Predicting Civil Jury Verdicts: How Attorneys Use (and Misuse) a Second Opinion

Jonas Jacobson, Jasmine Dobbs-Marsh, Varda Liberman, and Julia A. Minson*

When predicting potential jury verdicts, trial attorneys often seek second opinions from other attorneys. But how much weight do they give to these opinions, and how optimally do they use them? In a four-round estimation task developed by Liberman et al. (under review), pairs of law students and pairs of experienced trial attorneys estimated actual jury verdicts. When participants were given access to a partner’s estimates, participants’ accuracy improved in both groups. However, participants in both groups underweighted their partners’ estimates relative to their own, with experienced attorneys giving less weight to their partners’ opinions than did law students. In doing so, participants failed to reap the full benefits of statistical aggregation. In both groups, requiring partners to reach agreement on a joint estimate improved accuracy. This benefit was then largely retained when participants gave final individual estimates. In a further analysis, we randomly sampled estimates of various-sized groups. The accuracy of mean estimates substantially increased as group size increased, with the largest relative benefit coming from the first additional estimate. We discuss the implications of these findings for the legal profession and for the study of individual versus collective estimation.

I. Introduction

A. The Challenge of Predicting Civil Jury Verdicts

In deciding which cases to accept, which cases to settle, and which cases to take to trial, one of the factors that civil trial attorneys consider is the size of the potential jury verdict (Cooter et al. 1982; Gross & Syverud 1996). Accurate estimates of the jury award, provided that both sides make them, promote settlement and avoid the transaction costs of trial (Mnookin & Kornhauser 1979). Because most civil jury trials are personal injury cases (Langton &
Cohen 2008; Moller 1996), the ability to estimate injury awards is particularly relevant. Yet the magnitude of personal injury awards can surprise even experienced trial attorneys.

Awards are difficult to predict largely because of the noneconomic (pain and suffering) component of damages. While economic damages are based on relatively concrete information such as hospital bills and lost salary, jurors receive no specific guidance on how to decide noneconomic damages (Greene & Bornstein 2000). California jury instructions on noneconomic damages state: "No fixed standard exists for deciding the amount of these damages. You must use your judgment to decide on a reasonable amount based on the evidence and your common sense" (Judicial Council of California 2009). Given this vague guidance, jurors understandably struggle with the task and awards vary substantially even for identical injuries (Diamond et al. 1998; Schkade et al. 2000; Wissler et al. 1999).

Confronted with the difficulty of estimating jury verdicts, attorneys often make costly mistakes (Gross & Syverud 1992; Kiser et al. 2008; Rachlinski 1996; see review by Kiser 2010). Kiser et al. (2008) compared settlement offers attorneys rejected to the verdicts the same attorneys achieved at trial. They found that plaintiff’s attorneys rejected a favorable settlement offer in 61 percent of cases, with each “mistake” costing an average of $43,100. Defense attorneys rejected favorable offers in only 24 percent of cases, but at an average cost of $1,140,000.

B. The Value of Seeking Opinions from Other Attorneys

Even for experienced attorneys, it is difficult to accurately predict jury awards by relying on personal experience alone. Across many juries, average awards do vary consistently with injury severity—the more serious the injury, the higher the award (Diamond et al. 1998; Wissler et al. 1999; for a review, see Greene & Bornstein 2003). These data suggest that attorneys may be able to form a reasonable guess based on the average award in similar cases. However, even a highly experienced attorney will have seen a limited number of jury verdicts for any specific injury type, and the task of anticipating how various legally relevant and irrelevant considerations will influence a given award involves, at best, educated guesswork.

Beyond relying solely on personal experience, attorneys may use various other sources of information to improve their predictions of jury verdicts. The most elaborate method of award estimation is running mock trials or focus groups and having jury-eligible citizens deliberate to an award. Because the expense and effort required for these activities can be high, they are used primarily in higher-stakes cases (Posey & Wrightsman 2005). A second method is to study similar cases in a book or online database of past verdicts. The use of this method is limited by the difficulty of finding comparable cases and deciding what differences in facts and venue distinguish the case at hand. The fastest, cheapest, and likely

---

1Judges, the obvious alternative decisionmaker to jurors, also struggle with giving consistent noneconomic awards (Wissler et al. 1999).

the most common method to supplement an attorney’s personal estimate is simply to solicit the opinions of other attorneys. When we surveyed 28 experienced California trial attorneys, 82 percent reported consulting with colleagues either “always” or “frequently” when estimating jury verdicts.

Past research suggests that consulting other attorneys is a wise strategy. Both amateurs and experts can benefit from using the opinion of just one other person in a variety of estimation contexts, including sales forecasts, economic data, historical dates, and political opinions on the Middle East conflict (Liberman et al. under review; Minson et al. in press; Yaniv 2004; for a review, see Bonaccio & Dalal 2006). Accuracy gains are even more impressive when the estimates of more than two people are combined (Armstrong 2001; Yaniv 2004).

But what is the best way to combine estimates from multiple sources? Should one give equal weight to all available estimates or attempt to weight estimates by expertise, confidence, or some other indicator of relative accuracy? When individuals can reliably determine which estimate is more accurate, this question involves the relative statistical advantage of simple averaging versus that of attempting to unequally weight estimates in light of probable differences in accuracy. The ability of attorneys to make judicious weighting decisions is an empirical question to be addressed in our study.

Both weighted and unweighted averaging improve accuracy because in any distribution of estimates at least some pairs of estimates are likely to “bracket” the correct answer, with one estimate being too low and the other too high. When these bracketing estimates are averaged, mean accuracy increases. The greater the frequency of bracketing in a particular sample, the greater the degree to which averaging increases accuracy. The frequency of bracketing is, in turn, greatest when estimation errors are uncorrelated and no systematic biases make overestimates either more frequent or more severe than underestimates (Larrick & Soll 2006; Soll & Larrick 2009; Yaniv & Kleinberger 2000).

C. Why Do People Fail to Give Due Weight to Others’ Inputs?

The task of adjusting one’s own estimates in light of a peer’s estimates is frequently compromised because individuals tend to give too little weight to peers’ estimates. Across different estimation tasks, people working in dyads generally give their own estimates about twice as much weight as those of their partner, and individuals frequently give their partners’ estimates no weight at all. In only a minority of cases do people give their partners’ opinions equal or greater weight than their own (Bonaccio & Dalal 2006; Liberman et al. under review; Minson et al. in press; Soll & Larrick 2009).

A number of factors may be involved in the failure to give due weight to a partner’s input. One is the much documented overconfidence phenomenon, whereby people overestimate the likely accuracy of their judgments (Dunning et al. 1990; Kahneman & Tversky 1982). Another factor may be that people have mental access to the reasons for their own estimates and not to the reasoning of their partner, leading them to believe that their estimates are better founded (Yaniv & Kleinberger 2000). Even when individuals have access to their partners’ reasoning, “naive realism” (Ross & Ward 1995) may assert its influence: individuals assume that to the extent that others disagree with them, those others are subject to various cognitive and motivational biases (Pronin et al. 2004). In previous
research using the methodology employed here, but in a different domain, Minson et al.
(in press) showed that the underweighting of peer estimates is most apparent when
participants’ own estimates and their partners’ estimates differ widely. Moreover, this
underweighting was associated with participants’ belief that their partners’ estimates were
more biased than their own (Liberman et al. under review).

Even when people do not believe their estimates are more accurate or better justified
than those provided by their partners, they still underweight their partners’ estimates
relative to their own. This may be because people are still concerned with self-presentation
or self-perception, feeling that they should trust in themselves or not appear irresolute. The
phenomenon of anchoring (Tversky & Kahneman 1974) may also play a role, in that people
generate their estimates first and then do not adjust sufficiently toward their partners’
estimates. Finally, people may fail to give their partners’ inputs due weight, even when they
have no reason to assume their estimates to be more accurate than their partners’ estimates,
because they do not appreciate the statistical advantages of averaging (Larrick & Soll 2006).

D. The Present Research

Little previous research has examined the weight that experts give to input from peers with
similar expertise. In one study, business students estimated business salaries (Soll & Larrick
2009) and in another, competitive dancers estimated the scores their taped performances
would receive from professional judges (Minson et al. in press). In both studies, people
failed to give due weight to partners’ inputs, and only the business students showed
improvement in accuracy by using their partners’ opinions. However, unlike attorneys
estimating jury verdicts, these participants were not making estimates that are common in
their occupation. The dance score estimates were made on specific dimensions that are not
usually given separate scores. In neither study were participants making estimates that have
important financial implications for their livelihood and the fortunes of their clients.

The present research seeks to examine and compare the estimation accuracy and
estimate-weighting strategies of novices and experts in order to determine the role of a
second opinion in the accurate prediction of civil jury verdicts. We compare the perfor-
mance of attorneys to the performance of law students, who are engaged, but inexperi-
enced novices. We anticipate that the law students’ estimates will be less accurate than those
of practicing attorneys. However, it is an open question whether attorneys will differ in
willingness to yield to a partner’s input and, if so, how this difference will affect the accuracy
of their judgments. It is possible that attorneys have learned the value of averaging opin-
ions. Alternatively, they may have learned how to reliably discriminate the more accurate
estimates and give weight in proportion to presumed relative accuracy. Finally, it is possible
that the increased confidence (and perhaps ego) that comes with being a successful trial
attorney may lead attorneys to unwisely disregard the opinion of even an experienced
colleague.

\[\text{In the Minson et al. study, both dance partners typically overestimated their scores markedly. This systematic bias meant that any improvement in mean accuracy generally required the more optimistic partner to move relatively more in the direction of the less optimistic partner than vice versa.}\]
The current study uses a four-round estimation procedure developed in prior research by Liberman et al. (under review). In this procedure, dyad members first make individual estimates with no interaction with their dyad partner. In a second estimation round, dyad members have the opportunity to view each other’s estimates and revise their own estimates to any degree they consider appropriate. In the third estimation round, dyad members are obliged to agree on a joint estimate through discussion. In the fourth and final round, participants make another set of individual estimates and are free to give their partners’ prior inputs as much or as little weight as they wish.

As noted earlier, prior studies using this paradigm have shown that participants systematically failed to give due weight to their partners’ estimates and paid a price in terms of accuracy. Yet, second-round estimates were (with the notable exception of those provided by the dance partners in the Minson et al. study) more accurate than initial estimates, and agreed-upon joint estimates showed further improvement. Most notably, final independent estimates, although slightly less accurate than third-round (joint) estimates, were still more accurate than Round 2 estimates that were made prior to being compelled to discuss and reach agreement. Our study examined the extent to which jury verdict estimates would follow this pattern and whether the judgments of expert estimators (seasoned attorneys) would differ from those of relative novices (law students).

II. Method

A. Participants

A total of 42 law students (21 pairs) and 28 plaintiff’s personal injury attorneys (14 pairs) participated in the study. The law students were in their second and third year of law school, and few had experience with personal injury litigation. The lawyers were very experienced, reporting an average of 27 years of civil trial practice and an average of 52 civil trials each. To motivate precision and engagement in the task, participants began each round with a bonus of $50, which decreased by $1 for each percentage point that their estimate on each case differed from the actual verdict.

B. Materials

Each participant received a packet with short descriptions of six real personal injury cases tried in California. The cases ranged in severity from a car accident with minor injuries (a $35,000 verdict) to a wrongful death suit in which a five-year-old girl died from injuries due to sexual molestation (a $4,250,000 verdict). Each case description included the county the case was tried in, the basic facts of the incident, a short description of the plaintiff’s injuries and prognosis, and any economic damages (medical bills and lost wages) awarded. The case descriptions were found using VerdictSearch, a publisher of trials and

---

4Complete case descriptions can be found in the Appendix.
settlements. VerdictSearch summaries describe the parties, the major issues at trial, the plaintiff’s injuries, and the damages awarded. Multiple researchers have confirmed the reasonable accuracy of VerdictSearch reports (for a review, see Kiser et al. 2008:556).

To reduce the ambiguity of the task, we used only cases that resulted in nonzero verdicts for the plaintiff and that had no major disputes about the extent of the plaintiff’s injuries. The size of each verdict was judged to be relatively average by an independent group of experienced plaintiff’s attorneys.

C. Procedure

Participants were each paired with a partner to form dyads. They were told that all cases were verdicts for the plaintiff, and they were asked to estimate the full noneconomic award in each case, without any comparative negligence. For each case, participants made four rounds of estimates, following the estimation procedure established in Liberman et al. (under review). In Round 1, they made individual estimates for each verdict. They then had the opportunity to examine their partners’ estimates, without any discussion, and revise their initial estimates to any degree they considered appropriate (Round 2). In Round 3, each dyad was given 15 minutes to discuss and agree on a single joint estimate for each case. Finally, in Round 4, participants made a final set of individual estimates that could be the same or different from the estimates they had agreed on with their partners. In each round, participants indicated their confidence that their estimates were within 10 percent of the actual verdict, using a scale of 1 (not at all confident) to 5 (completely confident).

D. Measuring Estimation Error

To allow us to combine data from different cases in which both actual awards and estimates varied widely, the actual award amounts and participants’ estimates were transformed to base 10 logs. Estimation error was measured as the absolute deviation of the log of the estimate from the log of the actual award amount. For example, in a case where the actual verdict was $100,000, an estimate of $10,000 and an estimate of $1,000,000 would both produce a log error of 1. Our error measure allowed us to correct for the positive skew in error caused by the fact that damage awards are bounded at 0 at the low end but have no upper bound (Wissler et al. 1997:187).


6In California, the jury can freely divide responsibility between plaintiffs and defendants (Judicial Council of California 2009:Instruction No. 405). This system is known as “pure comparative negligence.” For example, a jury might assign 70 percent negligence to the defendant and 30 percent negligence to the plaintiff. The jury’s full damage award is then reduced by 30 percent to reflect the plaintiff’s share of responsibility. For a fuller discussion, see Schwartz (1978).

7For law students, confidence ratings were almost uniformly extremely low (M = 0.80, SD = 0.80). Thus it would be unreasonable to draw conclusions regarding the relationship between confidence and estimation accuracy or the relationship between confidence level and the distance students yielded to their partners’ estimates. The attorneys were somewhat more confident (M = 1.53, SD = 0.88). However, as in prior research, the relationship between confidence and either accuracy or yielding was not significant.
Despite clear instructions that all verdicts were nonzero awards for the plaintiff, three law students and one attorney estimated a verdict of zero dollars in one case. For purposes of analysis (and because we cannot assign a log value to a zero), we “winsorized” the data by replacing the zero estimates with the next lowest estimate offered by any other participant for that case (Wilcox 2010).

III. Results

A. The Accuracy of Individual Estimates

As expected, attorneys’ Round 1 estimates were substantially more accurate than the law students’ estimates. However, large estimation errors were the rule rather than the exception for both groups. For law students, the mean estimation error across the six cases ($M = 0.828$) was significantly higher than for attorneys ($M = 0.391$), $t(68) = 5.5$, $p < 0.001$. To provide a sense of the dollar value of this error, again consider a hypothetical $100,000 jury verdict. For this verdict, a mean estimation error of 0.828 is equivalent to a “too low” estimate of about $15,000 or a “too high” estimate of about $670,000. The attorneys’ mean estimation error of 0.391 corresponds to erroneous estimates of about $41,000 or $244,000.

Underestimates were more common than overestimates for both groups of participants (see Table 1). However, this bias was more pronounced for law students than for attorneys. For law students, 0.8 percent of estimates were exactly correct, 87.7 percent were too low, and only 11.5 percent were too high.\(^8\) By contrast, the attorneys’ estimates were

Table 1: Law Students’ and Attorneys’ Initial Accuracy in Estimating Jury Verdicts

<table>
<thead>
<tr>
<th>Case</th>
<th>Actual Verdict**</th>
<th>Mean Round 1 Estimate</th>
<th>Mean Round 1 Error (Logs)***</th>
<th>Percent of Underestimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Students</td>
<td>Attorneys</td>
<td>Students</td>
</tr>
<tr>
<td>1</td>
<td>$35,000</td>
<td>$16,214</td>
<td>$27,565</td>
<td>0.664</td>
</tr>
<tr>
<td>2</td>
<td>$357,000</td>
<td>$276,681</td>
<td>$439,286</td>
<td>0.695</td>
</tr>
<tr>
<td>3</td>
<td>$4,250,000</td>
<td>$1,333,262</td>
<td>$1,637,963</td>
<td>0.959</td>
</tr>
<tr>
<td>4</td>
<td>$516,250</td>
<td>$702,310*</td>
<td>$568,929</td>
<td>0.901</td>
</tr>
<tr>
<td>5</td>
<td>$2,000,000</td>
<td>$1,184,643</td>
<td>$1,949,107</td>
<td>0.780</td>
</tr>
<tr>
<td>6</td>
<td>$145,000</td>
<td>$56,080</td>
<td>$183,321</td>
<td>0.827</td>
</tr>
</tbody>
</table>

*More than 80 percent of the law students underestimated the actual award. Despite this fact, the mean estimate for all law students was higher than the actual award. This reflects the fact that two respondents grossly overestimated the verdict (with estimates of $3,650,000 and $20,000,000), pulling up the mean estimate. The median estimate was $90,000, a substantial underestimate.

**The actual verdicts are the noneconomic portion of the full award in each case. Participants were asked to estimate only the noneconomic portion of each award.

***Mean Round 1 error is the mean of the absolute deviations of the log of each Round 1 estimate from the log of the actual verdict in the relevant case.

\(^8\)Six of the exactly correct estimates (five by attorneys and one by a law student) were for Case 5, which had a “round number” verdict of $2,000,000; the other three exactly correct estimates, all made by attorneys, were for Case 1, a verdict of $35,000.
exactly correct in 4.2 percent of estimates, too low in 63.7 percent of estimates, and too high in 32.1 percent of estimates. Compared to law students, attorneys had both significantly fewer underestimates, $t(68) = 4.4, p < 0.001$, and significantly more overestimates, $t(68) = 4.4, p < 0.001$.

B. Weight Given to Partners’ Initial Estimates

In Round 2, after examining their partners’ initial estimates, participants had the opportunity to revise their own initial estimates, giving their partners’ estimates as much or as little weight as desired. We classified participants’ Round 2 estimates in terms of whether the individuals went “halfway” toward their partners’ Round 1 estimates, “less than halfway,” or “more than halfway.” We further noted the percentage of Round 2 estimates on which participants opted to “stand pat” (give their partners’ initial estimates no weight at all) or actually move farther away from their partners’ estimates. Finally, we noted the percentage of estimates on which participants moved “all the way” toward or beyond their partners’ estimates (giving no weight at all to their own estimates).\(^9\)

Table 2 shows that both law students and seasoned attorneys gave less weight to their partners’ initial estimates than their own estimates. However, this phenomenon was much more pronounced among the attorneys, who moved exactly halfway toward their partners’ estimates in only 5.5 percent of their estimates, less than halfway in 82.6 percent of their estimates, and more than halfway in only 11.9 percent of their estimates. By contrast, law students moved exactly halfway toward their partners’ estimates in 9.1 percent of their estimates, less than halfway in 65.0 percent of their estimates, and more than halfway in 25.9 percent of their estimates. When we examined the difference between the average frequency with which attorneys went less than halfway versus more than halfway, and contrasted it with the corresponding difference for the law students, the between-group difference was significant, $t(68) = 3.17, p < 0.003$.

Table 2: Relative Frequencies of Yielding Strategies Used by Law Students and Attorneys when Estimating Jury Verdicts

<table>
<thead>
<tr>
<th>Yielding Strategy*</th>
<th>Students</th>
<th>Attorneys</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stood pat</td>
<td>27.9%</td>
<td>53.1%</td>
</tr>
<tr>
<td>Moved less than halfway**</td>
<td>65.0%</td>
<td>82.6%</td>
</tr>
<tr>
<td>Moved exactly halfway</td>
<td>9.1%</td>
<td>5.5%</td>
</tr>
<tr>
<td>Moved more than halfway***</td>
<td>25.9%</td>
<td>11.9%</td>
</tr>
<tr>
<td>Moved all the way</td>
<td>4.4%</td>
<td>3.0%</td>
</tr>
</tbody>
</table>

*Yielding strategy classifies how far law students and attorneys moved toward their partners’ estimates, and thus reflects how much they yielded to their partners’ opinions.

**Including participants who stood pat.

***Including participants who moved all the way.

\(^9\)On five pairs of estimates, dyad partners offered identical first-round estimates. We excluded these data points from our analyses of partner influence, but retained them in our analyses of estimation accuracy.
Furthermore, attorneys were more likely to completely ignore their partners’ estimates than were law students. While law students opted to stand pat, giving no weight to their partners’ judgments, in only 27.9 percent of their estimates, attorneys chose to ignore their partner in a staggering 53.1 percent of their estimates.

As an additional measure of partner influence, we calculated the mean distance participants moved toward their partners’ Round 1 estimates. Overall, attorneys yielded an average of 19.9 percent of the distance between their own and their partners’ estimates. This figure was significantly short of the 50 percent average movement that would indicate a general willingness to give their partners’ inputs as much weight as their own, $t(27) = 11.3, p < 0.001$. Law students yielded an average of 33 percent of the distance toward their partners, a figure that was still significantly less than 50 percent, $t(41) = 6.3, p < 0.001$, but significantly more than the 19.9 percent figure for the attorneys, $t(68) = 3.3, p < 0.002$.

C. Round-to-Round Changes in Accuracy

In the following subsections we examine changes in accuracy for law students and attorneys across the four estimation rounds (see Figure 1).

**Figure 1:** Round-to-round changes in mean accuracy of law students’ and attorneys’ jury verdict estimates.

*Estimation error is the mean of the absolute deviations of the log of each estimate from the log of the actual noneconomic verdict in each case, across all six cases.

**The “R1: Dyad Mean” is the mean estimation error (in logs) that would result if all participants estimated the jury verdicts by using the average of their own Round 1 estimates and their partners’ Round 1 estimates.

Notes: This figure tracks law students’ and attorneys’ error in estimating jury verdicts across four rounds of estimation. In Round 1, participants made individual estimates. In Round 2, participants viewed the estimates of a partner and could choose to revise their own estimates to any degree they wished. In Round 3, participants negotiated joint estimates with their partners, and in Round 4, participants made final individual estimates.
1. The Benefits of Exposure to a Partner’s Initial Estimate

a. Law Students. Law students were able to significantly reduce estimation error from Round 1 ($M = 0.828$, $SD = 0.396$) to Round 2 ($M = 0.720$, $SD = 0.342$), $t(41) = 3.09$, $p < 0.005$. For a $100,000$ verdict, this improvement would correspond to a change from a “too high” estimate of $673,000$ to one of $532,000$, or from a “too low” estimate of $15,000$ to one of $19,000$.

To test whether law students’ improvement was due to the statistical benefits of averaging or an ability to discern which estimate was likely to be more accurate and assign weights accordingly, we categorized, for each pair of estimates, which estimate was more accurate (the “closer” partner) and which was less accurate (the “farther” partner). We then tested a hierarchical linear model using the percentage of distance yielded to partner as the dependent variable and whether the participant was the closer or the farther partner as the trial-level predictor variable. We found that the distance participants yielded to their partner did not vary significantly as a function of the relative accuracy of initial estimates ($z < 1$, ns).

Law students made some gains in accuracy as a result of giving some weight to each other’s estimates. However, in line with our prediction, they could have done better by consistently averaging their own and their partners’ estimates. Mean error for dyad members’ Round 2 estimates was significantly higher than the error that would have resulted had all dyad members simply moved exactly halfway toward their partners’ estimates ($M = 0.633$, $SD = 0.300$), $t(20) = 3.83$, $p < 0.002$.

b. Attorneys. Similarly to law students, the accuracy of attorneys’ estimates improved significantly from Round 1 ($M = 0.391$, $SD = 0.171$) to Round 2 ($M = 0.323$, $SD = 0.111$), $t(27) = 3.63$, $p < 0.002$. In a hypothetical $100,000$ case, attorneys improved from an estimate of $246,000$ to one of $210,000$, or from an estimate of $41,000$ to one of $48,000$. The magnitude of this improvement is notable given the relatively small amount attorneys yielded toward the initial estimates of their partners.

To assess the extent to which this increase in accuracy reflected attorneys’ ability to give more weight to the estimates of the more accurate partner, we again created a dichotomous variable indicating which partner’s estimate was closer to the actual jury verdict. We then tested a hierarchical linear model using the percentage of distance that participants yielded toward their partners’ estimates as the dependent variable and whether the participant’s estimate was the closer or farther one as a dichotomous trial-level predictor variable. This model revealed that while the less accurate partner moved 24.2 percent of the distance between the two estimates, the more accurate partner yielded only 16.0 percent of the relevant distance ($B = -0.082$, $z = -1.98$, $p < 0.05$). This modest but significant difference attests to the fact that attorneys, in contrast to law students, showed some ability to discriminate more accurate estimates from less accurate ones and weight them accordingly.

However, similarly to law students, attorneys could have reduced their error even further (to $M = 0.283$, $SD = 0.081$) by simply averaging their own and their partners’ esti-
mates, *t*(13) = 2.49, *p* < 0.03.\(^{10}\) That is, the statistical gains in accuracy achievable by simple averaging outweighed whatever benefit some pairs of attorneys achieved by assessing which partner was more accurate.

2. Effect of “Bracketing” on Accuracy Improvement

In considering the round-to-round changes in accuracy shown by law students and attorneys, it is useful to distinguish between instances where the dyad members’ initial estimates fell on opposite sides of the actual jury award (i.e., “bracketed”) and instances where both estimates were either too low or too high. In instances of bracketing, mean accuracy generally improved when dyad members moved toward each other’s estimates. In the absence of bracketing, increases in mean accuracy occurred only to the extent that the initially less accurate partner yielded relatively more than the more accurate partner.

Because law students tended to underestimate the relevant award, their bracketing rate was significantly lower (21.4 percent of initial estimates) than that of the attorneys (33.3 percent of initial estimates), *t*(68) = 2.16, *p* < 0.04. Thus, attorneys stood to benefit relatively more than law students from yielding to their partners’ estimates. However, attorneys failed to fully benefit from their higher bracketing rate because they yielded to each other’s estimates less than the more accurate partner.

3. Effect of a Second Opinion on the Elimination of Extreme Errors

Access to a second opinion was particularly beneficial in cases when initial estimates were particularly erroneous. Such estimates may have resulted from idiosyncratic assessments of the magnitude of the injury, a misunderstanding of jurors’ likely perceptions of the case, or misreading of the case facts. For our analysis, we first identified the largest estimation error made by law students and attorneys in each of the six cases and then determined how much each error was reduced as the result of exposure to peers’ estimates.

For law students, the relevant decrease was from \(M = 2.318\) in Round 1 to \(M = 1.717\) in Round 2.\(^{11}\) On a $100,000 case, this is equivalent to a decrease in an estimate from $20,800,000 to $5,200,000, or an increase from $480 to $2,000. Examination of the Round 1 versus Round 2 estimates made by the partners of those law students who produced the most erroneous estimates revealed a modest increase in the relevant errors (from \(M = 0.902\) in the first round to \(M = 1.400\) in the second round). However, this increase was far smaller than the corresponding decrease in error shown by their initially highly erroneous partners. In other words, the dyad (as a pair) fared better after being exposed to each other’s opinion.

\(^{10}\)To compare error on the second round of the task to the error that would have resulted from dyad members averaging their first-round estimates, the data were collapsed to the level of the dyad, with the degrees of freedom reflecting the number of dyads in the study, minus one.

\(^{11}\)In Case 1, there was a tie for the worst estimate, with two participants offering estimates of $400 on a case with a $35,000 verdict. In Case 2, there was a tie with four participants offering an estimate of $10,000 on a case with a verdict of $357,000. In both instances we used the average error of these participants’ errors in subsequent rounds as the error of the “worst estimate.”
For attorneys,\textsuperscript{12} the largest errors were smaller than those for the law students. However, we again see a marked reduction in the magnitude of such errors, from $M = 1.079$ in Round 1 to $M = 0.556$ in Round 2. For a $100,000$ verdict, this is equivalent to a change from an estimate of about $1,200,000$ to one of about $360,000$, or to a change from an estimate of about $8,300$ to one of about $28,000$. Furthermore, the mean error made by the partners of those making the largest initial errors hardly increased from Round 1 to Round 2 (from $M = 0.233$ to $M = 0.315$), suggesting that the more accurate attorneys recognized that their partners’ estimates were highly unrealistic and accordingly gave them little weight in revising their own estimates.

4. Benefits of Discussion and Reaching Agreement

In the third round of the task, each dyad was asked to discuss each case and agree on a single joint estimate. In this round, improvements in accuracy could reflect two (potentially competing) mechanisms. Discussion could allow dyad members to determine which partner was likely to be more accurate and give that partner’s input much greater weight in making their joint estimate. Alternatively, the requirement to reach agreement could force dyad members to bridge the gap between their initial estimates, giving them a chance to better reap the benefits of statistical averaging.

\textit{a. Law Students.} Law students’ mean estimation error in Round 3 ($M = 0.665$) was significantly lower than in Round 2 ($M = 0.72$), $t (20) = 2.70$, $p < 0.02$. The mean Round 3 error was not significantly different than the mean error that law students would have shown ($M = 0.633$) had they simply averaged their Round 1 estimates, $t (20) = 1.26$, $ns$.

\textit{b. Attorneys.} Attorneys similarly showed a significant decrease in mean error, from $M = 0.323$ in Round 2 to $M = 0.288$ in Round 3, $t (13) = 2.47$, $p < 0.03$. Like law students, attorneys’ mean Round 3 error was not significantly different than the mean error that would have resulted ($M = 0.283$, $SD = 0.081$) had the dyad members simply averaged their Round 1 estimates, $t (13) = 0.60$, $ns$.

These data confirm that making dyad members discuss and agree on a joint estimate results in significant accuracy gain. However, this accuracy gain was roughly the same as the increased accuracy that would have resulted from simple averaging. That is, whatever advantage dyad members derived from being able to weight their two prior estimates differentially on the basis of impressions about likely accuracy was roughly equal to the statistical benefits that could be achieved by simple averaging. This was true both for the amateur estimators (law students), whose initial estimates were extremely inaccurate and

\textsuperscript{12}Case 3 (in which the actual verdict was $4,250,000$) presented an eight-way tie for the least accurate estimate (seven participants estimated $250,000$ and the estimate of one participant who offered a zero estimate was winsorized to $250,000$). We used the average estimation error of these eight individuals on each of the subsequent rounds.
generally too low, and the more expert estimators (attorneys), whose initial estimates were comparatively more accurate and more evenly distributed between underestimates and overestimates.

5. Retained Benefits of Reaching Agreement

In the fourth and final round of the task, dyad members made individual estimates and were once again free to give as much (or as little) weight as they wished to their Round 3 joint estimates and discussion. Our analyses compared Round 4 and Round 3 accuracy, as well as Round 4 and Round 2 accuracy, with the latter comparison providing evidence for the retained benefits of reaching agreement with a dyad partner.

a. Law Students. As is apparent in Figure 1, the accuracy of law students’ estimates was slightly reduced in Round 4. The mean error on Round 4 \((M = 0.675)\) was slightly larger than in Round 3 \((M = 0.665)\), \(t = 0.66, ns\), but still significantly smaller than in Round 2, \((M = 0.720)\), prior to the requirement to reach agreement, \(t(20) = 2.17, p < 0.05\).

b. Attorneys. The mean estimation error for attorneys in Round 4 \((M = 0.293)\) was also only slightly higher than in Round 3 \((M = 0.288)\), \(t(13) < 1.00, ns\), and significantly lower than in Round 2 \((M = 0.323)\), \(t(27) = 2.03, p = 0.05\).\(^{13}\)

Once again, amateur estimators and experts showed the same patterns of change in estimation error. Both groups retained most of the benefits they derived from the requirement to reach agreement.

D. The Wisdom of the Crowd Versus the Wisdom of the Dyad

In addition to assessing how attorneys and law students can benefit from averaging their estimates with those of a single peer, we also examined how much participants’ estimates could improve by averaging many estimates and tapping into the “wisdom of crowds” (Surowiecki 2005). To address this question, we created “statisticized groups” (Lorge et al. 1958) by randomly aggregating the estimates made by each set of participants (law students and attorneys) into groups of 2, 4, 6, 8, 10, and so on, all the way to the total available sample (42 for law students and 28 for attorneys). We randomly sampled these group sizes with replacement, with 1,000 iterations of each sample size. We then averaged the Round 1 estimates of each sample and calculated the mean estimation error for each using the log transformation procedure described above.\(^{14}\)

\(^{13}\)Because both the second-round and final estimates were made individually, estimation error was calculated as the average error across six cases for each participant. Thus, like in the comparisons of Round 1 and Round 2 accuracy, the degrees of freedom reflect the number of participants in the study, minus one.

\(^{14}\)Because there was only one way to aggregate the estimates of all 42 law students and all 28 attorneys, the mean error presented in Figure 2 is the same as the mean error for the actual groups.
1. Law Students

Before any aggregation, the mean error for law students was 0.806—equivalent in the case of a $100,000 verdict to an estimate of $640,000 or one of $16,000. As illustrated in Figure 2, aggregating law student estimates into increasingly larger groups consistently reduced error, dramatically at first, then more gradually as the group sample size increased. For example, for groups of four students, the mean error fell to 0.510, equivalent to an estimate of $324,000 or $30,000 on a $100,000 case. For groups of eight, it fell to 0.436, equivalent to an estimate of $272,000 or one of $37,000. With the estimates of all 42 law students aggregated, the mean estimation error was 0.292, equivalent to an estimate of $196,000 or $51,000 on a $100,000 case.\footnote{Estimates based on the mean of 42 estimates yielded errors that were substantially lower than the mean error of individual estimates, despite the low bracketing rate for these law students. This is in part because while most individuals underestimated the award to the plaintiff, a small minority of law students grossly overestimated it, pulling up the mean estimate.}

2. Attorneys

Before aggregation, the average error for attorneys was 0.387. For a verdict of $100,000, this is equivalent to an estimate of $244,000 or $41,000. As illustrated in Figure 2, attorneys’ mean error also declined as the level of aggregation increased. Again, error fell dramatically at first and then more gradually. For groups of four attorneys, average estimation error dropped to 0.228, which on a $100,000 case is equivalent to an estimate of about $169,000 or $59,000. For groups of eight, the average error was 0.188, equivalent to an estimate of $154,000 or $65,000. For all 28 attorneys, estimation error was 0.130, equivalent to an estimate of $135,000 or $74,000, estimates within roughly 30 percent of the hypothetical $100,000 verdict.

Figure 2: Error in estimating jury verdicts as a function of the number of averaged estimates of law students and attorneys.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure2.png}
\caption{Error in estimating jury verdicts as a function of the number of averaged estimates of law students and attorneys.}
\end{figure}
In addition to attesting to the statistical value of aggregating estimates, Figure 2 illustrates two additional insights. First it shows that the mean estimate of 15 (or more) neophyte law students proved more accurate than a single experienced attorney. Second, it demonstrates the value of soliciting even a small number of additional opinions. Consistent with previous findings on aggregating expert estimates (Armstrong 2001), the relative benefit of additional opinions tailed off as group size increased. For the attorneys, averaging with a single peer reduced the mean error (in log terms) by more than a third of the error reduction that could have been achieved by averaging the estimates of the entire sample. In other words, the biggest “bang for the buck” in reducing error came from averaging with just one other individual.

IV. General Discussion

Our findings illustrate the value of giving weight to the opinion of a peer when estimating jury verdicts—a task with important professional and financial implications for practicing attorneys. Both experienced attorneys and law students improved their estimates significantly when exposed to their partners’ estimates. This improvement was partly due to some ability, especially among the attorneys, to discriminate which partner was more accurate and to adjust accordingly, particularly in cases where one of the two estimates was highly erroneous, but it mostly reflected the improvement that occurred when dyad members’ initial estimates “bracketed” the correct verdict and they adjusted their estimates at least somewhat in the direction of their partner. Relying on a second opinion in Round 2 was particularly valuable for participants whose initial estimates were highly erroneous. The hugely mistaken partner was able to avoid a catastrophic error, while the initially more accurate partner was left only slightly worse off.

However, both law students and attorneys failed to fully capture the value of a second opinion, giving less weight to their partners’ estimates than their own. Attorneys, who had more to gain from averaging because their estimates were more likely to bracket the actual jury verdict, actually yielded less to their partners than law students. Both groups could have done better through a strategy of simple averaging.

Encouragingly, our findings suggest that attorneys could benefit more from a second opinion by discussing their estimates with a peer until they reach agreement. On the third round of the task, when partners were compelled to agree on a joint estimate, they were able to achieve the accuracy attainable through simple averaging. Importantly, the benefits of having reached agreement were largely maintained in Round 4, when attorneys were once again free to give each other’s prior inputs as much or as little weight as they wished. This suggests that even if two attorneys reach agreement on a potential jury verdict and then return to their respective offices, they will carry most of the benefit of agreement with them.

Though having access to partners’ opinions significantly increased the accuracy of law students’ estimates, it could not fully make up for the vast experience difference between law students and practicing attorneys. Law students were significantly more willing to give weight to each other’s estimates than were the experienced attorneys, yet throughout the four estimation rounds their estimates remained far less accurate. In addition to
reflecting greater initial error, this also reflects the fact that law students were more prone than attorneys to underestimate the relevant verdict, resulting in a low bracketing rate. Finally, this phenomenon attests to the fact that the attorneys, unlike the law students, were somewhat able to more consistently give greater weight to the more accurate estimates. This is consistent with research demonstrating that those who are skilled in a domain are better able to assess their own competence relative to that of others (Kruger & Dunning 1999).

Our results also demonstrate the substantial benefits of seeking more than a single additional opinion, that is, the “wisdom of crowds” (Surowiecki 2005). When we averaged the estimates of randomly generated groups of different sizes, the estimation accuracy of both attorneys and law students improved dramatically as group size increased. The reduction in error was clear even in the case of law students, who were collectively biased toward underestimating verdicts. In fact, averaging the estimates of 15 student participants produced an estimation error that was smaller than the mean estimation error of the individual attorneys in our sample.

Though the accuracy gains from aggregating many estimates were impressive, it is important to note that the greatest relative benefit came from the first additional opinion. This is consistent both with statistical principles and with previous research on the aggregation of expert estimates demonstrating that the relative benefits diminish as the number of estimates increases (Armstrong 2001).

The benefits of collaboration and the accuracy of the crowd are particularly impressive given the brevity of the case descriptions. Attorneys had little or no information about the medical experts who testified at trial, the likability of the plaintiff, or the details of the plaintiff’s treatment and disabilities. The only information they had about the jury that decided the case was the county where the case was tried. Yet they were collectively able to estimate the size of the awards within approximately 30 percent of the actual verdicts.

Overall, our results show the potential real-world relevance of previous findings on the benefits of access to a second opinion. They also show the failure of individuals to fully reap those benefits by not giving sufficient weight to second opinions. Whereas the previous research we reviewed, with one exception, involved estimates of business facts, public opinion, and other matters irrelevant to the professional aspirations of the estimators, the present studies involved judgments in the domain most relevant to the present or future careers of our participants.

Our results also demonstrate that professional experience in a domain where numerical estimation is common does not lead to a greater appreciation for the value of averaging. Compared to novice law students, experienced attorneys were more inclined to underweight or completely disregard the input of a peer. Clearly, attorneys’ experience does not help them recognize the general benefits of giving more weight to the judgments of others. Law students, who presumably recognized their own lack of expertise, showed a greater willingness to yield than the lawyers. This is unfortunate in that the lawyers, because

16 However, the potential gain in accuracy from judicious weighting of estimates was offset by the attorneys’ general tendency to give less weight to their partners’ estimates than the law students did.
of their higher bracketing rate, actually stood to gain more from simple averaging than did the law students.

The practical implications for trial attorneys looking to improve jury verdict predictions and litigation decisions are clear. Our findings suggest that attorneys should seek out a second opinion and give that opinion more weight than they might be initially inclined to do. Better still, they could discuss the potential verdict and reach agreement. Even if one ultimately is going to rely on one’s own estimate, the process of discussing and reaching agreement on a joint estimate with a peer can improve the accuracy of verdict estimation and the overall quality of an attorney’s decision making.

Ideally, the best weighting strategy is simple averaging. Based on our data, it is reasonable for attorneys to think that they have some ability to discriminate more accurate estimates from less accurate estimates, and thus deviate from equal weighting in some instances. But it is also important to recognize that departing from the strategy of simple averaging has a statistical cost, and attorneys would be wise to consider whether any detectable difference in expertise is sufficient to justify that potential cost. For the attorneys in our study, it generally was not.

The benefits of averaging continue to mount with the aggregation of additional inputs. Thus, attorneys should seek as many opinions as possible. This could be done by phone, via attorney email lists, or using an online prediction market service that aggregates the “bets” of market participants. Even taking the mean estimate of a group of novice law students proved more accurate than the estimate of an individual attorney. However, in acknowledging the wisdom of the crowd, attorneys should not lose sight of the large benefit of increasing the size of one’s sample from one opinion to two. Taking a partner to lunch and discussing a possible verdict in concrete numerical terms or, better still, deliberating until you agree on a dollar amount, is likely to be an investment of time and money well worth making.

References


17See, for example, Inkling Markets <http://inklingmarkets.com/>. 

APPENDIX: CASE DESCRIPTIONS

Case 1

This is a personal injury case that was brought in San Luis Obispo County. The plaintiff is a 45-year-old woman who was rear-ended while driving. The defendant is the driver who ran into her car. In the accident, the plaintiff sustained cervical, thoracic, and lumbar sprain-strains. She received chiropractic treatment for three months. She is still having some pain in her neck, back, and left shoulder. Her medical costs are $8,763.38 and her lost earnings are $3,113.40 (she worked as a park ranger). So her total economic damages are $11,876.78. Please give your best estimate of only the noneconomic (pain and suffering) damages a jury would award.

Case 2

This is a personal injury case that was brought in Marin County. The plaintiff is a 15-year-old female high school student. The plaintiff hit her head on an un-padded pole on an outdoor basketball court. When she was taken to the athletic trainer’s office after hitting her head, she was left unattended and fell. The defendant is the school (it is a premises liability suit). From hitting her head on the pole, the plaintiff suffered a cut on her forehead and a concussion. Her fall in the trainer’s office knocked out one tooth and left two other teeth dangling by the roots. She spent the night in a pediatric intensive care unit, had two root canals, and had a dental bridge inserted (a bridge is a prosthetic replacement for lost teeth). She had an allergic reaction to the first bridge inserted and had to have another bridge inserted. When she is older, she will require further bridge work as well as implants. Following the accident, the plaintiff missed 10 days of school, performed poorly on her exams, and was cut from the basketball team because of low grades. The plaintiff’s past medical expenses are $25,000, and she will need between $80,000 and $100,000 for future dental expenses. So her total economic damages are between $105,000 and $125,000. Please give your best estimate of only the noneconomic (pain and suffering) damages a jury would award.

Case 3

This is a wrongful death case that was brought in Orange County. A five-year-old girl died from peritonitis (inflammation of the membrane which lines the inside of the internal organs) caused by a ruptured pancreas and duodenum resulting from severe physical and sexual abuse by the girl’s mother’s boyfriend. The mother and the mother’s boyfriend were both convicted on criminal charges. The plaintiff is the father of the girl. He is suing the county and a county social worker for failing to take appropriate action despite repeated warnings from various parties (including the plaintiff) that the girl was being abused. The
plaintiff is also suing the mother and the mother’s boyfriend. The plaintiff is not asking for economic damages; he is only seeking noneconomic damages for loss of care, comfort, and society. Please give your best estimate of only the noneconomic damages (loss of care, comfort, society) the jury would award.

Case 4
This is a personal injury case that was brought in Fresno County. The plaintiff, a five-year-old male kindergarten student, fell from the playground structure of his school and landed on his head. Safety codes called for soft-packed sand under playgrounds, but the sand the boy landed on was hard-packed. The defendant is the school, which is also a church that owns the playground. The boy sustained a left temporal parietal skull fracture, which led to a left middle meningeal artery bleed, ultimately causing an epidural hematoma (bleeding between the outer layer of the brain and the skull). The boy underwent emergency surgery, and the surgeon performed a craniotomy with an evacuation of the epidural hematoma. The boy was hospitalized for three days, and returned to school two weeks after the accident. A pediatric neurologist who evaluated him believes he has suffered a traumatic brain injury, causing some behavioral change. He scored lower than average on a standardized test administered after the accident, and has a risk (although not a certainty) of future cognitive problems. After repeating kindergarten, he is performing at grade level in the first grade. The only economic damages are medical expenses, which are $33,250.20. Please give your best estimate of only the noneconomic (pain and suffering) damages a jury would award.

Case 5
This is a wrongful death case that was brought in Los Angeles County. A 38-year-old, father of six, was working on engine troubles on the side of the road when his car was rear-ended by a drunk driver. He was killed instantly. The decedent was living in the United States illegally and working as a gardener. The plaintiffs are his mother, his brother, one adult child in his mid-20s, and five minor children, ages 3 to 15. The decedent was unmarried at the time of his death. The defendants are the driver, the fumigation company that the driver works for, and the rental car company where he rented his car. The plaintiffs are not asking for economic damages; they are only seeking noneconomic damages for loss of care, comfort, and society. Please give your best estimate of only the noneconomic damages (loss of care, comfort, and society) a jury would award. Please combine damages for all plaintiffs in your estimate.

Case 6
This is a personal injury case that was brought in Los Angeles County. The plaintiff, a 60-year-old woman, was rear-ended at a stoplight. The plaintiff was hit when she pulled out of a parking spot, crossed two lanes of traffic, and stopped at the left-turn lane of an intersection. At this point, she was struck from behind and her car was pushed into the car in front of her. The plaintiff fractured her right arm, requiring internal fixation surgery (insertion of a metal rod into the arm). She will need an additional surgery to remove the
hardware. The prognosis for her arm is good, but she will be at risk for swelling because she has no lymph nodes to drain fluid in her upper arm. (They were removed years ago during breast cancer treatment.) She missed two months of work due to the injury. Her past medical costs are approximately $18,000 and her future medical costs are approximately $9,000. Her lost earnings are approximately $9,000. So her total economic damages are estimated at $36,000. Please give your best estimate of only the noneconomic (pain and suffering) damages a jury would award.