CORNELL LAW SCHOOL

LEGAL STUDIES RESEARCH PAPER SERIES

From Meaning to Money: Translating Injury into Dollars

Valerie P. Hans*, Rebecca K. Helm, and Valerie F. Reyna*****

Cornell Law School
Myron Taylor Hall
Ithaca, NY 14853-4901

*Cornell Law School research paper No. 18-08
** University of Exeter
*** Cornell University

Running head: FROM MEANING TO MONEY

From Meaning to Money: Translating Injury into Dollars

Valerie P. Hans

Cornell Law School

Rebecca K. Helm

University of Exeter

Valerie F. Reyna

Cornell University

In press, *Law and Human Behavior*

Author Note

Valerie P. Hans, Cornell Law School, Cornell University; Rebecca K. Helm, Exeter Law School, University of Exeter; Valerie F. Reyna, Human Neuroscience Institute, Department of Human Development, and Center for Behavioral Economics and Decision Research, Cornell University.

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.

Correspondence concerning this article should be addressed to Valerie P. Hans, Cornell Law School, Ithaca, NY 14853. E-mail: valerie.hans@cornell.edu

Abstract

Legal systems often require the translation of qualitative assessments into quantitative judgments, yet the qualitative-to-quantitative conversion is a challenging, understudied process. We conducted an experimental test of predictions from a new theory of juror damage award decision making, examining how 154 lay people engaged in the translation process in recommending money damages for pain and suffering in a personal injury tort case. The experiment varied the presence, size, and meaningfulness of an anchor number to determine how these factors influenced monetary award judgments, perceived difficulty, and subjective meaningfulness of awards. As predicted, variability in awards was high, with awards participants considered to be "medium" (rather than "low" or "high") having the most dispersion. The gist of awards as low, medium, or high fully mediated the relationship between perceived pain/suffering and award amount. Moreover, controlling for participants' perceptions of plaintiffs and defendants, as well as their desire to punish and to take economic losses into account, meaningful anchors predicted unique variance in award judgments: A meaningful large anchor number drove awards up and a meaningful small anchor drove them down, whereas meaningless large and small anchors did not differ significantly. Numeracy did not predict award magnitudes or variability, but surprisingly, more numerate participants reported that it was *more* difficult to pick an exact figure to compensate the plaintiff for pain and suffering. The results support predictions of the theory about qualitative gist and meaningful anchors, and suggest that we can assist jurors to arrive at damage awards by providing meaningful numbers.

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

Public significance statement: Damage awards should reflect the underlying injury suffered by the plaintiff in civil cases, yet jurors find it challenging to decide on damage awards for an individual's pain and suffering. Controlling for case factors, this experiment shows that providing meaningful anchor numbers helps mock jurors choose damage amounts that correspond to their underlying sense – the gist – of an individual's pain and suffering.

From Meaning to Money: Translating Injury into Dollars

The conversion of qualitative to quantitative judgments is central to many legal decisions, including damage awards (Hans, Rachlinski, & Owens, 2011). A substantial amount of research documents the specific challenges people face in generating sound, reliable quantitative judgments (Cantlon & Brannon, 2006; Nieder, 2016; Peters, 2012; Reyna, Nelson, Han, & Dieckmann, 2009). That poses a special problem for the U.S. civil justice system, in which damage awards constitute a major remedy for civil wrongs. Despite the importance of better understanding the qualitative to quantitative translation process in the law, it is undertheorized and understudied. In particular, the role of meaning in numerical judgments is poorly understood. The current article breaks new ground by analyzing factors leading to difficulty in deciding on appropriate damages and testing predictions about the meaning of numerical information from a model of damage award decision making.

Jury Damage Awards

Jury damage awards are an important legal and policy matter in the U.S. Representative civil juries have the potential to provide an important contribution to civil justice. Combining jurors' diverse perspectives on the value of an injury can keep civil awards in line with the community's shifting assessments of appropriate compensation (Hans, 2014). Although the U.S. public is generally supportive of the jury system, some observers of the civil justice system pinpoint damage awards as an area of concern. Insurance agents, lawyers, and their clients worry about their ability to predict jury awards, which makes it difficult for them to assess in advance whether to settle or to take a case to trial (Daniels & Martin, 2015; Goodman-Delahunty et al., 2010; Hans, 2000; Hans & Eisenberg, 2011; Jacobson et al., 2011).

Juries (and judges as well) face an unusually ambiguous situation in determining money damages. The intangibility of many losses such as the pain of a severe injury or the death of a family member, uncertainty of future projections about the lifelong effects of an injury, and limited guidance all contribute to ambiguity (Diamond, Rose, Murphy, & Meixner, 2011; Sherwin, Eisenberg, & Re, 2012). An additional complication is that dollar awards are an unbounded scale, which contributes to variability (Hastie, 2011; Kahneman et al., 1998).

Analysis of jury damage awards provides a mixed picture. Reassuringly, more serious injuries typically receive greater compensation (Greene & Bornstein, 2003; Wissler et al., 1997; Wissler, Rector, & Saks, 2001). Yet jurors face substantial problems in coming up with damage award figures (Diamond et al., 2011; Kahneman, Schkade, & Sunstein, 1998; Mott, Hans, & Simpson, 2000). Sunstein et al. (2002) and Wissler et al. (1999) have found that people tend to show greater consensus in their qualitative assessements, such as evaluating the wrongfulness of a defendant's behavior or the severity of a plaintiff's injury, than in their dollar awards.

Economic damages often have straightforward referents such as lost income or medical expenses, a fact that may reduce difficulty and reliance on irrelevant factors in determining economic damages (e.g., Girvan & Marek, 2016). In contrast, pain and suffering and punitive damages do not typically have clear referents. The jury is well-suited to play an especially crucial role in judgments about intangible damages such as pain and suffering precisely because there are no clear referents. Combining the valuations of diverse members of the community, the jury award provides an approximation of the community's assessment of the value (Hans, 2014).

In sum, jurors deciding damage awards, and particularly awards for pain and suffering, encounter limited legal guidance, intangible injuries, uncertain projections, an unbounded scale for dollar awards, and low numeracy. Relatively little research has addressed the translation of injury severity to specific dollar awards. How individuals employ dollar scales to reflect their qualitative judgments is thought to be "possibly meaningless" rather than a systematic process to be studied (e.g., Kahneman et al., 1998, p. 67).

A New Model of Jury Damage Award Decision Making

With the aim of moving research and theory in this area forward, Hans and Reyna proposed (2011) and tested (Reyna, Hans, Corbin, Yeh, Lin, & Royer, 2015) a new model of damage award decision making. The goal of the model is to identify the process by which jurors move from the qualitative to the quantitative, from a judgment about injury to a decision about a dollar award. The current experiment tests predictions from the Hans-Reyna gist model.

The Hans-Reyna model draws on fuzzy-trace theory (FTT; Reyna, 2012a, b; Reyna & Brainerd, 2011). According to FTT, people encode two types of mental representations of information in parallel: verbatim and gist representations (e.g., Reyna, Corbin, Weldon, & Brainerd, 2016; Singer & Remillard, 2008). Verbatim representations are literal and detailed. In contrast, gist representations are the meaning a person derives from the information. Verbatim representations preserve the specific surface form of information, such as exact wording and precise numbers, compared to the bottom-line general meaning captured in gist representations. People engage in both gist and verbatim forms of thinking, but when they make judgments, gist tends to be more important (Hans & Reyna, 2011).

The Hans-Reyna model posits that jurors make both categorical (damages are warranted or not) and ordinal (the damages deserved are low or high) gist determinations, and search for dollar award amounts that fit the gist judgments (Hans & Reyna, 2011). Throughout the trial, the model assumes that jurors engage in mainly gist-based reasoning about the plaintiff's injury and the defendant's culpability to determine liability. The juror reaches a categorical (yes or no) judgment about whether damages are warranted. Jurors encode the subjective impression of the specific injury the plaintiff has suffered in terms of severity. In line with that ordinal gist of injury severity, they determine whether the damage award that is deserved is low or high. Intuitions about ordinal distinctions between low and high (but typically not specific numbers) are available in memory. If an award is judged as warranted, but it is neither low nor high, then the gist is less clear and it falls into a more nebulous hedging category of medium.

The ordinal gist judgment about the deserved damages (low, medium, or high) is mapped onto a number that corresponds to the gist of the award judgment. The juror will identify a number that is, to him or her, low, medium, or high, to match the deserved damages. Jurors are apt to rely on symbolic numbers from everyday life that already have meaning to them as low or high numbers. To most of us, one dollar is low; a million dollars is high. An individual's background will also shape his or her interpretation of low and high numbers. Jurors also obtain numbers from the case, including medical costs, lost wages, and attorney *ad damnum* requests, numbers that the attorneys have convincingly portrayed as having a meaningful interpretation in the context of the case (Hastie, 2011). The juror draws on these sources to generate an award amount that matches the ordinal (low, medium, high) gist of the deserved damages. The description of the process is presented as occurring in serial stages, but the model anticipates interactivity across these judgments. Thus, the decision making process is not linear. For example, judgments of the severity of an injury may influence attributions of responsibility (Robbennolt & Hans, 2016).

The model anticipates that psychological factors could affect damage award decision making. For example, cognitive style variables such as need for cognition (NC) and cognitive reflection, which figure promiennntly in contemporary theories of judgment and decision making (Frederick, 2005; Stanovich & West, 2008) and have been found in prior research to influence juror decisions (Bornstein, 2004; McAuliff & Kovera, 2008), may also affect damage award decisions. Earlier research has explored the role of NC in judgments of expert testimony. For example, Bornstein (2004) found that mock jurors who were low in NC were more responsive to anecdotal expert evidence; mock jurors high in NC did not favor anecdotal over experimental evidence. McAuliff and Kovera (2008) found that participants high in NC were sensitive to variations in the quality of expert evidence, but those low in NC were not. The high NC mock jurors were more likely to find for the plaintiff when her expert presented a valid study than an invalid study; low NC mock jurors were not influenced by the validity of the expert's study.

A fact finder's numeracy has been predicted to facilitate better and more consistent decision making involving numbers (Helm, Hans, & Reyna, 2017). Although numeracy itself has been extensively researched (for a review, see Reyna et al., 2009), its relationship to award decision making has received little attention (Rowell & Bregant, 2014; but see Bornstein, 2004, Experiment 2). In making decisions, people lower in numeracy are more likely to rely on irrelevant cues in lieu of using relevant numerical information (Peters et al., 2007; Reyna, Nelson, Han, & Pignone, 2015). The lack of "number sense" (Geary et al., 2011; Siegler & Booth, 2004) could lead to difficulty in producing a reliable gist of low or high dollar amounts. Literal numbers are not translated one-for-one into gist representations—they are *interpreted* in context as low or high. A gist model emphasizes the need to distinguish between number sense--the ability to meaningfully interpret numbers--and the ability to engage in numerical calculations (Liberali et al., 2012). That is important, because research has shown that people higher in computational numeracy may rate numerically inferior monetary gambles as more attractive than superior gambles (Reyna et al., 2009). The current experiment offers a new opportunity to explore how facility with numbers influences award judgments.

Measures of subjective numeracy assess self-perceptions of the ability to use numbers and generally correlate with performance-based objective numeracy measures (Fagerlin et al., 2007). However, like many self-perceptions of ability, they are likely to be subject to the Dunning-Kruger effect in which those with low ability overestimate their competence (Dunning, Heath, & Suls, 2004). Thus, subjective numeracy may be related to confidence and perceived difficulty of numerical judgments.

Previous Efforts to Test the Gist Model

Hans and Reyna (2011) examined whether predictions from the gist model were consistent with existing research findings about jury damage awards. Predictions derived from the model were that jurors search for meaningful numbers to reflect their gist of the cases; that damage awards overrepresent round numbers since round numbers are easier to map onto gist judgments; that there is ordinal regularity in damage award judgments; and that, because gist judgments are fuzzy, anchor figures provide context for jurors (allowing them to translate "low" and "high" into dollars) and thus influence damage awards. All four of these general predictions were borne out.

Following up, Reyna, Hans, et al. (2015) tested specific predictions from the Hans-Reyna model. Participants read a scenario of a personal injury motor vehicle accident trial and were asked to give a specific dollar award for pain and suffering. A statement made by the hypothetical foreman varied three dimensions of an anchor number: the amount of the anchor, the meaningfulness of the anchor, and the relative perception of the anchor number. In the meaningful anchor conditions, the foreman offered a meaningful number that could be used to judge the gist of damages (e.g., the national median annual income). In contrast, in the meaningless anchor conditions, the foreman mentioned the same number but in an irrelevant context (the cost of courtroom renovations). A relative perception manipulation aimed to shift the interpretation of the gist of the number as comparatively low or high. For example, the foreman observed, “To begin the conversation, I was thinking $40,000 would be a good award because that is way above the national median annual income.”

Research on anchoring (for summaries see Gilovich, Griffin, & Kahneman, 2002; Kahneman, 2011) has been conducted almost exclusively with meaningless anchors. Reyna, Hans, et al. (2015) expected that even a meaningless anchor would provide some amount of context and would influence award decisions because people were more likely to have fuzzy gists in mind rather than concrete numbers. However, they hypothesized that the meaningful anchor number would be more influential than the meaningless anchor. A meaningful anchor could help jurors by providing information about the magnitude of relevant numbers; it might also resonate with case facts. The gist-based model predicts that jurors will anchor onto meaningful numbers that match their own gist of the case and use these numbers to generate a damage award amount. Jurors should anchor more strongly to meaningful anchors, hence those awards should be less variable than awards in meaningless anchor conditions. As for the relative perception manipulation, the foreman’s suggestion that a specific anchor is a relatively small amount of money should provide context for what a low amount means, leading participants to award higher amounts of money, compared to participants who are told that the anchor amount is a great deal of money. It should be noted that all of the anchors were manipulated as relative comparisons; meaningful small versus large anchors were not directly investigated.

The size, meaningfulness, and perception manipulations influenced the awards in ways consistent with gist theory. The anchor size, anchor meaning, and relative perception interacted. Meaningful large and small anchors differed significantly whether or not the relative perception of the anchor numbers varied. The relative perception manipulation was meant to drive awards either upwards or downwards by making the anchor appear to be a small or a large number. It was effective too, but only in conditions with meaningless anchors. When a meaningful anchor was present, it drove the damage award. When it was absent, the contextual cue provided through the relative perception measure drove the damage award. The results supported another prediction, that jury awards would be more variable in the meaningless anchor condition, because in the meaningful anchor condition participants' awards would center more tightly around the relevant anchor. Meaningful anchors led to damage awards of lower variability compared to awards in the meaningless anchor conditions and to a no anchor control condition.

Current Experiment: Refining and Expanding Measures and Testing Predictions

The current experiment builds upon this base developed by Hans, Reyna, and their colleagues to test the model by (a) directly manipulating small versus large anchors that vary in meaningfulness without relative comparisons, (b) explicitly measuring perceived meaningfulness of anchors, (c) examining the "fuzziness" or malleability of award judgments, (d) extending the model to consider theoretically important individual differences (e.g., in numeracy, cognitive reflection, and need for cognition) and (e) analyzing factors that affect perceived difficulty of making pain and suffering award judgments. The experiment manipulates the presence, size, and meaningfulness of an anchor number provided in a mock juror scenario study. It uses a different approach to the anchor manipulation, poses questions that directly test aspects of the Hans-Reyna model including perceived meaningfulness, and examines the impact of numeracy and cognitive style. In addition to effects on award judgments, we examine effects of these factors on rated difficulty of and confidence in awards. Ultimately, we are able to model award judgments controlling for perceptions of case factors, such as perceptions of plaintiffs and defendants, along with individual differences in the ability and desire to process numbers, isolating unique variance associated with manipulating the meaningfulness of numerical anchors.

Hypotheses 1-3: Meaningfulness and size of anchor numbers

The first three hypotheses tested in the experiment relate to the meaningfulness and size of the anchor numbers. We predicted that (1) income measures (meaningful anchors) would be perceived as more meaningful than the courtroom renovations measure (meaningless anchor); that (2) numerical anchors, in addition to perceptions of the case and individual differences, would shift damage awards in the direction of the anchors; and that (3) the amount and the meaningfulness of the anchor would interact. Meaningful large anchors were predicted to drive awards up, compared to meaningless anchors; in contrast, meaningful small anchors were predicted to drive awards down, compared to meaningless anchors.

The current experiment examines anchor meaning in more detail and measures the participants' subjective evaluations of meaning. This experiment investigates the direct effect of meaningful versus meaningless anchors on award amounts, as opposed to effects of relative comparisons to anchors, as was done in the previous study on meaningful anchors (Reyna, Hans, et al., 2015). As described above, the Reyna, Hans, et al. (2015) study incorporated the anchor manipulations into the foreman's remarks, varying whether the dollar figure the foreman mentioned was an income measure or the cost of courtroom construction. When it came to pain and suffering, average income was assumed to be more meaningful to participants than construction costs of a courtroom because the former provides a benchmark for judging whether the magnitude of an award is small or large. However, the perceived subjective meaningfulness of an anchor is important to study directly, and the current experiment does so.

Hypotheses 4 amd 5: Ordinality and variability of awards

The model also led to the predictions that (4) gist is roughly ordinally coherent; awards perceived as low versus high would differ in magnitude in the appropriate direction; and that (5) jurors who give an award that they consider to be a “medium” amount would show the most variability in award judgments, while those who give an award that they consider to be “low” or “high” would show less variability in award judgments. The Hans-Reyna theory assumes that participants make ordinal judgments about the appropriate award as low, medium or high, and then select numbers that fit the ordinal gist. To test these hypotheses about ordinality and variability of awards, the current experiment asks specifically whether the award amounts that participants have given are, in their view, nothing (nil), low, medium, or high. We then compare the low, medium, and high awards and examine their central tendencies and their dispersion.

Hypotheses relating to role of individual difference variables in award decisions

The experiments permit a test of the effects on damage award decision making of individual-difference variables found in prior research to be important in juror decisions and in numerical judgments, including objective numeracy, subjective numeracy, and aspects of cognitive style. The experiment extends mock jury research on need for cognition (NC) by examining whether NC is associated with damage award judgments. We predicted that both (6) cognitive style and (7) numeracy would affect damage award decisions. Higher numeracy and NC are usually expected to reduce susceptibility to anchoring and other biases (see Reyna et al., 2009). However, if measures capture sensitivity to the gist of numbers, then more numerate and high NC jurors would be predicted to be more responsive to the relevant cue of meaningful anchorsin arriving at damage awards, compared to jurors low in numeracy and in NC.

Hypotheses relating to predictors of difficulty in arriving at a damage award

The experiment permits us to examine the perceived difficulty in arriving at a damage award amount and confidence in the decision. Participants are asked how difficult it was for them to determine an appropriate award in the case, and how confident they are in that judgment. We predict that (8) anchors, especially meaningful anchors, (9) cognitive style, and (10) numeracy will decrease the perceived difficulty of making an award and increase confidence in award judgments. In sum, the current experiment has the potential to make theoretical progress by testing elements of the Hans-Reyna damage award model, and examining the role of numeracy and cognitive style in damage award judgments.

Method

Participants

A total of 154 undergraduate students (77% female; M_{Age} = 20.01, *SD* =1.67) took part in the experiment. Participants received course credit in Psychology or Human Development courses at a northeastern university. It was a racially diverse group: 70% of participants identified themselves as White, 18% as Asian, 3% as mixed ethnicity, 7% as African American and 2% selected other. Furthermore, 5% of the participants identified themselves as having Hispanic, Latino or Spanish origin. All participants indicated that they were 18 years of age or older and could communicate effectively in English. Our prior work with this population (Reyna, Hans, et al., 2015) showed that the majority fulfilled key requirements for jury service in New York (i.e., at least 18 years of age, reside in the area, U.S. citizens, and facility with English). An absence of felony convictions is also a jury service requirement, but we did not ask about participants' criminal history.

Procedure

Cornell University's Institutional Review Board for Human Participants reviewed and approved the study procedure and all materials as qualifying for an exemption from full IRB review. The experiment was a 2 (anchor amount: large anchor, small anchor) X 2 (anchor meaning: anchor is meaningful, anchor is meaningless) between-subjects experiment with an additional no anchor control group. The experiment was conducted online through Qualtrics. After agreeing to participate, the participants were randomly assigned to one of the five experimental conditions. They read a 1,272 word case scenario, decided on a damage award, and answered questions about the case. Participants answered more general questions tapping psychological tendencies and provided their demographic characteristics.

Materials

The study materials began with a set of instructions about informed consent and what to expect in the study. It then proceeded to present a trial summary in the case of *Monroe v. Rumson*, a negligence case involving a personal injury based on the actual case of *Abbinante v. O'Connell* (1996) and employed in previous research projects on damage awards (McAuliff & Bornstein, 2010; Reyna, Hans, et al., 2015). Reflecting the usual sequence of events at trial, the summary presented the case facts and then short judicial instructions. The summary specified that the defendant had been found liable for the plaintiff's injuries and that economic damages had already been paid to the plaintiff. The judge explained that their task was to arrive at a dollar figure to compensate the plaintiff for pain and suffering. The judge also explained that there are no definite mathematical formulas or instructions for determining pain and suffering awards. The instructions directed jurors to use their best judgment in coming up with an award amount. 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, which varied whether the anchor was large or small and meaningful or meaningless. Participants in the large anchor conditions saw $1.5 million as their anchor. This was described as either the median lifetime income (large meaningful condition) or the cost of courtroom renovations (large meaningless condition). Participants in the small anchor conditions saw $50,000 as their anchor. This was described as either the median annual income (small meaningful condition) or the cost of courtroom renovations (small meaningless condition). For example, in the meaningless large anchor condition, the question read as follows: "The courtroom is under construction, and this project will cost about $1.5 million. Do you think the award amount to the plaintiff Rebecca Munroe for her pain and suffering should be above or below this number?" Two options (above; below) were given. (Just 11% of participants in the large anchor conditions thought the award should be above $1.5 million, compared to 42% in the small anchor conditions who thought the award should exceed $50,000.) In the no anchor condition, the question was omitted.

Anchor numbers were selected based on previous research and the plausibility of the numbers as estimates of median annual income and median lifetime income. The control group participants in Reyna, Hans, et al. (2015) reading the same case awarded an average of $224,288, an amount which fell between the two anchor numbers. The U. S. Census Bureau estimated annual median household income rose to $55,775 in 2015 (U. S. Census Bureau, 2016). Multiplying the annual median income figure that we used ($50,000) by 30 years of such income produces the large anchor amount of $1.5 million.

After being asked for their award amount, participants were asked to state whether they considered their award amount to be nil (basically nothing), low, medium, or high, to state their highest acceptable award, to state their lowest acceptable award, to provide confidence judgments, and to describe the reasons for the award amount. Confidence judgments were evaluated on 1 (not at all confident) to 10 (completely confident) scales. (We retained scale values for judgments used in prior work.) Participants were then asked whether they agreed that the median annual income, 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 to 7 (1 = strongly disagree, 7 = strongly agree).

The participants rated the severity of the plaintiff's injuries and the severity of the plaintiff's pain and suffering on a scale from 1 to 7 (1 = low, 7 = high). The participants also rated their perceptions of the plaintiff and of the defendant on a scale from 1 to 7 (1 = extremely negative, 7 = extremely positive). In addition, participants rated the extent to which the defendant was negligent and the extent to which the defendant contributed to the plaintiff's injuries, making both ratings on a scale from 1 to 11 (1 = not at all, 11 = extremely). Participants were further asked to estimate how difficult it was to decide on an award, how motivated they were while reading the case, and how much cognitive effort they exerted. These estimates were made on a scale from 1 to 7 (1 = very little, 7 = great amount). Participants were asked if their award decision was affected by their desire to punish the defendant, on a scale from 1 to 7 (1 = not at all, 7 = a great deal). Subjects were also asked if they took economic damages (medical bills, lost wages, etc.) into account, and if so, the extent to which that influenced their decision, on a scale from 1 to 7 (1 = not at all, 7 = a great deal). The questions and scale values were consistent with those used in prior research (Reyna, Hans, et al., 2015).

Participants were given tasks which measure numeracy and thinking styles. Specifically, participants completed the 15-item expanded numeracy scale, which 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 added by Peters et al. (2007), which test familiarity with simple arithmetical operations (e.g. multiplication), basic probability and related ratio concepts (e.g., fractions, decimals, proportions, percentages, and probability), and ability to keep track of class-inclusion relations (Liberali et al., 2012; Reyna & Brainerd, 2008). They also completed the 8-item subjective numeracy scale (Fagerlin, Zikmund-Fisher, Ubel, Jankovic, Derry, & Smith, 2007), which assesses individuals' beliefs about their mathematical skills and preferences for numerical information, and the cognitive reflection test (Frederick, 2005), which tests the ability to inhibit intuitive responses in favor of deliberative responses to mathematical problems. Finally, participants completed the 18-item need-for-cognition scale (Caccioppo, Petty, & Kao, 1984), which measures how much individuals enjoy engaging in effortful thinking. Questions tapping austistic tendencies were included, but are not analyzed in the current article.

Results

Income Measures Were Perceived As More Meaningful than Renovations

We expected that income measures (our meaningful anchor conditions) would be perceived as more meaningful than the courtroom renovations measure (our meaningless anchor condition). This prediction was supported by our results.

Regardless of which anchor was presented to the participant as a manipulation, we asked all participants to rate the meaningfulness of the three different types of anchors – courtroom renovations, median lifetime income, and median annual income, on a scale from 1 (strongly disagree that it is meaningful) to 7 (strongly agree that it is meaningful). A repeated-measures ANOVA with meaningfulness rating as the dependent variable and anchor type (courtroom renovations, median lifetime income, median annual income) as a within-subjects factor revealed a main effect of anchor type ($F(2, 152) = 96.21$, $p < .001$, $\eta_p^2 = .56$). Participants found the median annual income significantly more meaningful than either the median lifetime income

($t(153) = 4.57$, $p < .001$, $d = 0.33$, 95% CI [0.15, 0.51]) or the cost of courtroom renovations ($t(153) = 13.83$, $p < .001$, $d = 1.43$, 95% CI [1.25, 1.61]) ($M_{annual\ income} = 4.52$, $SD = 1.56$; $M_{lifetime\ income} = 3.99$, $SD = 1.67$; $M_{\text{courtroom renovations}} = 2.25$, $SD = 1.63$). The median lifetime income was rated higher than the meaningless courtroom renovations anchor ($t(153) = 11.23$, p < .001, $d = 1.06$, 95% CI [0.88, 1.24]). Both income anchors means are close to the midpoint of the scale, suggesting that they were not considered to be strongly meaningful.

In order to check that participants who received our meaningful anchor found the anchor they saw more meaningful than those who saw our meaningless anchor, we also conducted an ANOVA with "meaningfulness of anchor seen" as the dependent variable. This variable was the meaningfulness rating each participant gave to the anchor that he or she had seen, with higher numbers indicating greater meaningfulness. For instance, for participants who saw an anchor that was described as the median lifetime income, the dependent measure was their rating of the meaningfulness of the median lifetime income anchor. Manipulated anchor meaning (meaningful, meaningless) and the amount of anchor (high, low) were entered into the ANOVA as factors.

This analysis revealed a main effect, as expected, of anchor meaning. Participants found the meaningful anchors (median lifetime income in the high condition and median annual income in the low condition) more meaningful than the meaningless anchors (courtroom renovations) ($M_{\text{meaningful}} = 3.97$, $SD = 1.58$; $M_{\text{meaningless}} = 2.56$, $SD = 1.73$; $F(1, 117) = 22.25$, $p < .001$, $d = 0.85$, 95% CI [0.56, 1.14]). There was no significant main effect of anchor amount ($F(1, 117) = 1.97$, $p = .16$, $d = 0.26$, 95% CI [-0.05, 0.58]) and no significant interaction between anchor meaning and anchor amount ($F(1, 117) = .94$, $p = .56$, $\eta_p^2 = .00$). In sum, both analytic approaches confirmed expectations that the meaningful anchors would be perceived as more meaningful than the meaningless anchors.

Numerical Anchors, and the Amount and Meaning of those Anchors, Influenced Damage Award Decisions

We predicted that numerical anchors, in addition to perceptions of the case and individual differences, would influence damage awards. We also predicted that the influence of anchor on damage award decisions would depend on the numerical amount of the anchor and the meaningfulness of the anchor. Meaningful large anchors should drive awards up, compared to meaningless anchors; in contrast, meaningful small anchors should drive awards down, compared to meaningless anchors. In order to test this prediction, we first conducted an initial ANOVA examining the effects of the manipulated factors (i.e., anchor amount and meaningfulness). We then examined mock jurors' individual differences and perceptions of the case that we planned to control for. Finally, we examined the effect of our manipulations of anchor meaning and amount controlling for these individual differences and case perceptions.

Overall ANOVA examining mean differences. We first conducted an analysis of variance with award amount as the dependent measure, award type (lowest acceptable award, award amount, and highest acceptable award) as a within-subjects factor and anchor type (no anchor, meaningless small anchor, meaningless large anchor, meaningful small anchor, meaningful large anchor) as a between-subjects factor. This analysis revealed a main effect of award type, such that participants gave the lowest amount for their lowest acceptable award (M = \$93,003.72, SE = \$16,852.92), an intermediate amount for their actual award amount (M = \$231,941.26, SE = \$35,188.42), and the highest amount for their highest acceptable award (M = \$528,067.13, SE = \$88,325.41) ($F(2,145) = 20.03$, $p < .001$, $\eta_p^2 = .22$). All of these differences were large (lowest acceptable award – award amount $t(152) = 5.89$, $p < .001$, $d = 0.40$, 95% CI

[0.16, 0.64]; lowest acceptable award – highest acceptable award $t(150) = 5.26$, $p < .001$, d = 0.54, 95% CI [0.31, 0.76]; award amount – highest acceptable award $t(151) = 4.50$, $p < .001$, d = 0.34, 95% CI [0.12, 0.57]. Figure 1 displays the enormous variability in these acceptable awards.

Analyses also revealed a main effect of condition ($F(4, 146) = 5.76$, $p < .001$, $\eta_p^2 =.14$). Participants gave the least in the meaningful small (*M* = \$103,722.28, *SE* = \$96,809.07), meaningless small (*M* = \$105,805.41, *SE* = \$98,464.04), and no anchor conditions (*M* = \$166,815.63, *SE* = \$93,734.98). The meaningless large amount (*M* = \$404,526.88, *SE* = \$95,234.83) was significantly higher than the meaningless small amount ($t(59) = 3.23$, $p = .03$, $d = .78$, 95% CI [0.26, 1.30]) and the meaningful small amount ($t(59) = 3.27$, $p = .03$, $d = 0.79$, 95% CI [0.26, 1.31]). The meaningful large amount (*M* = \$640,816.67, *SE* = \$98,464.04) was significantly higher than the no anchor amount ($t(61) = 2.80$, $p = .001$, $d = .63$, 95% CI [0.13, 1.14]), the meaningful small amount ($t(59) = 3.06$, $p < .001$, $d = 0.74$, 95% CI [0.22, 1.26]), and the meaningless small amount ($t(59) = 3.04$, $p < .001$, $d = .74$, 95% CI [0.22 – 1.26]).

We expected that the raw award mean for the meaningful large award condition would be greater than the meaningless large award mean (and that the meaningful small award condition would be lower than the meaningless small award). Although the means were in the predicted direction, these amounts were not significantly different from each other (before controlling for other factors; see below) ($M_{meaningfulsmall}$ = \$78,483.33, *SD* = \$91,037.55; $M_{meaninglesssmall}$ = \$80,117.33, *SD* = \$105,675.27, $t(58) = .06$, $p = .95$, $d = .02$, 95% CI [-0.49, 0.52]; $M_{meaningfullarge}$ = \$552,433.33, *SD* = \$819.515.66; $M_{meaninglesslarge}$ = \$350,566.67, *SD* = \$429,392.81, $t(120) = 1.10$, $p = .24$, $d = 0.31$, 95% CI [-0.20, 0.82]).

Finally, there was a significant two-way interaction between condition and award type ($F(8, 292) = 2.86$, $p = .004$, $\eta_p^2 = .07$), displayed in Figure 1. (Note: Mauchly's Test of Sphericity showed that the sphericity assumption was violated; correcting for the sphericity violation with the Greenhouse-Geisser correction produced the same significant results.) As one can observe by inspecting Figure 1, and consistent with the significance tests, the three types of awards were more similar to each other in the control condition than in the anchor conditions. In the no anchor condition, there were no differences among award types (although this result just missed statistical significance) ($F(2, 30) = 3.25$, $p = .053$, $\eta_p^2 = .18$). In the small and large anchor conditions, award types differed significantly (Meaningless small $F(2, 27) = 7.66$, $p = .002$, $\eta_p^2 = .36$, Meaningful small $F(2, 28) = 8.46$, $p = .001$, $\eta_p^2 = .38$, Meaningless large $F(2, 29) = 12.95$, $p < .001$, $\eta_p^2 = .47$, Meaningful large $F(2, 27) = 5.13$, $p = .01$, $\eta_p^2 = .28$).

Understanding Individual Differences and Case Perceptions. The descriptive statistics for individual-difference measures and questions about the case are contained in Table 1. In order to understand differences in how jurors perceived and processed the case, we ran a factor analysis (technically, a principal component analysis) using varimax rotation, with individual-difference measures for each subject (objective numeracy, subjective numeracy, cognitive reflection test, and need for cognition) and with the subjects' responses to questions that we asked them about the case (questions about the plaintiff, the defendant, difficulty picking a figure, role of punishment and economic losses in making an award judgment, and motivation and cognitive effort related to the task). A factor analysis using Promax (oblique) rotation was also performed, with similar results; we report the Varimax rotation results because the Varimax rotation isolates unique sources of variance. The results of our analysis using oblique rotation are reported in the the supplemental materials (Table 5 (S2)).

Five independent factors with eigenvalues over 1.0 were extracted, accounting for 59.41% of the variance. Table 2 displays all factor loadings, bolding those above .4.

Interpretations of the factors were as follows: Factor 1 represents numeracy and difficulty picking a figure; Factor 2 represents motivation and effort in determining the award; Factor 3 represents perceptions of the plaintiff; Factor 4 represents perceptions of the defendant; and Factor 5 represents the extent to which punishment and economic loss were taken into account.

Influences on Damage Award Amounts. In order to test for the predicted interaction between anchor meaning and anchor amount while controlling for individual-difference and case perception factors, we ran a linear regression using our experimental manipulations (anchor amount [large, small] and anchor meaning [meaningful, meaningless]), the interaction between anchor amount and anchor meaning, and the case perception and individual-difference factors from our factor analysis to predict damage award amounts.

Table 3 presents the results of this regression. Four variables were statistically significant predictors of participants' award amounts: the interaction between anchor amount and anchor meaning, the perception of the plaintiff; the perception of the defendant; and the extent to which punishment and economic loss were taken into account.

The significant interaction between anchor amount and anchor meaning provided some support for our prediction that meaning would boost the impact of the anchor amount. As Figure 2 illustrates, the meaningful anchor significantly boosted the impact of the large anchor amount, relative to both small anchors. As expected, differences in the meaningless condition were smaller. Indeed, in the meaningless anchor conditions, awards in the small and large anchor conditions did not differ significantly (and small anchors did not differ significantly).

As for the other significant predictor variables, perceptions of the plaintiff and the defendant influenced awards. People who perceived the plaintiff more positively gave higher awards, and those who perceived the defendant more negatively gave higher awards. The participants' ratings about the extent to which they took punishment and economic loss into account also influenced award judgments. People who said they took punishment and economic loss into account gave higher awards. We also ran this regression controlling for the gender of participants. Significant and nonsignificant results remained the same and gender was not a significant predictor (B = -24110.59, SE = 104632.32, β = -.019, t(145) = -.23, p = .82). Therefore, we collapsed across gender in the analyses.

To test the prediction that participants with high numeracy and NC would lead them to be more responsive than other participants to meaningful anchors, we added an interaction term between the Numeracy/difficulty factor and anchor meaning to an expanded regression. However, the interaction term was not significant (B = .07, t (145) = -.56, p = .58).

In response to a reviewer's suggestion, we ran the regression analysis using the square root of the award amount (as opposed to the raw award amount) as the dependent variable, to control for a positive skew in our award amount data: skewness of 3.72 (SE = .22). This analysis produced very similar results, except that the interaction between anchor amount and anchor meaning missed significance (B =188.76, SE = 109.89, β = .24, t (145) = 1.72, p = .09) and the main effect of anchor amount (small/large) was significant (B =163.48, SE = 77.99, β = .24, t (145) = .2.10, p = .04).

Serial Mediation Analysis Regarding Perception of the Severity of Pain and Suffering, Gist Judgments, and Award Amounts

Regression analysis was used to investigate the hypothesis that a participant's categorization of his or her award amount as nil, low, medium, or high would mediate the relationship between perception of the severity of pain and suffering and award amount. When entered separately in linear regressions, the participant's rated severity of pain and suffering was a significant predictor of his or her award amount (*B* = 84634.62, *SE* = 735819.68, β = .21, *t*(152) = 2.36, *p* = .02), and the gist judgment of the award amount was also a significant predictor of the award amount (*B* = 234411.86, *SE* = 61088.33, β = .33, *t*(152) = 3.84, *p* = .001). Participants' pain and suffering perceptions were also significant predictors of the gist judgment (*B* = .14, *SE* = .04, β = .27, *t*(152) = 3.45, *p* = .001).

We then ran a regression examining the effect of severity of pain and suffering rating on award amount controlling for the participant's gist of the award amount as a mediator. In this regression, the severity of pain and suffering rating was no longer a significant predictor of award amount (*B* = 46085.26, *SE* = 33993.24, β = .12, *t*(151) = 1.36, *p* = .18). This suggests that the relationship between the participant's pain and suffering severity rating and the award amount is fully mediated by the gist judgment of the award amount as nil, low, medium, or high. We used the Sobel Test (Sobel, 1982) to test whether gist of award amount significantly mediated the relationship between the pain and suffering rating and the award amount, and results indicated that it did (Sobel Test Statistic = 2.67, *SE* = 12,071.05, *p* = .008).

We also tested an alternative model in which the relationship between severity of pain and suffering and gist judgment of award was mediated by the participant's award amount, as might happen if the participants engaged in post-hoc adjustments following the award judgment. In this model, the award amount was only a partial mediator of gist judgment. When predicting the gist judgment of award using both the pain and suffering severity rating and the award amount, both predictors were significant (*B* = 4.16, *SE* = <.001, β = .29, *t* (151)= 3.39, *p* = .001; and *B* = .10, *SE* = .05, β = .19, *t*(151) = 2.22, *p* = .03 respectively).

Variability in Awards in the "Muddy Middle"

We predicted that jurors who gave an award that they considered to be a "medium" amount would show the most variability in award judgments, while those who gave an award they considered to be "low" or "high" would show less variability in award judgments as predicted by the Hans-Reyna model. Our results supported this prediction. The majority of participants (*n* = 96) characterized their awards as "medium," 2 considered their award "nil," 30 considered their award "low," and 22 considered their award "high" (one participant did not respond and the three who gave zero awards were not included in this analysis).

To test the ordinality prediction, an ANOVA with award amount as the dependent variable and subjective ordinal gist of award (as nil, low, medium, or high) as a between-subjects factor was conducted. Since sample sizes for our groups were unequal and we expected different variances between groups, we confirmed results using the Brown-Forsythe test to ensure results were not influenced by violation of the assumption of homogeneity of variance. This analysis showed that dollar awards differed across these gist categories ($F(3, 146) = 9.05$, $p = .006$, $\eta p^2 = .16$). We conducted a subsequent analysis eliminating the nil group (because of small numbers), again using award amount as the dependent variable and subjective ordinal gist of award (as low, medium, or high) as a between-subjects factor. This analysis also showed that dollar awards differed by the subjective ordinal gist of the award ($F(2,145) = 13.30$, $p < .001$, $\eta p^2 =.16$), as predicted.

Followup pairwise comparisons using Least Significant Difference corrections showed that participants who rated their awards as low gave significantly less than those who considered their awards medium (M_{low} = \$20,524.00, *SD* = \$35,333.02, M_{medium} = \$203,447.92, *SD* = \$342,435.02, $F(1, 124) = 8.47$, $p = .04$, $d = 0.61$, 95% CI [0.19, 1.03]) and those who considered their awards high (M_{high} = \$606,250.00, *SD* = \$796,677.53, $F(1, 50) = 16.29$, $p < .001$, $d = 1.13$, 95% CI [0.53, 1.72]); and participants who considered their awards medium gave significantly less than those who considered their awards high ($F(1, 116) = 13.77$, $p < .001$, $d = 0.88$, 95% CI [0.40, 1.35]). The kurtosis of the distribution for those who considered their awards low was 8.35, the kurtosis of the distribution for those who considered their awards medium was 12.43, and the kurtosis of the distribution for those who considered their awards high was 8.26.

To show the distinctive patterns in the low, medium, and high conditions, Figure 3 displays the percentage of participants who categorized their awards as low, medium, or high and who gave specific dollaw awards (participants who considered their awards nil are not included due to the small sample size). One can observe that awards judged to be low and high are distinctive (clustered at the low and high ends of the dollar awards, respectively), but participants who judged their awards to be medium had a broader and flatter distribution of values.

Higher Numeracy is Associated With More Difficulty Reaching a Damage Award Amount

We predicted that individual differences in need for cognition, cognitive reflection, and numeracy would affect participants' subjective judgments of the difficulty of making an award recommendation, such that greater need for cognition, cognitive reflection, and numeracy would reduce perceived difficulty of deciding on a damage award amount.

To investigate this prediction, we modified our initial factor analysis by removing the participants' ratings of the difficulty they reported determining a pain and suffering award (because that would become the outcome variable). Then, we used the remaining items in a second factor analysis. Five factors with eigenvalues over 1.0 were extracted, accounting for 62.44% of the variance. Table 4 (S1) displays all factor loadings, bolding those above .4. The factors in this factor analysis replicated the five factors from our initial factor analysis, except that factor 2 now represented numeracy only, as the difficulty picking a figure item was no longer included in the analysis. Comparison of Table 2 and Table 4 (S1) shows strong overlap in the factors.

We then entered these factors along with our experimental manipulations (anchor amount and anchor meaning) and the interaction between anchor amount and anchor meaning into a linear regression predicting difficulty reporting an exact figure (1 = extremely easy; 7 = extremely difficult). In this regression, just one predictor was statistically significant – our numeracy factor ($B = .29$, $SE = .10$, $\beta = .25$, $t(145) = 2.80$, $p < .05$). However, the direction of the numeracy effect was surprising. The results showed that as participants became more numerate, they found it *more* difficult to pick an exact figure to compensate the plaintiff for her pain and suffering. We also ran this regression controlling for the gender of participants. Significant and nonsignificant results remained the same and gender was not a significant predictor ($B = .04$, $SE = .27$, $\beta = .02$, $t(144) = .16$, $p = .88$).

Discussion

Summary of the Results

Award judgments varied dramatically as a function of whether numerical anchors were presented. When the anchor amounts were large, simply mentioning a relevant or irrelevant amount of money—as opposed to encouraging relative comparisons as in prior work—nudged awards significantly in the direction of the anchor. Mock jurors' exact numerical judgments of damage awards were malleable, replicating the impact of anchors found in prior research (Chapman & Bornstein, 1996; McAuliff & Bornstein, 2010; Reyna, Hans, et al., 2015).

Information about magnitude of income (annual or lifetime) was perceived as more meaningful to gaining a sense of the size of damage awards than arbitrary values, such as courtroom renovations. Participants generally perceived the gist of their awards as medium rather than low or high, and award amounts for the latter two categories clustered in opposite directions, as expected. Thus, numerical judgments varied but they had some coherence within gist categories of low versus high damages. In contrast, medium awards had a relatively flat distribution, covering a wide range of judgments that differed by orders of magnitude. Serial mediation analysis revealed that gist judgments fully mediated the relationship between the perceived severity of the plaintiff's injury and the damage award amount. Each of these results is in line with our hypotheses and predictions from the gist model of damage awards (Hans & Reyna, 2011; Reyna, Hans, et al., 2015).

Meaningfulness interacted with the size of the anchor (diverging for meaningful small vs. large anchors) once other factors were controlled, such as the perceptions of the plaintiff and defendant and individual differences in numeracy and cognitive control. Validating the meaningfulness manipulation, ratings of meaningfulness differed for the meaningful versus meaningless anchors—meaningful anchors were rated higher—but meaningfulness ratings remained modest on average.

Together with prior research, there are now two studies demonstrating the distinctive effect of meaningful, as opposed to arbitrary, numerical anchors on damage awards. These anchors are meaningful numbers in that they convey a sense of the magnitude of damages in the context of the case. Consistent with this interpretation, effects of anchors accounted for unique variance controlling for case factors. The effect of anchors on award judgments was demonstrated here using a more traditional anchoring manipulation. Results of the current experiment are consistent with a theoretical framework that emphasizes qualitative representations that govern quantitative judgments, in conjunction with independent effects of perceptions of case factors. Although multiple social, political, and personality characteristics of jurors likely contribute to perceptions of the plaintiff's and defendant's cases, these seem to be integrated into two opposing factors that, along with economic damages and desire to punish, determined award judgments. Research varying such characteristics may show additional effects.

We also examined cognitive factors that many theorists have argued should influence numerical judgments and decisions, including objective performance measures of quantitative competence and subjective perceptions of such competence (e.g., Fagerlin et al., 2007); cognitive reflection, which was introduced to capture executive processes involved in inhibiting incorrect numerical responses that correlates with general intelligence (e.g., Frederick, 2005); and need for cognition, which is the desire to actively engage in cognitive processing as distinct from the ability to do so, which has also been associated with superior numerical judgments (e.g., Liberali et al., 2012; Stanovich & West, 2008). Although these measures all loaded on the same numeracy/difficulty factor in principal component analyses, this factor did not predict award judgments. The numeracy/difficulty factor also did not interact with anchor meaning in the regression. Thus, given the same case facts, more cognitively able individuals were not more likely to give smaller, larger, or more variable pain and suffering awards, , suggesting that calls for technically proficient jurors—as opposed to ordinary people—will not necessarily decrease variability and uncertainty in pain and suffering damage awards.

Our results also show that quantitative competence predicts perceived difficulty in award judgments, but in a surprising way opposite to our hypothesis. Reviews of the numeracy literature suggest that higher numeracy should generally improve numerical judgments, although there are important exceptions to this rule (Peters, 2012; Reyna et al., 2009). In a typical civil trial in which jurors might need to calculate losses due to injury or perform other complex numerical assessments, numeracy should increase jurors' competence and confidence in their damage award judgments. However, our project asked participants to assess money damages only for pain and suffering, and we found that numeracy was *negatively* related to perceived difficulty. More numerate people may have recognized the difficulty of making pain and suffering award judgments to a greater extent than less numerate people.

Limitations of the Present Research

The ability to test theoretical ideas in a highly controlled setting comes with some clear limitations. Our sample was limited to college students, and although some prior research reassures us that college students and other groups often respond comparably to mock trial stimuli (Bornstein, 2017; Bornstein, Golding, Neuschatz, Kimbrough, Reed, Magyarics, & Luecht, 2017), we believe it is important to test the theoretical predictions with broader samples in the future. Because of the potential impact of numeracy, we think it is especially important to study samples with broader ranges of subjective and objective numeracy. Despite the level of education of this sample, however, they displayed enormous variability in award judgments.

Although we were able to manipulate the meaningfulness of the anchor, so that low and high meaningful anchors produced significant differences in rated meaning, the “high” meaningful anchor was rated as close to the midpoint rather than at the high end of the meaningfulness scale. It appears that a more powerful manipulation of meaning is needed that better conveys the gist of the magnitude of injury. Rose and Diamond (2017) report that the civil juries whose deliberations they analyzed were more responsive to dollar values that resonated with the specific facts of the case and the plaintiff’s injury than to general dollar amounts drawn from everyday life. In addition, relative comparison of magnitudes—not implemented in this study—may be an important component to achieve larger effects of meaningfulness of anchors, as Reyna, Hans, et al. (2015) found. So both magnitude comparisons and tighter links between suggested dollar values and the case facts might boost the effects of meaningful numbers.

Then, too, despite the fact that the specific dollar anchor amounts in the current experiment were selected based on prior research with a very similar population (Reyna, Hans, et al., 2015), the small anchor conditions did not produce awards that differed significantly from the no anchor control condition. About half the participants thought the damage award should be above and half thought it should be below the small anchor number. That suggests the small anchor number was insufficiently distinctive to participants to create anchoring effects through contrast or assimilation. Using the same case facts but a different population of 35 jury-eligible community members, McAuliff and Bornstein (2010) obtained an average judgment of $61,992 in their no anchor pilot study, a number that is not much different from the $50,000 number we used for our small anchor conditions, but is substantially different from the Reyna, Hans, et al. (2015) average award in the no anchor condition of $224,288. This difference offers another illustration of the variability of what is considered to be an appropriate damage award, and underscores the challenge of empirical research on anchoring and awards.

To allow a clearer test of the hypotheses, we informed respondents that liability had been established, that economic losses had already been covered, and that their job was only to decide on pain and suffering. Focusing on a single damages element is an approach employed in other jury damages research (e.g., McAuliff & Bornstein, 2010; Sunstein et al., 2002). But that approach obviously differs from the situation confronting jurors in real trials, who decide on liability, compensatory economic and noneconomic damages, and sometimes punitive damages. Prior research indicates that there is interactivity among these decisions (Robbennolt & Hans, 2016), so the singular focus on pain and suffering damages, while it avoids confounding, is a limitation of this experiment in terms of ecological validity. Other limitations include the reliance on a relatively short written trial stimulus and the absence of jury deliberation. We hope in future research to address these and other limitations by testing theoretical predictions using broader samples and more realistic jury decision making paradigms.

One final limitation stems from our decision to present the case scenario and independent variables and ask for participants' damage awards first, before obtaining responses on scales measuring individual differences such as numeracy and cognitive style. Although the latter are assumed to be relatively fixed tendencies, we cannot rule out that exposure to the stimulus materials affected participants' responses on the individual difference measures. However, presenting the scales first has the potential to contaminate the experimental manipulations. This issue of order of tasks should be addressed in future research by counterbalancing the individual difference scales and case scenarios.

Directions for Future Research

The results suggest some promising directions for testing Hans and Reyna's model and for exploring practical applications to real-world jury decisionmaking. Pain and suffering, as well as other non-economic damages, are inherently subjective and, we would argue, qualitative in that they fundamentally involve appreciating the magnitude of injury as low or high. The exact dollars that are attached to this appreciation of the extent of injuries will naturally vary in different time periods and jurisdictions, but this contextual flexibility is a strength of using the gist of damages to make award judgments. Consistent with the model, the results of the experiment suggest that numerical award judgments will vary and are subject to biases from minimal manipulations, such as meaningless and meaningful yardsticks. However, the model implies that appreciating the gist of damages as low or high imposes some ordinal coherence in numerical judgments (although perceptions of "medium" damages elicit a wider spread of numbers), and could be the focus of interventions for jurors.

Jurors high in numeracy are likely to perform well with some types of damage award judgments, for example, those requiring mathematical calculations. But mathematics instruction or relying on expert jurors may be less helpful when it comes to assessing pain and suffering and other intangible damages. Thus, future research should continue to explore numeracy, but should also focus on ways to provide insight into the subjective magnitude of low or high damages to imbue numbers with meaning.

Laypeople's perceived difficulty in making awards and confidence in their awards are also important judgments from a policy perspective and worthy of additional study. We discovered a surprising inverse relationship between numeracy and perceived difficulty; no other case or individual factors predicted the perceived difficulty of arriving at a damage award, but more research on this neglected topic is needed. The civil justice system permits judges and jurors without special training to decide on money damages. The just and fair allocation of monetary compensation for wrongful injury is central to the operation of the U.S. civil justice system. When jurors report that they find it difficult, we should look for ways to help them. The current experiment suggests we might explore various approaches to guidance. For example, Ball (2011) recommends that attorneys employ a scaling approach, instructing jurors to differentiate among low, medium and high injuries and guiding jurors about how to match their sense of the severity, length, and consequences of an injury to dollar amounts. Wissler, Saks, Hart, and Evans (1997) also propose presenting different forms of information to jurors to help guide their award decisions.

McAuliff and Bornstein (2010) argue compellingly that "Public perceptions of injustice are bound to increase if like cases are not treated alike by our legal system" (p. 173). Their results showed that damage award recommendations that were objectively identical (a seemingly paltry $10/hour vs. a seemingly excessive $7,300/month) produced opposite effects. As in this experiment, the dollar amounts in civil damages would seem to be easily manipulated. However, our results go beyond the important findings of malleability and variability by indicating that jurors have a strong sense of the magnitude of injury, at least within broad categories of low and high damages, and that they are seeking a meaningful metric by which to adjust dollar amounts to reflect that sense of injury. Jurors may indeed treat “like cases… alike” if similarity is defined with respect to the underlying gist of the case. Such a definition would imply that legal practices and policies should focus on helping jurors translate the qualitative essence of pain and suffering into meaningful numbers, harnassing intuition rather than defeating it.

References

Abbinante v. O'Connell, 277 Ill. App. 3d 1046, 662 N. E. 2d 126 (1996).

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

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. (2004). The impact of different types of expert scientific testimony on mock jurors' liability verdicts. *Psychology, Crime & Law, 10*, 429–446. https://doi.org/10.1080/1068316030001629292

Bornstein, B. H., Golding, J. L., Neuschatz, J., Kimbrough, K., 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

Cacioppo, J. T., Petty, R. E., & Kao, C. F. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment, 48*(3), 306-307. https://doi.org/10.1207/s15327752jpa4803_13

Cantlon, J. F., & Brannon, E. M. (2006). Shared system for ordering small and large numbers in monkeys and humans. *Psychological Science, 17*, 402–407. https://doi.org/10.1111/j.1467-9280.2006.01719.x

Chapman, G. B., & Bornstein, B. H. (1996). The more you ask for, the more you get: Anchoring in personal injury verdicts. *Applied Cognitive Psychology, 10*, 519-540. doi: 10.1002/(SICI)1099-0720(199612)10:6<519::AID-ACP417>3.0.CO;2-5

Daniels, S., & Martin, J. (2015). *Tort reform, plaintiffs' lawyers, and access to justice*.

Lawrence, Kansas: University of Kansas Press.

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

Dunning, D., Heath, C. & Suls, J.M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest*, *5(3)*, 69-106. https://doi.org/10.1111/j.1529-1006.2004.00018.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. *Medical Decision Making, 27*(5), 672-680. doi:10.1177/0272989X07304449

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*(4), 25-42. https://doi.org/10.1257/089533005775196732

Geary, D. C., Berch, D. B., Boykin, A. W., Embretson, S., Reyna, V., & Siegler, R. S. (2011). Learning mathematics: Findings from the national (United States) mathematics advisory panel. In N. Canto (Ed.), *Issues and proposals in mathematics education* (pp. 175-221). Lisbon, Portugal: Gulbenkian.

Gilovich, T., Griffin, D., & Kahneman, D. (2002). (Eds.). *Heuristics and biases: The psychology of intuitive judgment.* Cambridge: Cambridge University Press.

Girvan, E. J., & Merek, H. (2016). Psychological and structural bias in civil jury awards. *Journal of Aggression, Conflict, & Peace Research, 8,* 247-257. https://doi.org/10.1108/JACPR-09-2015-0190

Goodman-Delahunty, J., Granhad, P. A., Hartwig, M., & Loftus, E. F. (2010). Insightful or wishful: Lawyers' ability to predict case outcomes. *Psychology, Public Policy, and the Law, 16*, 133-157. https://doi.org/10.1037/a0019060

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

Hans, V. P. (2000). *Business on trial: The civil jury and corporate responsibility*. New Haven, CT: Yale University Press.

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

Hans, V. P., & Eisenberg, T. (2011). The predictability of juries. *DePaul Law Review, 60,* 375-396.

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

Hans, V. P., & Reyna, V. F. (2011). To dollars from sense: Qualitative to quantitative translation in jury damage awards. *Journal of Empirical Legal Studies, 8* (Special Issue 1), 120-147. https://doi.org/10.1111/j.1740-1461.2011.01233.x

Hastie, R. (2011). The challenge to produce useful "legal numbers." *Journal of Empirical Legal Studies, 8* (Special Issue 1), 6–20. https://doi.org/10.1111/j.1740-1461.2011.01224.x

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

Jacobson, J., Dobbs-Marsh, J., Liberman, V., & Minson, J. (2011). Predicting civil jury verdicts: How attorneys use (and misuse) a second opinion. *Journal of Empirical Legal Studies, 8*

(Special Issue 1), 99-119. https://doi.org/10.1111/j.1740-1461.2011.01229.x

Kahneman, D. (2011). *Thinking, fast and slow*. NY: Farrar, Strauss and Giroux.

Kahneman, D., Schkade, D., & Sunstein, C. R. (1998). Shared outrage and erratic awards: The psychology of punitive damages. *Journal of Risk and Uncertainty, 16*, 49–86. https://doi.org/10.1023/A:1007710408413

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. https://doi.org/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*, 37–44. http://dx.doi.org/10.1177/0272989X0102100105

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. https://doi.org/10.1007/s10979-009-9178-8

McAuliff, B. D., & Kovera, M. B. (2008). Juror Need for Cognition and sensitivity to methodological flaws in expert evidence. *Journal of Applied Social Psychology, 38*, 385-408. https://doi.org/10.1111/j.1559-1816.2007.00310.x

Mott, N. L., Hans, V. P., & Simpson, L. (2000). What's half a lung worth? Civil jurors' accounts of their award decision making. *Law and Human Behavior*, *24*, 401-419. https://doi.org/10.1023/A:1005540229224

Nieder, A. (2016). The neuronal code for number. *Nature Reviews Neuroscience, 17*, 366-382. doi:10.1038/nrn.2016.40

Peters, E. (2012). Beyond comprehension: The role of numeracy in judgments and decisions. *Current Directions in Psychological Science, 21*(1), 31-35. https:/doi.org/10.1177/0963721411429960

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

Reyna, V. F. (2012b). Risk perception and communication in vaccination decisions: A fuzzy-trace theory approach. *Vaccine, 30*, 3790-3797. https://doi.org/10.1016/j.vaccine.2011.11.070

Reyna, V. F., & Brainerd, C. J. (2008). Numeracy, ratio bias, and denominator neglect in judgments of risk and probability. *Learning and Individual Differences, 18,* 89–107. http://dx.doi.org/10.1016/j.lindif.2007.03.011

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

Reyna, V. F., Nelson, W. L., Han, P. K., & Pignone, M. P. (2015). Decision making and cancer. *American Psychologist, 70*, 105-118. http://dx.doi.org/10.1037/a0036834

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

Rose, M., & Diamond, S. (2017, June 21). Patterned variability: What deliberations reveal (and cannot reveal) about "pain and suffering" damage decisions. Paper presented at the International Meeting on Law and Society, Mexico City, MX.

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

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*, 966–972. http://dx.doi.org/10.7326/0003-4819-127-11-199712010-00003

Sherwin, E., Eisenberg, T., & Re, J. R. (2012). *Ames, Chafee, and Re on damages*. NY: Foundation Press.

Siegler, R. S. & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development, 75*, 428-444. https://doi.org/10.1111/j.1467-8624.2004.00684.x

Singer, M., & Remillard, G. (2008). Veridical and false memory for text: A multiprocess analysis. *Journal of Memory and Language, 59*, 18-35. https://doi.org/10.1016/j.jml.2008.01.005

Stanovich, K. E., & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology, 94*, 672-695.

doi:10.1037/0022-3514.94.4.672

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.

U. S. Census Bureau (2016). Household income: 2015. Report number acsbr/15-02. Accessed at https://www.census.gov/library/publications/2016/acs/acsbr15-02.html.

Wissler, R. L., Evans, D. L., Hart, A. J., Morry, M. M., & Saks, M. J. (1997). Explaining "pain and suffering" awards: The role of injury characteristics and fault attributions. *Law and Human Behavior, 21*, 181-207. https://doi.org/10.1023/A:1024878329333

Wissler, R. L., Hart, A. J., & Saks, M. J. (1999). Decisionmaking about general damages: A comparison of jurors, judges, and lawyers. *Michigan Law Review, 98*, 751-772. https://doi.org/10.2307/1290315

Wissler, R. L., Rector, K. A., & Saks, M. J. (2001). The impact of jury instructions on the fusion of liability and compensatory damages. *Law and Human Behavior, 25,* 125-139. https://doi.org/10.1023/A:1005689230013

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

Table 1

Descriptive Statistics for Judgments and Individual Differences

Scale	Mean	*S.D.*	Minimum	Maximum
Severity of injuries	4.71	1.00	2	7
Severity of pain and suffering	4.90	1.25	2	7
Perception of P	4.88	1.13	2	7
Perception of D	3.84	1.02	1	7
Extent D negligent	7.12	1.94	1	11
Extent D caused P's injuries	8.57	2.15	2	11
Difficulty picking figure	5.45	1.17	2	7
Motivation to determine award	4.67	1.39	1	7
Cognitive effort to determine award	4.62	1.24	1	7
Role of punishment	2.61	1.47	1	7
Role of economic losses	3.40	1.92	1	7
Objective Numeracy (α = .55)	12.33	2.06	4	15
Subjective Numeracy (α = .81)	30.27	6.06	13	43

CRT (α = .59)	1.92	1.07	0	3
Need for Cognition (α = .90)	14.10	18.68	-35	57

Note. P = Plaintiff; D = Defendant. CRT = Cognitive Reflection Test. Higher numbers reflect greater severity of injuries and pain and suffering, more positive perceptions of plaintiff and defendant, greater defendant negligence and causation of plaintiff injuries, more difficulty picking figure, more motivation and cognitive effort to determine award, greater role of punishment and economic losses in determining award, and higher scores on scales of Objective Numeracy, Subjective Numeracy, CRT, and Need for Cognition.

Table 2

Factor Solution for Individual Differences and Case Questions

Measures	Numeracy /Difficulty	Motivation/ Effort	Perception of P	Perception of D	Punishment/ Economic Loss
Severity of injuries	-.09	.01	**.80**	.03	.11
Severity of pain and suffering	-.11	.01	**.80**	.03	-.08
Perception of P	.09	.26	**.65**	.11	-.08
Perception of D	.07	.23	.02	**-.84**	.03
Extent D negligent	.07	.30	.14	**.79**	-.12
Extent D caused P’s injuries	.31	.33	.27	**.42**	.02
Difficulty picking figure	**.52**	-.29	.08	.03	-.22
Motivation to determine award	.04	**.88**	.13	.18	-.08
Cognitive effort to determine award	.07	**.86**	.07	-.10	-.11
Role of punishment	-.16	.01	-.07	-.04	**.82**
Role of economic losses	.04	-.17	.03	-.07	**.77**
Objective Numeracy	**.75**	.15	-.02	.01	.05

Subjective Numeracy	**.66**	.03	.09	-.05	.01
CRT	**.60**	.03	-.11	.25	-.20
Need for Cognition	**.52**	.08	-.25	-.05	.02
Eigenvalue	2.82	2.13	1.50	1.34	1.20
-- (% variance)	(18.83)	(14.12)	(10.03)	(8.93)	(8.00)

Note. P = Plaintiff; D = Defendant. CRT = Cognitive Reflection Test. Higher numbers reflect greater severity of injuries and pain and suffering, more positive perceptions of plaintiff and defendant, greater defendant negligence and causation of plaintiff injuries, more difficulty picking figure, more motivation and cognitive effort to determine award, greater role of punishment and economic losses in determining award, and higher scores on scales of Objective Numeracy, Subjective Numeracy, CRT, and Need for Cognition.

Table 3

Multiple Regression Predicting Award Amount

Variables	B	SE	β	*t*
Anchor meaning	-134,036.95	117,043.82	-.13	-1.15
Anchor amount	163,498.95	114,449.81	.16	1.43
Meaning x amount	334,934.40	161,418.08	.29	2.08*
Motivation / effort	57,389.05	40,784.56	.11	1.41
Numeracy / difficulty	-74,052.46	412,111.56	-.14	-1.80
Perception of plaintiff	91,878.80	43,700.12	.17	2.10*
Perception of defendant	130,320.87	410,000.38	.25	3.18*
Punishment / economic loss	132,307.16	40,844.57	.26	3.24*

Note. Anchor meaning: 0 = meaningless; 1= meaningful); anchor amount: 0 = low; 1 = high. Higher numbers indicate greater motivation, greater numeracy and difficulty picking award, more positive perception of the plaintiff and greater perceived severity of pain and suffering, more negative perception and greater perceived negligence of the defendant, and more role for punishment and economic loss in determining award. Regression model summary: $R = .57$, $R^2 = .33$, Adjusted $R^2 = .28$, $SE = 430{,}921.65$. $^*p < .05$.

Table 4 (S1)

Factor Solution for Individual Differences and Case Questions without Difficulty Item

Measures	Motivation / Effort	Numeracy	Perception of P	Perception of D	Punishment / Economic Loss
Severity of injuries	.01	-.07	**.81**	.03	.09
Severity of pain and suffering	.00	-.09	**.82**	.03	-.10
Perception of P	.32	.04	**.61**	.11	-.05
Perception of D	.23	.07	.01	**-.84**	.04
Extent D negligent	.31	.07	.13	**.79**	-.13
Extent D caused P's injuries	**.40**	.25	.22	**.44**	.06
Motivation to determine award	**.88**	.04	.11	.16	-.08
Cognitive effort to determine award	**.87**	.05	.04	-.12	-.10
Role of punishment	-.01	-.14	-.06	-.04	**.81**
Role of economic losses	-.14	.02	.02	-.06	**.79**
Objective Numeracy	.14	**.77**	-.01	.02	.04
Subjective Numeracy	-.03	**.72**	.14	-.06	-.05
CRT	.00	**.63**	-.09	.25	-.23

Need for Cognition	.06	**.54**	-.24	-.04	.01
Eigenvalue	2.81	2.03	1.47	1.31	1.18
-- (% variance)	20.10	14.50	10.49	9.34	8.43

Note. P = Plaintiff; D = Defendant. CRT = Cogntive Reflection Test. Higher numbers indicated greater severity of injuries and pain and suffering, more positive perceptions of plaintiff and defendant, greater defendant negligence and causation of plaintiff injuries, more motivation and cognitive effort to determine award, greater role of punishment and economic losses, and higher scores on scales of Objective Numeracy, Subjective Numeracy, CRT, and Need for Cognition.

Table 5 (S2)

Factor Solution with Oblique (Promax) Rotation for Individual Differences and Case Questions

Measures	Numeracy / Difficulty	Motivation/ Effort	Perception of P	Perception of D	Punishment/ Economic Loss
Severity of injuries	-.09	.09	**.79**	.10	.10
Severity of pain and suffering	-.10	.10	**.80**	.12	-.09
Perception of P	.13	.35	**.67**	.21	-.11
Perception of D	.02	.16	-.04	**-.80**	.08
Extent D negligent	.17	.40	.24	**.84**	-.19
Extent D caused P's injuries	.38	**.42**	.33	**.50**	-.05
Difficulty picking figure	**.51**	-.21	.05	.07	-.26
Motivation to determine award	.13	**.91**	.23	.28	-.11
Cognitive effort to determine award	**.15**	**.86**	.15	.01	-.13
Role of punishment	-.23	-.06	-.09	-.13	**.83**
Role of economic losses	-.05	-.21	-.01	-.15	**.77**
Objective Numeracy	**.76**	.21	-.02	.08	-.03

Subjective Numeracy	**.65**	.09	.07	.01	-.06
CRT	**.64**	.10	-.09	.30	-.28
Need for Cognition	**.51**	.09	-.25	-.03	-.02
Eigenvalue	2.83	2.13	1.50	1.34	1.20
-- (% variance)	18.86	14.17	10.03	8.93	8.00

Note. P = Plaintiff; D = Defendant. CRT = Cognitive Reflection Test. Higher numbers reflect greater severity of injuries and pain and suffering, more positive perceptions of plaintiff and defendant, greater defendant negligence and causation of plaintiff injuries, more difficulty picking figure, more motivation and cognitive effort to determine award, greater role of punishment and economic losses in determining award, and higher scores on scales of Objective Numeracy, Subjective Numeracy, CRT, and Need for Cognition.

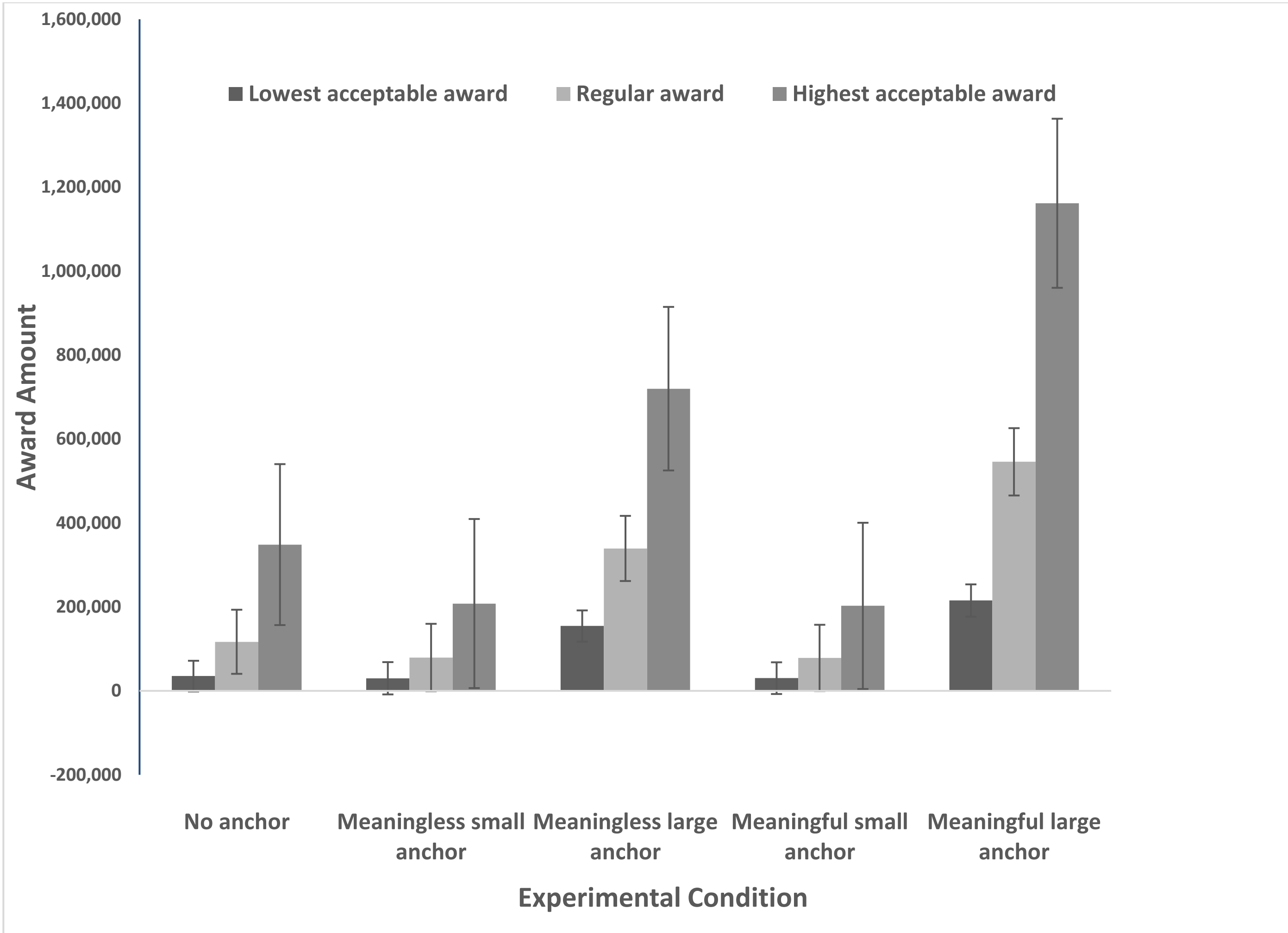

Figure 1. Significant two-way interaction between anchor condition (no anchor, meaningless small anchor, meaningless large anchor, meaningful small anchor, and meaningful large anchor) and award type (lowest acceptable award, regular award, highest acceptable award). Error bars represent + / - 1 standard error.

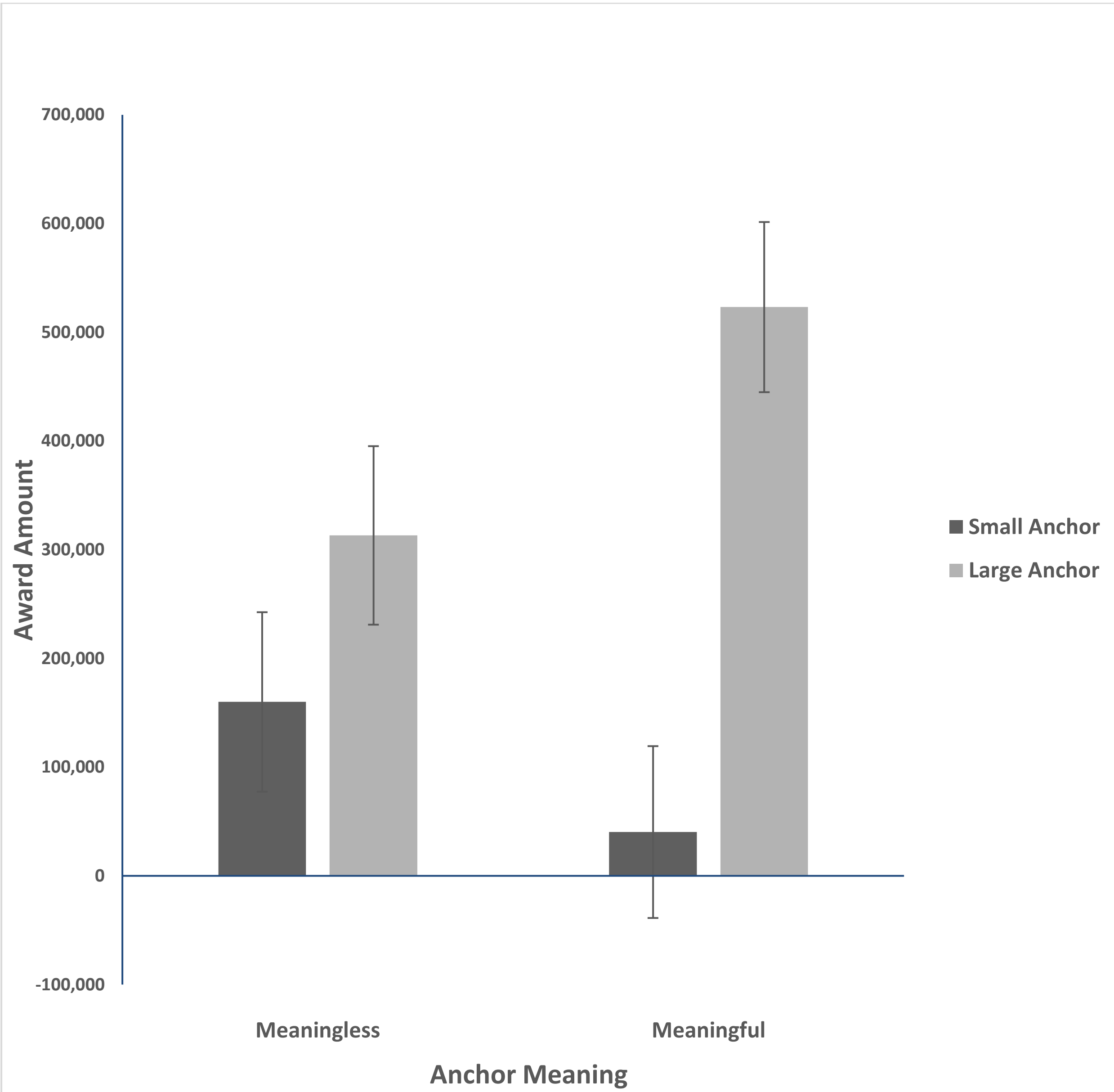

Figure 2. Significant interaction between anchor meaning and anchor amount (controlling for factors representing numeracy / difficulty, motivation / cognitive effort, perception of the plaintiff, perception of the defendant, and extent punishment and economic loss were taken into account). Error bars represent +/- 1 standard error.

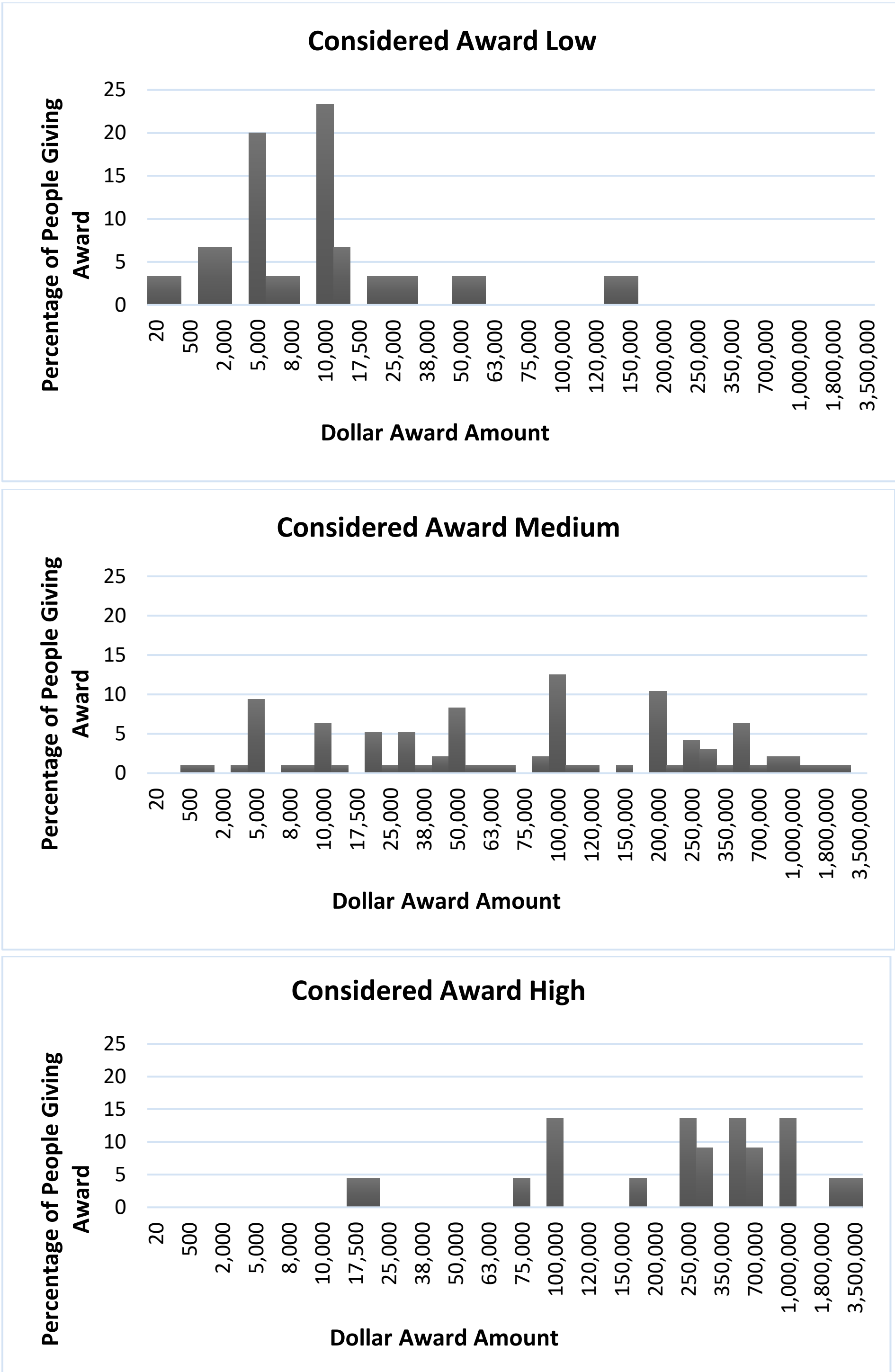

Figure 3. Distribution of awards for participants who considered their award to be low, medium, and high.