Numeracy in the Jury Box:
Numerical Ability, Meaningful Anchors, and Damage Award Decision Making

Rebecca K. Helm

University of Exeter Law School

Valerie P. Hans

Cornell Law School

Valerie F. Reyna

Cornell University

Krystia Reed

Cornell Law School

Author Note

Preparation of this article was funded by National Science Foundation grant SES-1536238: "Quantitative Judgments in Law: Studies of Damage Award Decision Making" to Valerie P. Hans and Valerie F. Reyna. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation

Correspondence concerning this article should be addressed to Rebecca K. Helm, University of Exeter Law School, Exeter, UK, EX4 4RJ. Email: r.k.helm@exeter.ac.uk.

Abstract

This project employs an experimental design to test theoretical predictions regarding how numeracy can assist jurors in determining damage awards to compensate a plaintiff for pain and suffering, and how the use of meaningful numerical anchors may produce similar benefits. Mock jurors (N=345) reviewed a legal case and were asked to give a dollar award to compensate the plaintiff for pain and suffering. The presence and nature of a numerical anchor and the duration of pain and suffering were manipulated. Participants' numeracy was measured. Results provided support for predictions. Jurors higher in numeracy gave awards that more appropriately reflected the duration of pain and suffering, and showed less variability in awards. Similar benefits were obtained by exposing jurors to meaningful numerical anchors to help them contextualize dollar amounts. Thus, introducing meaningful anchors to jurors may provide similar benefits to numeracy, without the drawbacks associated with selecting only numerate jurors.

Keywords: jury decisions; damage awards; fuzzy-trace theory; numeracy; anchoring effects

Numeracy in the Jury Box:
Numerical Ability, Meaning, and Damage Award Decision Making

Damage awards are one illustration of a common type of decision the legal system requires of its human actors, the translation of qualitative assessments into quantitative judgments (Greene & Bornstein, 2003; Hans, Rachlinski, & Owens, 2011; Helm, Hans, & Reyna, 2017). Specifically, jurors must assess the pain and suffering of a claimant, which is not the same thing as economic damages and is non-numerical. Jurors must then determine a damage award, the number of dollars that will compensate the claimant for this pain and suffering. Damage award decisions are important; they are the major way that wrongfully injured plaintiffs are compensated (Robbennolt & Hans, 2016). Yet there are multiple challenges in determining damage awards, including the intangibility of many losses, the unpredictability of the lifelong impact of an injury, and the limited guidance provided by the courts (Diamond, Rose, Murphy, & Meixner, 2011).

The challenges are most daunting when jurors are asked to determine awards when there is no concrete referent, as in pain and suffering awards (Reyna, Hans, Corbin, Lin, & Royer, 2015). Consider, for example, a case in the state of Georgia where a jury was asked to decide on an award for a woman raped by an apartment complex's security guard (for more information, see Brumback, 2018). How is the jury to determine an award commensurate with the pain and suffering of a rape? Observers of the Georgia civil trial, and even the victim's lawyer, were surprised by the jury's one billion dollar award in that case, an award that dwarfs typical jury awards in the US (Cohen, 2009; Robbennolt & Hans, 2016). How jurors (and judges) arrive at money damage awards remains something of an enigma, albeit one with significant power to delight or dismay the parties in a lawsuit.

Most states allow trial attorneys in civil cases to make recommendations about appropriate damage awards, and research has found that such recommendations have substantial impact on damage award amounts (Campbell, Chao, & Robertson, 2017). Nonetheless, jurors report difficulty making such judgments and their damage award decisions are often criticized as being based on irrelevant factors and being highly variable (Bornstein & Greene, 2017).

One factor that has been shown to improve a person's ability to make numerical judgments in other contexts is numeracy: the ability to use and understand numbers (e.g., Peters, Västfjäll, Slovic, Mazzocco, & Dickert, 2006; for a review, see Reyna, Nelson, Han, & Dieckmann, 2009). It is therefore likely that numeracy will help jurors making damage award decisions (see Helm et al., 2017). This article presents the results from an experiment examining the role of a juror's numeracy in damage award decision making. Relying on a model of damage award decision-making proposed by Hans and Reyna (2011), we develop and test predictions about whether and how numeracy may help jurors to arrive at appropriate damage awards. In particular, the study explores whether numeracy can help jurors convert their qualitative understanding of the severity of an injury into a money damage award. We also test a theoretically-designed intervention using numerical anchors to explore how the benefits of numeracy might be emulated in jurors generally, thereby minimizing the need to select highly numerate jurors and the associated drawbacks including reduced juror representativeness (Helm et al., 2017) and potential increases in associated biases (e.g. Liberali et al., 2012).

Numeracy

Psychological research on the individual characteristic of numeracy confirms that in many situations, numeracy enhances a person's ability to handle numerical information in decision making (Fagerlin et al., 2007; Reyna et al., 2009). Some assessments of numeracy measure people's actual performance on numerical problems, which is referred to as "objective numeracy." Other tests ask people about their self-perceived ability and preference for numbers, such as preference for numbers over words, which is referred to "subjective numeracy."

Objective numeracy. Objective numeracy is a performance measure of people's numerical ability. Researchers have examined how objective numeracy is related to performance in tasks involving numbers and have found some evidence that higher levels of objective numeracy are associated with more accurate and less biased performance on such tasks. For example, in one study Peters and her colleagues found that people who were lower in objective numeracy were more affected by irrelevant factors and more influenced by framing effects than people who were higher in objective numeracy (Peters et al., 2006). In another project, people lower in objective numeracy were more affected by the way numbers were presented: People lower in objective numeracy estimated risk as lower when the risk information was presented as a percentage rather than a frequency, while people higher in numeracy showed no differences (Peters et al., 2011). That said, there is also evidence that numeracy can have paradoxical effects, such as when people who are better at performing computations nonetheless rate numerically inferior money gambles as more attractive than superior gambles (Reyna et al., 2009; Peters, Fennema, & Tiede, 2018; see also Kahan et al., 2012).

Although numeracy itself has been the subject of psychological study and research suggests generally beneficial effects of numeracy on numerical judgments, its effects on legal judgments in general and damage award decision making in particular have not been extensively examined (but see Bornstein, 2004; Rowell & Bregant, 2014). Thus far, neither objective nor subjective numeracy have been associated with higher or lower damage award amounts (Hans et al., 2018; Reyna et al., 2015). In a mock juror study examining the differential impact of meaningful and meaningless anchors on juror damage award decisions, Hans et al. (2018) measured objective numeracy of jurors and found that even though damage awards were unrelated to mock jurors' numeracy, more numerate mock jurors reported *more* difficulty in determining a money damage award for pain and suffering. The authors speculated that more numerate jurors were better able to recognize the challenge of attaching a dollar value to the intangible nature of an individual's pain.

Subjective numeracy. One problem with objective numeracy is that as a performance measure, it may be difficult to assess in applied settings. Few people would relish taking a math quiz as part of jury duty or a trip to the doctor's. However, in these settings, they might be more willing to answer questions about their preference for and comfort with numbers. Therefore, researchers have developed a subjective numeracy scale that asks people to subjectively rate their comfort and ability with numbers (Fagerlin, et al., 2007). Subjective numeracy incorporates both preference and ability components; those who prefer numerical information tend to be better able to process it, but preference and ability are nevertheless distinct constructs.

Subjective numeracy can serve as an approximate measure of objective numeracy. Typically, people who believe they are better at numeracy and express a preference for numerical information tend to be objectively better at numerical calculations. But it is not a perfect measure of actual numerical facility. Responses to objective and subjective tests are correlated, but far from perfectly (about .40 in community and student samples; Fagerlin et al., 2007; Liberali et al., 2012)). Preference relates to emotional reactions to numbers and to confidence in using them (Peters & Bjalkebring, 2015). Preferences for numerical information and one's own assessment of perceived ability are vulnerable to self-reporting biases such as the Dunning-Krueger overestimation-of-ability effect (Reyna et al., 2009). Objective numeracy, as a performance measure, is likely to be a more accurate measure, but the ease and simplicity of asking people for their subjective judgments of how good they are with numbers and how much they prefer to work with numerical information makes it a more feasible approach in applied contexts, such as the courtroom. Due to the distinct benefits of each measure, we examine the relationship between both constructs and damage award decisions in this study.

Psychological theory can provide insight into when and how numerical ability is likely to influence jurors and how legal procedures may be able to facilitate the benefits of numeracy in the jury box. This is important since there are drawbacks to selecting only more numerate jurors. Most importantly, because numeracy is associated with other individual characteristics such as gender and education, a strategy of selecting numerate jurors may compromise the representativeness of jurors making damage award decisions (Helm et al., 2017). In addition, evidence suggests that jurors with higher levels of numeracy are likely to be subject to predictable biases in decision-making (e.g. Liberali et al., 2012) and, as noted above, find the task of awarding a dollar damage award amount more difficult (Hans et al., 2018).

Fuzzy-Trace Theory and the Hans-Reyna Model

How might an individual fact finder's numeracy influence his or her determination of damages? First, let's consider the process by which fact finders interpret case evidence to arrive at a dollar damage award. The Hans-Reyna model of damage award decision-making draws on an established model of cognition called Fuzzy-Trace Theory (FTT; Reyna, 2012; Reyna & Brainerd, 2011). According to FTT, people encode two types of mental representation of information in parallel: verbatim and gist (Reyna et al., 2016). Verbatim representations capture the surface form of information (e.g., exact numbers), while gist representations capture its meaning or interpretation (based on context, and other factors known to affect meaning) (Reyna et al., 2009). When adults make judgments, research has shown that gist tends to be more important (this is known as a "fuzzy-processing preference" and has been demonstrated using mathematical models and experimental tasks, see Reyna, 2012). Hans and Reyna (2011) drew on this basic verbatim-gist distinction to propose a model of juror damage award decision making. When jurors are making decisions about damages, Hans and Reyna theorized, jurors encode their subjective impression of the injury that the plaintiff has suffered in terms of severity (e.g., nil, low, medium, high). In line with that gist of injury severity, jurors make an ordinal gist judgment regarding whether the damage award that is deserved is nil, low, medium, or high, and then identify a number that is, to them, nil, low, medium, or high, to match the deserved damages. This model has been tested in and supported by two reported experimental studies (e.g. Hans et al., 2018; Reyna et al., 2015).

In this model, numbers are primarily relevant in the last stage of damage award decision-making, when the ordinal gist judgment (i.e., nil damages are warranted, low damages are warranted, medium damages are warranted, high damages are warranted) is translated into a numerical amount Thus, since the numerical component of damage award decision-making arises at this last stage of converting an ordinal judgment into a numerical amount, we would expect any benefit of numeracy to arise at this stage. The decisions of more numerate jurors are not automatically assumed to be normatively better, particularly given demonstrated paradoxical effects of numeracy (Reyna et al., 2009). However, numeracy is more likely to result in decision-making that achieves horizontal and vertical equity in legal cases (i.e., awarding categorically similar amounts in similar cases and categorically different amounts in cases that differ in severity) because appropriate numbers are attached to mental representations of the gist of damages. This is because those with higher numeracy are likely to be able to convert the gist into a numerical figure more easily due to a greater ability with numbers, and might also have a better sense of the magnitude of awards, too. Technical numerical ability (i.e. advanced objective mathematical ability beyond very basic number sense; National Mathematics Advisory Panel, 2008) is unlikely to help with the conversion of gist to a damage award number (although damage award decisions that require the calculation of economic damages should benefit from jurors with greater numeracy; see Bajwa, 2014).

In the context of allocating damages for pain and suffering, when jurors are tasked with assigning a numerical value to a qualitative assessment of pain and suffering, any benefits of numeracy are likely to be due to a better understanding of numbers per se—such as what number constitutes a "large" number of dollars. Through having greater familiarity and facility with numbers, jurors with greater numeracy are more likely to know what is an objectively "large" amount or "small" amount. This understanding should assist jurors in appropriately allocating numerical figures to ordinal gist judgments. This is independent of knowledge of awards in prior legal cases, which is unlikely to be influenced by numeracy.

Predicted Benefits of Numeracy

We use three concepts, validity, reliability, and verbatim-gist correspondence, to provide insight into the appropriate allocation of numerical figures to pain and suffering. Validity refers to the accuracy of a judgment, or whether a judgment reflects what it is supposed to measure (in the present experiment, whether juror judgments adequately capture the distinction between a short-term and long-term injury). Reliability refers to the extent to which judgments are consistent (in this context whether there is some consistency, and thus less variability, in the numerical amounts allocated by jurors). Verbatim-gist correspondence refers to the extent to which numerical awards covary appropriately with ordinal gist judgments (e.g. jurors who rate their awards as “high” awarding higher numerical amounts than jurors who rate their awards as “low”). If, as predicted by the Hans-Reyna model, numeracy helps jurors where it provides a basic number sense that enables the appropriate conversion of ordinal gist into a numerical award amount, then numeracy is likely to benefit damage award decisions in each of these three ways.

First, it should lead to damage award decisions that vary appropriately with the severity of an injury and duration of pain and suffering in a case, since numbers are likely to increase appropriately as the extent of pain and suffering increases (i.e., have greater validity). Specifically, when jurors are more numerate, we expect jurors hearing high severity cases to more consistently award more than jurors hearing low severity cases. Importantly, this is the case even though the actual jurors hearing the cases are different because jurors will typically only give an award in one case (although note arguments for providing guidance based on comparable cases, Bavli, 2017). This is because, through having a better sense of what is a “high” number and what is a “low” number, more numerate jurors will be more likely to select a number that is objectively “high” once they have identified a case as high severity (or “low” once they have identified a case as “low” severity). One important way this is done is through a better understanding of what is a high or low number, rather than simply through comparing two cases and awarding more in the case identified as more severe.

Second, it should produce damage award decisions that are less variable, since judgments should converge more reliably on an appropriate number to allocate to an ordinal gist (i.e. have greater reliability). This is not because that there is necessarily a normatively correct damage award in a case, but simply that when jurors all have an understanding of the meaning of a monetary amount in context (whether it is large or small compared to other relevant metrics) this reduces one source of variability - variability based on an understanding of the relative value of a monetary amount. Reducing this source of variation will reduce overall variability, despite the fact that other sources of variability (e.g. variability due to differing interpretations of the severity of pain and suffering) are likely to remain. Reducing variability resulting from a lack of understanding of numbers is particularly important in ensuring horizontal equity, meaning that cases involving the same level of pain and suffering are treated alike (Saunders, 2006).

Third, it should generate damage award amounts that more closely correspond with ordinal gist (i.e., have greater verbatim-gist correspondence). That is, the damage award amount should increase when ordinal gist judgments increase from nil to low to medium to high.

Replicating Predicted Benefits of Numeracy

This insight into potential beneficial effects of numeracy can also facilitate understanding of how any benefits of numeracy might be generated without selecting only numerate jurors. This is important since, as described above, there are drawbacks to selecting only numerate jurors.

If the benefits of numeracy are in providing insight into the meaning of numbers that facilitate allocating an appropriate number to an ordinal gist judgment, we should be able to achieve some of the effects of numeracy through exposing jurors to meaningful anchor numbers. Although all numbers can convey some meaning, we use the term meaningful to describe relatively more meaningful anchors and meaningless to describe relatively less meaningful anchors, in line with previous work (see Reyna et al., 2015). This is because such anchors provide insight into the value of a number by providing a relevant frame of reference (Diamond et al., 2011; Sugden, Zheng, & Zizzo, 2013). For example, if a number equates to the median *lifetime* income, this is likely to be seen as a large number since it is the amount of money a person might expect to make over an entire lifetime. So, if jurors are told the median lifetime income is $1.5 million, this gives them insight into the meaning of that number, as a large number, which can help with allocating a specific figure to the ordinal gist judgment.

The effect of numerical anchors, which render subsequent judgments biased towards them (see Tversky & Kahneman, 1974), have been demonstrated in many domains including in the courtroom (Furnham & Boo, 2010). For example, both criminal sentences (Englich, Mussweiler, & Strack, 2006; Enough & Mussweiler, 2001) and civil damage awards (Hastie, Schkade, & Payne, 1999; Marti & Wissler, 2000) have been shown to be higher when attorneys' initial requests are higher. Traditionally, anchoring effects have been shown to exist when the provided anchor is irrelevant to a decision (e.g. Brewer, Chapman, Schwartz, & Bergus, 2007). However, the Hans-Reyna model suggests that when anchors provide meaningful information to decision-makers, they are likely to have greater influence on decision-making through helping impart meaning and context to numbers and facilitate the allocation of a number to a gist judgment. This is supported by literature showing stronger effects of anchoring in terms of consumer purchases when the anchor used is a plausible price (Sugden et al., 2013), and greater reliance on relevant compared to irrelevant anchors when multiple anchors are presented (Caputo, 2014; Whyte & Sebenius, 1997). In the courtroom, jurors have been found to rely less on damage award requests that are not supported by evidence (Diamond et al., 2011), and meaningful anchors have been shown to have a greater influence on damage award decision-making than meaningless anchors (Reyna et al., 2015; Hans et al., 2018).

The predictions of the Hans-Reyna model, and supporting research, suggest that through conveying important contextual information, anchors, like numeracy, should help jurors convert ordinal gist judgments into dollar award amounts. We would expect a weaker effect, or no effect, with meaningless anchors since they do not convey meaningful information about numbers. In sum, since numeracy and meaningful anchors are both helpful in providing insight into the meaning of numbers, presenting jurors with meaningful anchors may provide a way to replicate the benefits of numeracy.

The Current Study

In this study, we examine hypotheses regarding numeracy and anchors using an experimental design in which participants were asked to act as mock jurors and make judgments to compensate a plaintiff for pain and suffering. We investigate two classes of questions. First, how does a participant's numeracy influence his or her damage awards for pain and suffering? Second, can some of the anticipated beneficial effects of numeracy be achieved through exposing jurors to meaningful numerical anchors?

We tested the following predictions informed by psychological theory and previous research discussed above. Note, our numeracy predictions relate specifically to objective numeracy as a more accurate construct in representing numerical ability. However, since subjective numeracy does reflect objective numeracy (although imperfectly) we expect similar patterns when using subjective numeracy, although these may be less strong due to other biases reflected in the measure. We nevertheless test predictions with subjective numeracy as well as objective numeracy measures due to the potential advantages of subjective numeracy measures in applied contexts such as jury selection.

Hypothesis 1: More numerate jurors will show greater *validity* in their awards. Specifically, compared to less numerate jurors, more numerate jurors will be more likely to reach damage award decisions for pain and suffering that vary appropriately with the duration of pain and suffering. Jurors seeing a long-term case of pain and suffering will award more than jurors seeing a short-term case of pain and suffering, even though jurors themselves do not view both cases and thus cannot compare them directly.

Hypothesis 2: Compared to less numerate jurors, more numerate jurors will show greater *reliability* in their awards. Specifically, more numerate jurors will show *less variability* in their damage award amounts.

Hypothesis 3: Compared to less numerate jurors, more numerate jurors will show *greater verbatim-gist correspondence* between their ordinal judgments of the deserved damages (nil, low, medium, high) and dollar award amounts.

Hypothesis 4: Based on previous research findings (Hans et al., 2018), we expect that despite their advantages in handling numerical information, more numerate jurors will *report finding it more difficult* to reach a damage award to compensate for pain and suffering than their less numerate counterparts.

Hypothesis 5: The three identified benefits of numeracy will be emulated in all jurors by providing meaningful anchors to assist in their decision making. These benefits will not be present (or will at least be significantly reduced) when an anchor provided is meaningless.

Method

Participants

A total of 345 undergraduate students (78.3% female, M_{Age} = 19.81, *SD* = 1.28) took part in the experiment. We selected sample size based on previous work testing the Hans-Reyna model (e.g. Reyna et al., 2015; Hans et al., 2018). Participants received course credit in Psychology, Human Development, or undergraduate Law courses at a northeastern university. The group of participants was racially diverse, with 54% of participants identifying themselves as White, 32% as Asian, 5.2% as Black/African American, and 8.1% as mixed ethnicity/other. Additionally, 11.1% were of Hispanic, Latino, or Spanish origin. Participants were all over 18 and able to communicate effectively in English.

Materials and Procedure

The university's Institutional Review Board approved the study. The experiment was a 3 (anchor type: no anchor, meaningless anchor, meaningful anchor) X 2 (case: short-term injury case, *Monroe v. Rumson*, or long-term injury case, *Jeansonne v. Landau*) between-subjects experiment. We identified the *Monroe* case as a case in which lower damages were warranted due to a relatively short duration of the plaintiff's pain and suffering (two years), compared to the much longer duration of similar pain and suffering in the *Jeansonne* case (anticipated for the rest of the plaintiff's life). Previous research with these two cases also showed that on average larger pain and suffering awards were given in the *Jeansonne* case than in the *Monroe* case (Reyna et al., 2015). The experiment was conducted online. After agreeing to participate, the participants were randomly assigned to one of the six experimental conditions.

The study began by presenting a trial summary of, depending on experimental condition, the auto accident case of *Monroe v. Rumson* or *Jeansonne v. Landau.* Both cases were negligence cases involving a personal injury, had one plaintiff and a noncorporate defendant, and involved a plaintiff who bore no responsibility for the accident. Both cases had been used in previous research on damage award decision making (see Greene, Johns, & Smith, 2001; Hans et al., 2018; McAuliff & Bornstein, 2010; Reyna et al., 2015).

After reading the case facts, subjects read instructions from the judge. They were told that the defendant had been found liable for the plaintiff's injuries and that economic damages had been paid to the plaintiff. The instructions explained that their task was to arrive at a dollar award amount to compensate the plaintiff for pain and suffering. The instructions directed jurors to use their best judgment in coming up with an award amount, observing that "the law prescribes no definite standard or method of mathematical calculation by which to fix reasonable compensation for pain and suffering…" They were instructed that "punitive damages intended to punish the defendant should not be taken into account when determining the plaintiff's award."

The anchor manipulation was included in the first question that participants answered. Participants in the anchor conditions were asked whether they thought that the award amount in the case they saw should be above or below $1.5 million. In the meaningful anchor (MF) condition, they were told this was the median lifetime income. In the meaningless anchor (ML) condition, they were told this was the cost of courtroom renovations. In contrast, participants in the no-anchor condition received no such question.

All participants were then asked to give an award amount for the pain and suffering in the case. After being asked for their award amount, participants were asked to give an ordinal gist judgment of the award amount. Specifically, they were asked to state whether they considered their award amount to be nil (basically nothing), low, medium, or high. They also answered a question about how difficult they found it to reach a damage award decision on a scale from 1 (not at all difficult) to 7 (extremely difficult), and some other questions about their perceptions of the case. All the participants were then asked whether they agreed that the median lifetime income and cost of courtroom renovations were meaningful when determining the damage award. They answered each of these questions on a scale from 1 (strongly disagree) to 7 (strongly agree).

Participants then answered additional questions, including an expanded objective numeracy scale, a subjective numeracy scale, several tasks that are not part of this paper, and demographic items. The 15-item expanded objective numeracy scale is composed of the 3-item general numeracy scale (with minor variations) created by Schwartz, Woloshin, Black and Welch (1997), an additional 8 items proposed by Lipkus, Samsa and Rimer (2001), and four items from Peters et. al. (2007), which test familiarity with simple arithmetical operations such as multiplication, basic probability and related ratio concepts, and ability to keep track of class-inclusion relations (Liberali et al., 2012). The 8-item subjective numeracy scale (Fagerlin et al., 2007) is a measure of participants' perceptions of their own numeracy and is made up of two subscales: numerical ability (measuring perceived ability with numbers, e.g., How good are you at working with fractions?) and numerical preference (measuring preference for receiving information in numerical form, e.g., When people tell you the chance of something happening, do you prefer that they use words ["it rarely happens"] or numbers ["a 1% chance"]?).

Results

Initial Descriptive Statistics

Instructions in our cases indicated that a prior jury already had established the defendant's liability for damages. Therefore, in line with the approach taken by Wissler, Hart, and Saks (1999) and Reyna et al. (2015), the five respondents who gave an award of zero were dropped from analyses involving award amounts, leaving a total of 340 participants. Two participants did not complete our numeracy questions, leaving a total of 338 participants in analyses involving numeracy. Overall, four participants classified their awards as nil, 60 as low, 231 as medium, and 45 as high. Mean award amounts in each of our experimental conditions, split by high and low objective numeracy, are presented in Table 1.

Descriptive statistics for the numeracy scales and numeracy median-splits are presented in Table 2. We used our median split to separate participants into “high numeracy” and “low numeracy” groups, although note that these groups were comparative in nature (based on the median split) so that those with scores above the median were classified as “high” numeracy and those with scores below the median were classified as “low” numeracy. The correlation between objective and subjective numeracy was significant, and moderately strong in a positive direction (such that higher objective numeracy scores were associated with higher overall subjective numeracy scores (r = .41, p < .001)) (objective numeracy and subjective numeracy ability subscale r = .35, p < .001; objective numeracy and subjective numeracy preference subscale r = .38, p < .001).

In order to ensure that our meaning manipulation had worked and participants found our meaningful anchor more meaningful than our meaningless anchor, we conducted a paired samples *t*-test that compared meaningfulness ratings for the median lifetime income and the cost of courtroom renovations. As predicted, participants evaluated the median lifetime income as more meaningful in determining a damage award than the cost of courtroom renovations (M_{income} = 4.07, SD = 1.64, $M_{\text{renovations}}$ = 2.33, SD = 1.49, p = .002, d = .86).

Validity of Awards (Hypotheses 1 and 5)

We predicted that among jurors with higher numeracy levels (but not jurors with lower numeracy levels), jurors seeing a long-term case would award more than jurors seeing a short-term case would, despite each only seeing one type of case. To test this prediction, we analyzed the relationships between the experimental variables and our numeracy variables using two approaches: an analysis of variance (ANOVA) with the high and low numeracy groups, and a follow-up regression with the full-scale numeracy scores to check the robustness of the results with the dichotomous numeracy grouping.[1] In our analyses, the dependent measure was the log award amount (lg10). Employing log award amounts rather than raw award amounts allowed us to correct skewing (award amount values had a skewness of 5.20, *SE* = .13; log award amount values had a skewness of -.70, *SE* =.13).

Objective Numeracy

ANOVA Analysis. In our ANOVA**,** between-subjects factors were case (*Monroe*, the short-term injury case, *Jeansonne*, the long-term injury case), anchor condition (no anchor, meaningless anchor, meaningful anchor), and objective numeracy group of the participant (low numeracy, high numeracy). The dependent measure was the log award amount. The ANOVA tested for the significance of all factors, and all possible interactions between them.

The ANOVA revealed a main effect of the participant's objective numeracy group. Participants in the high objective numeracy group gave higher awards overall than participants in the low objective numeracy group ($M_{highnumeracy}$ = 5.03, *SE* = .09, $M_{lownumeracy}$ = 4.70, *SE* = .09, $F(1, 249) = 6.28$, $p = .013$, $\eta_p^2 = .03$). There was also a main effect of case such that participants

---

[1] We focus primarily on the ANOVA, since the use of high and low numeracy groups allows for a fully-crossed design and the results are more easily interpreted in light of our hypotheses. This is particularly true when considering hypothesis 1, since although we expect higher awards in a long-term case (compared to a short-term case) where jurors have a certain level of numeracy, we do not expect that the difference in awards between short-term and long-term cases will continue to increase as numeracy increases. That is, we do not expect very high numeracy jurors to give an extraordinarily low amount in a short-term case and an extraordinarily high award in a long-term case.

who saw the long-term injury case (*Jeansonne*) gave larger awards than participants who saw the short-term injury case (*Monroe*) ($M_{long\text{-}term\ injury}$ = 5.01, SE = .09, $M_{short\text{-}term\ injury}$ = 4.72, SE = .09, $F(1, 249) = 4.83$, $p = .03$, $\eta_p^2 = .02$).

As predicted, results showed that in the high objective numeracy group (but not the low numeracy group) jurors who saw a long-term injury case awarded more than jurors who saw a short-term injury case. Specifically, there was a statistically significant two-way interaction between the participant's numeracy group and the case the participant decided ($F(1, 249) = 5.01$, $p = .03$, $\eta_p^2 = .02$). As Figure 1 shows, participants in the low objective numeracy group did not differ significantly in the awards they gave based on the case that they saw ($M_{long\text{-}term\ injury}$ = 4.70, SE = .14, $M_{short\text{-}term\ injury}$ = 4.71, SE = .15, $F(1, 133) = .001$, $p = .98$, $\eta_p^2 = .00$) whereas participants in the high numeracy group gave significantly more when they had seen the long-term injury case ($M_{long\text{-}term\ injury}$ = 5.31, SE = .11, $M_{short\text{-}term\ injury}$ = 4.74, SE = .11, $F(1, 116)$ = 13.89, $p < .001$, $\eta_p^2 = .11$).[2]

As predicted, among participants who saw a meaningful anchor (but not among those who saw a meaningless anchor or no anchor) jurors who saw a long-term injury case awarded more than jurors who saw a short-term injury case. Specifically, there was a statistically significant two-way interaction between the anchor condition and case ($F(2, 249) = 5.24$, p = .006, $\eta_p^2 = .04$) (see Figure 2). Participants who saw no anchor or a meaningless anchor did not differ significantly in the awards they gave based on the case they saw. As Figure 2 shows, participants in the no anchor group actually awarded slightly, although nonsignificantly, more on

[2] Note that although this interaction was not moderated by any higher order interactions our descriptive statistics for each group (see Table 1) show that even participants in our high numeracy group did not award more in the long-term case without seeing an anchor (although they still awarded more than the low numeracy jurors in the long-term case and less than the low numeracy jurors in the short-term case). This is considered further in our discussion.

average in the short-term injury case than in the long-term injury case (No Anchor: $M_{long\text{-}term\ injury}$ = 4.23, *SE* = .14, $M_{short\text{-}term\ injury}$ = 4.46, *SE* = .13, $F(1, 85) = 1.41$, $p = .24$, $\eta_p^2 = .02$). Corresponding means for the Meaningless Anchor were: $M_{long\text{-}term\ injury}$ = 5.33, *SE* = .18, $M_{short\text{-}term\ injury}$ = 5.05, *SE* = .16, $F(1, 83) = 1.37$, $p = .25$, $\eta_p^2 = .02$). In contrast, participants who saw a meaningful anchor gave significantly more when they had seen the long-term injury case ($M_{long\text{-}term\ injury}$ = 5.45, *SE* = .16, $M_{short\text{-}term\ injury}$ = 4.66, *SE* = .18, $F(1, 81) = 11.39$, $p = .001$, $\eta_p^2 = .12$). There were no other significant main effects or interactions in this analysis.

Regression Analysis. We conducted a comparable regression predicting log award amount using case, anchor type (coded from least meaningful to most meaningful: no anchor, meaningless anchor, meaningful anchor), numeracy score, the interaction between numeracy score and case, and the interaction between anchor type and case, in order to examine our predictions involving numeracy using the entire numeracy scales (as opposed to dichotomized variables). Regression results generally were in line with the results from the ANOVA analysis. Specifically, the same interactions between objective numeracy and case and anchor type and case were present. The interaction between anchor type and case was again significant ($B = .43$, $SE = .13$, $\beta = .68$, $t = -3.24$, $p = .001$) and the interaction between objective numeracy score and case was marginally significant ($B = .09$, $SE = .05$, $\beta = .58$, $t = 1.89$, $p = .06$). Full regression results are presented in Table S1.

Subjective Numeracy Overall Scale. Next, we repeated the analyses described above, replacing objective numeracy with subjective numeracy. The ANOVA revealed main effects of case and numeracy that mirrored our objective numeracy analysis. Specifically, participants in the high subjective numeracy group gave higher awards overall than participants in the low subjective numeracy group ($M_{\text{highnumeracy}}$ = 5.07, *SE* = .08, $M_{\text{lownumeracy}}$ = 4.67, *SE* = .08, $F(1, 311)$

= 12.70, $p < .001$, $\eta_p^2 = .04$); participants who saw the long-term injury case (*Jeansonne*) gave larger awards than participants who saw the short-term injury case (*Monroe*) ($M_{long\text{-}term\ injury}$ = 5.01, *SE* = .08, $M_{short\text{-}term\ injury}$ = 4.73, *SE* = .08, $F(1, 311) = 6.14$, $p = .01$, $\eta_p^2 = .02$). This analysis also replicated our significant interaction between case and anchor type ($F(2, 311) = 5.67$, p = .004, $\eta_p^2 = .04$) (see Figure S1).

Unlike the ANOVA with objective numeracy, the analysis with subjective numeracy did not produce the predicted interaction between numeracy and case. In other words, those who were identified as having high or low subjective numeracy did not differentiate between the cases in the way that the high and low objective numeracy participants did.

Additionally, the ANOVA with subjective numeracy group revealed a main effect of anchor type ($F(2, 311) = 23.93$, $p < .001$, $\eta_p^2 = .13$). Follow-up pairwise comparisons revealed that awards were significantly lower in the no anchor condition ($M = 4.33$, SE =.10) than in the meaningless anchor condition ($M = 5.23$, $SE = .10$, $p < .001$) and the meaningful anchor condition ($M = 5.05$, SE =.10, $p < .001$). The meaningful and meaningless anchor conditions were not significantly different from one another ($p = .18$).

In the follow-up regression, as in the ANOVA, the interaction between subjective numeracy score and case was not significant ($B = .056$, $SE = .121$, $\beta = .125$, $t = .459$, $p = .646$). Full regression results are presented in Table S2.

Subjective Numeracy Preference Subscale. Our ANOVA using the preference subscale, measuring participants' preference for numbers over words, revealed a main effect of case and a main effect of anchor type, replicating the patterns shown in the overall subjective numeracy analysis. This analysis also revealed a main effect of subjective numeracy preference ($F(1, 325) = 3.90$, $p = .05$, $\eta_p^2 = .01$) such that participants with high subjective numeracy preference gave significantly more than those with low subjective numeracy preference (*M* = 5.04, *SE* = .09, compared to *M* = 4.79, *SE* = .08). In this analysis the predicted interaction between case and subjective numeracy preference was not significant.

Similarly, in the follow-up regression the interaction between subjective numeracy preference and case was not a significant predictor ($B = .048$, $SE = .029$, $\beta = .443$, $t = 1.618$, $p = .107$). Full regression results are presented in Table S3.

Subjective Numeracy Ability Subscale. Our ANOVA using the ability subscale, measuring participants' perceptions of their numerical ability, revealed a main effect of anchor type, a main effect of case, a main effect of numeracy, and a significant interaction between anchor type and case, replicating the patterns shown in our overall subjective numeracy analysis. However, no other effects were significant. In particular, the predicted interaction between the case and the participant's subjective numeracy ability – where in the high numeracy group (but not the low numeracy group) jurors who saw a long-term injury were expected to award more than jurors who saw a short-term injury - was not significant. In the follow-up regression, as in the ANOVA, the interaction between subjective numeracy ability score and case was not significant ($B = -.007$, $SE = .024$, $\beta = .125$, $t = -.066$, $p = .775$). Full regression results are presented in Table S4.

Reliability of Awards (Hypotheses 2 and 5)

We predicted that jurors high in numeracy and jurors who have seen a meaningful anchor should show less award variability. To test our numeracy prediction, we conducted an analysis of variance with the absolute deviation from condition mean as the dependent variable, a measure that captured the average variability in each condition. To find this, we calculated the absolute distance between each award and the mean award in the relevant experimental condition (the AMD). One challenge for award variability analyses is that there is naturally greater variability as the award amounts increase, so we included the award amount as a covariate in the analysis to control for this.

Objective Numeracy

ANOVA Analysis. In this analysis, between-subjects predictors were the case (short term injury, long term injury), the anchor condition (no anchor, meaningless anchor, meaningful anchor), and the objective numeracy group of the participant (low numeracy, high numeracy). Award amount was entered as a covariate. These factors were used to predict AMD.

First, and as expected, our ANOVA analysis revealed a main effect of the covariate, the participant's award amount in the case. As participants' awards increased, AMD increased ($F(1, 252) = 898.71, p < .001, \eta_p^2 = .78$). The analysis also revealed a marginally significant main effect of objective numeracy group, such that participants in the high objective numeracy group showed less deviation from their condition mean than participants in the low objective numeracy group ($M_{highnumeracy}$ = \$460,496.26, SE = \$36,585.26, $M_{lownumeracy}$ = \$556,084.82, SE = \$34,768.68, $F(1, 252) = 3.57, p = .06, \eta_p^2 =. 01$). Finally, the analysis revealed a significant main effect of anchor type ($F(2, 252) = 5.175, p = .006, \eta_p^2 = .04$). Participants showed the highest AMD in the meaningless anchor condition (M = \$623,769.24, SD = \$44,196.50), followed by the no anchor condition (M = \$465,727.52, SD = \$43,448.96), and the lowest AMD in the meaningful anchor condition (M = \$444,703.10, SD = \$44,703.10). The AMD in the meaningful anchor condition and the no anchor condition were both significantly less than the AMD in the meaningless condition ($p = .003$ and $p = .01$, respectively). The AMD in the meaningful anchor condition and the no anchor condition did not differ significantly ($p = .63$).

Regression Analysis. We followed up on ANOVA results using a linear regression predicting AMD using the entire numeracy scales (as opposed to the median split). Predictors were case (short term injury, long term injury), anchor type (coded as a continuous variable from least meaningful to most meaningful, i.e. no anchor, meaningless anchor, meaningful anchor), numeracy score, and award amount. In this regression, objective numeracy score was a significant predictor. As predicted, those with higher numeracy showed less AMD (B = -26,573.85, SE = 9,142,89, β = -.07, t = -2.91, p = .004). This analysis also revealed a main effect of award amount, such that those who gave higher awards showed greater AMD (B = .75, SE = .02, β = .90, t = 38.03, $p < .001$). There were no other significant effects in this analysis.

Subjective Numeracy Overall Scale. Next, we repeated the ANOVA described above, replacing objective numeracy with subjective numeracy. This analysis replicated the main effect of anchor type and main effect of award amount revealed in our objective numeracy analysis. However, the prediction that people high in numeracy would show lower award variability was not confirmed using this scale. No other effects were significant.

We also performed a regression analysis using the full subjective numeracy scale (as opposed to the median split), as with objective numeracy above. The only significant predictor in this analysis was award amount. Again, those who gave higher awards showed greater AMD (B = .75, SE = .02, β = .90, t = 37.26, $p < .001$). Subjective numeracy was not a significant predictor of AMD in this analysis.

Subjective Numeracy Subscales. The results of our ANOVA and regression analyses using the subjective numeracy subscales replicated the corresponding analyses using the overall subjective numeracy scale. Neither of the subscales were significant predictors in either analysis.

Verbatim-gist correspondence of Awards (Hypotheses 3 and 5)

We predicted that jurors high in objective numeracy and jurors who have seen a meaningful anchor should show a stronger association between ordinal gist judgments of their award amount and dollar award amounts, compared to participants with lower numeracy and in the no anchor or meaningless anchor conditions. To assess this while minimizing the number of predictors in regressions, we first ran a regression focused on testing our numeracy prediction. This regression analysis attempted to predict the log award amount with numeracy score, anchor type (coded from no anchor to meaningful anchor), case, ordinal gist judgment, and the interaction between numeracy score and ordinal gist judgment as predictors. We expected a significant interaction between numeracy score and ordinal gist judgments, such that as juror numeracy increased, the relationship between ordinal gist judgment and dollar award would get stronger.

Objective Numeracy. The results of this regression are reported in Table 3. The regression revealed a significant main effect of ordinal gist judgment (as ordinal gist judgment moved from nil to low to medium to high, the log award amount increased), a main effect of case (awards were higher in the long-term injury case than in the short-term injury case), a main effect of anchor type (as anchors increased in meaning log award amounts increased), and a main effect of objective numeracy score (as objective numeracy score increased, the log award amount increased). The predicted interaction between numeracy and ordinal gist judgment was marginally significant (p = .08, see Table 3).

We followed up on the marginally significant interaction between numeracy and ordinal gist judgment by looking at the correlations between ordinal gist judgment and log award amount in our low objective numeracy group and our high objective numeracy group. In both the high and low numeracy groups there were significant positive correlations between ordinal gist and dollar awards such that dollar awards increased as ordinal judgments moved from nil to low to medium to high. This correlation was qualitatively stronger in the high objective numeracy group ($r = .49$, $p < .001$) than in the low objective numeracy group, $r = .38$, $p < .001$), but analyses using the Fisher *r*-to-*z* transformation did not indicate a statistically significant difference between these correlations ($z = 1.1$, $p = .14$).

Subjective Numeracy. We repeated these analyses using subjective numeracy in place of objective numeracy (and, in subsequent analyses, the subjective numeracy ability subscale and subjective numeracy preference subscale). When using either overall subjective numeracy, subjective numeracy ability, or subjective numeracy preference, the interaction between numeracy and ordinal judgment was not significant (although note that when using the preference subscale, the interaction just missed significance [$B = -.04$, $SE = .02$, $\beta = -.58$, t = 01.80, $p = .07$]). Full regression results are reported in Tables S5 (overall subjective numeracy), S6 (subjective numeracy preference), and S7 (subjective numeracy ability).

Anchor Type Predictions. We then ran a regression focused on testing our anchor type prediction. Specifically, we ran a regression predicting log award amount, with anchor type (coded as a continuous variable from no anchor to meaningful anchor), case, ordinal gist judgment, and the interaction between anchor type and ordinal gist judgment as predictors. We expected a significant interaction between anchor and ordinal gist judgments, such that as the meaningfulness of anchor increased (from no anchor to meaningless anchor to meaningful anchor) the relationship between ordinal gist judgment and dollar award amount would get stronger. The results of this regression are reported in Table 4. As in the previous regression, this regression revealed a significant main effect of ordinal gist judgment (as ordinal gist judgment moved from nil to low to medium to high, log award amount increased). The predicted interaction between anchor type and ordinal gist judgment missed significance ($B = -.18$, $SE =$

.10, β = -.44, *t* = -1.75, *p* = .081). We followed up on this result by looking at correlations between ordinal gist judgment and log award amount in each of our anchor type groups. This correlation was non-significant in our no anchor group (*r* = .10, *p* = .298) but was significant in our meaningless anchor group (*r* = .23, *p* = .015) and our meaningful anchor group (*r* = .26, *p* = .006).

Numeracy, Anchor, and Difficulty Reaching an Award Amount (Hypotheses 4 and 5)

As predicted and in line with previous research, a significant relationship was observed between a participant's objective numeracy and the perceived difficulty of determining money damages for pain and suffering. As a participant's objective numeracy score increased, the participant also became more likely to report that it was difficult to determine damages (*r*(346) = .18, *p* = .001). This was also confirmed using *t*-test comparisons between our dichotomous groupings. Those in the high objective numeracy group reported that they found it significantly more difficult to determine a damage award amount for pain and suffering than those in the low objective numeracy group ($M_{\text{highnumeracy}}$ = 5.60, *SD* =1.24, $M_{\text{lownumeracy}}$ = 5.27, *SD* = 1.26, *F*(1, 267) = 4.44, *p* = .04, η_p^2 = .02). This relationship was not found for subjective numeracy. Correlational analyses and pairwise comparisons revealed no significant relationship between subjective numeracy (or its subscales) and perceived difficulty reaching an award for pain and suffering. Results of correlational analyses showed no significant relationship between the meaningfulness of an anchor (from no anchor to meaningless to meaningful) and difficulty reaching an award. T-tests showed a similar result. Specifically, there was no significant difference in perceived difficulty of determining damages between our three anchor groups.

Discussion

In this study, we tested predictions of FTT regarding the influence of juror numeracy and meaningful anchors on pain and suffering awards. We hypothesized that more numerate jurors and those who saw meaningful anchors would have three advantages when deciding on dollar award amounts for pain and suffering. Specifically, more numerate jurors and those exposed to meaningful anchors would show greater validity and reliability in their awards and greater verbatim-gist correspondence between ordinal and dollar award judgments. Results provide support for these benefits of objective numeracy and meaningful numerical anchors. These results differ from previous work on anchoring, which has shown an effect of anchors either pulling award amounts up or down, through showing a beneficial effect of anchors in helping jurors contextualize award amounts.

Objective Numeracy and Damage Award Decisions (Hypotheses 1-4)

First, more numerate jurors were predicted to show greater validity in their awards, demonstrated by the fact that in the high numeracy group (but not the low numeracy group), jurors who saw a long-term injury would award more than jurors who saw a short-term injury (*Hypothesis 1*). Results using an objective numeracy measure were consistent with this hypothesis. Specifically, as predicted, participants in a high objective numeracy group were able to appropriately discriminate between the short-term and long-term case in their pain and suffering awards, awarding more in the long-term case, while participants in the low objective numeracy group were not. Second, more numerate jurors were predicted to show greater *reliability* in their awards, demonstrated by less variability in award amounts (*Hypothesis 2*). Again, results using an objective numeracy measure were consistent with this hypothesis. Dollar award amounts were less variable in the high numeracy group compared to the low numeracy group. Third, more numerate jurors were predicted to show stronger verbatim-gist correspondence between ordinal gist judgments and dollar award amounts (*Hypothesis 3*). We found some modest support for this prediction in a marginally significant interaction between numeracy and ordinal gist judgment in predicting award amounts. Although this result just missed significance, it is consistent with an implication of the Hans-Reyna model that it is especially difficult for jurors lower in numeracy to match numerical judgments—dollars awarded—with their gist intuitions about the magnitude of damages. Despite the observed benefits of numeracy, participants with greater objective numeracy reported that they found it more difficult to determine a damage award amount for pain and suffering (*Hypothesis 4*). This likely reflects the better appreciation of those with higher objective numeracy that determining damages for such intangibles is inherently difficult (Hans et al., 2018).

Replicating Effects Using a Subjective Numeracy Measure

We also examined the effects of numeracy when measured using a subjective numeracy scale, since subjective numeracy questions would be easier to ask jurors during voir dire. However, unsurprisingly because of self-report biases and other sources of measurement error in assessing numeracy per se, the subjective numeracy measure fell short in replicating the effects observed using our objective numeracy measure. Subjective numeracy preferences, confidence, or assumed associated emotions related to processing numbers did not reflect improved numerical judgments.

Emulating Benefits Using Meaningful Numerical Anchors (Hypothesis 5)

As predicted, we were able to emulate benefits similar to those associated with higher objective numeracy in jurors with a range of numeracy scores by providing them with meaningful anchors to assist in contextualizing dollar amounts. In the group that had seen a meaningful anchor prior to making an award (but not in the other groups), jurors who saw a long-term injury awarded more than jurors who saw a short-term injury, as indicated by a significant interaction between anchor type and case in predicting award amounts. Furthermore, awards to some extent showed greater verbatim-gist correspondence between ordinal gist judgments and dollar award amounts, as indicated by a significant correlation between ordinal gist judgments and dollar award amounts in our meaningful anchor group but not in our no anchor group. In terms of variability, our results showed that there was less variability in decisions when a meaningful anchor was given or no anchor was given, compared to when a meaningless anchor was given. Results partially confirmed predictions since a meaningful anchor was associated with less variability than a meaningless anchor, but the similar variability of the meaningful and no anchor conditions was not predicted. Future research should follow up on this result.

Limitations and Future Directions

This study provides initial support for predictions based on FTT regarding the benefits of juror numeracy and providing meaningful numerical anchors to jurors. Implementing a relatively simple experimental and observational design allowed us to test theoretical ideas in a highly controlled setting. However, our findings must be interpreted in light of some limitations.

First, our sample was limited to college students. Although prior research suggests that college students and other groups respond comparatively to mock trial stimuli (Bornstein, 2017; Bornstein et al., 2017), our sample's levels of education are likely to be associated with high levels of numeracy and less of a range in numeracy levels. However, both mean subjective and objective numeracy and standard deviations in these measures were comparable to unpublished results obtained in a community sample completing the same numeracy measures described in this paper (available from the fourth author) ($\text{ObjectiveNumeracy}_{\text{Community,}}$ M = 11.67, SD = 2.58;

$ObjectiveNumeracy_{StudySample}$ $M = 12.26$, $SD = 2.39$; $SubjectiveNumeracy_{Community,}$ $M = 37.33$, $SD = 7.65$; $SubjectiveNumeracy_{StudySample}$ $M = 32.71$, $SD = 7.60$). It is worth noting that college students generally have higher numeracy than the average person; Reyna & Brainerd, 2007. Thus, our comparisons are relative rather than absolute, as is typical in the numeracy literature more generally. To maximize the generalizability of results, future work should study samples with broader ranges of numeracy. Second, future work should test hypotheses in a more realistic mock trial format. Jurors in our study did not have the opportunity to deliberate, which may influence results in real cases. In addition, in our study participants were informed that liability had already been established, that economic losses had been covered, and that their only job was to decide on pain and suffering. Although focusing on one element of damages is an approach often employed in jury damages research (e.g., McAuliff & Bornstein, 2010; Sunstein et al., 2002), this differs from the task facing jurors in real trials who have to make a series of decisions in a case. Third, in assessing damage awards, we were attempting to assess an actual legal determination that jurors would be asked to make. It is possible, though, that we were measuring another type of attitude. Previous researchers have argued that monetary responses can be considered as noisy and highly variable attitude measures (Kahneman, Ritov, Schkade, Sherman, & Varian, 1999). However, this would not explain our results as it is unclear how presenting anchors would influence underlying attitudes. In addition, we did not test the temporal nature of predictions in our study. We have interpreted results based on theoretical predictions as suggesting that those with higher levels of objective numeracy are better at allocating a numerical amount to their ordinal judgment. However, results could also support the opposite temporal prediction – that those with higher levels of objective numeracy are better at allocating an ordinal judgment to their chosen numerical amount. Follow-up research might test the temporal nature of the relationships with numeracy to confirm that numeracy helps convert gist judgments into dollar award amounts, rather than dollar award amounts into gist judgments.

Finally, we saw some results in this study that were not predicted and should be followed up further. First, although we did see that overall high numeracy jurors awarded more when they saw the long-term case and less when they saw the short-term case, this was not the case in the no anchor condition (although they still awarded more than the low numeracy jurors in the long-term case and less than the low numeracy jurors in the short-term case). This may suggest that even numerate jurors may benefit from the use of relevant anchors. Second, in this study, the AMD was greater in the meaningless anchor condition than in the meaningful anchor condition and the no anchor condition. We predicted that the meaningful anchor would reduce variability compared to the meaningless anchor, in line with previous research findings (Reyna et al., 2015), and it did. However, we expected, also in line with the results in Reyna et al. (2015), that a meaningful anchor would reduce variability compared to the no anchor control, and it did not, perhaps because of the very large size of the anchor number. Future research should follow up on this finding since it could be important in the context of current cases in which anchors are often introduced in the courtroom.

Policy Implications

Subject to confirmation in future work, this study has important policy implications. Results suggest that selecting jurors with greater numeracy or providing jurors with meaningful numerical anchors have important advantages in terms of ensuring damage awards appropriately distinguish between different cases, reducing the variability of damage awards, and enhancing the relationship between judgments of harm and dollar award amounts. Ensuring that awards are higher in more severe cases, even though each jury only makes a decision in one case and does not see comparable cases, is particularly important in the legal system, since a guiding principle in the tort system is that damage awards should be generally proportionate to the harms done (Sherwin & Bray, 2018). Plaintiffs deserve to be fairly compensated for their losses (Robbennolt & Hans, 2016). More numerate participants and participants who had seen meaningful numerical anchors appeared to be able to better put these fairness principles into practice.

But it is important to note potential limitations in an applied context. Although participants' objective numeracy does appear to hold benefits for deciding pain and suffering awards at the individual juror level, these may not be translated into gains at the group or jury level because of group processes (e.g., social conformity) that dilute rather than amplify effects of expertise. What is more, jurors higher in objective numeracy reported that it was more difficult to reach a damage award amount, suggesting more numerate jurors may struggle more with the task of assigning an award amount to pain and suffering. Additionally, as we noted earlier, even if jurors' objective numeracy is a positive benefit for damage award decision making, there are some major disadvantages to selecting citizens on that basis for jury service. The first is a practical challenge. How might the selection process include questions about numeracy? One can imagine a judge, presiding over a complex trial with extensive economic or statistical evidence, permitting a question or two during jury selection asking about a prospective juror's comfort with and preference for numerical information. It is difficult to imagine a judge allowing lawyers to give prospective jurors math problems to solve.

Our findings about the differences between objective and subjective numeracy are important to consider in light of this practical challenge for jury selection, and perhaps other applied contexts. As noted above, subjective numeracy is more easily measured than objective numeracy in an applied context (such as during voir dire). However, in this study we found that although objective and subjective numeracy were correlated, our predictions were only confirmed when using the objective numeracy measure. This is not surprising since the correlation between objective numeracy and subjective numeracy was only moderate, reinforcing the notion that subjective numeracy is an imperfect reflection of numerical ability. Thus, our results suggest that current measures of subjective numeracy may not be effective to measure numeracy in this context. Additional research might focus on further developing measures of objective numeracy that could be used during voir dire.

A second challenge to considering numeracy is that jurors higher in numeracy are different from jurors lower in numeracy in ways other than just their numeracy scores (Helm et al., 2017). If the lawyers' jury selection strategies favored more numerate jurors, it could result in a more numerate yet overall less representative slice of the population. That would interfere with the jury's ability to reflect the conscience of the community (Hans, 2014). Therefore, providing jurors with meaningful numerical anchors may be a more promising way to assist jurors in reaching damage awards for pain and suffering. Providing such anchors may provide similar benefits to numeracy but without the associated drawbacks. Providing participants with a meaningful anchor does not require different jury selection strategies that could compromise jury representativeness. The ideal way to provide such anchors may be to incorporate them in judicial instructions. However, previous work suggesting jury instructions designed to improve consistency in damage-award decision-making (e.g. Saks, Hollinger, Wissler, Evans,& Hart, 1997) has been met with resistance in the legal system. Therefore, the most practically feasible way to introduce meaningful anchors may be to encourage attorneys to find a way to incorporate meaningful anchors into the trial. Expert testimony, opening statements, and closing arguments are all possible moments during trials to introduce meaningful anchor numbers in order to put the dollars in damage awards in context (Ball, 2011). It should be noted that although plaintiffs' attorneys are only likely to provide such numbers when they are likely to result in an increased award for a client, meaningful anchors have the virtue of having a rationale (i.e., of making sense). This means that when they are introduced, provided they are well-chosen and well-explained, they have the potential to improve (rather than bias) juror decision-making. Thus, even if they are introduced inconsistently, this is better than not being introduced at all.

References

Bajwa, S. (2014). Apple v. Samsung: Is it time to change our patent trial system? *Global Business and Development Law Journal, 27*(1), 77-106.

Ball, D. (2011). *David Ball on damages* (3rd ed.) Louisville, CO: National Institute for Trial Advocacy.

Bavli, H. J. (2017). The logic of comparable-case guidance in the determination of awards for pain and suffering and punitive damages. *University of Cincinnati Law Review, 85,* 1–31.

Bornstein, B. H. (2004). The impact of different types of expert scientific testimony on mock jurors' liability verdicts. *Psychology, Crime & Law, 10,* 429–446. doi: 10.1080/10683160030001629292

Bornstein, B. H. (2017). Jury simulation research: Pros, cons, trends, and alternatives. In M. B. Kovera (Ed.), *The psychology of juries* (pp. 207-226). Washington, DC: American Psychological Association. http://dx.doi.org/10.1037/0000026-010.

Bornstein, B. H., Golding, J. L., Neuschatz, J., Kimbrough, C., Reed, K., Magyarics, C., & Luecht, K. (2017). Mock juror sampling issues in jury simulation research: A meta-analysis. *Law & Human Behavior, 41*, 13-28. doi: 10.1037.lhb0000223.

Bornstein, B. H., & Greene, E. (2017). *The jury under fire: Myth, controversy, and reform*. Oxford, UK: Oxford University Press.

Brewer, N. T., Chapman, G. B., Schwartz, J. A., & Bergus, G. R. (2007). The influence of irrelevant anchors on the judgment and choices of doctors and patients. *Medical Decision Making, 27,* 203-211. doi: 10.1177/0272989X06298595

Brumback, K. (2018, May 23). Georgia jury awards $1 billion after guard rapes teen. *Associated Press*, https://apnews.com/68f43d3f40af4b948f07fbbfad5bef50.

Campbell, J., Chao, B., & Robertson, C. (2017). Time is money: An empirical assessment of non-economic damages arguments. *Washington University Law Review, 95,* 1–52.

Caputo, A. (2014). Relevant information, personality traits and anchoring effect. *International Journal of Management and Decision Making, 13*, 62-76.

Cohen, T. H. (2009). *Tort bench and jury trials in state courts, 2005*. Washington, DC: Bureau of Justice Statistics, https://www.bjs.gov/index.cfm?ty=pbdetail&iid=2132.

Diamond, S. S., Rose, M. R., Murphy, B., & Meixner, J. (2011). Damage anchors on real juries. *Journal of Empirical Legal Studies, 8*, 148-178. doi: 10.1111/j.1740-1461.2011.01232.x

Diamond, S. S., Saks, M. J., & Landsman, S. (1998). Juror judgments about liability and damages: Sources of variability and ways to increase consistency. *DePaul Law Review, 48*(2) 301-326.

Englich, B., Mussweiler, T., & Strack, F. (2006). Playing dice with criminal sentences: The influence of irrelevant anchors on experts' judicial decision making. *Personality and Social Psychology Bulletin, 32*, 188-200. doi: 10.1177/0146167205282152

Enough, B., & Mussweiler, T. (2001). Sentencing under uncertainty: Anchoring effects in the courtroom. *Journal of Applied Social Psychology, 31*, 1535-1551. doi: abs/10.1111/j.1559-1816.2001.tb02687.x

Fagerlin, A., Zikmund-Fisher, B. J., Ubel, P. A., Jankovic, A., Derry, H. A., & Smith, D. M. (2007). Measuring numeracy without a math test: Development of the Subjective Numeracy Scale (SNS), *Medical Decision Making, 27,* 672-680. doi: 10.1177/0272989X07304449.

Frederick, S., & Fischhoff, B. (1998). Scope (in)sensitivity in elicited valuations. *Risk Decision and Policy, 3*, 109-123. doi: 10.1080/135753098348239

Furnham, A., & Boo, H. C. (2011). A literature review of the anchoring effect. *The Journal of Socio-Economics, 40*, 35-42. doi: 10.1016/j.socec.2010.10.008

Greene, E., & Bornstein, B. H. (2003). *Determining damages: The psychology of jury awards.* Washington DC, US: American Psychological Association. doi: 10.1037/10464-000.

Hans, V. P. (2014). What's it worth? Jury damage awards as community judgments. *William & Mary Law Review, 55*, 935-969.

Hans, V. P., Helm, R. K., & Reyna, V. F. (2018). From meaning to money: Translating injury into dollars. *Law and Human Behavior*, 42, 95-109. Doi: 10.1037/lhb0000282

Hans, V. P., Rachlinski, J. J., & Owens, E. G. (Eds.) (2011). Judgment by the numbers: Converting qualitative to quantitative judgments in law. *Journal of Empirical Legal Studies, 8* (Special Issue 1), 1-5. doi:10.1111/j.1740-1461.2011.01222.x

Hastie, R., Schkade, D. A., & Payne, J. W. (1999). Juror judgments in civil cases: Effects of plaintiff's requests and plaintiff's identity on punitive damage awards. *Law & Human Behavior, 23*, 445-470. doi: 10.1023/A:1022312115561

Helm, R. K., Hans, V. P., & Reyna, V. F. (2017). Trial by numbers. *Cornell Journal of Law and Public Policy, 27,* 107-143.

Kahan, D. M., Peters, E., Wittlin, M., Slovic, P., Larrimore Oullette, L., Braman, D., & Mandel, G. (2012). The polarizing impact of science literacy and numeracy on perceived climate change risks. *Nature Climate Change, 2,* 732-3733. doi: 10.1038/nclimate1547.

Kahneman, D., Ritov, I., Schkade, D., Sherman, S. J., & Varian, H. R. (1999). Economic preferences or attitude expressions?: An analysis of dollar responses to public issues. In B. Fischhoff & C. F. Manski (Eds.), *Elicitation of Preferences* (pp. 203-242). New York City, NY: Springer. doi: https://doi.org/10.1007/978-94-017-1406-8_8

Liberali, J. M., Reyna, V. F., Furlan, S., Stein, L. M., & Pardo, S. T. (2012). Individual differences in numeracy and cognitive reflection, with implications for biases and fallacies in probability judgment. *Journal of Behavioral Decision Making, 25*, 361-381. doi: 10.1002/bdm.752

Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical Decision Making*, *21*(1), 37-44. doi: 10.1177/0272989X0102100105

Marti, M. W., & Wissler, R. L. (2000). Be careful what you ask for: The effect of anchors on personal-injury damages awards. *Journal of Experimental Psychology: Applied, 62,* 91-103. doi: 2000-03676-001

McAuliff, B. D., & Bornstein, B. H. (2010). All anchors are not created equal: The effects of per diem versus lump sum requests on pain and suffering awards. *Law and Human Behavior, 34*, 164-174. doi: 10.1007/s10979-009-9178-8.

National Mathematics Advisory Panel (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel,* U.S. Department of Education: Washington, DC.

Peters, E., & Bjalkebring, P. (2015). Multiple numeric competencies: When a number is not just a number. *Journal of Personality and Social Psychology, 108*, 802-822. doi: 10.1037/pspp0000019

Peters, E., Fennema, M. G., & Tiede, K. E. (2018). The loss-bet paradox: Actuaries, accountants, and other numerate people rate numerically inferior gambles as superior. *Journal of Behavioral Decision Making*. doi: 10.1002/bdm.2085

Peters, E., Västfjäll, D., Slovic, P., Mertz, C. K., Mazzocco, K., Dickert, S. (2006). Numeracy and decision making. *Psychological Science, 17*, 407-413. doi: 10.1111/j.1467-

9280.2006.01720.x

Peters, E., Dieckmann, N. F., Dixon, A., Slovic, P., Mertz, C. K., & Hibbard, J. H. (2007). Less is more in presenting quality information to consumers. *Medical Care Research and Review*, *64*(2), 169-190. doi:10.1177/10775587070640020301

Reyna, V. F. (2012). A new intuitionism: Meaning, memory, and development in fuzzy-trace theory. *Judgment and Decision Making, 7*, 332-359.

Reyna, V. F., & Brainerd, C. J. (2011). Dual processes in decision making and developmental neuroscience: A fuzzy-trace model. *Developmental Review, 31*, 180-206. doi: 10.1016/j.dr.2011.07.004

Reyna, V. F., Corbin, J. C, Weldon, R. B., & Brainerd, C. J. (2016). How fuzzy-trace theory predicts true and false memories for words, sentences, and narratives. *Journal of Applied Research in Memory and Cognition, 5*(1), 1-9. doi:10.1016/j.jarmac.2015.12.003

Reyna, V. F., Hans, V. P., Corbin, J. C., Yeh, R., Lin, K., & Royer, C. (2015). The gist of juries: Testing a model of damage award decision making. *Psychology, Public Policy, and Law, 21,* 280-294. doi:10.1037/law0000048

Reyna, V. F., Nelson, W., Han, P., & Dieckmann, N. F. (2009). How numeracy influences risk comprehension and medical decision making. *Psychological Bulletin, 135*, 943-973. doi: 10.1037/a0017327

Robbennolt, J. K., & Hans, V. P. (2016). *The psychology of tort law*. NY: NYU Press.

Rowell, A., & Bregant, J. (2014). Numeracy and legal decision making. *Arizona State Law Journal, 46*, 13-29.

Saks, M. J., Hollinger, L. A., Wissler, R. L., Evans, D. L., & Hart, A. J. (1997). Reducing variability in civil jury awards. *Law and Human Behavior, 21*(3), 243-256. doi:

10.1023/A:1024834614312.

Saunders, J. (2006). Why do proposals designed to control variability in general damages (generally) fall on deaf ears? (And why this is too bad). *DePaul Law Review, 55,* 489-516.

Schwartz, L. M., Woloshin, S., Black, W. C., & Welch, H. G. (1997). The role of numeracy in understanding the benefit of screening mammography. *Annals of internal medicine*, *127*(11), 966-972.

Sherwin, E., & Bray, S. (2018). *Ames, Chafee, & Re on Remedies: Cases and Materials*. St. Paul, MN: Foundation Press.

Sugden, R., Zheng, J., & Zizzo, D. J. (2013). Not all anchors are created equal. *Journal of Economic Psychology, 39*, 21-31. doi: 10.1016/j.joep.2013.06.008

Sunstein, C. R, Hastie, R., Payne, J. W., Schkade, D. A., & Viscusi, W. K. (2002). *Punitive damages: How juries decide*. Chicago: University of Chicago Press.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124-1131. doi: 10.1126/science.185.4157.1124

Welsh, M. B., Delfabbro, P. H., Burns, N. R., & Begg, S. H. (2014). Individual differences in anchoring: Traits and experiences. *Learning and Individual Differences, 29*, 131-140. doi: 10.1016/j.lindif.2013.01.002.

Whyte, G., & Sebenius, J. K. (1997). The effect of multiple anchors on anchoring in individual and group judgment. *Organizational Behavior and Human Decision Processes, 69*, 74-85. doi: 10.1006/obhd.1996.2674

Wissler R. L., Hart A. J., & Saks M. J. (2001). Decision making about general damages: A comparison of jurors, judges, and lawyers. *Michigan Law Review*, *98*, 751–772.

Tables

Table 1

Mean Award Amounts in Experimental Conditions, Split by High and Low Objective Numeracy Group

Case	Anchor Condition	Numeracy Group	*n*	*M*	*SD*
Short-term	No Anchor	Low	23	231,847.83	481,601.20
		High	25	142,880.00	402,226.86
	Meaningless	Low	29	560,275.86	713536.42
		High	21	701,142.86	1,260,452.07
	Meaningful	Low	20	318,725.00	556,132.99
		High	17	235,302.94	242087.66
Long-term	No Anchor	Low	20	81,405.00	223,850.58
		High	21	110,371.43	158,959.23
	Meaningless	Low	19	347,821.11	362,898.85
		High	18	1,651,000.00	2,377,736.76
	Meaningful	Low	28	865,107.25	1,873,278.11
		High	20	778,250.00	666,635.85

Note. This table only includes participants who gave a non $0 award amount and thus are included in our analyses involving award amount.

Table 2

Numeracy Descriptive Statistics

Scale	*M*	*SD*	Observed Range (Theoretical Range)	Median	*n* (low group	*n* (high group)
Objective Numeracy	12.30	2.39	1-15 (0-15)	13	139	122
Subjective Numeracy	32.68	7.52	8-48 (8-48)	33	159	164
Subjective Numeracy Ability	15.33	4.79	4-24 (4-24)	15	149	168
Subjective Numeracy Preference	17.35	3.82	4-24 (4-24)	18	168	142

Note. This table only includes participants who gave a non $0 award amount and thus are included in our analyses involving award amount. Participants with the exact median score were excluded from median split numeracy groups. The number of participants excluded for this reason were as follows, 77 in objective numeracy analyses, 14 in subjective numeracy analyses, 20 in subjective numeracy ability analyses, and 27 in subjective numeracy preference analyses.

Table 3

Regression Assessing Verbatim-gist Correspondence Predicting Log Award Amount (Objective Numeracy)

Predictor	*B*	*SE*	β	*t*	*p*
Case (1=Short-term injury, 2=Long-term injury)	.22	.10	.10	2.15	.03
Anchor Type (1 = No Anchor, 2 = Meaningless Anchor, 3= Meaningful Anchor)	.34	.06	.26	5.40	<.001
Objective Numeracy Score	.15	.07	.52	2.21	.03
Ordinal Gist Judgment (1=Nil, 2=Low, 3=Medium, 4=High)	1.41	.38	.77	3.71	<.001
Objective Numeracy Score x Ordinal Gist Judgment	-.04	.02	-.58	-1.80	.08

Table 4

Regression Assessing Verbatim-gist Correspondence Predicting Log Award Amount	*B*	*SE*	β	*t*	*p*
Case (1 = Short-term injury, 2 = Long-term injury)	.19	.10	.09	1.86	.06
Anchor Type (1 = No Anchor, 2 = Meaningless Anchor, 3 = Meaningful Anchor)	.20	.30	.15	.64	.52
Ordinal Gist Judgment (1 = Nil, 2 = Low, 3 = Medium, 4 = High)	1.12	.22	.61	5.02	<.001
Anchor Type x Ordinal Gist Judgment	-.18	.10	-.44	-1.75	.08

Figures

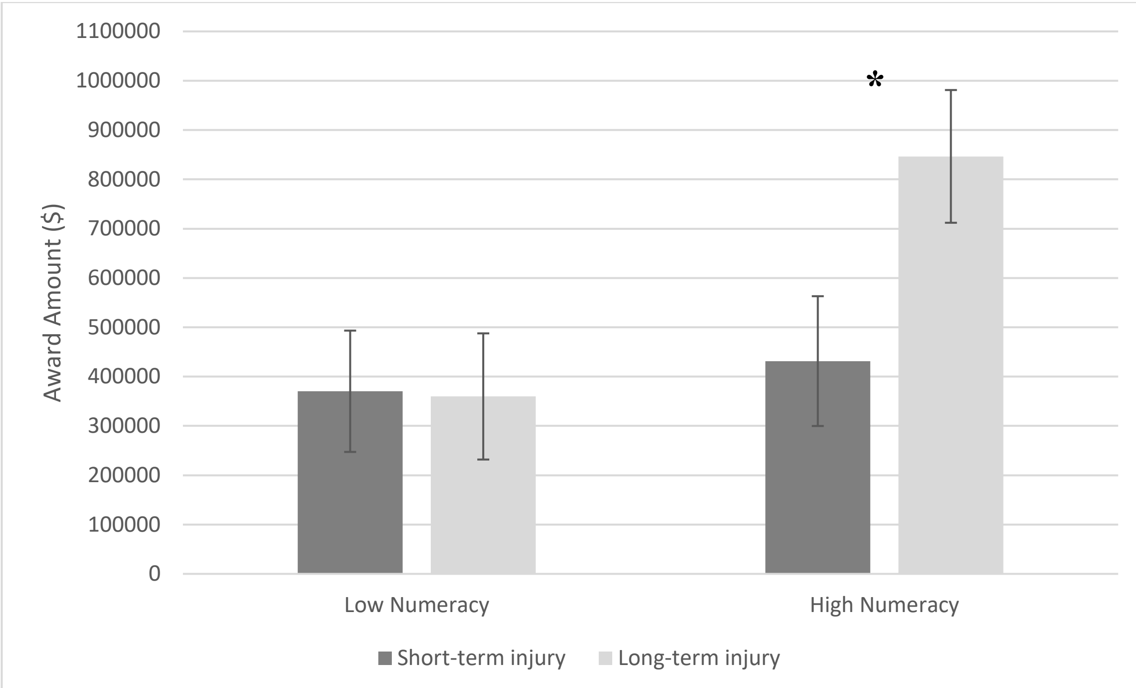

Figure 1: Significant two-way interaction between numeracy and case when predicting log award amount (note that log values have been transformed back to raw award values for the purpose of this figure). Error bars represent +/- 1 standard error. **p* < .05.

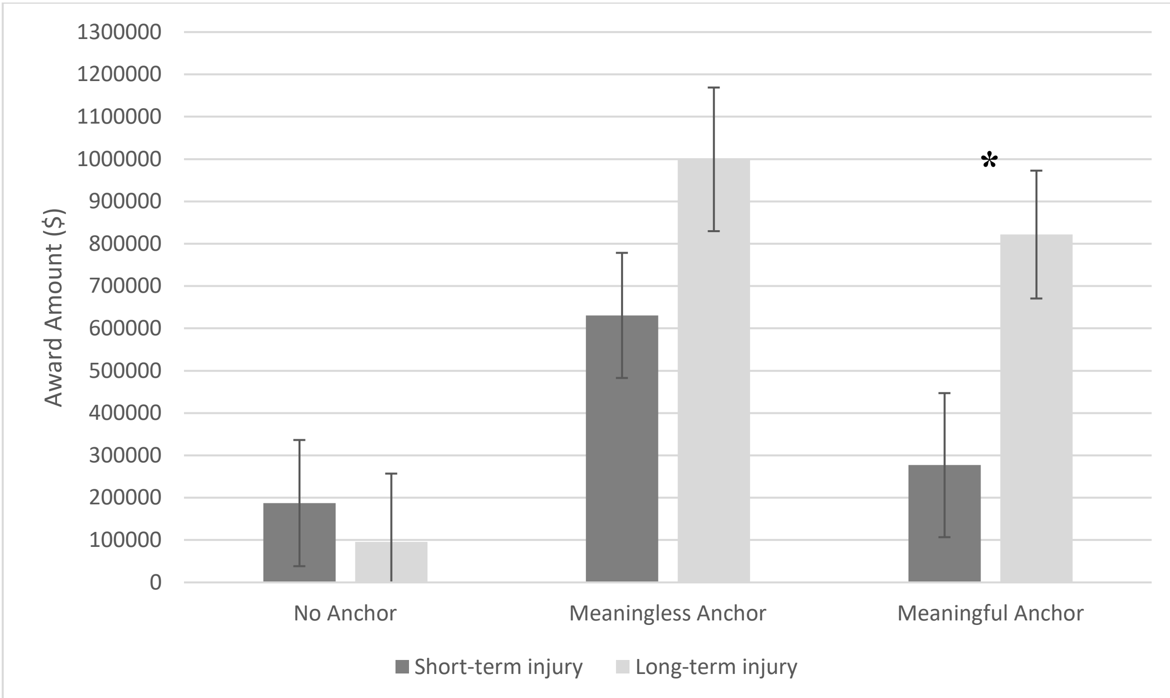

Figure 2: Significant two-way interaction between anchor condition and case when predicting log award amount (note that log values have been transformed back to raw award values for the purpose of this figure). Error bars represent +/- 1 standard error. *p < .05.

Supplemental Materials

The results below are included as supplemental materials in order to allow a thorough review of results for all analyses, while avoiding repetition in the main article body.

Table S1

Regression Assessing Validity Predicting Log Award Amount (Objective Numeracy)

Predictor	*B*	*SE*	β	*t*	*p*
Case (1= Short-term injury, 2 = Long-term injury)	-1.70	.67	-.78	-2.55	.01
Anchor Type (1 = No Anchor, 2 = Meaningless Anchor, 3 = Meaningful Anchor)	-.29	.21	-.22	-1.38	.17
Objective Numeracy Score	-.05	.08	-.12	-.71	.47
Anchor Type x Case	.43	.13	.68	3.23	.001
Objective Numeracy Score x Case	.09	.05	.58	1.89	.06

Table S2

Regression Assessing Validity Predicting Log Award Amount (Subjective Numeracy)

Predictor	*B*	*SE*	β	*t*	*p*
Case (1= Short-term injury, 2 = Long-term injury)	-.75	.59	-.35	-1.28	.20
Anchor Type (1 = No Anchor, 2 = Meaningless Anchor, 3 = Meaningful Anchor)	-.26	.21	-.20	-1.19	.23
Subjective Numeracy Score	.01	.03	.13	.45	.65
Anchor Type x Case	.40	.14	.62	2.93	.004
Subjective Numeracy Score x Case	.06	.12	.13	.46	.66

Table S3

Regression Assessing Validity Predicting Log Award Amount (Subjective Numeracy Preference)

Predictor	*B*	*SE*	β	*t*	*p*
Case (1= Short-term injury, 2 = Long-term injury)	-1.4	.60	-.63	-2.26	.02
Anchor Type (1 = No Anchor, 2 = Meaningless Anchor, 3 = Meaningful Anchor)	-.26	.21	-.20	-1.22	.22
Subjective Numeracy Preference Score	-.03	.05	-.11	-.66	.51
Anchor Type x Case	.40	.14	.63	2.95	.003
Subjective Numeracy Preference Score x Case	.05	.03	.44	1.62	.11

Table S4

Regression Assessing Validity Predicting Log Award Amount (Subjective Numeracy Ability)

Predictor	*B*	*SE*	β	*t*	*p*
Case (1= Short-term injury, 2 = Long-term injury)	-.43	.48	-.20	-.91	.37
Anchor Type (1 = No Anchor, 2 = Meaningless Anchor, 3 = Meaningful Anchor)	-.26	.22	-.20	-1.20	.23
Subjective Numeracy Ability Score	.04	.04	.17	.99	.32
Anchor Type x Case	.40	.14	.62	2.92	.004
Subjective Numeracy Ability Score x Case	-.01	.02	-.07	-.29	.78

Table S5

Regression Assessing Verbatim-gist Correspondence Predicting Log Award Amount (Subjective Numeracy)

Predictor	*B*	*SE*	β	*t*	*p*
Case (1= Short-term injury, 2 = Long-term injury)	.22	.10	.10	2.09	.04
Anchor Type (1 = No Anchor, 2 = Meaningless Anchor, 3 = Meaningful Anchor)	.33	.06	.25	5.36	<.001
Subjective Numeracy Score	.06	.03	.44	1.98	.049
Ordinal Gist Judgment (1 = Nil, 2 = Low, 3 = Medium, 4 = High)	1.24	.35	.68	3.59	<.001
Subjective Numeracy Score x Ordinal Gist Judgment	-.02	.01	-.45	-1.49	.14

Table S6

Regression Assessing Verbatim-gist Correspondence Predicting Log Award Amount (Subjective Numeracy Preference)

Predictor	*B*	*SE*	β	*t*	*p*
Case (1 = Short-term injury, 2 = Long-term injury)	.22	.10	.10	2.15	.03
Anchor Type (1 = No Anchor, 2 = Meaningless Anchor, 3 = Meaningful Anchor)	.34	.06	.26	5.40	<.001
Subjective Numeracy Preference Score	.15	.07	.52	2.21	.03
Ordinal Gist Judgment (1= Nil, 2 = Low, 3 = Medium, 4 = High)	1.41	.38	.77	3.71	<.001
Subjective Numeracy Preference Score x Ordinal Gist Judgment	-.04	.02	-.58	-1.80	.07

Table S7

Regression Assessing Verbatim-gist Correspondence Predicting Log Award Amount (Subjective Numeracy Ability)

Predictor	*B*	*SE*	β	*t*	*p*
Case (1 = Short-term injury, 2 = Long-term injury)	.20	.10	.09	1.94	.05
Anchor Type (1 = No Anchor, 2 = Meaningless Anchor, 3 = Meaningful Anchor)	.33	.06	.25	5.22	<.001
Subjective Numeracy Ability Score	.07	.05	.31	1.41	.16
Ordinal Gist Judgment (1 = Nil, 2 = Low, 3 = Medium, 4 = High)	.99	.26	.54	3.87	<.001
Subjective Numeracy Ability Score x Ordinal Gist Judgment	-.02	.02	-.26	-.98	.33

Figure S1

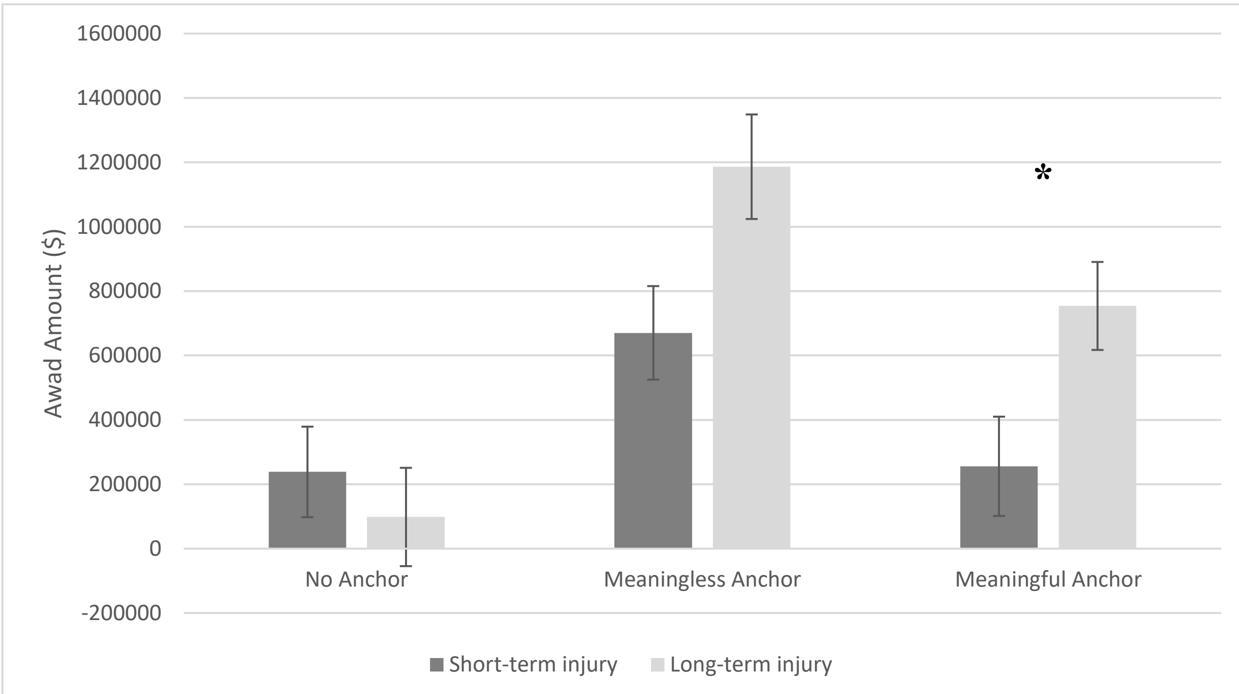

Figure S1: Significant two-way interaction between anchor condition and case when predicting log award amount in ANOVA including subjective numeracy (note that log values have been transformed back to raw award values for the purpose of this figure). Error bars represent +/- 1 standard error. *$p < .05$