March 1, 2014

Why Don't Juries Try Range Voting?

F.E. Guerra-Pujol, Barry University

Available at: https://works.bepress.com/f_e_guerra_pujol/34/
Why don’t juries try range voting?

Enrique Guerra-Pujol
Barry University School of Law
Orlando, Florida 32807, USA
eguerra@barry.edu

Abstract: In this paper, the author proposes the use of a simple “range voting” method by juries in which jurors would rate or score on a scale of zero to ten (or some other specified scale) the evidence presented by the parties at trial. The jury’s verdict would thus consist of a numerical value, either the average or the sum total of all the individual scores, which the author refers to as a “range verdict.” Range voting by juries thus produces a numerical verdict, a range verdict, consisting of an average value or total sum, and a plaintiff or other moving party would prove his case as a matter of law only if the average value or sum total of the jury’s collective score exceeded some critical threshold value. In addition, the author explains how range voting solves several problems endemic to juries, including holdouts, strategic jurors, and ignorant or partially informed jurors. Lastly, the author also explains how range voting improves jury accuracy and examines the relation between his proposed model of range voting in juries and the diversity prediction theorem in mathematics or “wisdom of crowds” effect.

Keywords: jury accuracy, juries, range voting, range verdicts, probabilistic verdicts, holdouts, strategic voting, ignorance, incomplete information, diversity prediction theorem

JEL Codes: C71 (cooperative games), K41 (litigation process)

Date: 28 February 2014
In this paper, we propose the use of a simple “range voting” method by juries in which jurors would rate or score on a scale of zero to ten, or some other specified scale, the evidence presented by the parties at trial. The jury’s verdict would thus consist of a numerical value, either the average or the sum total of all the individual scores, which we refer to as a “range verdict,” and a plaintiff or other moving party would prove his case as a matter of law only if the average value or sum total of the jury’s collective score exceeded some critical threshold value. In addition, the author explains how range voting solves several problems endemic to juries, including holdouts, strategic jurors, and ignorant or partially informed jurors. Lastly, we explain how range voting improves jury accuracy, and we also examine the relation between his proposed model of range voting in juries and the diversity prediction theorem in mathematics or “wisdom of crowds” effect.

In summary, this paper is organized as follows: 1. Overview and mechanics of range voting by juries, 2. Holdouts and stubborn jurors, 3. Strategic or biased jurors, 4. Ignorant or partially informed jurors and the diversity prediction theorem (or “wisdom of crowds” effect), and 5. Conclusion.

1. **Range voting by juries** (overview and mechanics)

The type of jury voting we are proposing here is “jury scoring” and is often called “range voting” in the literature on the voting systems. (See Smith, 2000; Hillinger, 2005; Poundstone, 2008, ch. 14. For a glossary of different voting procedures, see Poundstone, pp. 287-289.) Specifically, range voting by juries would work as follows:

At the end of a jury trial, jurors would be instructed to rate or score on a scale of zero to ten (or some other specified scale) the evidence presented by the parties at trial. The jury’s verdict would then consist of a numerical value, either the average of each juror’s score or the sum or total of all the individual scores. There are thus two ways one could express the jury’s collective score or “range verdict”—one could calculate the average value of the jurors’ ratings, or in the alternative, one could simply sum or add up the jurors’ scores. In either case (average score or total score), range voting by juries produces not a binary or qualitative verdict but a numerical or quantitative verdict consisting of an average score or total score. Also, since the jury’s verdict is a function of a range voting procedure, we refer to such a numerical verdict as a range verdict.
The scoring system can take many different forms. At a minimum, jurors would have to score the evidence presented by the moving party—the plaintiff in a civil case; the prosecutor in a criminal case. The plaintiff or prosecutor would prove his case as a matter of law only if the average value or sum total (as the case may be) of the jury’s collective score exceeded some critical threshold value. In the alternative, jurors could be asked to score the evidence presented by the defendant. In this version of range rating, the side with the highest average score or highest total score wins. (In criminal cases, where the moving party’s burden of proof is much higher than in a civil case, one could require the moving party’s average or total score to exceed the defendant’s score by some specified margin.) We, however, prefer the simplest possible procedure in which jurors are only required to score the moving party’s case on a scale of zero to ten. The moving party wins only if the jury’s average score exceeds some threshold, such as 5 in a civil case or a higher value in a criminal case.

In this approach, the judge (or some other institution) would have to establish the threshold or “cut-off” value for criminal convictions or for the imposition of civil liability. That is, to convict a defendant of a crime or to impose civil liability on a defendant, the sum total of the jurors’ scores (the additive method), or the average value of all the jurors’ scores (the average-value method), as the case may be, would have to exceed some minimum threshold or cut-off value. Otherwise, when the jury’s collective score falls below this threshold value, the defendant would be immune from civil or criminal liability.

Consider, by way of example, a standard civil case, such as a personal injury or wrongful death lawsuit. In a civil case, where the plaintiff’s standard of proof is based on the “preponderance of the evidence” or “more likely than not” standard, the threshold for imposing civil liability could be set at any value greater than 5, if the average-value method is used, or at any value greater than 30 or 60, depending on the total number of jurors (six or twelve), if the additive approach is used.

Although our unconventional proposal may sound strange or exotic, range voting is actually very common in many areas of life, especially on the World Wide Web. In the words of one commentator, for example: “YouTube and Amazon allow users to rate videos and books on a five-point scale. The Internet Movie Database (IMDb) has ten-point ratings of movies.” (Poundstone, 2008, p. 233; see also Hillinger, 2006.) If ordinary people are thus accustomed to range voting in their everyday activities (such as rating movies and restaurants), why not extend this method to juries? Our proposal, in a nutshell, would do just that—extend range voting to juries.

[Before proceeding, it is also worth noting that range voting works best when the same group of people rates all the candidates or products. [WHY?] (See, e.g., Poundstone, 2008, p. 233.) Range voting is thus an especially appropriate method for jurors, since this is precisely how the jury system works: the same
group of people (the jury) must rate the strength of the evidence presented at trial.

In addition to its simplicity and user-friendly nature, our proposed method of range verdicts has another advantage: it allows stubborn jurors or “holdouts” to freely express their individual assessments of the case.

2. Idiosyncratic jurors (the problem of “holdouts” or stubborn jurors)

_Holdouts can be oddballs and misfits. But they can also be people devoted to the principle that a just conviction requires true unanimity, not compromise. So lawyers ... say the legal risks are great and the strategy uncertain when there are hints of a holdout._ (Glaberson, 2010, p. A13.)

In most cases, jury verdicts must be unanimous. Holdouts (i.e. stubborn jurors who strongly disagree--usually for reasons of principle or for other idiosyncratic reasons--with the collective decision of their fellow jurors) thus pose a serious problem: the possibility of jury deadlock or a “hung jury.” In addition to the deadlock problem, holdout jurors also often face hostility and threats. (See, e.g., Glaberson, 2010; Schwartz, 2010.) Legal commentators have thus noted “the stubborn place of the holdout in the legal system and the roiling emotions that can come to play behind the jury room door.” (Glaberson, 2010, p. A13.) One holdout juror, for example (Ruth Jordan, Juror No. 4 in the 2004 trial of two Tyco executives in New York City), once described the atmosphere in her jury room as “‘positively poison,’ * * * ‘They shouted at me,’ * * * ‘One of the jurors threatened me. He turned on me in a fury, just a fury, and he said, “I’m going to spend the rest of my life destroying you.”’” (Ibid.)

Under our proposed range voting procedure, however, holdouts would not result in deadlock or unleash negative emotions. Under range voting holdouts would not pose a problem at all because holdouts would be able to freely vote their conscience without affecting the jury’s collective verdict. To illustrate our argument, consider the following scenario involving a range verdict by a six-man jury in a hypothetical civil case:

**Example #1**

<table>
<thead>
<tr>
<th>Juror</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>9</td>
</tr>
<tr>
<td>B</td>
<td>8</td>
</tr>
<tr>
<td>C</td>
<td>8</td>
</tr>
<tr>
<td><strong>D</strong></td>
<td>0</td>
</tr>
<tr>
<td>E</td>
<td>9</td>
</tr>
<tr>
<td>F</td>
<td>8</td>
</tr>
</tbody>
</table>

In this example, Juror D is our lone holdout. (He or she scores the plaintiff’s case a zero.) Jurors A, B, C, E, and F, by contrast, are reasonably confident in the strength of the plaintiff’s case, given their high scores. Now, consider the jury’s
collective score or “range verdict” in this case, either the sum total or the average value of all the jurors’ scores. Here, the sum total is 42, since $9 + 8 + 8 + 0 + 9 + 8 = 42$, while the average value is 7.0, since $42 ÷ 6 = 7.0$. In both cases, the jury’s collective score (or range verdict) easily exceeds the minimum threshold value for imposing civil liability on the defendant ***

The stubborn behavior of sincere holdouts, however, leads to a more sinister scenario, the possibility of biased or “sleeper” jurors, i.e. jurors who vote strategically, not sincerely. This possibility of strategic voting in juries poses an important question. What happens when a juror votes strategically by artificially assigning an extremely high or an extremely low value on his individual scorecard in order to manipulate the jury’s collective score? More to the point, how does our method of range voting deal with strategic or biased jurors and strategic voting generally?

3. **Strategic jurors** (the problem of extreme values in scoring)

To illustrate the problem of strategic voting in juries, consider the following example of a range verdict by a six-man jury in a hypothetical civil case:

**Example #2**

Juror A = 9  
Juror B = 8  
Juror C = 8  
Juror D = 0  
Juror E = 9  
Juror F = 8

Notice that, from a purely **quantitative** perspective, this example of strategic voting is totally indistinguishable from the example above involving a holdout. Although the **qualitative** reasons for Juror D’s score may vary in both scenarios— in the previous scenario, Example #1, Juror D is presumably basing his score on reasons of principle, while in this scenario, Example #2, Juror D is voting strategically in order to manipulate the outcome of the case—, the effect on the jury’s collective score in both cases is the same: negligible at best.

In other words, a lone strategic juror (or a lone holdout) does not pose a problem under range voting *** But what happens when a small subset of jurors (say, two jurors in a six-man jury) vote strategically? The short answer is, “it depends.” Specifically, it depends on whether these strategic jurors share the same biases, i.e. whether they vote as a unified bloc or not.

If these strategic jurors see the case differently, i.e. if they do not share the same biases about the case, then their strategic scores will simply cancel each other out. To see this, consider the following example of strategic voting:
Example #3

Juror A = 7
Juror B = 7
Juror C = 5
Juror D = 0
Juror E = 10
Juror F = 7

In this example, the scores of Jurors A, B, and F indicate that these jurors are fairly confident in the plaintiff’s case. Juror C is undecided, while Jurors D and E appear to be voting strategically, given the extreme values of their scores, especially in relation to the scores of the other jurors. In particular, notice that the sum total of their strategic scores is 10 (out of a possible high score of 20), while the average value of their strategic scores is 5 (out of a possible high 10). In other words, when strategic jurors see the same case differently, their strategic votes simply cancel each other out.

But what happens these strategic jurors actually share the same biases about the cases and vote as a unified bloc? For example:

Example #4

Juror A = 7
Juror B = 7
Juror C = 5
Juror D = 10
Juror E = 10
Juror F = 7

In this particular example, the jury’s collective score exceeds the threshold value for the imposition of civil liability, since the sum total of the jury’s range verdict in this example is 46 and the average value of the scores is 7.666. It thus appears that a small subset of strategic jurors is indeed able to manipulate the outcome of the jury’s verdict. But, in reality, this conclusion is unwarranted. The threshold would have still been exceeded in this case even if these strategic voters had voted sincerely instead of strategically. That is, even if both jurors had assigned a score of 7 to the plaintiff’s case, like three of their fellow jurors, or even a score of 6, the plaintiff in this case would have still won ***. (This same analysis applies to strategic scoring in other direction as well.)

Of course, if a larger proportion of jurors conspire to vote strategically in the same direction, say three or four jurors in a six-man jury, then, yes, the extreme scores of such strategic jurors will affect the jury’s final verdict. But as perverse as this may sound, this is as it should be. Why? Because even a strategic range verdict (i.e. a range verdict that is the product of strategic voting) still tells us
which side of the case the juror believes more strongly in! (Cf. Poundstone, 2008, p. 241.)

Thus far, we have described our proposed method of range voting in juries (or “range verdicts”) and analyzed the problem of “holdouts” or stubborn jurors as well as the problem of strategic or biased jurors. But what about ignorant or poorly informed jurors? Does range voting still work (i.e. produce accurate results) even when jurors are willfully ignorant or poorly informed?

4. Ignorant jurors (the relation between jury accuracy and juror accuracy)

We conclude with a conjecture. Our conjecture is that range voting produces a higher level of jury accuracy than “binary verdicts” do (e.g. “guilty” or “not guilty”), even when individual jurors are ignorant or partially informed. Why? Because range voting produces a “wisdom of crowds” effect when certain conditions are met (which we discuss in greater detail below), conditions generally inapplicable to binary verdicts.

Before proceeding, however, we wish to clarify the meaning of the term “ignorant,” since this word often (though not always, see, e.g., Firestein, 2012) carries a negative connotation. By “ignorant” we simply mean “partially informed.” Thus our intent in using this term is not to criticize the intelligence or moral character of the average juror or jurors generally. Instead, we wish to emphasize a frequent and formidable phenomenon that occurs in every legal trial: the exclusion of relevant evidence from the jury’s consideration. A juror might thus be ignorant (or partially informed) because relevant evidence has been excluded from his consideration on hearsay or policy grounds, such as the constitutional privilege against self-incrimination or the attorney-client privilege, for example. Such a juror, when deciding how to rate or score the strength of the evidence in the case, is thus basing his rating or score on partial or incomplete information. In sum, then, when we speak of an ignorant or partially informed juror, we refer to a juror’s access any relevant evidence or information.

Since jurors are thus ignorant or partially informed in every legal trial due to the exclusion of relevant but inadmissible evidence, our range voting method can help increase jury accuracy, even when jurors are ignorant or partially informed. Specifically, building on the work of Scott Page (2007), we begin this last section of our paper by examining the relation between collective accuracy and individual accuracy. Specifically, in the context of jury trials, we consider the collective accuracy of the jury as a whole (jury accuracy) in relation to the accuracy of the individual members of the jury (juror accuracy) ***

Broadly speaking, we should expect a more accurate verdict from a jury the more accurate the individual members of the jury are, but according to Scott Page (2007, ch. 8), we should also expect a higher rate of collective accuracy from a jury the more diverse the individual jurors are, a result Page calls the “diversity prediction theorem.” By diversity, Page refers to the amount of variance among
the individual predictions made by the members of a group. In the case of a jury trial, a jury’s verdict is nothing but a prediction or educated guess about whether the defendant is guilty or not guilty. Ideally, we would like the jury’s verdict or prediction to be as accurate as possible. Our claim is that a “range voting” procedure produces the most accurate verdicts.

Stated simply, Page’s “diversity prediction theorem” states that the group accuracy is equal to individual accuracy minus diversity or variance. Group accuracy refers to the collective error of a group body (like a jury), or the distance from the crowd to the truth. Individual accuracy refers to the average error of each person in the group, or how far people are from group average. And diversity refers to the variation in each person’s prediction about the truth.

Before proceeding, notice the relevance of this theorem to jury trials. In essence, a jury’s verdict is nothing but a prediction or educated guess about whether the defendant is guilty or not guilty, or put another way, the individual jurors are essentially attempting to predict whether a defendant is guilty or not guilty. Group accuracy in the context of a trial jury refers to the accuracy of the jury’s collective verdict or prediction. Similarly, individual accuracy in this context refers to the accuracy of each individual juror’s prediction, while diversity refers to the variation or variance among the individual jurors’ predictions.

Stated formally, our trial jury version of the diversity prediction theorem can be expressed mathematically as follows:

\[
(J - T)^2 = \frac{1}{n} \sum_{i=1}^{n} (j_i - T)^2 - \frac{1}{n} \sum_{i=1}^{n} (j_i - J)^2
\]

In words, jury accuracy, or the distance of the jury’s collective verdict \(J\) from the true value \(T\), is equal to the average of each individual juror’s accuracy, or \(j_i - T\), minus the variance in prediction errors among jurors, or \(j_i - J\), where \(J\) is the collective error rate of the jury; \(T\) is the true value the jury is trying to discover, either 1 in the case of guilty or 0 in the case of not guilty; \(j_i\) is the individual juror \(i\)’s prediction or best guess about the true value; and \(n\) is the number of jurors.

This mathematical result is useful because it tells us that the collective error term will be small even when the average individual error term is large so long as the diversity or variance term is also large (Page, ***). That is, in the context of a jury trial, even if some jurors are biased, ignorant, or poorly informed (resulting in high individual errors on average), allowing each individual juror to engage in range voting by scoring the evidence presented on some scale (such as 0 to 1, or the more familiar scale of 0 to 10) should generally produce a large amount of diversity or variance. Why is a large variance important? Because when individual errors are high, the diversity prediction theorem tells us that a large
amount of variance will produce a “wisdom of crowds” effect, i.e. lead to small collective errors ***

To understand this insight, consider the following simple example:

[insert and discuss numerical examples from my Orange Notebook]

Our method of range voting for juries thus captures the “wisdom of crowds” effect (or small collective errors, or in the alternative, high collective accuracy) when average errors and variance are large ***

[Lastly: discuss madness of crowds -- REVISE OR DELETE -- What about “herd jurors” or jurors who all think alike? In short, no voting system is likely to provide accurate or reliable results under these circumstances, since the accuracy or reliability of biased, ignorant, and irrational jury verdicts will most likely be no better than the accuracy of random verdicts or verdicts produced by a random-decision mechanism, such as a coin toss in the case of binary verdicts or the throw of a ten-sided die in the case of a range verdict (one based on a ten-point scale). The problem in these cases is not with the voting rules per se but rather with the procedures for the selection of jurors and with the rules of evidence.]

A few words about range voting and probabilistic verdicts

Before concluding, we wish to say a few words in passing about the relation between range voting and the concept of “probabilistic verdicts.” In a previous paper (Guerra-Pujol, 2012, pp. 118-120), we presented a simple thought-experiment involving the possibility of probabilistic verdicts. (Imagine, for example, a legal system in which jurors would estimate the probability of defendant’s guilt from 0 to 1.) We also contrasted the idea of probabilistic verdicts with the traditional system of “binary verdicts” (i.e. “guilty” or “not guilty”). Here, we compare and contrast the concept of probabilistic verdicts with our method of range voting (or “range verdicts”).

In summary, in a range voting system each juror must answer the following question: how strong is the plaintiff’s case? Or more precisely, on a scale of 0 to 10, how strong or persuasive is the evidence presented at trial? A range verdict is thus the jury’s collective evaluation of the strength (or “persuasiveness”) of the evidence in the case, since under a range voting system each juror is rating or scoring the evidence in the case. Probabilistic verdicts, by contrast, pose a slightly different question: given the evidence presented at trial, how likely is it that the defendant committed the wrongful act alleged in the plaintiff’s complaint or the grand jury’s indictment? A probabilistic verdict is the jury’s collective guess or probability estimate about the defendant’s alleged conduct. To render a
probabilistic verdict, each juror must convert his belief about the strength of the evidence in the case into a probability estimate.

Despite these superficial differences, however, both types of verdicts are equivalent because the probability of guilt in a given legal case should reflect the strength of the evidence in the case. [Notice, however, that the reverse is not always true: the strength of the evidence in a case is not always a function of the defendant’s guilt. Some evidence might point to a non-guilty person.] That is, whether a juror is making a probability estimate or whether he is simply rating or scoring the strength of the evidence in the case, in either case such a probability estimate or numerical score (as the case may be) should reflect his belief about the strength of the evidence against the defendant. This observation about the relation between range voting and probabilistic verdicts is therefore significant because range voting thus provides a practical method for implementing our “probabilistic verdict” concept * * *

Conclusion

In this paper we have proposed the use of “range voting” in juries. Unlike a binary verdict, in which jurors are limited to voting up or down (guilty or not guilty), range voting allows jurors to rate or score the evidence presented in the case. In summary, jurors would rate or score the evidence presented by the parties at trial on a scale of zero to ten or some other specified scale. The jury’s collective verdict would consist of a numerical value, either the average or the sum total of all the individual scores, and the plaintiff, prosecutor, or other moving party would prove his case as a matter of law only if the average value or sum total of the jury’s collective score exceeded some critical threshold value.

Range voting also solves several problems endemic to juries, including holdouts and strategic jurors. In essence, holdout behavior is just a special case of strategic voting, since a holdout, like a strategic juror, is one who rates or scores every case either a perfect ten or a zero (or whatever the prescribed maximum and minimum scores are). Range voting minimizes the ability of holdouts and strategic jurors to manipulate the outcomes of verdicts since deadlocks cannot occur under range voting *** [summarize additional reasons] *** Lastly, range voting can increase jury accuracy, even when jurors are ignorant or partially informed *** [conclude by summarizing our analysis of ignorant jurors above]
Bibliography

Kenneth Arrow (1951), *Social choice and individual values*, Yale University Press.


William Poundstone (2008), *Gaming the vote: why elections aren’t fair (and what we can do about it)*, Hill & Wang.

