (3) by a preponderance of evidence.

The most stringent standard of proof (and the standard also used in criminal cases) is “beyond a reasonable doubt.” In criminal law, this strict criterion is weighted in the defendant’s favor and reflects the severe consequences of mistakenly convicting an innocent person. Studies that measured quantitative interpretations of this standard show wide inter-individual differences (Dhami, 2008) with probabilities ranging from .50 to 1.00 and a mean of approximately .90–.92 (Hastie, Penrod, & Pennington, 1983; McCauliff, 1982). This standard is relevant to decisions about liability for punitive awards in just one state, Colorado (ATRA, 2007).

The “preponderance of evidence” standard, commonly used in judging liability for compensatory damages in civil trials, is the most lenient. It requires the plaintiff to prove that his or her claims are more likely to be true than not true, i.e., that the probability of a disputed fact is greater than .50. This standard allocates the risk of error roughly equally between plaintiffs and defendants (Addington v. Texas, 1979). Nineteen states use this standard for deciding liability for punitive damages (ATRA, 2007).

The standard “clear and convincing evidence” is positioned between the other two, though converting this standard into a probability estimate presents difficulty (Friedman, 2001). Allen and Jehl (2003) suggest that it refers to a probability estimate higher than .75, one appellate court quantified the “clear and convincing evidence” standard for admission of evidence as .80 (United States v. Fatico, 1979), and surveys of judges suggest that the probability is approximately .70. or .80 (McCauliff, 1982). Thirty states as well as the District of Columbia use this standard of proof in decisions regarding liability for punitive damages (ATRA, 2007).

The three standards of proof direct triers of fact to use different levels of certainty in making their decisions. This is problematic for at least three reasons. First, as noted, subjective judgments as to the level of certainty imposed by any given standard differ widely. In an early psychological review of these standards, Levine (1974) noted that “the probabilistic decision criterion is unabashedly subjective” (p. 672; see also Hyman & Newhouse, 1965, for an early legal review) and judges vary amongst themselves in interpretations of what the standards require (McCauliff, 1982). Second, even if decision makers could agree on the level of certainty required before making a particular judgment, they may have difficulty applying these standards to a controversial set of facts (see Kagehiro & Stanton, 1985; Reardon, O’Neil, & Levett, 2007). Applying the appropriate standard may be especially problematic in the realm of punitive damages, where, according to some theorists (e.g., Sunstein, Kahneman, Schkade, & Ritov, 2002), jurors who share a sense of moral outrage at the defendant’s
conduct have no agreed-upon method of converting that sentiment into a punitive damage award. Finally, when standards of certainty vary from “slightly more certain than not” (the preponderance of evidence standard) to “certain beyond a reasonable doubt,” the implications for civil plaintiffs and defendants could be enormous. Presumably, plaintiffs are most likely to receive punitive damages in jurisdictions with a preponderance of evidence standard and least likely to do so in the jurisdiction with a reasonable doubt standard.

Qualitative and Quantitative Standards of Proof

Standards of proof can also be presented in different formats: qualitative (sometimes termed “legal” because they tend to be written in legalistic language), quantitative, or both qualitative and quantitative. For example, a qualitative definition of beyond a reasonable doubt is “proof that leaves you firmly convinced of the defendant’s guilt” (Federal Judicial Center, 1987, p. 28). A quantitative definition is the following: “the defendant is presumed innocent unless the evidence against him has at least 91% probability of truth.”

Do jurors’ interpretations and applications of standards of proof depend on the format in which such standards are presented? Apparently, yes. Studies of how jurors interpret the reasonable doubt standard show that quantitative standards (or quantitative wording followed by qualitative wording) are generally more effective than qualitative standards alone in informing jurors of the stringent nature of this criterion (Dhami, 2008; Kagehiro, 1990).

Jurors’ Abilities to Apply Standards of Proof

Legal standards of proof, particularly those used in civil trials, are arguably arcane concepts to many laypeople. Whereas the public may be familiar with the concept of reasonable doubt from media attention to criminal trials, the concepts of preponderance of evidence and clear and convincing evidence are undoubtedly less well known. In their study of mock jury decisions in a product liability case, Landsman, Diamond, Dimitropoulos, and Saks (1998) found that jury-eligible adults had significant difficulty understanding what the standard of proof meant. Similarly, when Greene and Johns (2001) asked mock jurors in a negligence case about the standard of proof, only approximately half the jurors knew what percentage of evidence equals a preponderance.

Some research has clarified the circumstances under which standard of proof instructions are likely to be used correctly. For example, Kagehiro and Stanton (1985) evaluated mock jurors’ ability to apply standards of proof in deciding civil liability. They asked jurors to read instructions involving different standards of proof (i.e., either preponderance of evidence, clear and convincing evidence, or beyond a reasonable doubt) in either qualitative or quantitative terminology and to apply them to a set of facts involving an accident. Researchers measured the likelihood of verdicts favorable to the plaintiff as well as perceptions of the plaintiff’s difficulty in meeting the standard of proof, and found that only those jurors who read quantified

---

1 It is unclear to what decision, exactly, the standard of proof applied. Because these were accident cases and jurors were asked to render a verdict in favor of either the plaintiff or defendant, they presumably determined whether the defendant was liable for negligence, although they were not instructed on the elements of negligence.
instructions used the standards of proof in accordance with legal expectations. Specifically, among this group, as the standard of proof became more stringent, the likelihood of a verdict favoring the plaintiff decreased, and the perceived difficulty in meeting the standard of proof increased. But neither qualitative nor combined qualitative and quantitative standards led to this result.

These results apply only to jurors’ decisions regarding liability for compensatory damages and not to decisions regarding punitive damages. Because punitive damage awards require jurors to make a different legal decision, namely whether the defendant has shown willful or wanton disregard for others, we cannot presume that findings regarding how jurors use standard of proof instructions on compensatory liability would also apply to punitive liability judgments. Thus, in this study we assess whether jurors systematically consider the standard of proof when translating sentiments about the defendant’s conduct into decisions about liability for punitive damages. In particular, we ask how variations in the standard of proof (involving increasing levels of stringency and both qualitative and quantitative formats) influence mock jurors’ judgments regarding liability for punitive damages.

Hypotheses

Based on the findings of Kagehiro and Stanton (1985), we hypothesized that jurors would correctly interpret and apply the increasingly stringent standards of proof only when those standards were presented in quantitative fashion. We expected that, when the standards provided numerical information regarding the level of proof required, jurors would be able to distinguish more lenient from less lenient standards and adjust their judgments concerning liability for punitive damages accordingly. More precisely, we hypothesized that jurors with quantitative standards would be more likely to find in favor of the defendant (i.e., declining to impose a punitive damage award) as the rigor of the standard increased from preponderance of evidence to beyond a reasonable doubt. We further hypothesized that verdicts from jurors with qualitative standards would not differ as the rigor of the standard increased. We also hypothesized that jurors who read quantitative standards would be more confident than those with qualitative instructions that they followed instructions. Finally, as the standard of proof rose for jurors with quantitative standards, we expected that participants would have increasing confidence in their liability decisions and would report higher ratings of the degree to which the evidence favors the plaintiff when the defendant is found liable.

METHOD

Participants and Design

Participants included 626 undergraduate students at Western universities who participated as part of a course requirement. We treated all participants in accordance with American Psychological Association ethical principles (American Psychological Association, 2002).

The design was a 3 (standard of proof: preponderance of evidence, clear and convincing evidence, reasonable doubt) × 2 (format of standard of proof: qualitative,
quantitative) between subjects factorial. There were between 95 and 113 participants in each of six conditions, although, as noted below, only 333 participants produced usable data.

**Materials**

**Trial Stimuli**

Participants read summaries of two actual cases. One involved a claim of injury against the owner of a mall after a rape was committed in the parking lot, and the second involved a wrongful death claim brought by the survivors of 39 seamen who died when their vessel sank (*Jardel Co., Inc. v. Hughes*, 1987, and *In re Marine Sulphur Queen*, 1972, respectively). These cases were used in previous research involving jurors’ decisions of liability for punitive damages (Hastie et al., 1998). The summaries, between 1000 and 1500 words long, were based on the courts’ summary statements and provided closing arguments in favor of and against punitive damages from the plaintiff and the defense, respectively.

**Sexual Assault Injury Case**

The plaintiff was a woman who had been sexually assaulted as she left the mall after an evening shift at the retail store where she worked. Jardel Co., Inc., the owner and operator of the mall, had hired a security contractor to provide security for the mall. The contractor repeatedly recommended that Jardel increase the number of guards but Jardel, citing budgetary concerns, declined to do so, even though there had been numerous police responses to criminal activity in the parking lot in the previous two and a half years. When the assault occurred, the single guard on duty was in his patrol vehicle on the other side of the mall. The plaintiff was awarded compensatory damages for economic losses and pain and suffering, and sought punitive damages. In the original trial, the jury found Jardel Co., Inc. liable for punitive damages and awarded $250,000 in punitive damages. On appeal, the Delaware Supreme Court reversed the jury’s decision to award punitive damages.

**Wrongful Death Case**

The wrongful death case was brought against Marine Sulphur Transport Corp., the owner of a tanker refitted to transport molten sulfur, by the heirs of 39 seamen who lost their lives when the ship sank in mild seas. The refitting had involved some untested

---

2 We performed a series of power analyses to determine the sample sizes required to have sufficient power to detect main effects (i.e., differences between three standards of proof and between two formats) and the 3 x 2 interaction. Studies of jury decisions in cases involving punitive damage awards often result in small to medium effect sizes (see, e.g., Matire & Kemp, 2009; McAuliff, Kovera, & Nunez, 2009). Using an alpha level of .05, and an f of .25, 52 participants in each of six cells are required to produce power of .80 for detecting 3 x 2 interactions (Cohen, 1988).

3 Both cases were referred to a higher court for review of the decision to allow punitive damages and, although the higher courts denied the request for punitive damages, disagreements existed among judges (Hastie, Schkade, & Payne, 1999; Vidmar, 1999). The conflicting rulings reflect the difficulties inherent in resolving the cases and suggest that the fact patterns could be useful in evaluating the effects of varying standards of proof.
conversion techniques, and a Coast Guard naval engineer testified that the conversion of the ship “was a calculated risk.” The ship appeared to have capsized suddenly, perhaps due to poor rolling capability related to the conversion or to excessive cargo loading. According to the plaintiffs, Marine Sulphur Transport Corp. knew that the refitting could present problems but did not adequately test the system, repeatedly allowed the ship to sail while overloaded, and did not have an adequate warning system for the crew. Marine Sulphur Transport Corp. paid compensatory damages to the plaintiffs, who then sought punitive damages because the company’s decisions “were entirely dominated by profit considerations.” The original case was decided by a United States District Court, which found the ship’s owners liable for compensatory and punitive damages. On appeal, the Second Circuit Court of Appeals supported the owners’ request for denial of punitive damages.

**Instructions and Standards of Proof**

The 500-word-long punitive damages jury instruction was identical to that used by Hastie et al. (1998). It (a) defined and listed the punitive and deterrent purposes of punitive damage awards, (b) defined malicious and reckless conduct, (c) provided four grounds for recklessness (the defendant must be aware of a risk, evaluate the risk, disregard the risk, and do so to a degree that grossly deviates from the standard of a reasonable person), and (d) differentiated between recklessness and negligence.

The qualitative versions of the standards of proof were as follows: “Preponderance of evidence is such evidence as, when weighed with that opposed to it, has more convincing force and the greater probability of truth”; “Clear and convincing evidence is that measured degree of proof which will produce in the mind of a juror a firm belief or conviction as to the truth of the allegations sought to be established”; “Reasonable doubt is a doubt based on reason and common sense which arises from a fair and rational consideration of all of the evidence, or lack of evidence, in the case. It is a doubt which is not vague, speculative or imaginary doubt, but such a doubt as would cause reasonable people to hesitate to act in matters of importance to themselves” (Colorado Supreme Court Committee on Civil Jury Instructions, 1999).

In quantitative format, each standard of proof gave the percentage of evidence that must favor the plaintiff before participants could find the defendant liable for punitive damages. The preponderance of evidence standard instructed jurors to find for the plaintiff if 51% of the evidence favored the plaintiff; the values for clear and convincing evidence and beyond a reasonable doubt were 71% and 91%, respectively.

**Case Questionnaire**

After reading each case summary and the relevant jury instructions, participants answered a series of questions. The first was a “yes/no” question as to whether the defendant was liable for punitive damages. The next question asked participants to rate their confidence in that decision using a 100-point scale (0%, *not at all confident*; 100%, *completely confident*). They also used 100-point scales to rate the degree to which the evidence favored the defendant or the plaintiff (0%, *completely favored defendant*; 100%, *completely favored plaintiff*), and the confidence that they had followed the instructions (0%, *not at all confident*; 100%, *completely confident*).
Procedure

Prior to reading each case summary, participants were informed that the defendant had previously been found negligent, that the plaintiff had already received compensatory damages, and that their task was to decide whether the defendant was liable for punitive damages. After they read the case summary and jury instructions regarding punitive damages, they read the appropriate standard of proof in either qualitative or quantitative language. They then answered the case questionnaire and went on to the second case. Participants had the same standard of proof instruction for both cases. At the conclusion of the study and as a manipulation check, participants answered a multiple-choice question about which standard of proof they had been instructed to use. We debriefed all participants.

RESULTS

Manipulation Check

Only 333 (53.2%) of the 626 participants correctly answered the manipulation check question requiring them to identify the standard of proof they were instructed to use. To evaluate effects of the standard of proof and the form of instructions on participants’ responses, we used a binary logistic regression with a correct response to the manipulation check question as the dependent variable and standard of proof (preponderance of evidence, clear and convincing evidence, and beyond a reasonable doubt) and instruction format (qualitative or quantitative) as predictor variables. The model was significant, \(-2\text{ log likelihood} = 822.24\) (df = 623), \(p = .008\), Cox and Snell \(R^2 = .02\). Standard of proof significantly affected participants’ likelihood of correctly answering the manipulation check question, Wald = 11.21, \(p = .004\), but the format of instructions was not significantly associated with response accuracy (Wald = .35, \(p = .56\)). More specifically, the manipulation check question was correctly answered by 49.5% of participants in the preponderance of evidence condition, 51.5% of participants in the clear and convincing evidence condition, and 65.2% of participants in the reasonable doubt condition. Analyses of dummy-coded variables in the logistic regression equation indicated that participants using the reasonable doubt standard were more likely to correctly identify that standard than participants in the preponderance of evidence condition (\(p = .002\)) or participants in the clear and convincing condition (\(p = .007\)).

Resultant Power

All subsequent analyses involve responses from only the 333 participants who correctly identified the standard of proof they were instructed to use. After removing participants who did not accurately answer this question, we retained sufficient power to detect main effects and interactions; in particular we had between 52 and 64 participants in each of six cells to detect small interactions.

4 Although this finding may, in itself, be instructive regarding jurors’ attention to the standard of proof issue, we suspect that it may be an artifact of the simulated nature of the jurors’ task, as we note in our discussion.
Case Questions

Liability Decisions

In the sexual assault case, 188 (56.5%) participants found the defendant liable for punitive damages, and in the wrongful death case, 269 (80.8%) participants found the defendant liable. The difference between cases was significant, $X^2(1) = 46.68, p < .001$, $\phi = .27$, CI$_{.95} = .19, .33$. Because of this difference, we analyzed other dependent variables separately by case.

We expected that participants would be less likely to find the defendant liable for punitive damages as the standard of proof became increasingly strict, but only when the standard of proof was expressed quantitatively. Table 1 shows the number and percentage of participants who found the defendant liable for punitive damages as a function of case type, the format of the instructions and the standard of proof. We used two logistic regression equations (i.e., one for each case) to evaluate the effects of these variables and their interaction on participants’ liability decisions. The equation for the sexual assault case was not significant, $-2 \log L = 450.37$ (df = 327), $p = .34$, and did not reveal any significant main effects or interactions (all $p > .11$). The equation for the wrongful death case was also not significant, $-2 \log L = 314.61$ (df = 328), $p = .16$, and no main effects or interactions were significant (all $p > .08$). There was a small significant effect for the interaction of form and a specific dummy-coded variable for standard of proof, however, such that in the preponderance of evidence condition participants who read qualitative instructions were slightly more likely to find the defendant liable than were participants who read quantitative instructions, $\text{Wald} = 3.88, p = .049$.

Confidence in Liability Decisions

We expected that mock jurors’ confidence in their liability decisions would increase with increasingly stringent decision criteria because one needs more certitude to be convinced beyond a reasonable doubt than to be convinced by a preponderance of evidence. To evaluate participants’ perceptions of confidence in their liability decisions, we used two ANOVAs (i.e., one for each case) with standard of proof and form of instructions as independent variables. For the 188 participants who found the defendant liable in the sexual assault case, there were no significant differences in confidence as a function of standard of proof and form of instructions (all $p > .31$, all $\eta^2_p < .02$); the same pattern of results existed for the 269 participants who found the defendant liable in the wrongful death case (all $p > .17$, all $\eta^2_p < .02$). Although increasingly stringent standards of proof require jurors who find the defendant liable to be increasingly confident in their decisions, jurors in the present study did not behave accordingly, and the format of instructions did not affect their confidence ratings.

Weight of Evidence

We evaluated participants’ perceptions of the degree to which the evidence favored the parties as a function of their punitive liability decisions. Participants used a percent scale (0%, evidence completely favored defendant; 100%, evidence completely favored plaintiff) to rate the extent to which the evidence favored the parties. We expected participants who found the defendant liable to perceive the evidence as more favorable to
Table 1. Total N and percent of participants who found the defendant liable for punitive damages as a function of case, format of instructions, and standard of proof

<table>
<thead>
<tr>
<th>Standard of proof</th>
<th>Sexual assault case</th>
<th>Wrongful death case</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Qualitative</td>
<td>Quantitative</td>
</tr>
<tr>
<td>Preponderance of evidence</td>
<td>29 (33.7%)</td>
<td>30 (55.6%)</td>
</tr>
<tr>
<td>Clear and convincing evidence</td>
<td>35 (66.0%)</td>
<td>23 (44.2%)</td>
</tr>
<tr>
<td>Beyond a reasonable doubt</td>
<td>38 (59.4%)</td>
<td>33 (58.9%)</td>
</tr>
<tr>
<td>Totals</td>
<td>102 (59.6%)</td>
<td>86 (53.1%)</td>
</tr>
</tbody>
</table>
the plaintiff than did participants who did not find the defendant liable. For the sexual assault case, participants who found the defendant liable perceived the evidence to be more in favor of the plaintiff ($M = 79.96\%, \ SD = 14.65$) than did participants who did not find the defendant liable ($M = 40.88\%, \ SD = 22.97$), independent $t(331) = 18.83$, $p < .001$, $d = 2.08$, CI$_{95} = 1.81, 2.35$. The same pattern emerged in the wrongful death case: participants who found the defendant liable perceived the evidence to be more favorable to the plaintiff ($M = 84.26\%, \ SD = 12.83$) than did participants who did not find the defendant liable ($M = 44.51\%, \ SD = 19.80$), independent $t(222) = 19.72$, $p < .001$, $d = 2.76$, CI$_{95} = 2.38, 3.13$.

We also analyzed perceptions of the evidence among only those participants who found the defendant liable for punitive damages. We expected participants who found the defendant liable using increasingly stringent standards of proof to say that the evidence increasingly favored the plaintiff. These data are shown in Table 2. Using an ANOVA with the standard of proof and the form of the instructions as independent variables, we analyzed data from the 188 participants who found the defendant liable for punitive damages in the sexual assault case. There were no significant main effects or interactions (all $p > .44$, all $\eta^2 < .01$). In the wrongful death case, the analysis also failed to reveal significant main effects or interactions (all $p > .14$, all $\eta^2 < .02$). Among jurors who theoretically “applied” the standard of proof to determine that the defendant was liable, neither the particular standard of proof nor the form of the instruction affected their estimates of the extent to which the evidence favored the plaintiff.

**Jurors’ Confidence that they Followed Instructions**

All jurors reported their perceived confidence that they obeyed the judicial instructions using a 0–100% scale. We used ANOVAs to evaluate the effects of standard of proof and form of instructions. The ANOVA for the sexual assault case did not reveal significant differences (all $p > .08$, all $\eta^2 < .02$). The ANOVA for the wrongful death case demonstrated that participants who read qualitative instructions ($M = 87.02, \ SD = 13.37$) were slightly more confident that they followed directions than were participants who read quantitative instructions ($M = 83.30, \ SD = 16.42$), $F(1, 326) = 4.83$, $p = .03$, $\eta^2 = .01$, CI$_{95} = .00, .05$.

**DISCUSSION**

The standard of proof relevant to punitive damages is important for a number of reasons. Beyond its impact in specific trials, the standard of proof involving liability for punitive damages can influence decisions about whether to go to trial and venue selection, and raising the standard is touted as a way to reform jury trials. Courts expect standards of proof to guide jurors’ decisions regarding liability for punitive damage awards. Hypothetically, more stringent standards of proof should result in (a) more decisions in favor of the defendant, (b) higher ratings of the extent to which the evidence favors the plaintiff when the defendant is found liable, and (c) greater confidence among jurors in their liability decisions. If jurors struggle to use instructions written in legal language (so-called “qualitative instructions”) they may be able to better use quantitative instructions (Kagehiro & Stanton, 1985).
<table>
<thead>
<tr>
<th>Standard of proof</th>
<th>Sexual assault case</th>
<th>Wrongful death case</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Qualitative</td>
<td>Quantitative</td>
</tr>
<tr>
<td>Preponderance of evidence</td>
<td>80.69 (11.52)</td>
<td>76.90 (11.55)</td>
</tr>
<tr>
<td>Clear and convincing evidence</td>
<td>78.77 (16.60)</td>
<td>81.52 (13.01)</td>
</tr>
<tr>
<td>Beyond a reasonable doubt</td>
<td>82.32 (11.87)</td>
<td>79.58 (19.60)</td>
</tr>
<tr>
<td>Means</td>
<td>80.63 (13.79)</td>
<td>79.16 (15.67)</td>
</tr>
</tbody>
</table>
Our findings suggest, though, that even among jurors who know which standard to apply, the standard of proof had relatively little impact on their decisions regarding punitive liability. Jurors were no more likely to decide that the defendant was liable for punitive damages when they had to be convinced by a mere preponderance of evidence than when they had to be convinced beyond a reasonable doubt. Furthermore, quantitative instructions did not improve jurors’ performance over instructions written in qualitative, legal language. For jurors in the wrongful death case only, the marginal difference in liability rates as a function of form of instructions for jurors in the preponderance of evidence condition (i.e., participants in the qualitative condition were slightly more likely to find the defendant liable for punitive damages) reflects a very small effect in one condition in one case for the variable with the most statistical power and did not persist across conditions or extend to the other case.

As expected, mock jurors who found the defendant liable for punitive damages rated the evidence as more favorable to the plaintiff than did jurors who found in favor of the defendant. Yet among participants who found the defendant liable, there were no differences in the rated extent to which the evidence favored the plaintiff as a function of increasingly strict standards of proof. Nor were mock jurors with stringent standards any more confident in their decisions than were participants who used less stringent standards. In short, none of our main hypotheses was confirmed.

We suspect that these findings are not attributable to mock jurors’ uncertainty about which standard to use to assess liability because all the relevant data came from participants who correctly identified which standard they were instructed to use. (Importantly, actual jurors take no test to determine whether they are aware of the relevant standard, so it is quite likely that some enter the deliberation room with little understanding of the fact that they must reach a decision by attending to a particular decision criterion. In fact, in a similar study, when participants were asked to freely recall the required standard of proof, only 14.6% could do so correctly [Rudebeck, Ness, Greene, & Woody, 2001]). Nor do we believe that the results reflect a lack of statistical power to find small effects.

Rather, we suspect the findings stem from the fact that standards of proof are unfamiliar and murky concepts to lay jurors. Applying novel concepts like “such evidence as, when weighed with that opposed to it, has more convincing force...” and “that measure of degree of proof which will produce... a conviction as to the truth of the allegations...” demands effortful cognitive processing: slowing reading speed, parsing dense syntactic structures, and extracting meaning from complex language.

Dual process models of information processing (e.g., Cognitive–Experiential Self-Theory [Epstein & Pacini, 1999]) propose that in situations in which decision-makers lack sufficient motivation or ability to process information carefully and systematically and in which the information is uncertain or ambiguous, people will rely on less effortful, experiential processing that reduces the complexity of the task (Chaiken, Liberman, & Eagly, 1989). Yet understanding and applying the relevant standard of proof demands deliberative, analytical, and effortful processing that may elude many jurors. Interestingly, participants in this study were most likely to recognize the standard they were instructed to use when that standard (“beyond a reasonable doubt”) matched a concept that, due to its ubiquity in both news and popular media related to crime, is already accessible and meaningful to them.

When legal concepts inherent in standards of proof are unfamiliar to laypeople, some of them may simply opt to rely on their naïve, “everyday” concepts of the law—their commonsense notions—in determining whether the defendant acted with reckless
or malicious disregard of others. As a result, they pay little heed to the level of certitude that should frame the judgment because that requirement may not comport with their commonsense ideas about justice and fairness (Finkel, 2000). Hastie et al. (1998) made the apt observation that “[jurors’] everyday habits of assessing blame do not match the legal procedures of verifying that a series of requisite ‘elements’ have each been established to a specified (and unfamiliar) ‘standard of proof’” (p. 307).

We acknowledge that the simulated nature of this study might also have contributed to the apparent disregard of the standard of proof. In particular, had jurors come together to discuss the case as a jury, they might have been more likely to review the jury instructions. Hastie et al. (1998) reported that during mock jurors’ deliberation on a civil case 14.6% of statements related to legal issues, which included judicial instructions. This jury discussion could have focused attention on the requisite standard of proof; small group discussion might also have enhanced that focus if one or two jurors with awareness and understanding could inform the entire group. Yet, as Kagehiro and Stanton (1985) noted, if anything, a simulation study that involves condensed evidence and jury instructions, followed immediately by questions about the required standard of proof, should allow for the standard of proof to have its maximum impact. We suspect that the nonuse of standards of proof in the present study stems from the abstract nature of the requirement to assess various elements of a claim against a vague standard of certainty.

Consequences of Nonuse

What consequences result if jurors do not appropriately use standards of proof to decide liability for punitive damages? First, their decisions in cases involving punitive damages may be influenced by pre-existing beliefs and attitudes about the law (e.g., a belief in the existence of a litigation crisis) as well as by individual differences in need for cognition, rationality, and strength of will (Vinson, Costanzo, & Berger, 2008).

Second, although attorneys assume that one of the most important factors in determining civil litigation outcomes is the choice of forum (Borchers, 2010), our data suggest that plaintiffs or defendants who painstakingly consider the costs and benefits associated with filing cases in forums with more favorable standards of proof may be expending energy unnecessarily. Seeking or resisting a forum with a particular standard of proof rests on the assumption that jurors appropriately use these standards, and these subtleties may be lost on many jurors (Lieberman, 2009).

Finally, our findings suggest that reformers might reevaluate their calls for heightened standards of proof as an avenue for reform. Although courts have suggested that more stringent standards of proof could reduce the frequency of punitive awards (Apelbaum & Ryder, 1999; Pacific Mutual Life Insurance Company v. Haslip et al., 1991), any changes will be rendered moot if jurors pay little regard to the requisite standard of proof.

Limitations and Strengths

Our study employed fairly standard mock juror methodology. It can be argued that various methodological choices (e.g., to use student jurors and condensed evidentiary material and to measure responses at the predeliberation stage) limit the external validity of our findings; for example, Fox, Wingrove, and Pfeifer (2011) found differences in punitive damage awards recommended by community and student jurors. On the
other hand, Rose and Ogloff (2001) state “for assessing the comprehensibility of jury instructions, it may be acceptable to use undergraduate students as participants, to use individual participants without group deliberation, and to employ written stimulus materials” (p. 409). Indeed, our intention was to evaluate whether jurors attend to and employ the standard of proof they are instructed to use, and one might anticipate that attention to that standard would be heightened when other trial-relevant information is condensed and when the crucial decision (i.e., whether the defendant is liable for punitive damages) follows shortly after the instructions are presented. In addition, some research has shown that, when compared with representative mock jurors from the community, student jurors tend to use a more rational style of information processing (McCabe, Krauss, & Lieberman, 2010), enhancing the likelihood that they would consider and correctly apply the standard of proof.

This suggests that, had the relevant standard of proof actually framed jurors’ decisions and served as an important criterion against which to evaluate the plaintiff’s claims, we should have seen differences in punitive liability decisions as a function of varying standards of proof, even in a simulation study. Yet there were no discernible differences in judgments when the standards were presented in legal language, and quantitative instructions did little to improve jurors’ ability to apply the standard. These findings highlight the need to develop clearer and more useful directives related to the burden of proof, consider alternative ways to inform jurors of this vital decisional benchmark, and review whether jurors have, in fact, evaluated the plaintiff’s claims against the imposed standard of proof.

We do not wish to overstate the implications of our findings, and certainly encourage further analysis of this issue to assess whether and to what extent standards of proof for punitive liability are outcome determinative. Yet research has already shown that laypeople are largely insensitive to differences in standards of proof for decisions in realms as varied as manslaughter (Dhami, 2008), child protection investigations (Levine, 1998), and mental retardation and mental illness in capital cases (Reardon et al., 2007), and we believe that our findings are consistent with the general consensus of these studies. Without careful attention to the standard of proof, jurors’ judgments in civil cases involving punitive damages may be influenced by their biases and commonsense interpretations of the law.

REFERENCES

Standards of proof


