DEDUCTIVE MODELING TO DETERMINE AN OPTIMUM JURY SIZE AND FRACTION REQUIRED TO CONVICT*

by Stuart S. Nagel** and Marian Neef***

I. The Basic Problem

In 1970, the United States Supreme Court held in *Williams v. Florida*[1] that due process is not violated if a state chooses to conduct criminal cases with six-person juries rather than twelve-person juries. Two years later, in *Apodaca v. Oregon*,[2] the Court held that due process is not violated if a state chooses to have juries decide criminal cases by a majority of ten out of twelve jurors. Those cases were followed by a substantial literature,[3] but there is still no systematic

* This Article is based on a longer paper presented in the workshop on "Science and Technology in Criminal Justice" at the 1975 annual meeting of the American Society for Public Administration. The authors appreciate the helpful suggestions made by Allen Barton of the Columbia Bureau of Applied Social Research, Peter Aranson of Georgia Institute of Technology. W.G. Leight of the National Bureau of Standards, John Murphy of the University of Cincinnati Law School, Richard Lempert of the University of Michigan Law School, and Sarah Schramm of George Washington University. The authors are also grateful to the LEAA National Institute, the Ford Foundation Public Policy Committee, and the University of Illinois Law and Society Program for financing the research of which this Article is a part, although none of these groups are responsible for the ideas advocated here.

** Professor of Political Science, University of Illinois; member of the Illinois Bar. B.S., 1957, J.D., 1958, Ph.D., 1961, Northwestern University.

*** Ph.D. Candidate, Department of Political Science, University of Illinois. B.A., 1973, M.A., 1974, University of Illinois.

1. 399 U.S. 78 (1970). In Colgrove v. Battin, 413 U.S. 149 (1973), the Court held that a jury of six persons satisfies the seventh amendment guarantee of trial by jury in civil cases.

2. 406 U.S. 404 (1972). On the same day, the Court upheld the Louisiana provisions that permitted conviction of defendants in certain noncapital criminal cases by a majority of nine out of twelve jurors. Johnson v. Louisiana, 406 U.S. 356 (1972).

3. *See, e.g.*, Institute of Judicial Administration, A Comparison of Six- and Twelve-Member Civil Juries in New Jersey Superior and County Courts (1973) (containing bibliography); *Symposium—The Jury*, 10 Trial, Nov., Dec. 1974, at 11; Zeisel, . . . *And Then There Were None: The Diminution of the Federal Jury*, 38 U. Chi. L. Rev. 710 (1971). There have been several important studies of the effects of changing jury size. *E.g.*, Institute of Judicial Administration, *supra*; E. Beiser, The Trial Jury: Empirical and Normative Considerations (paper presented at the annual meeting of the American Political Science Association, 1973); Mill, *Six-Member and*

analysis of the fundamental issue raised by *Williams* and *Apodaca*: What is the relation between jury size, or a requirement of unanimity, and conviction of the guilty and acquittal of the innocent?

A. *Inability of Empirical Data to Indicate Effects of Jury Size*

Prior studies have mainly considered the effect of the Supreme Court decisions on the representativeness of juries or the probability of conviction rather than examining the probability of convicting the innocent or acquitting the guilty. These studies have been especially concerned with actual decisions, comparing decisions by twelve-person juries with those by six-person juries or comparing decisions by unanimous juries with those by nonunanimous juries.[4] Such comparisons tend to have limited meaning, however, because there is no way to control for differences in the cases that are presented to the alternative kinds of juries.

Despite the lack of experimental controls, significant data can be collected from some empirical research. For instance, the University of Chicago Jury Project took a nationwide sample of 3,576 jury trials in which twelve-person juries had been required to give unanimous verdicts. The Project found that in 64.2 percent of the trials, the juries returned verdicts of guilty, resulting in the defendants' convictions; in 30.3 percent, the juries returned verdicts of not guilty, resulting in the defendants' acquittals; in the remaining 5.5 percent, the juries were hung juries, that is, unable to reach unanimous decisions.[5] If a state conformed to the national average before the *Williams* case, what would it mean if, after *Williams*, the state adopted six-person juries and its juries still convicted 64 percent of the time?

The same conviction rate would not mean that a change in jury size had no effect on the probability of conviction. Presumably, defense counsel would be more willing to bring their weak cases before a twelve-person jury than before a six-person jury. Defendants' weak cases are those in which the jury is more likely to convict. On the other hand, prosecutors would probably be more willing to bring their weak cases before a six-person jury than before a twelve-person jury. Prosecutors'

Twelve-Member Juries: An Empirical Study of Trial Results, 6 U. MICH. J.L. REFORM 671 (1973). For important deductive mathematical models of the effects of changing jury size, including references, see note 30 *infra*.

4. *See* sources cited in note 3 *supra*.

5. H. KALVEN & H. ZEISEL, THE AMERICAN JURY 56 (1966) [hereinafter cited as KALVEN & ZEISEL].

weak cases are those in which the jury is more likely to acquit. These two effects may offset each other in such a way that the cases brought before six-person juries result in the same 64-percent conviction rate observed in a different sample of cases brought before twelve-person juries. Even if we are certain that smaller juries are easier to persuade and hence more likely to convict—all other things being equal—they may in fact convict only 40 percent of the time because defense counsel plea-bargain weak cases and bring only especially strong cases before the small six-person juries. We cannot interpret actual changes, or predict future changes, simply by observing the juries themselves.

The situation is further complicated by the possibility that in jurisdictions permitting the election of jury size less severe cases may tend to go before six-person juries and more severe cases before twelve-person juries. On the one hand, juries may have a greater tendency to convict in more serious cases because juries want those cases resolved or because prosecutors prepare better for them. On the other hand, juries may be less likely to convict in more serious cases because juries are more afraid of making a mistake or because defense counsel are better prepared. These conflicting forces make empirical comparisons between cases heard by twelve-person and six-person juries virtually meaningless, at least with regard to the important issue: What is the likelihood that an average defendant—whether truly guilty or innocent—will be convicted if he is brought to trial before a twelve-person rather than a six-person jury? In other words, what is the probability, given various assumptions, that a jury of a given size will render a correct verdict? If we could randomly assign a large number of cases to twelve-person juries and to six-person juries, we could then compare the results, because the large sample sizes would tend to make two random samples virtually identical. Any differences in the results would then be attributable to the difference in jury size.[6] This kind of randomization, even if practical, might be unconstitutional, however, because one or the other group of defendants would be favored without adequate justification.

As an alternative to this experiment, Zeisel and Diamond suggest having criminal cases decided simultaneously by twelve-person and six-person juries.[7] The decision of only one of the two juries would determine the defendant's fate. The jurors on the nondeterminative jury

6. *See* Zeisel & Diamond, *"Convincing Empirical Evidence" on the Six Member Jury*, 41 U. CHI. L. REV. 281 (1974).
7. *Id.*

might be affected, however, by the knowledge that their verdict will have no practical effect. But if there is randomization in the choice of which jury will be determinative, and the jurors are not told, then in effect the fate of defendants in criminal cases would be decided by the flip of a coin, which would be unconstitutional.[8]

There are other difficulties in comparing empirically the results of juries of different sizes and conviction requirements. First, there is little likelihood that actual juries of all possible sizes (from six to twelve members) and fractional conviction requirements could be found; hence empirical data would be limited. With only a few data points, assumptions can be made about the missing data, but such an analysis is a step away from an empirical approach and begins to become a deductive approach. Second, and more important, an empirical approach cannot tell us how many innocent defendants are likely to be convicted by juries of various sizes or how many guilty defendants are likely to be released. It is the number of mistakes rather than the gross number of

8. As an alternative method of handling the different kinds of cases that twelve-person and six-person juries are likely to decide, a group of knowledgeable lawyers could assign conviction probabilities, or PAC's (*see* Table 1 and accompanying text), based on twelve-person juries to both sets of cases for comparison purposes. More specifically, one could transcribe 50 randomly selected twelve-person cases and 50 randomly selected six-person cases. Then five lawyers could read all 100 cases or randomly allocate 20 cases to each of themselves. Each lawyer could indicate a PAC for each case if the case were to go to a twelve-person jury. If the twelve-person cases are a representative sample, then the actual conviction percentage should be about 64 percent. If lawyers are accurate perceivers of the probabilities, then their PAC's for the twelve-person cases should average about 0.64, or whatever the actual conviction percentage was for those cases. If the lawyers perceive the six-person cases as having an average PAC of 0.45, this would tend to indicate that the cases going to six-person juries are those in which a conviction is more difficult to obtain than the cases going to twelve-person juries.

If the six-person juries actually convicted 45 percent of the defendants in those cases, this would tend to show that six-person juries have the same propensity to convict as twelve-person juries. If, however, the six-person juries actually convicted more than 45 percent of the defendants in those cases, this would tend to indicate that six-person juries have that much higher a propensity to convict than twelve-person juries. Even experienced trial lawyers, however, may misestimate PAC's, as is indicated by the number of cases going to court that should have been settled if the experienced lawyers on both sides of the case could have agreed on the PAC and the sentence (or damages) likely to be awarded upon conviction (or finding of liability). In some cases, of course, experienced lawyers may go to trial because they are being paid to do so, even though they perceive that the last-offered plea bargain (or out-of-court settlement) is favorable to their perception of the probability of success on trial. In spite of the possible defects of this kind of empirical approach, a moderately good empirical approach plus a moderately good deductive approach is better than the bad empirical approach of comparing actual twelve-person jury results with six-person jury results.

convictions which concerns us, and this information cannot be derived from empirical studies of actual jury behavior.[9]

B. *Deductive Analysis*

Another means of analyzing the effect of jury size and unanimity is the deductive technique of model-building.[10] A model is simply a formal way of expressing the rules for reasoning from premises to a conclusion. Although these rules may be expressed in complex mathe-

9. Mock or experimental juries represent another way of measuring the effects of jury size on verdicts. Although this approach has the advantage of making it possible to present exactly the same case to a set of twelve-person juries and to a set of six-person juries, the approach has the disadvantage of lacking realistic jury trial procedure, personnel, and variety that may affect the comparisons. In addition, it requires many expensive mock trials to detect the possibly small but important differences in the propensities of six- and twelve-person juries to convict. The experimental studies that have been made with considerable concern for realism and number of trials are inconclusive about the effects of jury size on conviction probabilities. *See* A. Singer & A. Barton, Interim Report: Experimental Study of Decision-Making in the 12- Versus 6-Man Jury under Unanimous Versus Non-Unanimous Decisions, May 1975 (unpublished mimeographed report available from Columbia Bureau of Applied Social Research); Davis, Kerr, Atkin, Holt & Meek, *The Decision Processes of 6- and 12-Person Mock Juries Assigned Unanimous and Two-Thirds Majority Rules*, 32 J. PERSONALITY & SOCIAL PSYCHOLOGY 1, 11-12 (1975). Generalizations are difficult to make from these experimental studies because none used more than a few different trials, although each trial was presented to many twelve-person and six-person juries. Professor Davis has concluded on the basis of his deductive model that six-person juries have a slightly higher propensity to convict than twelve-person juries but that the difference is too small to be detectable in any feasible mock jury experiment. *See* note 30 *infra*.

10. The questions concerning jury size and unanimity pose an especially appropriate opportunity for deductive modeling to aid in prescribing policy decisions. At least four conditions must be met for a prescriptive model to be meaningful. First, the premises of the model concerning normative goals should be the goals that policymakers are likely to have. Second, the assumptions about reality should be empirically validated, or at least consistent with empirical knowledge. In any model, conclusions are drawn from these assumptions about goals and facts. Third, these conclusions should follow from premises in accordance with rules of logical deduction. Fourth, a deductive model should be capable of indicating how any changes in its normative and empirical premises affect the conclusions.

There are other important, but less essential, criteria for good deductive models. The conclusions should have breadth in time, geography, and conditions. The conclusions are also more useful if they explain how one might more effectively achieve given goals (a prescriptive model) and if they show the causes of events (a descriptive model). Thus, for example, to determine what jury size maximizes the number of guilty convicted plus the number of innocent not convicted, we have to know something about the causal relation between jury size and the conviction of the innocent and the guilty. An additional important goal is that the total structure of the model should be as simple and understandable as possible.

matical terms, a model may be as simple as a syllogism: (1) smaller juries are more likely to convict than larger juries; (2) six-person juries are smaller than twelve-person juries; (3) therefore, six-person juries are more likely to convict than twelve-person juries.

The problem to be solved requires answering the following question: If 64 percent is the probability that an average defendant will be convicted by a twelve-person jury deciding unanimously, what is the probability that an average defendant will be convicted by (1) a six-person jury deciding unanimously, (2) a twelve-person jury deciding by a vote of ten out of twelve, and (3) other possible combinations? Then, having determined the probability of convicting an average defendant, we can proceed to determine the probability of wrongly convicting an innocent defendant. And finally, it may be possible to determine the jury size and the fraction required to convict that would best meet society's goal of convicting the guilty without wrongly convicting the innocent.

C. *Preliminary Assumptions*

There are three general kinds of influence on an individual juror's decision. He or she may be affected (1) by the case itself: the evidence, its presentation, and the judge's instructions; (2) by forces unique to the individual juror: biases, history, mood, health, and other factors; or (3) by the interaction with other jurors. We hope that the juror is principally influenced by the case itself, and as the evidence considered below indicates, this is in fact probably the principal influence. To the extent that jurors make up their minds solely on the basis of the case presented to them and the judge's instructions, it is clear that the size of the jury, or the proportion of jurors required for a verdict of guilty, will have no effect on the average verdict; the outcome is determined by the merits of the case, and not by the jury acting arbitrarily.

However desirable such a state of affairs might be, we know the merits of the case are not the sole determining factor, or at least that jurors are affected differently by the case presented to them. In addition to the forces applied by prosecution, defense, and court, which act from outside the jury, there are forces acting within the jury and within the jurors themselves. We know that such forces exist because juries are not always unanimous.

If the size of a jury makes any difference then, it must be because of these internal forces. Of the two kinds of internal forces—those within the jury and those within the juror—the independent reaction of the juror is the more amenable to analysis because of the very independence of such reactions. If jurors respond to the case presented in ways that cannot be predicted from the nature of the case itself, or from the votes of the other jurors, their behavior is just the kind that mathematical statistics and probability theory are designed to handle. Twelve truly independent jurors behave like twelve coins, and there are excellent and simple means of predicting the result of flipping large numbers of coins. A juror who tends to vote for conviction, regardless of the case presented to him or the reactions of fellow jurors, would be analogous to a coin that is very heavily weighted on one side. In the usual language of statistics, we say that an ordinary coin has a 50 percent, or 0.50, probability of landing heads up; a juror voting independently might have a 96.4 percent, or 0.964, probability of voting for conviction. If all jurors acted in this fashion, a twelve-person jury would convict defendants about 64 percent of the time, which is 0.964 multiplied by itself twelve times. The smaller the jury, the higher the probability of conviction would be.[11] Thus, to the extent that jurors behave like coins, jury size, or the proportion needed to convict, will clearly have an important effect on the likelihood that a defendant will be convicted. In reality, such independent decisions account for only a small proportion of jury behavior, no more than about eleven percent, as we shall see. Nevertheless, the extent to which jury behavior is analogous to flipping twelve or six coins may determine the differences between large and small juries and between unanimous and nonunanimous voting.

The second kind of internal jury force, the behavior of the other jurors, is much more difficult to analyze or predict. For instance, we know that in some cases a single juror or a minority of the jury will sway the majority to its position. It may be, however, that the jury usually acts as a kind of averaging machine, converting a majority view into a unanimous verdict. Since there are occasionally hung juries and instances of minorities persuading the majorities, we know that this averaging machine effect does not happen all the time. Nevertheless, as we shall see later, juries actually do convert the majority's view to a unanimous one in most cases, and to this extent, juries act as averaging machines, responding as units to the case presented. Just as we observed when the jurors are directly influenced by the case, the size of the

11. For example, the probability that a six-person jury would convict an average defendant would then be $(0.964)^6 = 0.80 = 80$ percent.

jury would again be unimportant, and we could construct our model of jury behavior by saying that regardless of size, the average jury convicts 64 percent of the time if it contains average jurors.

It may be, therefore, that although the juror's independent reactions, in other words, his idiosyncrasies, are not very important—at least relative to the averaging perspective—in determining the overall number of convictions, these independent reactions may nonetheless be decisive in answering our questions about the effects of jury size and unanimity. This observation should not be surprising, since we are concerned primarily with the mistakes that different-sized juries may make, and the subjective characteristics of the jurors will certainly play a large role in whatever mistakes are made.

To define more precisely the effects of the jurors' independent behavior, we can construct a simple model that considers this coin-flipping aspect of juror behavior alone. Once this model has been constructed, it can be modified to allow mathematical expression of the impact that outside forces—the attorneys, the court, the evidence—have on the jury as a whole. This modified expression will also represent the interactions among the jurors to the extent that a jury acts as an averaging machine, the function of which is unaffected by its size. We will return to the validity of these assumptions at the close of the discussion.

II. Basic Data and Assumptions

In their University of Chicago Jury Project, Professors Kalven and Zeisel found that twelve-person unanimous juries convict 64 percent of the defendants brought before them.[12] This finding will be our reference point for constructing a jury model reflecting the independent juror behavior that is analogous to coins being flipped.

A. *The Probability of Convicting an Average Defendant*

Before proceeding further, it will be helpful to have a set of simple symbols. We will use NJ to refer to the number of jurors on a jury and NC to refer to the number of jurors needed to convict. The symbol PAC, with capital letters, will be used to refer to the probability of an average defendant being convicted by an average *jury*, and the symbol pac, with small letters, to refer to the probability of an average defendant receiving a conviction vote from an average *juror*.[13] The symbol

12. *See* note 5 *supra* and accompanying text.

13. We will frequently add a subscript to the probability symbols to indicate the size

PAN refers to the probability of an average defendant *not* being convicted by an average jury, and the symbol pan refers to the probability of an average defendant receiving a nonconviction vote from an average juror. PAC is thus the complement of PAN, since both must add up to 1.0, and pac is the complement of pan. Applying the Kalven and Zeisel data, PAC = 0.64 and PAN = 0.36. If we assume that jurors act as if they were coins being flipped, what would be the value of pac and pan? In other words, how are the coins weighted?

We know that to get a conviction, all twelve coins must come up "convict." The probability of flipping twelve ordinary coins and having them all come up heads is calculated by taking 0.5, the probability of a head on any one flip, and multiplying it by itself twelve times, which is expressed as $(0.5)^{12}$. In general, if the flips are independent of each other, the probability of a head, or a "favorable" outcome, raised to the nth power (where n is the number of consecutive flips) gives the probability that the flips will be "unanimous." If we were dealing with a jury, this general rule could be stated as a formula, $\text{PAC} = (\text{pac})^{\text{NJ}}$. Using the analogy of an ordinary coin, pac would be 0.5, NJ would be 12, and PAC, the probability of conviction, would be only 0.0002.

From Kalven and Zeisel's data, however, we already know that PAC = 0.64, and thus the problem becomes one of solving for pac in the equation $0.64 = (\text{pac})^{12}$. Raising both sides of the equation to the 1/12th power changes the equation to $\text{pac} = (0.64)^{1/12}$ or $\text{pac} = (0.64)^{0.083}$, and pac = 0.964. In other words, if the average juror behaves as if he were a coin being flipped, the coin would be a heavily weighted one, since there is a 0.964 probability of the juror's voting for conviction. The complement of this probability, 0.036, would then represent the probability of a juror voting against conviction.

---

of the jury or the fraction required to convict. Thus, $\text{PAC}_{12/12}$ or simply PAC_{12} refers to the probability of an average defendant being convicted by a twelve-person jury. The symbol PAC_{NJ} refers to the probability of an average person being convicted by a jury of NJ size or with N jurors. Similarly, $\text{PAC}_{11/12}$ refers to the probability of an average defendant being convicted with a twelve-person jury when only eleven jurors are needed to convict (*i.e.* the fraction required to convict is 11/12). The symbol $\text{PAC}_{\text{NJ-1/NJ}}$ refers to the probability of an average defendant being convicted with a jury of NJ size when one less than NJ is needed to convict. PAC without any subscript generally refers to the probability of an average person being convicted given varying jury sizes and varying fractions required to convict. The particular context, however, may indicate that the unsubscripted probability refers to an NJ of twelve operating under a unanimity rule.

B. *The Probability of Convicting an Innocent or Guilty Defendant*

These are the values of PAC, PAN, pac, and pan that we would predict if jurors voted quite independently in an average case. We suspect, however, that a truly innocent[14] defendant would be less likely to be convicted than an average defendant. Thus, if 64 percent of all defendants are convicted, the proportion of truly innocent defendants convicted would be smaller than 64 percent. But because truly innocent defendants will not usually be brought to trial unless they appear to be guilty, the probability of their being convicted may not be very far below the average of 64 percent. We will say, for the sake of discussion, that 40 percent of innocent defendants are convicted, although we will experiment later with other conviction probabilities.

A truly guilty defendant, on the other hand, would probably be more likely to be convicted than an average defendant. If most defendants are actually guilty, the probability of convicting the truly guilty defendant would not be much higher than the probability of convicting the average defendant. Seventy percent can be used for the sake of discussion, although again, we can later experiment with other conviction probabilities.

We can represent the probability that a truly innocent defendant will be convicted by an average jury as PIC and its complement, the probability that an innocent defendant will not be convicted, as PIN. PGC represents the probability that a truly guilty defendant will be convicted by an average jury, and PGN is the complement. Our assumptions can be expressed in terms of these new symbols: $PIC = 0.40$, $PIN = 0.60$, $PGC = 0.70$, and $PGN = 0.30$.

The probability that an average juror will vote to convict an innocent defendant and the probability that an average juror will vote to convict a guilty defendant can be determined as we determined pac, by solving for pic in the equation $PIC = (pic)^{NJ}$ or $0.40 = (pic)^{12}$, and solving for pgc in the equation $PGC = (pgc)^{NJ}$ or $0.70 = (pgc)^{12}$. Doing so yields a pic of $(0.40)^{0.083}$ or 0.926, and a pgc of $(0.70)^{0.083}$ or 0.971. These numbers are a prediction that the average juror votes for conviction more than 90 percent of the time. We know that this is not the case, however, because Kalven and Zeisel's data show that about 30

14. Throughout this discussion we will assume that innocence and guilt are factual matters, that is, that a defendant may actually be "guilty" although mistakenly acquitted by a jury.

percent of all juries are unanimous for acquittal. The prediction that jurors vote for acquittal less than ten percent of the time stems from the temporary assumption, for the sake of calculation, that only the juror's independent behavior is important. This assumption will be modified later.

C. *The Proportion of Innocent and Guilty Defendants*

We now must make some assumptions about the proportion of innocent and guilty defendants among the total number of criminal defendants whose cases are submitted to American juries. It seems reasonable to assume that truly innocent defendants account for only a small percentage. Yet this percentage is not likely to be zero since 30.3 percent of defendants are unanimously acquitted, and 5.5 percent of the juries can reach no verdict.[15] The requirement of finding a defendant guilty "beyond a reasonable doubt" can be interpreted[16] as meaning that the evidence should weigh against a defendant at about a 0.95 probability level before that defendant can be convicted.[17] Let us, therefore, temporarily assume that about 95 percent of the defendants tried by juries are in fact guilty and about 5 percent are innocent. For these five percent, there is, according to our previous assumptions, a 40-percent chance of conviction. To facilitate calculation, the proportion of inno-

---

15. *See* note 5 *supra* and accompanying text.

16. *See* Simon & Mahan, *Quantifying Burdens of Proof*, 5 LAW & SOC'Y REV. 319 (1971).

17. The 0.95 probability level is customarily used in social science to determine whether an hypothesis has been confirmed; in discussing that level, statisticians sometimes analogize to criminal case decision-making. *E.g.*, T. WONNACOTT & R. WONNACOTT, INTRODUCTORY STATISTICS 167, 171-74 (1972). The Blackstone Standard, which would allow ten guilty people to go free to save one innocent person from conviction, *see* note 18 *infra* and accompanying text, can be translated into a probability of 10/11 or 0.91. Thus, the probability of a defendant's guilt would have to exceed 0.91 before the defendant could be convicted. *See generally* Kaplan, *Decision Theory and the Factfinding Process*, 20 STAN. L. REV. 1065, 1071-77 (1968).

The 0.95 probability level is equivalent to a 0.05 probability that a given result was due to chance. The judicial system, perhaps without realizing it, has sometimes applied this 0.05 level of statistical significance in contexts other than "beyond a reasonable doubt." *See, e.g.*, Ulmer, *Supreme Court Behavior in Racial Exclusion Cases: 1935-1960*, 56 AM. POL. SCI. REV. 325 (1962). Ulmer pointed out that the Supreme Court was consistently finding racial discrimination in jury selection when, given the percentage of blacks in the population, there was less than a 0.05 probability that a jury panel randomly selected would have consisted of as few blacks as were actually present on the panel. When the probability was greater than 0.05, however, the Court consistently found no racial discrimination.

cent and guilty defendants will be expressed as 50 and 950, as if there were only 1000 defendants in all, and these two numbers will be represented by the symbols #I and #G.

D. *Summary of the Coin-Flipping Model*

The symbols, data, and assumptions of the model we have constructed, the formulas for using these data and assumptions to predict juror and jury behavior, and the results of our predictions, are summarized in Table 1. The reader is reminded, once again, that the numerical predictions will be modified by other factors known to be operating on juries. The coin-flipping model is used simply to examine one aspect of juror behavior in isolation. A rough check on the model's internal consistency can be made by operating it in reverse, that is, by trying to predict, from the conclusions we have derived, the actual number of convictions brought by an average jury—which is the fact with which we began, PAC= 0.64. The model gives PGC (#G) + PIC (#I) = PAC = 685/1000 = 0.685, reasonably close to our starting point.

TABLE 1. SUMMARY OF THE COIN-FLIPPING MODEL OF A TWELVE-PERSON UNANIMOUS JURY

TYPE OF DEFENDANT	TYPE OF DECISION-MAKER	CONVICTION PROBABILITY (C)	NON-CONVICTION PROBABILITY (N)
Innocent (I)	Jury	PIC = 0.40*	PIN = 0.60
Innocent (i)	Juror	pic = 0.926**	pin = 0.074
Guilty (G)	Jury	PGC = 0.70*	PGN = 0.30
Guilty (g)	Juror	pgc = 0.971	pgn = 0.029

Basic formulas:

$\text{pic} = (0.40)^{1/12}$; $\text{pgc} = (0.70)^{1/12}$; $\text{PIC} = \text{pic}^{\text{NJ}}$; $\text{PGC} = \text{pgc}^{\text{NJ}}$

PIC + PIN = PGC + PGN = 1.0

pic + pin = pgc + pgn = 1.0

Number of jurors = NJ = 12

#G = 950 guilty defendants per 1000 defendants

#I = 50 innocent defendants per 1000 defendants (#I = 1000 − #G).

* Assumed values.

** Jury probabilities are given to two decimal places, and the juror probabilities are given to three decimal places.

III. Optimizing Jury Size and Fraction Required for Conviction

A. *Effects of Changes in Jury Size on Jury Errors*

This simple model reflecting a juror's independent behavior can now be used to explore the effects of jury size. In particular, what changes in the jury's overall error rate will result from changes in the size of the jury? We can explore this question quite simply by substituting various numbers other than twelve for NJ in the expressions given in Table 1, and then recalculating all of the model's predictions, while holding other assumptions constant. The results of such calculations for a wide range of jury sizes are presented in Table 2.

TABLE 2. UNANIMOUS CONVICTION PROBABILITIES AND NUMBERS OF ERRORS FOR VARIOUS JURY SIZES WITH 1000 DEFENDANTS
(Independent Probability Submodel)

Jury Size (NJ)	Prob. Innocent Conv. (PIC)	Number Innocent Conv. (#IC)	Prob. Guilty Conv. (PGC)	Prob. Guilty Not Conv. (PGN)	Number Guilty Not Conv. (#GN)	Unwtd. Sum of Errors (USE)	Weighted Sum of Errors (WSE)
15	.32	16	.64	.36	339	355	497
14	.34	17	.66	.34	321	338	491
13	.37	18	.68	.32	302	320	486
12	.40	20	.70	.30	283	302	481
11	.43	22	.72	.28	263	281	477
10	.46	23	.74	.26	242	265	474
9	.50	25	.77	.23	221	246	471
8	.54	27	.79	.21	199	226	470
*7	.58	29	.81	.19	177	206	468*
6	.63	32	.84	.16	154	185	469
5	.68	34	.86	.14	130	164	470
4	.74	37	.89	.11	106	142	473
3	.79	40	.92	.08	80	120	477
2	.86	43	.94	.06	54	97	483
1	.93	46	.97	.03	28	74	491

Formulas for the columns (discrepancies are due to rounding):
PIC = $(0.926)^{NJ}$. #IC = (PIC)(50). PGC = $(0.971)^{NJ}$.
PGN = 1.0 − PGC. #GN = (PGN)(950).
USE = #IC + #GN. WSE = 10(#IC) + (#GN).

* NJ where WSE is minimized (*i.e.* NJ_0*, the optimum NJ with 0 dissents).

Column 7 of the table shows the sum of the two kinds of errors that a jury is capable of making: the number of innocent defendants that the jury convicts plus the number of guilty defendants that it fails to convict. These are the errors we are most concerned to minimize. Column 8 shows a weighted sum of errors. We give the two kinds of errors different weights because we assume that society does so. According to Blackstone's well-known statement of this view, "It is better that ten guilty persons escape than that one innocent suffer."[18] Having regard for the great influence of Blackstone on the authors of the Constitu-

18. 4 W. BLACKSTONE, COMMENTARIES *358.

tion,[19] we will also use a trade-off weight of ten, multiplying the number of innocent defendants wrongly convicted by ten before adding it to the number of guilty defendants who escape conviction, to obtain a weighted sum of all jury errors.[20]

The table shows that as the jury size decreases, the probability of convicting either an innocent or a guilty defendant increases. This result is directly attributable to the prediction, based on our coin-flipping model, that individual jurors vote for conviction more than 90 percent of the time, regardless of the guilt or innocence of the defendant. When jury size is reduced, unanimity becomes easier to obtain, causing the risk of wrongful conviction to increase while the risk of failing to convict a guilty defendant decreases. Therefore, the sum of the two kinds of errors, appropriately weighted, first declines and then rises, as shown in Figure 1.

B. *The Optimum Jury Size*

The most important aspect of Table 2 and Figure 1 is the point at which the weighted sum of errors is least. This point is reached at a jury size somewhere between six and eight; the nearest whole number is seven.[21] The model therefore predicts that a jury of seven members will

---

19. For discussion of Blackstone's influence on the founding fathers, see F. AUMANN, THE CHANGING AMERICAN LEGAL SYSTEM 30 (1940); FREDERICK WILLIAM MAITLAND READER 129-30 (V. Delany ed. 1957); R. POUND, THE FORMATIVE ERA OF AMERICAN LAW 8-9 (1938).

20. By providing a trade-off weight between errors of convicting the innocent and errors of not convicting the guilty, we avoid the extremely difficult problem of trying to express how much each of those kinds of errors are worth in a common unit of measurement like dollars. In other words, we are only concerned with the relative value of disutility of the kinds of errors, not their absolute value.

21. The minimum value for the weighted-error function can be calculated directly, of course, without reference to the table. If WSE = 10(#IC) + (#GN), then by substituting values for #IC and #GN and simplifying, we have

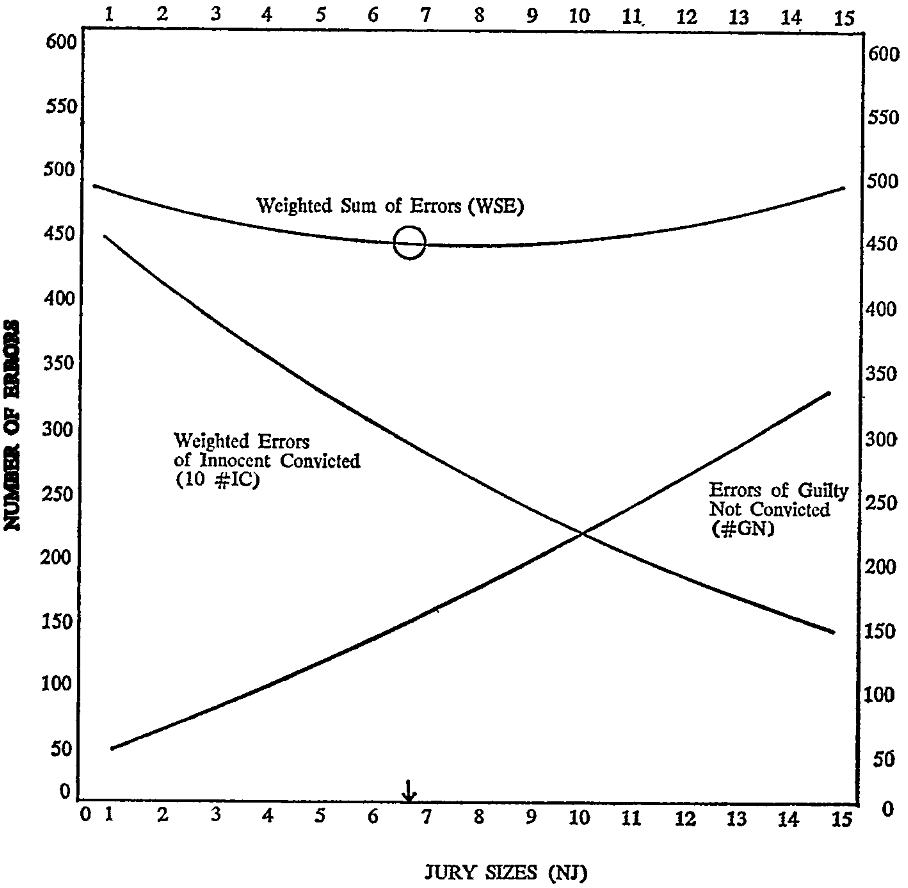
$$\begin{aligned} \text{WSE} &= 10(0.926)^{\text{NJ}}(50) + [1.0 - (0.971)^{\text{NJ}}](950) \\ &= 500(0.926)^{\text{NJ}} + 950 - 950(0.971)^{\text{NJ}}. \end{aligned}$$

Given that relation between WSE and NJ, we can say that the slope of WSE with respect to NJ at any point on the WSE curve equals

$$500(0.926)^{\text{NJ}}(\text{LN } 0.926) + 0 - 950(0.971)^{\text{NJ}}(\text{LN } 0.971),$$

where LN X means the natural logarithm of X (the logarithm to the base e). This slope of WSE to NJ follows from the rule that if $Y = b^x$, then the slope of Y to X is (b^x) (LN b), and from the rule that if Y is a constant, then the slope of Y to X is zero. To solve for NJ, we set that slope expression equal to zero. We then have one equation with one unknown, which we can solve by the rules for dealing with exponents encountered by most attorneys in high school algebra. Doing so reveals that

FIGURE 1. GRAPHING THE NUMBER OF ERRORS FOR VARIOUS JURY SIZES



the optimum NJ, where the slope of WSE reaches its lowest point, is NJ = 6.7, or 7 if we round to the nearest integer.

For further detail on what is involved in finding the bottom point on a valley-shaped total cost curve or the top point on a hill-shaped total benefit curve, see W. Baumol, Economic Theory and Operations Analysis 1-69 (1965); M. Brennan, Preface to Econometrics 1-192 (1973); S. Nagel, P. Wice & M. Neef, The Policy Problem of Doing Too Little or Too Much: Pre-Trial Release as a Case in Point (Sage Professional Papers in Administrative & Policy Studies No. 03-037 1976); S. Richmond, Operations Research for Management Decisions 6-10, 57-58, 87-88 (1968). These references also discuss the rules for finding the slope of exponential functions similar to those which are involved in the jury optimization problem. For a more general discussion of the jury size optimization problem, including the general equation for finding NJ*, see note 31 *infra* and accompanying text.

minimize errors in the fashion we assume would be optimum. Subject to the limitations on the coin-flipping model noted above, we can refer to a seven-member jury as the optimum jury size for unanimous juries. This prediction will be of importance when we consider a more complete model of the jury.

C. *Optimizing the Fraction Required to Convict*

Table 3 shows the probabilities that guilty or innocent defendants will be convicted by juries of varying sizes operating under rules requiring either unanimity or the indicated fraction for conviction. The table is constructed in a manner similar to Table 2, except for the calculation of the probabilities of conviction (PIC and PGC) when divided juries are permitted. In such cases, several possible outcomes will lead to conviction, and the probability of each must be calculated and the results added together. This is slightly more difficult to do than would first appear. We can easily calculate the probability of all the jurors in any given jury arriving at the same decision, but the probability that at least eleven out of the twelve will agree is more complex, because any one of the twelve jurors may disagree without affecting the outcome. If we continue to assume that all the jurors are acting independently of each other, however, then there are well-known mathematical formulas that will yield such probabilities.[22] If the jurors are completely independent

22. The binomial probability formula for the probability of exactly R favorable, successful, or designated outcomes out of N attempts is

$$ {}_{N}P_{R} = \frac{N!(P)^{R}(Q)^{N\text{-}R}}{R!(N\text{-}R)!}, $$

where P is the probability of a successful outcome on one attempt or trial and Q is the complement of P. To determine the probability of at least R successful outcomes out of N attempts, one must sum the probabilities of exactly N out of N, N-1 out of N, N-2 out of N, down to and including R out of N. If we assume the coin is evenly balanced so that the probability of a head on one flip is 0.5, and if we use the binomial probability formula to calculate the probability of eleven flips coming up heads out of twelve coin flips, the expression is

$$ P = \frac{12!(0.5)^{11}(0.5)^{12\text{-}11}}{11!(12\text{-}11)!}. $$

If we carry out the arithmetic, we find that the probability for exactly eleven heads out of twelve coin flips is $12(0.5)^{11}(0.5) = 0.002928$. Since twelve out of twelve heads has an exact probability of $(0.5)^{12}$ or 0.000244, the probability of at least eleven heads out of twelve flips is 0.002928 plus 0.000244, or 0.003172.

of each other, the final vote tallies will be arranged in a skewed Poisson distribution, and tables giving such distributions are in many standard statistical reference works.[23]

A factor of ten is again used in Table 3 to weight the number of errors resulting in wrongful convictions, and the weighted sum of errors again shows a decline as jury size decreases. The probability of convicting any defendant increases when either the size of the jury or the number required for conviction is reduced, as we would expect, and the probability of failing to convict a guilty defendant decreases under the same conditions. Curves, similar to those shown in Figure 1, could be plotted to compare these effects and the weighted sums of errors would

---

If our coin is not evenly balanced, but instead has a probability of coming up heads—or conviction of an innocent defendant—of 0.926, then the probability of exactly eleven out of twelve proconviction votes would be

$$\text{PIC} = \frac{12!(0.926)^{11}(0.074)^{12\text{-}11}}{11!(12\text{-}11)!} = 12(0.926)^{11}(0.074) = 0.38.$$

The 0.074 in the equation is pin, the complement of pic, from Table 1. Because we began with the assumption that the probability of twelve out of twelve jurors voting to convict an innocent defendant was 0.40, the probability of at least eleven conviction votes out of twelve jurors is 0.38 plus 0.40, or the 0.78 that is shown on row 5 of Table 3. Similarly, the probability of exactly eleven out of twelve jurors voting to convict a guilty defendant would be

$$\text{PGC} = \frac{12!(0.971)^{11}(0.029)^{12\text{-}11}}{11!(12\text{-}11)!} = 12(0.971)^{11}(0.029) = 0.25.$$

Adding 0.25 and 0.70, the probability that twelve out of twelve jurors will convict a guilty defendant, we get a cumulative or "at least" probability of the 0.95 shown in Table 3.

We could follow this procedure to calculate all the PIC's and PGC's in columns 2 and 4. Doing so, however, would be rather laborious. To save time, a binomial probability table might be consulted, although no published binomial probability table that has been found included single juror probabilities like 0.074 and 0.029. The smallest single-flip probabilities given were 0.05.

For further detail on binomial probability and related probability matters, see H. BLALOCK, SOCIAL STATISTICS (1972); M. BRENNAN, *supra* note 21; J. KEMENY, J. SNELL & G. THOMPSON, INTRODUCTION TO FINITE MATHEMATICS 113-77, 148 (1957); S. RICHMOND, *supra* note 21, at 127-78, 162.

23. The mathematical table used in preparing Table 3 is that given in S. RICHMOND, *supra* note 21, at 973-76. To use the table, we must compare (pin)(NJ) or 0.074(NJ) with the Poisson table, read off the probabilities corresponding to various numbers of nonconviction votes, and add those probabilities to get the PIC's. We then compare (pgn) (NJ) or 0.029(NJ) to the Poisson table in a similar manner to get the PGC's. Some interpolation within the table may be needed. The more precise binomial probability formula was used to calculate the probabilities for the various fractions required to convict in the vicinity of the critical region where the weighted sum of errors reaches a minimum.

TABLE 3. CONVICTION PROBABILITIES AND NUMBER OF ERRORS FOR VARIOUS FRACTIONS REQUIRED TO CONVICT
(Independent Probability Submodel)

Fraction Required to Convict (NC/NJ)	Prob. Innocent Conv. (PIC)	Number Innocent Conv. (#IC)	Prob. Guilty Conv. (PGC)	Prob. Guilty Not Conv. (PGN)	Number Guilty Not Conv. (#GN)	Unwtd. Sum of Errors (USE)	Weighted Sum of Errors (WSE)
12/12	.40	20	.70	.30	283	302	481
10/10	.46	23	.74	.26	242	265	474
8/8	.54	27	.79	.21	199	226	470
6/6	.63	32	.84	.16	154	185	469
*11/12	.78	39	.95	.05	48	87	438*
9/10	.83	42	.96	.04	38	80	453
7/8	.89	44	.97	.03	28	72	474
**10/12	.92	46	.98	.02	19	63	479**
5/6	.93	46	.98	.02	19	65	484
8/10	.94	47	.98	.02	19	66	489
9/12	.97	48	.99	.01	10	58	494
6/8	.98	49	.99	.01	10	59	500
7/10	.98	49	.99	.01	10	59	500
8/12	.99	50	1.00	0	0	50	495
4/6	1.00	50	1.00	0	0	50	500
5/8	1.00	50	1.00	0	0	50	500
6/10	1.00	50	1.00	0	0	50	500
7/12	1.00	50	1.00	0	0	50	500

Formulas for the columns (discrepancies are due to rounding):
PIC and PGC are determined by use of a Poisson probability table (see text).
#IC = (PIC)(50). PGN = 1.0 − PGC. #GN = (PGN) (950).
USE = #IC + #GN. WSE = 10(#IC) + (#GN).

* Value of NJ where WSE is minimized with one dissent allowed (*i.e.* NJ_1^*).

** NJ where WSE is minimized with two dissents allowed (*i.e.* NJ_2^*).

show the same rise and fall, but nothing of importance would be added to the qualitative conclusions already noted. Minimum values in the table are noted by asterisks; among the combinations considered,[24] these would be optimum values.

IV. Completing the Model

A. *Combining the Independent-and Collective-Mind Perspectives*

The model of independent juror decisions has now been developed in some detail, and it is possible to include additional considerations. At the outset, we noted the three determinants of any juror's decision: the case presented to him, the influence of other jurors, and the factors

24. To keep the table within reasonable bounds, only a few possible jury sizes and conviction fractions have been used. It is unlikely that juries larger than twelve or smaller than six will be used in the near future, however, and the table gives most

unique to that juror. This last factor, the juror's idiosyncrasy, or independence, is represented in the model we have constructed. The remaining task is to determine the extent to which this coin-flipping form of behavior determines jury decisions as a whole and to find means of expressing the effects of the other two influences.

It is not difficult to formulate a simple expression to describe the influence that acts equally and to the same effect on every juror. Each juror would vote as all the others did, and our model becomes one of simple identity: PAC = pac, PIC = pic, and PGC = pgc. If this were in fact true, as it would be if jurors always made decisions based solely on the merits of the case presented, there would be no hung juries, but only unanimous votes for acquittal or conviction. Jury size would be irrelevant. This identity model is clearly not the reality of jury function, but we can determine the extent to which it reflects juror behavior.

Since both independent probability and the collective effects of the case are influences on each juror, we can compare the two and determine their relative importance. This can be done if we assume, for the moment, that only these two influences act on each juror and then modify our independent probability model to incorporate the uniform effects of the case presented to the collective jurors. The simplest and most direct way of combining these two influences is simply to average them, although there is nothing in what we know about the true situation that dictates this form of expression.

The two models will be distinguished, for convenience, by calling them the "independent-mind" and "collective-mind" models. The latter is simply an identity, PAC = pac. Returning to the Kalven and Zeisel data, we know that the average proportion of convictions is 0.64 = PAC. The "independent-mind" formula is PAC = pac^{NJ}. A simple average of the two PAC's would be PAC = $1/2(\text{pac}^{\text{NJ}} + 0.64)$ or $1/2(0.964^{\text{NJ}} + 0.64)$. Similarly, we would estimate combination PIC = $1/2(\text{pic}^{\text{NJ}} + 0.40)$ or $1/2(0.926 + 0.40)$ and the combination PGC = $1/2(\text{pgc}^{\text{NJ}} + 0.70)$ or $1/2(0.971^{\text{NJ}} + 0.70)$.

There is no particular reason to suppose that the two models have equal weights in influencing jury behavior, however, and so we should determine a weighted average instead of a simple average. But what weighting factor is correct? The answer is not difficult to determine. The Kalven and Zeisel data supply us with the total number of juror

values between these two extremes, provided that at least a majority will always be required for conviction.

votes for conviction or acquittal, including that which resulted in hung juries rather than convictions: 67.7 percent of all individual juror's votes were for conviction and 32.3 percent were for acquittal. These figures reflect the following empirical facts: 64.2 percent of all juries voted unanimously for conviction; an average of 7.5 votes were for conviction and 4.5 for acquittal on the 5.5 percent of all juries that could reach no decision; and the remaining 30.3 percent of juries voted unanimously for acquittal.[25] We can view the 67.7 percent votes for conviction as an empirically determined pac rather than one determined from either separate model. If the individual juror's propensity to convict is determined by the combination of the two models that we have discussed, then pac must be equal to the weighted average of the independent-mind and collective-mind values for pac.

In the case of a twelve-person unanimous jury, this can be expressed as

$$\text{empirical pac}_{12} = \frac{(\text{WE}(0.64)^{1/12} + 0.640)}{\text{WE} + 1}$$

since $0.64^{1/12}$ (or 0.964) and 0.640 were the values assigned to pac in the two different models. If we set pac_{12} in this equation equal to 0.677, our "empirical" pac, and solve for WE, we will determine the correct weighting factor for averaging the two values of pac, or for combining the two models. WE, in short, is the factor we can use to combine the two separate and incomplete models in such a way that together they will be in agreement with the observed probability that an average juror will vote to convict an average defendant 67.7 percent of the time. Solving the above equation for WE, we find WE to be equal to 0.13. Thus, while the weight of the collective-mind model is 1.00, the weight of the independent model is only 0.13, for a total weight of 1.13. This means that the independent-mind, coin-flipping model we have developed accounts for about eleven percent of the total (*i.e.* 0.13/1.13), or that the independent probability propensities of jurors account for about eleven percent of jury behavior.[26]

25. *See* KALVEN & ZEISEL 460.

26. When we say that jury decisionmaking is about eleven percent in conformity with the independent probability model and 89 percent in conformity with the collective mind model, we are not saying that eleven percent of the jurors on an average jury follow the independent probability model and 89 percent follow the collective mind model. Rather, we are saying that jurors (like human beings in most group decision situations) simultaneously manifest some degree of independence and some degree of willingness to go along with the average person in the group. The average juror thus has about an

A single dissident juror may be more able to maintain his position against the smaller number of antagonists in a six-person jury; yet he may find more allies on a twelve-person jury. Some people, in some circumstances, may be capable of swaying an opposing majority; others may not. There is no way of predicting all such interactions in advance, and to this extent, no model can completely predict the effects of changing jury size or unanimity requirements in a specific case, as opposed to an average case or a class or category of cases.

In by far the majority of cases, the ultimate verdict is the one favored by a majority of the jury from the outset; juries that are unable to come to agreement because of a single holdout are extremely rare.[27] In most instances, a jury behaves like an averaging machine, converting a majority vote into a unanimous vote. Although there are certainly interactions among the jury, the effect of the interactions is simply to amplify the key outside influence, namely the case and instructions presented to the jurors. With any given jury, therefore, this kind of common stimuli and interaction is the collective-mind behavior already discussed. Jurors in most cases would vote unanimously, and the residue of hung juries would be attributed in part to the independent-mind effects and in part to other unpredictable interactions idiosyncratic to specific cases, juries, or jurors.

The averaging-machine interactions may have a very large effect on jury outcomes; only 30 percent of verdicts are brought without deliberation, and in only three percent does deliberation result in a minority persuading the majority.[28] In about two-thirds of all cases, the majority view prevails through deliberation. Unfortunately, we know very little about this process or how it might be affected by changes in jury size. Will minorities be more or less resistant to persuasion in smaller juries? Will lone dissidents find allies in large juries, or will they be emboldened in small ones? These unpredictable factors seem to balance out with a

eleven percent independence orientation and an 89 percent collective-mind orientation. Similarly, the average jury composed of twelve average jurors also has an eleven percent independence orientation in its decisionmaking.

27. *See* KALVEN & ZEISEL 460.

28. The data for comparing first-ballot votes with final-ballot votes is given in KALVEN & ZEISEL 491. From a sample of 225 cases, the minority on the first ballot won out on the final ballot in only six cases. In 69 of the 225 cases, the first ballot produced a unanimous decision. In 13 of the cases, the jury was unable to reach a unanimous decision. In the remaining 137 cases, the initial majority won out on the final ballot (*i.e.* the position of the initial majority became the unanimous position).

twelve-person unanimous jury and do not have a large effect on overall conviction rates, as we have seen. But would they result in unbalanced changes in these rates if jury size or the requirement of unanimity were changed? We simply do not know, although the infrequency of hung juries prompts us to assume that the averaging machine is a powerful one, operating effectively in most circumstances. Only 5.5 percent of verdicts are indeterminate, and only three percent of verdicts are the result of a minority persuading a majority.[29] We can only assume that an averaging machine this effective will continue to operate under altered circumstances. In other words, we assume that the relative weights of 0.13 and 1.00 for the independent probability model and the collective-mind model will roughly prevail for different jury sizes.

Because the weighted average of the two models seeks to incorporate all three influences on the juror, the combination model is thus a reasonably complete model of jury behavior. To complete this model numerically, we need only make some modest alterations in the independent-mind model already developed to accommodate the weighting factor for combining the independent-mind model with the collective-mind model. Of the influences we are considering, only the independent-mind factor (and representativeness of sample size variations) call for any altered result in juries of different sizes. It is, therefore, not surprising that the independent-mind model determines the effects of altering jury size or unanimity rules. The *magnitude* of such changes, however, will be small, in accord with the low weighting factor of 0.13 already calculated.[30]

29. *See* KALVEN & ZEISEL 460.

30. An overemphasis on the independent probability perspective is the main defect in the previous mathematical models that have attempted to relate jury size to conviction probabilities. *E.g.*, Friedman, *Trial by Jury: Criteria for Convictions, Jury Size, and Type 1 and Type 2 Errors*, 26 AM. STATISTICIAN, April, 1972, at 21-23; Note, *The Effect of Jury Size on the Probability of Conviction: An Evaluation of Williams v. Florida*, 22 CASE W. RES. L. REV. 529 (1971).

The student author assumes that the majority viewpoint on the jury becomes the winning viewpoint, which is nearly always true. To compare the probability of a six-person jury convicting a defendant with that of a twelve-person jury convicting a defendant, the author calculates these probabilities, using a multiplicative model, for various fractions of jurors inclined to consider a defendant guilty prior to deliberations. *Id.* at 540-47. His assumption that the majority always wins makes it impossible to test the effects of reducing the fraction required to convict. Data is also unavailable on predeliberation propensities as contrasted to propensities manifesting themselves on the initial or final ballot. The author's model seems to run contrary to justifiable common sense in saying that there will be more acquittals with a six-person jury than with a twelve-person jury when the fraction of jurors inclined to convict is greater than 0.5. *Id.* at 547. Part

B. *Revised Data and Results*

1. *Jury Size*

Table 4 is a revised version of Table 2 and shows conviction probabilities and numbers of errors for various jury sizes, calculated by using the completed combination model. The method for calculating #IC,

---

of the problem is his failure to consider hung juries (which occur more often with twelve-person juries) as being closer to acquittals than convictions. In addition, the model provides no way of reconciling the Kalven-Zeisel data, which show that juries convict 64 percent of the time and that jurors vote to convict 67.7 percent of the time.

Friedman's model involves jury outcomes (*i.e.* PAC's) as a function, not of juror propensities, but rather of the defendant's appearance of guilt (the X variable). His model is also a purely multiplicative one in that for a twelve-person unanimous jury, PAC $= X^{12}$, and for a six-person unanimous jury PAC $= X^{6}$, but the model is capable of considering a reduction in the fraction required to convict (as well as the jury size) by not making any assumptions about the majority winning. The X variable in Friedman's model, however, is even more difficult to verify empirically than the predeliberation propensity variable of the model above. Neither model seeks to come to grips with the problems involved in determining an optimum jury size.

Mathematicians such as Laplace, Condorcet, Poisson, and Cournot have also used jury decisionmaking to illustrate mathematical rules of independent probability, but without seeking to determine the extent to which the independent probability model actually fits jury decisionmaking. Some of the relevant work of Laplace and Condorcet is discussed in Ulmer, *Quantitative Analysis of Judicial Decisions*, 28 LAW AND CONTEMP. PROB. 164 (1963).

For more recent deductive jury-size models, see B. Grofman, Not Necessarily Twelve and Not Necessarily Unanimous: Conviction, Acquittal, and Hung Juries as a Function of Jury Size and Jury Decision Rule (mimeographed paper available from the author at State University of New York at Stony Brook, 1974). Grofman adds to the previous literature by speaking in terms of the probability of convicting an innocent defendant and the probability of not convicting a guilty defendant. He also assumes, however, an independent probability model and further assumes that PIC equals PGN. Partly because of these assumptions, which he later questions, his deductive model is the only one that arrives at no conclusions about the effects of jury size. Like the student author above and unlike the Supreme Court, *see* text following note 60 *infra*, Grofman concludes "that changes from a unanimity rule to say a 11-1 or 10-2 or even a 9-3 rule would have little impact on conviction rates." B. Grofman, *supra* at 21.

Davis adds to the previous literature by combining deductive modeling with empirical testing in experimental juries. *See* Davis, *Group Decision and Social Interaction: A Theory of Social Decision Schemes*, 80 PSYCHOLOGICAL REV. 97, 104-06 (1973); Davis, Bray & Holt, *The Empirical Study of Social Decision Processes in Juries*, in LAW, JUSTICE, AND THE INDIVIDUAL IN SOCIETY (J. Tapp & F. Levine eds., in press 1976). Davis also combines deductive modeling with a social psychological perspective that does not assume the majority always wins. His model has the disadvantage of being the most complicated of the deductive models discussed above, and of involving a kind of ad hoc multiplicative probability that does not involve recognized binomial or Poisson distributions, although it does make use of multinomial concepts. His multinomial categories refer to the three, rather than two, categories of convict, acquit, or hung within a jury. He concludes, as the present paper does, that "the theoretically possible difference be-

PGN, GN, USE, and WSE is the same as that used in Tables 2 and 3. The formulas for calculating PIC and PGC are changed to take into consideration our new combination of the independent-mind and collective-mind models. In the independent-mind model, we calculated PIC for various NJ's by determining $(0.926)^{NJ}$ where $0.926 = (0.40)^{1/12}$. Using the collective-mind model, PIC is constant at 0.40, regardless of the size of the jury, as long as the jury consists of average jurors who would vote to convict an innocent person 40 percent of the time (the original assumption for PIC). Under the combination approach, PIC represents the weighted average of those two approaches, with the independent probability model receiving a weight of 0.13 and the collective-mind model receiving a weight of 1.00. This assumes that the weights assigned to collective and independent elements of jury decision-

TABLE 4. REVISED CONVICTION PROBABILITIES AND NUMBER OF ERRORS FOR VARIOUS JURY SIZES
(Combined Independent and Collective Model)

Jury Size (NJ)	Prob. Innocent Conv. (PIC)	Number Innocent Conv. (#IC)	Prob. Guilty Conv. (PGC)	Prob. Guilty Not Conv. (PGN)	Number Guilty Not Conv. (#GN)	Unwtd. Sum of Errors (USE)	Weighted Sum of Errors (WSE)
15	.390	19.5	.693	.307	291.2	310.7	486.2
14	.393	19.7	.696	.304	289.1	308.8	486.1
13	.396	19.8	.698	.302	287.0	306.8	485.0
12	.400	20.0	.700	.300	285.0	305.0	485.0
11	.403	20.2	.703	.297	282.4	302.6	484.4
10	.407	20.4	.705	.295	280.1	300.5	484.1
9	.412	20.6	.708	.292	277.6	298.2	483.6
8	.416	20.8	.710	.290	275.1	295.9	483.2
*7	.421	21.1	.713	.287	272.6	293.7	483.1*
6	.427	21.3	.716	.284	269.9	291.2	483.2
5	.432	21.6	.719	.281	267.2	289.8	483.4
4	.439	21.9	.722	.278	264.3	286.3	483.6
3	.445	22.3	.725	.275	261.4	283.7	484.1
2	.453	22.6	.728	.272	258.5	281.1	484.8
1	.461	23.0	.731	.269	255.4	278.4	485.7

Formulas for the columns (discrepancies are due to rounding):
$PIC = [0.13(.926)^{NJ} + 0.40] / 1.13$. $\#IC = (PIC)(50)$.
$PGC = [0.13(.971)^{NJ} + 0.70]/1.13$. $PGN = 1.0 - PGC$.
$\#GN = (PGN)(950)$.
$USE = \#IC + \#GN$. $WSE = 10(\#IC) + (\#GN)$.
* Value at which WSE is minimized.

tween juries of six and 12 is at a *maximum only about 8%* for the simple majority social decision rule, and under most conditions even smaller. Thus, *very* large samples of exceptionally 'noise-free' data would be required to pick up this difference." Davis, Bray & Holt, *supra* (emphasis in original).

making are the same, whether the jury is judging an average defendant, an innocent defendant, or a guilty defendant. There is no reason to assume that the characteristics of a defendant would affect the balance of those two elements. The combination PGC is then a weighted average between the independent probability of $(0.971)^{NJ}$ and the collective probability of 0.70.

In Table 4, as in Tables 2 and 3, when the jury size decreases, the probability of an innocent person being convicted increases and the probability of a guilty person not being convicted decreases. Since there are likely to be many more guilty defendants in the system than innocent defendants, the unweighted sum of errors also keeps decreasing as the jury size decreases. The most important column is again the last, showing the weighted sum of errors. Because the number of innocent defendants convicted is weighted as ten times more important than the number of guilty defendants not convicted, the weighted sum of errors decreases, until we reach a jury size of seven persons at the minimum WSE level, and then begins increasing. This is precisely the optimum jury size determined by the independent-mind model alone.

If we were to make a graph like Figure 1 of our new WSE curve, the new WSE curve would reach its lowest point close to an NJ of a seven-person jury, just as did the WSE curve based only on the independent probability model. We can confirm that result by determining the slope of WSE relative to NJ, using our new formulas shown at the bottom of Table 4. When we set that slope equal to zero and solve for NJ, we find that the solution is the same in the independent and complete models.[31]

31. The procedure for finding the minimum point of the curve is essentially that given in note 21 *supra*, because the weighting factors in the completed model drop out when we set the expression for the slope of the weighted-sum-of-errors (WSE) curve equal to zero. This is another way of saying that the independent probability model determines the shape of the curve. The complete derivation of the formula for the value of NJ at which WSE is minimized is given below. (Note that the W here is the weight assigned to erroneous convictions, not the weight of the independent-mind model.) To simplify the expressions, the weight of the independent-mind model is assumed to be one-half (the two models are simply averaged). We begin with the expression for WSE in terms of NJ:

1. $$WSE = (W)\ \frac{pic^{NJ} + PIC_{12}}{2}\ (\#I) + (\#G) - \frac{pgc^{NJ} + PGC_{12}}{2}\ (\#G)$$

(Substituting the combination perspective for the independent probability perspective)

2. $$WSE = (W)\ (\tfrac{1}{2}\ pic^{NJ} + \tfrac{1}{2}\ PIC_{12})\ (\#I) + (\#G) - (\tfrac{1}{2}\ pgc^{NJ} + \tfrac{1}{2}\ PGC_{12})\ (\#G)$$

(Dividing both parts of the numerator in the compromise perspective by 2)

In contrast to both Table 2 and Figure 1, however, our new WSE curve is quite flat. Table 2 showed a WSE as high as 497 with an NJ of 15, and as low as 468 with an NJ of seven. The more accurate Table 4, however, shows a WSE only as high as 486.2 with an NJ of 15 and only as low as 483.1 with an NJ of seven. The numbers in Table 4 were carried out to an extra digit to show more clearly the trends in each column and the minimum value of WSE. Since the combination approach of Table 4 comes closer to the collective-mind approach, it follows that there is less relation between jury size and the sum of errors. If a purely collective-mind approach had been used, there would have been no relation at all between jury size and the weighted sum of errors.

Because the results are virtually the same for all jury sizes from six to twelve, one might conclude from Table 4 that jury size has very little effect on the weighted sum of errors. The calculated differences observed might be attributed to our reasonable but unverifiable premises that $PIC_{12} = 0.40$, $PGC_{12} = 0.70$, or $\#G = 950/1000$. In fact, there may be little or no difference between juries of six or twelve,

---

3. $WSE = (W)(\frac{1}{2})(pic)^{NJ}(\#I) + (W)(\frac{1}{2})(PIC_{12})(\#I) + (\#G) - (\frac{1}{2})(pgc)^{NJ}(\#G) - (\frac{1}{2})(PGC_{12})(\#G)$
(Removing the brackets by multiplying)
4. $(W)(\frac{1}{2})(pic)^{NJ}(\#I)(LN\ pic) + 0 + 0 - (\frac{1}{2})(pgc)^{NJ}(\#G)(LN\ pgc) - 0 = 0$
(Setting the slope of WSE with respect to NJ equal to zero)
5. $(W)(\frac{1}{2})(pic)^{NJ}(\#I)(LN\ pic) = (\frac{1}{2})(pgc)^{NJ}(\#G)(LN\ pgc)$
(Adding to both sides and removing the zero slopes)
6. $\dfrac{(pic)^{NJ}}{(pgc)^{NJ}} = \dfrac{(\frac{1}{2})(\#G)(LN\ pgc)}{(W)(\frac{1}{2})(\#I)(LN\ pic)}$
(Dividing both sides)

The ½'s cancel out on the right side, and we have an expression identical to that which would be derived from the independent probability model alone. Solving for NJ, we get:

7. $NJ\log(pic) - NJ\log(pgc) = \log K$
(Substituting K for the right side and taking log of both sides)
8. $NJ(\log pic - \log pgc) = \log K$
(Factoring out NJ)
9. $NJ = \dfrac{\log K}{(\log pic - \log pgc)}$
(Dividing both sides)
10. $NJ^* = \dfrac{\left[\log \dfrac{(\#G)(LN\ pgc)}{(W)(\#I)(LN\ pic)}\right]}{(\log pic) - (\log pgc)}$

NJ* is the value of NJ at which WSE is minimized, the optimum jury size. Substituting the appropriate numerical values from Table 2 or 4 yields an optimum value of seven (to the nearest integer).

although some authors have hypothesized or purported to find significant differences.[32] A conclusion that jury size makes no difference might lead one to preserve the status quo of twelve jurors or it might lead one to advocate a switch to a jury of six on the theory that when in doubt, one should choose the simplest or least costly alternative.

The data presented in Table 4, however, do not necessitate a conclusion that jury size is irrelevant to jury errors. First, much larger differences can easily be obtained by simply changing the normative trade-off weight. For example, if we are quite prosecution minded, we could give a much greater weight to an error of not convicting the guilty than to an error of convicting the innocent. Conversely, if we are quite defense minded, we could give a much greater weight to an error of convicting the innocent. Either way, we could justify six- or twelve-person juries on the basis of substantial differences in the weighted sum of errors. Working backwards from the jury size to trade-off weight, we find that the use of a twelve-person unanimous jury implies a trade-off weight that is 13,[33] rather than Blackstone's ten. This indicates (assuming our empirical premises are reasonable) that society by supporting the twelve-person unanimous jury considers convicting the innocent 13 times worse than not convicting the guilty and is impliedly willing to let 13 guilty persons go free to save one innocent person from conviction.

A second, closely related way to see big differences in Table 4 is to note that in the course of 1000 trials, 21.3 innocent defendants are likely to be convicted by a six-person jury, while only 20 innocent

32. It is not so unusual to find that hotly-disputed alternative legal policies have little difference in their effects. *See, e.g.*, R. DIXON, DEMOCRATIC REPRESENTATION: REAPPORTIONMENT IN LAW AND POLITICS 574-81 (1968) (equal apportionment may be irrelevant to actual legislative outcomes); Rosenberg, *Comparative Negligence in Arkansas: A "Before and After" Survey*, 13 ARK. L. REV. 89 (1959) (contributory negligence rules may be irrelevant to issue of court congestion).

33. In note 31 *supra*, we derived a general equation for determining the optimum NJ for a jury deciding unanimously, given W, #G, #I, pgc, and pic:

$$NJ^* = \frac{\log \dfrac{(\#G)(\text{LN pgc})}{(W)(\#I)(\text{LN pic})}}{(\log \text{pic}) - (\log \text{pgc})}.$$

If we set NJ* equal to twelve, substitute the appropriate numerical values for #G, #I, pgc, and pic from Tables 2 and 4, and solve for W, we find that W equals $10^{1.1}$ or 13. In other words, if the trade-off weight is 13, a twelve-person jury is the optimum jury size in terms of achieving the minimum weighted sum of errors.

defendants are likely to be convicted by a twelve-person jury. That 1.3 difference represents a 6½ percent increase over 20 defendants, or a six percent decrease from 21.3 defendants. It does sound socially undesirable to increase the number of innocent defendants convicted by 6½ percent. It sounds even worse to note that the probability of convicting an innocent defendant goes up from 0.40 to 0.427 when a jurisdiction changes from a twelve-person jury to a six-person jury. That is almost a seven percent increase in the probability of an innocent person being convicted (0.027/0.40). Conversely, someone concerned about not convicting the guilty might note that the probability of a guilty defendant *not* being convicted rises from 0.284 for a six-person jury to 0.297 for a twelve-person jury. That represents an increase of almost five percent (0.013/0.284).

2. *Fraction Required to Convict*

Larger differences in the weighted sum of errors result from varying the fraction required to convict as well as the jury size. This would involve creating a Table 5 analogous to Table 3. In that new table, we would calculate a combination PIC by the formula

$$\text{PIC} = \frac{0.13(\text{ALP}) + 0.40}{1.13}$$

where ALP stands for the "at least" probability (for example, at least eleven out of twelve) that is calculated by using the binomial probability formula or the Poisson probability table. Since the ALP's have already been calculated in columns 2 and 4 of Table 3, we need only insert those numbers into the above formula and calculate our new combined PIC. For example, Table 3 gives 0.78 as the probability of convicting an innocent defendant with a fraction required to convict of 11/12. Therefore, the combination PIC would be

$$\text{PIC} = \frac{0.13(0.78) + 0.40}{1.13} = 0.44.$$

Similarly, since Table 3 gives 0.95 as the probability of convicting a guilty defendant with a fraction required to convict of 11/12, the combination PGC would be

$$\text{PGC} = \frac{0.13(0.95) + 0.70}{1.13} = 0.73.$$

Table 5 was prepared in the same manner, using the probabilities from Table 3 to calculate the revised data for our combination model.[34]

TABLE 5. REVISED CONVICTION PROBABILITIES AND NUMBER OF ERRORS FOR VARIOUS FRACTIONS REQUIRED TO CONVICT WITH 1000 DEFENDANTS
(Combined Independent and Collective Model)

Fraction Required to Convict (NC/NJ)	Prob. Innocent Conv. (PIC)	Number Innocent Conv. (#IC)	Prob. Guilty Conv. (PGC)	Prob. Guilty Not Conv. (PGN)	Number Guilty Not Conv. (#GN)	Unwtd. Sum of Errors (USE)	Weighted Sum of Errors (WSE)
12/12	.400	20.0	.700	.300	285.0	305.0	485.0
10/10	.407	20.4	.705	.295	280.6	301.0	484.1
8/8	.416	20.8	.710	.290	275.2	296.0	483.3
6/6	.426	21.3	.716	.284	269.7	291.0	483.0
*11/12	.444	22.2	.729	.271	257.7	279.9	479.6*
9/10	.449	22.5	.730	.270	256.6	279.1	481.3
7/8	.456	22.8	.731	.269	255.5	278.3	483.7
**10/12	.460	23.0	.732	.268	254.4	277.4	484.3**
5/6	.461	23.0	.732	.268	254.4	277.4	484.9
8/10	.462	23.1	.732	.268	254.4	277.5	485.5
9/12	.466	23.3	.733	.267	253.3	276.6	486.1
6/8	.467	23.3	.733	.267	253.3	276.6	486.7
7/10	.467	23.3	.733	.267	253.3	276.6	486.7
8/12	.468	23.4	.735	.265	252.2	275.6	486.1
4/6	.469	23.5	.735	.265	252.2	275.7	486.7
5/8	.469	23.5	.735	.265	252.2	275.7	486.7
6/10	.469	23.5	.735	.265	252.2	275.7	486.7
7/12	.469	23.5	.735	.265	252.2	275.7	486.7

Formulas for the columns (discrepancies are due to rounding):
PIC = [0.13 (Table 3 Prob.) + 0.40] / 1.13. #IC = (PIC)(50).
PGC = [0.13 (Table 3 Prob.) + 0.70]/1.13. PGN = 1.0 − PGC.
#GN = (PGN) (950).
USE = #IC + #GN. WSE = 10(#IC) + (#GN).
* Value of NJ where WSE is minimized, with one dissent allowed.
** Value of NJ where WSE is minimized, with two dissents allowed.

34. For ease in calculating the probabilities in Table 5, the PIC formula can be given as $\frac{0.13(\text{Table 3 Prob.}) + 0.40}{1.13}$, and the PGC formula as $\frac{0.13(\text{Table 3 Prob.}) + 0.70}{1.13}$, Similarly, the PIC formula in Table 4 could be given as $\frac{0.13(\text{Table 2 Prob.}) + 0.40}{1.13}$, and the PGC formula in Table 4 as $\frac{0.13(\text{Table 3 Prob.}) + 0.70}{1.13}$.

Table 5 shows that the smallest weighted sum of errors occurs when only one dissent out of twelve votes is allowed. When two dissents are allowed, a jury of twelve—as opposed to a smaller jury—still produces the smallest weighted sum of errors. These are the same results that appeared in Table 3; the WSE curve bottoms out at the same point, even though the curve is now flatter.

Although there is less variance in WSE values under the combination model, the values vary slightly more when we change the fraction required to convict than when we change the size of the jury as in Table 4. Thus, Table 5 shows that within the minimum constraints of a majority rule and a six- to twelve-person jury, the WSE can rise to 486.7 at a fraction required to convict of 7/12 and fall to 479.6 at a fraction required to convict of 11/12. Table 4 shows that within the minimum constraints of a unanimity rule (conviction requires no dissents) and a six- to twelve-person jury, the WSE can only rise to 485.0 at an NJ of twelve and fall to 483.1 at an NJ of seven. Because PIC is not dependent on any of the premises about W, #G, or PGC_{12}, a comparison of the range of PIC values in Table 4 with the range in Table 5 shows most clearly the difference between changing jury size and changing the fraction required to convict. Lowering NJ from twelve to six raises PIC from 0.40 to 0.43, an increase of eight percent; lowering the fraction required to convict from 12/12 to 7/12, however, raises the PIC from 0.40 to 0.47, an increase of 18 percent. Even if we assumed a PIC other than 0.40 for a twelve-person unanimous jury, lowering the fraction required to convict from 12/12 to 7/12 would still produce a substantial increase in the probability of an innocent person being convicted.

V. Effects of Changes in Assumptions or Data

Having constructed a reasonably complete model of jury behavior that predicts the jury size that will produce the minimum weighted sum of errors—the optimum jury—we should test the sensitivity of the model to changes in the empirical data and in the assumptions used in constructing it.[35] The model is stable with respect to most of its premises,

35. This will test our model's conformity to the first and fourth criteria established for a good deductive model: whether its premises are reasonable and how sensitive the conclusions are to changes in those premises. *See* note 10 *supra*. Sensitivity testing traditionally serves another function, however; it indicates the degree to which the model depends on data or premises that may be subject to considerable change. The *sensitivity*

but its predictions vary quite drastically according to the assumption we make about the proportion of all defendants who are truly guilty. This discovery exemplifies one of the strengths of model building: the identification of critical variables that are not intuitively obvious. In this case we can conclude with some assurance that the choice of an optimum jury size depends heavily on the assumption made about the proportion of truly guilty defendants among all defendants who receive jury trials.

A. *Effects of Changing the Normative Premises on the Optimum Unanimous Jury Size*

1. *Normative Premises Other than the Trade-Off Weight*

Our model assumes that the optimum jury should minimize the weighted sum of the number of innocent defendants convicted plus the number of guilty defendants not convicted. Another goal would be to minimize the number of innocent convicted. This choice can be expressed by giving #IC an infinite weight over #GN. Such a goal would result in an infinitely large optimum jury size deciding by unanimous decision, because only then would we never convict any innocent people—although we would never convict any guilty people, either. A more realistic goal would be to minimize the number of innocent convicted, with the constraint of a maximum jury size of twelve. As already noted, the use of a jury of twelve is the equivalent of giving #IC a weight of 13 (rather than Blackstone's weight of ten or a weight of infinity) and then minimixing the sum of #IC plus GN.[36]

At first glance, one might think a reasonable goal would be to minimize the weighted sum of PIC and PGN, in other words, the probabilities, rather than the number, of errors. It is not possible, however, to minimize the two kinds of errors without including both the probabilities and the number of innocent and guilty defendants out of 1000 or N defendants. Otherwise, we may have the irrational outcome of minimizing a percentage figure (PIC + PGN) but not the actual number of erroneous convictions (PIC (#I) + PGN (#G)), which equals each probability weighted by how often it is applied.

---

of a model is equivalent to its reliability in changed circumstances. To test the *validity* of a model, its predictions are matched against empirical results. As we noted earlier, predictions about the numbers of jury errors are inherently difficult to test empirically, and thus the deductive models proposed in this Article are also inherently difficult to validate completely.

36. *See* note 33 *supra* and accompanying text.

Some apparently different goals turn out to be identical to the minimum WSE with which we have worked. For instance, we might try to maximize the weighted sum of the guilty convicted plus the innocent not convicted or maximize the weighted difference of the guilty convicted minus the innocent convicted. Both goals are equivalent to minimizing the weighted sum of errors.

There are, however, some quite different criteria for determining optimum jury size. Such criteria, or goals, include representation of the community. A randomly selected twelve-person jury is more likely to be representative than a randomly selected six-person jury since, by chance, six jurors are all more likely to be unusual in some demographic dimensions than twelve jurors. For example, if blacks constitute ten percent of the population of a community, as they do in the United States as a whole, then the probability of obtaining twelve out of twelve whites on a jury is $(0.90)^{12}$ or 0.28. With only a six-person jury, the probability of obtaining an all white jury increases to $(0.90)^{6}$ or 0.53, or more than 50 percent. We could change our normative goal to minimizing

$$(\#IC)^{W_1}\,(\#GN)^{W_2}\,(PEM)^{W_3},$$

where PEM stands for the probability of completely excluding some minority from the jury, given the percentage the minority constitutes within the larger population from which jurors are drawn. Each of the three factors gets its own exponential weight, and the factors are multiplied, rather than added, because they are not all measured in the same units. A model such as this might give very different results from the one considered here.

Other goals are sometimes mentioned in the controversy over the desirability of jury trials versus bench trials. These goals, however, relate more to the question whether jury trials should be abolished in general, or in certain kinds of cases, than to the more limited question of what size a jury should be. For example, some people argue that the jury system produces excessive delay, but some of the delay attributed to juries has nothing to do with their size. Jury-selection and deliberation are probably somewhat longer with a twelve-person jury than with a six-person jury, but these considerations seem minor. The jury system is often defended because it gives the public a greater sense of involvement in the judicial process and an opportunity to inject communal opinion into the law. These considerations, however, are probably not substantially enhanced by changing the size of the jury. Similarly, the fraction required to convict has little effect on the goals of reducing court delay and increasing public participation, although a smaller fraction may speed deliberation and give more of a sense of majority control.[37]

2. *The Trade-off Weight*

The normative assumption that seems most subject to change is that a trade-off weight of ten is appropriate in evaluating the relative undesirability of convicting the innocent and not convicting the guilty. The figure ten is used because it was suggested by Blackstone, who influenced the authors of the Constitution, and also because it is a nice round number to work with. One alternative would be to use the trade-off weight that makes a twelve-person jury the optimum size for a unanimous jury. Since a twelve-person jury implies a trade-off weight of 13,[38] our tentative assumption of ten as a trade-off weight is probably a reasonable one.

The trade-off weight implicit in any given jury size, or the optimum jury size for any given trade-off weight, can be determined by inserting different values in the equation we previously developed for determining the minimum WSE as a function of NJ.[39] Experimenting with different trade-off weights (W) shows that when W is less than 9.5, the optimum jury size falls below six, which is the smallest jury that the Supreme Court has permitted in criminal cases; when W is greater than 13, the optimum jury size is larger than twelve, which is probably the maximum jury size that is politically feasible. This experimentation both confirms the reasonability of our choice of $W = 10$ and demonstrates that other choices would not greatly change the outcome of our calculations.

B. *Effects of Changing the Empirical Premises on the Optimum Unanimous Jury Size*

1. *The Premise that PGC_{12} is 0.70*

Because the equation[40] for minimizing WSE expresses the relation between optimum NJ and PGC_{12}, PIC_{12}, and #G, the equation can

37. For a discussion of the relevance of the jury system to public participation and court delay, see H. Zeisel, H. Kalven & B. Bucholz, Delay in the Court 71-109 (1959); Broeder, *The University of Chicago Jury Project*, 38 Neb. L. Rev. 744 (1959); Pabst, *Statistical Studies of the Costs of Six-Man Versus Twelve-Man Juries*, 14 Wm. & Mary L. Rev. 326 (1972).

38. *See* note 33 *supra* and accompanying text.

39. *See* notes 21, 31, 33 *supra.*

40. *See id.*

also be used to test the effects on optimum jury size of substituting different values for those three basic parameters. First with regard to PGC_{12}, we estimated the probability of a guilty defendant being convicted by a twelve-person unanimous jury to be 0.70. This probability cannot be less than 0.64, since that is the known probability of an average defendant being convicted by a twelve-person unanimous jury, and presumably a truly guilty defendant is at least as likely to be convicted as an average defendant. Nor can the probability be greater than 1.00, which is equivalent to certainty. In fact, PGC probably cannot be much higher than 0.70 unless the proportion of truly guilty defendants is much smaller than the 95 percent we have assumed. If 95 percent were truly guilty, a PGC of 0.71 would result in 695 out of 1000 defendants being convicted, but the Kalven and Zeisel data tell us that only about 640 out of 1000 defendants are convicted.

Testing the effect of varying PGC between 0.64 and 0.70 while holding the other variables constant, we find that as PGC goes down, the optimum jury size also goes down. In other words, as the guilty become harder to convict, smaller juries are needed because they are more likely to convict. A PGC below 0.68 would yield an optimum jury size below six. Since juries smaller than six are not now acceptable, however, if PGC were, in fact, much smaller than we have assumed, our normative goals would have to be revised. We could either remove the normative constraint excepting juries smaller than six persons or change the trade-off weight to make the weighted sum of errors fall between optimum NJ's of six and twelve.

In summary, PGC cannot be higher than the 0.70 value we have chosen and is unlikely to be much lower; to some extent, however, its value depends on the value of #G.

2. *The Premise that PIC_{12} is 0.40*

With respect to PIC_{12}, the probability of an innocent defendant being convicted before a twelve-person unanimous jury, we estimated a value of 0.40. PIC_{12} must be smaller than 0.64, since that is the known probability of the average defendant being convicted, and presumably a truly innocent defendant is less likely to be convicted than an average defendant. Nor can PIC fall below zero, since it is impossible to have a negative number of defendants convicted. If the number of truly innocent defendants is small, however, PIC, unlike PGC, can vary substantially without having much effect on the total number of defendants convicted. In other words, if only about 50 defendants out of 1000 are innocent, then PIC could be 0.60, rather than 0.40, and would add only 30 convicted innocent defendants (0.60 times 50), rather than 20 (0.40 times 50), to the 665 convicted guilty defendants (0.70 times 950). Whatever the trade-off weight given to #IC, changes in PIC have very little effect on our calculation of optimum jury size. With a 0.40 value for PIC, we get an optimum jury of seven which is in accord with Tables 2 and 4. For PIC = 0.20, the optimum jury is eight (after rounding to the nearest integer). This small increase in optimum jury size provides further confirmation that the model is not subject to wild fluctuations in its conclusions from small changes in its probability premises.

3. *The Premise That #G is 950*

The number of guilty defendants per 1000 defendants, #G, was estimated as 950.[41] If PGC is held constant, #G cannot vary greatly without causing the number of defendants convicted to vary substantially from the known figure of 640 convictions per 1000 defendants. For example, if #G were 600, then 420 guilty and 160 innocent defendants would be convicted, for a total of 580 convictions,[42] which is far below the known figure of 640. Similarly, if #G, were 980, then 687 guilty and 8 innocent defendants would be convicted, for a total of 695 convictions, which is far above the known figure of 640. Not only do such changes in #G create substantial variations from the known number of convictions, but even relatively small variations in #G, can cause large variations in optimum jury size. Thus, if #G = 900, the optimum jury size would be 23, which is the size of grand juries, rather than trial juries. As #G goes down, the optimum jury size rises; to offset the greater number of innocent defendants who are then being tried and need to be protected from conviction, a larger jury, which is less likely to err in that direction, is needed. Unlike the variables W, PGC_{12}, and PIC_{12} the variable #G adds an unstable element to the model. It is unstabilizing because it can vary widely, and such varia-

41. Note that the variable #I is not really a separate variable, since it is a function of #G by virtue of the definitional relation, #I = 1000 − #G or #G = 1000 − #I.

42. (#G)(PGC) + (#I)(PIC) = (600)(0.70) + (400)(0.40) = 420 + 160 = 580.

tions will produce large changes in the calculated optimum jury size.

Perhaps we have been too lenient, however, in allowing #G, to vary so widely; we may have needlessly allowed too much looseness in the model. What we should require is that PGC_{12}, PIC_{12}, and #G be estimated in such a way that jointly they (1) make theoretical sense in light of relevant aspects of the criminal justice system, such as the prosecutor's incentives to avoid prosecuting innocent persons, and (2) mathematically generate exactly 640 convictions without any upward or downward leeway. Because PGC and #G are mutually dependent,[43] empirical research would be required to determine their values when taken alone. It may be that the proportion of truly guilty defendants is lower than we assumed, but that juries only very rarely fail to convict guilty defendants. Nevertheless, the values of PGC and #G assumed for the model do jointly produce reasonable results in accord with available empirical data, which is the best we can do given the present state of knowledge. Perhaps the identification of these parameters as crucial will stimulate further empirical research on the behavior of juries, the relative number of truly guilty defendants, and the relations among these and other variables affecting jury errors.[44]

C. *Effects of Changing the Premises on the Optimum Nonunanimous Fraction Required to Convict*

Testing the effect of changes in our assumptions or empirical data is more difficult when we attempt to establish the optimum jury size under rules other than a rule of unanimity. This difficulty is due to the greater complexity that probability expressions introduce into the formula relating jury size to the other variables.[45] The calculations needed

43. PGC and #G are together constrained by the known proportion of 640 convictions per 1000 cases.

44. The reader can extend this sensitivity analysis by determining the effect that a change in one of the variables W, PIC_{12}, PGC_{12}, or #G would have on one or more of the others, rather than on the optimum jury size. The equation derived in note 31 *supra* could be used with NJ* set equal to twelve and all but two of the parameter variables set at their Table 1 values. One of the two parameters would then be changed to see how changing it affects the other parameter. Such an analysis demonstrates the interdependence of the variables in the model.

45. A generalized expression giving NJ as a function of WSE, comparable to the one developed in note 31 *supra*, would contain a series of probability expressions in the form of the binomial probability expression. *See* note 22 *supra*. Ordinary algebraic techniques are inadequate to arrive at a general solution for such an equation in terms of NJ, which appears as part of a series of complicated exponential expressions. Estab-

to test variations of all the variables would be extensive and difficult to perform, and would add very little to the qualitative discussion. To find the overall optimum result, we would first determine the optimum fraction for any given jury size, and then repeat the process for each possible jury size. From what we have already discovered, however, we know that there are feasibility and consistency constraints. Within those constraints, changes in the estimated PGC_{12} and especially PIC_{12} have little effect on the optimum fraction of votes needed for a conviction, just as feasible changes in those values have little effect on optimum jury size, although changes in #G can have substantial effects, depending on how much change is allowed. A more extensive analysis would show that the effects of changes in the normative premises, particularly with regard to the size of the trade-off weight, are somewhat narrowly confined if the other estimated variables are held constant. The direction, rather than the magnitude, of the direct or inverse relation between W, PGC_{12}, PIC_{12}, and #G on the one hand and the optimum number of votes on the other hand would be the same as in the relation with optimum jury size under a rule of unanimity.[46]

VI. VARIATIONS ON THE BASIC MODEL

The basic model, with its average jurors, juries, and cases, can be varied to consider dissimilarities among jurors and juries, different types of cases, criminal or civil, and the general effect of group size on the quality of group decisionmaking.

A. *Jury Representativeness with Respect to Conviction Probabilities*

1. *The Problem and the Formula for Resolving It*

Smaller juries are less likely than larger juries to be representative of the average jury. We assumed in both our independent-mind and collective-mind models that all twelve jurors were roughly alike and average. In our combination model, we likewise accept these assump-

---

lishing a minimum value for such an expression would be a difficult problem in generating reiterative approximations, and its difficulty would be far out of proportion to the value of the solution.

46. To determine the new optimum values when W, PGC_{12}, PIC_{12}, and #G are changed from the values we have been using of 10, 0.70, 0.40, and 950 out of 1000, we can apply the binomial probability formula, *see* note 22 *supra*, to create a new Table 3. From the last or WSE column of that new Table 3, we can observe when WSE is minimized and thus observe what the new optimums for different jury sizes are.

tions about the average jurors on juries of all sizes. Tables 4 and 5 might change, however, if we abandoned our assumption of homogeneity of juror characteristics.

The characteristic of jurors most important for predicting the outcome of criminal cases is not race, sex, age, political party affiliation, urbanism, or any other demographic characteristic, although such characteristics may have some predictive or explanatory value.[47] Nor is it an attitudinal characteristic such as economic liberalism, civil libertarianism, or even tolerance of criminal behavior.[48] Rather, the most important predictive characteristic of a juror is his propensity to convict. We know from the Kalven and Zeisel data that the average juror has a propensity to convict of 0.677, that is, if we examine an average sample of 100 juries or 1200 votes, we would find about 812 votes for conviction and about 388 votes for acquittal.[49] Thus, our key question is how the size of a jury affects the jury's representativeness with respect to juror propensity to convict.[50]

---

47. Some studies indicate that judicial background characteristics correlate with judicial decisions in divided cases. J. HOGARTH, SENTENCING AS A HUMAN PROCESS 211-28 (1971); Nagel, *Judicial Backgrounds and Criminal Cases,* 53 J. CRIM. L.C. & P.S. 333 (1962), *reprinted in* S. NAGEL, THE LEGAL PROCESS FROM A BEHAVIORAL PERSPECTIVE, ch. 18 (1969). Background characteristics tend to be especially explanatory when they are analyzed in combination rather than one at a time. *See* Goldman, *Voting Behavior on the U.S. Courts of Appeals Revisited,* AM. POL. SCI. REV. (1975); Nagel, *Multiple Correlation of Judicial Backgrounds and Decisions,* 2 FLA. ST. U.L. REV. 258 (1974), *reprinted in* S. NAGEL, IMPROVING THE LEGAL PROCESS: EFFECTS OF ALTERNATIVES, ch. 12 (1975). If the cases could be held constant, we would probably find an even stronger relation between jurors' backgrounds and their decisions than with judges. Jurors are freer to inject their backgrounds into their decisions because they do not have to write opinions justifying their decisions, they deal with more subjective factual issues, and they are not expected to follow precedent as closely.

48. *See, e.g.,* Becker, et al., *The Influence of Judges' Values on their Verdicts: A Courts and Politics Experiment,* SW. SOCIAL SCI. Q. 130 (1965); Boehm, *Mr. Prejudice, Miss Sympathy, and the Authoritarian Personality: An Application of Psychological Measuring Techniques to the Problem of Jury Bias,* 1968 WIS. L. REV. 734; Nagel & Weitzman, *Sex and the Unbiased Jury,* 56 JUDICATURE 108 (1972).

49. *See* note 25 *supra* and accompanying text.

50. Even though we can better predict how a jury will decide if we know the conviction propensities of each juror, it may still be useful to know the relation between decisions and backgrounds and attitudes. Backgrounds and attitudes can serve as indirect measures or predictors of conviction propensities. If we can show that the sex of jurors or juries correlates with case outcomes, we can more easily seek to have juries include more women. Demographic representativeness can also be justified for its symbolic significance in giving various demographic groups a feeling of participation in the judicial process, which possibly increases both their respect and that of others for the legal system.

If we have an infinitely large population of jurors, and the average juror in that population has a 0.677 propensity to convict, then we can determine approximately the range of the averages or means of 100 sets of twelve-person juries, or 100 sets of ten-, eight-, and six-person juries, drawn at random. We know that not all of the 100 sets of twelve-person juries would have 0.677 as the average of their individual jurors' propensities to convict. One set of twelve drawn from that infinite population might have a 0.823 average, and another set might have a 0.591 average. Presumably, if we added the averages for each twelve-person set and divided by 100, that overall average would be about 0.677, but the interesting question is how far from this overall average would the individual juries fall. What would be the range, for instance, in which we would expect to find the 50 sets of juries closest to the average?

To answer that last question involves some relatively simple arithmetic. Fifty percent of the juries are likely to have an average that falls within a range around 0.677 of plus or minus the square root of the quotient $\frac{0.677\ (1.0 - 0.677)}{NJ-1}$, with that square root multiplied by what is known as the critical t-score at the 0.50 level. Suppose, for example, that NJ is twelve. If we consult a comprehensive t-score table which is available in many statistics textbooks,[51] we find that at the 0.50 level with NJ - 1 equal to eleven, the t-score is 0.697. This means that 50 percent of our twelve-person juries would have pac averages within the range of 0.677 + or - 0.697 $\sqrt{0.219/11}$ where the 0.219 = (0.677) (1 - 0.677). In other words, 50 percent of our twelve-person juries would have averages between 0.579 and 0.775. Conversely, 50 percent of our twelve-person juries would have averages outside that range.

2. *Applying and Interpreting the Results of the Formula*

If we apply the same probabilistic reasoning to the drawing of 100 sets of six-person juries, then we would expect to find 50 percent of our six-person juries with pac averages within the range between 0.525 and 0.829. As the jury size becomes smaller, the range into which 50 percent of the juries are likely to fall becomes larger. Thus, while the

51. The t-table used is the one available in T. WONNACOTT & R. WONNACOTT, *supra* note 17, at 481, which includes the 0.50 level. For discussion of how to use the table with proportions, see *id.* at 174-78. A t-table is used because sample sizes of twelve and six are too small to use the more common normal-curve table.

50 percent range for twelve-person juries involves a 0.20 spread between 0.58 and 0.78, the range for six-person juries involves a 0.30 spread between 0.53 and 0.83, a substantial increase in both absolute and percentage terms.

We could apply the same probabilistic reasoning to the individual juror's propensity to convict an innocent defendant (pic) or the individual juror's propensity to convict a guilty defendant (pgc). For pic, 50 percent of the twelve-person juries fall within the range 0.36 to 0.57, a spread of 0.21.[52] With a six-person jury, the spread is 0.32. For pgc, we find a spread of 0.19 with a twelve-person jury and a spread of 0.29 with a six-person jury.

In short, when we go from a twelve-person, randomly selected jury, to a six-person, randomly selected jury, we can expect a substantial deviation from the average pac, pic, and pgc with which we may be working, whatever they might be. The probability figures shown in columns 2 and 4 of Tables 4 and 5 could now be shown as ranges rather than as points. The ranges will be substantially wider, and thus prediction will be substantially more difficult, as the jury becomes smaller. In other words, changing from twelve-person juries to six-person juries is likely to change substantially the probability of error in a given case, regardless of which of our models is used.[53] Specifically, a drop from twelve to six in jury size means about a 50 percent increase[54] in the range

52. If we assume (1) the independent probability pic is 0.926, as indicated in Table 1, (2) the collective mind pic is 0.400, also as indicated in Table 1, and (3) the relative weight of the independent probability approach is 0.13, as previously discussed, then the true pic is 0.461 by virtue of the formula:

$$\text{true pic} = \frac{0.13(0.926) + 0.400}{1.13}$$

With a twelve-person jury, we are thus asking what is 0.461 + or − 0.697 $\sqrt{0.248/11}$, where the 0.248 equals (0.461)(1.0 − 0.461).

53. These calculations also show the need for making jury selection more representative of important demographic and attitudinal characteristics, or at least more random, to minimize the erratic spread of average jury propensities. This kind of substantial change is associated with representativeness rather than the statistician's "type 1" errors of convicting the innocent, or "type 2" errors of not convicting the guilty. As such, this kind of substantial change does not consistently increase or reduce either kind of error. It simply makes the predictability of both kinds of errors and of jury decision-making in general more erratic and unpredictable.

54. Individual jury propensities with a spread of about 0.20 increase to a spread of about 0.30. KALVEN & ZEISEL 42. Although Kalven and Zeisel report acquittal percentages, approximate conviction percentages can be obtained by taking the complement of the figures given and subtracting six percent for hung juries, which are more like acquittals than convictions. One can obtain similar conviction percentages from the U.S.

around the mean in which cases are likely to fall. That seems to be a high cost to pay in predictability unless the benefits of smaller juries are worth this cost, in terms of reducing the number of guilty not convicted or attaining other values maximized by a smaller jury.

B. *Other Variations*

1. *Varying the Crime*

Our model is based on conviction data for major crimes, the kind of cases in the Kalven-Zeisel study. If, however, one wished to prepare a table like Tables 2 and 3 for certain specialized crimes, the Kalven-Zeisel data provide conviction probabilities for about 15 different major crimes.[55] For example, rape has a conviction percentage in the Kalven-Zeisel data of about 54 percent, less than the average conviction percentage of 64 percent. If a women's or other group were especially concerned about convicting the guilty in rape cases, that group might want to prepare a version of Tables 2 and 3 just for rape cases. Preparing the table would require determining a PGC_{12} that would be slightly greater than 0.54, a PIC_{12} that would be substantially less than 0.54, a pgc equal to $(PGC_{12})^{1/12}$, and a pic equal to $(PIC_{12})^{1/12}$. A different proportion of truly guilty defendants could also be assumed. Perhaps the lower conviction percentage for rape indicates a lower proportion of guilty defendants, although different crimes have different degrees of provability. If there is any extra importance to convicting the guilty in rape cases, it is also possible that the group might attach a smaller normative weight than ten to #IC.

The greater ease with which twelve-person unanimous juries convict defendants in minor criminal cases than in major criminal cases may indicate that society places a relatively lower weight on #IC in the minor cases. Nevertheless, it may be politically unfeasible, or even unconstitutional, to have different-sized juries for different kinds of crimes, except for very gross classifications, such as felonies versus misdemeanors. The model has enough flexibility to consider all categories of crimes, however, with their accompanying differences in PGC_{12}, PIC_{12}, #G, and the trade-off W. Different crimes may also have different empirical weights for the relative importance of the

BUREAU OF THE CENSUS, JUDICIAL CRIMINAL STATISTICS (1945), which are also reported in KALVEN & ZEISEL 42.

55. KALVEN & ZEISEL 63. The Kalven-Zeisel data also show that in personal injury cases alone, the average defendant was found liable 56 percent of the time. *Id.* at 64.

independent-mind model versus the collective-mind model, depending on the ability of the crime to stir up dissension and thus cause more independence and less collective action on the part of jurors.

2. *Civil Cases*

The model can also be applied to civil cases, although the symbolism would need changing: one might talk in terms of PLL and PIL, where PLL is the probability of a truly liable defendant being found liable and PIL is the probability of a truly innocent defendant being found liable. Kalven and Zeisel reported that in their sample of civil cases, the average defendant was found liable 59 percent of the time. Presumably, a truly liable defendant would have a probability of being found liable somewhat higher than 0.59, while a truly innocent defendant would have a probability that would be substantially lower than 0.59.

The normative weight attached to minimizing the number of innocent defendants found liable (#IL) should probably be substantially less than the weight attached to minimizing the number of innocent defendants convicted in criminal cases. By imposing a higher standard of proof in criminal cases and providing a right to free counsel only for indigent criminal defendants, society clearly indicates that it is more important to save the innocent from conviction in criminal cases than the innocent from liability in civil cases.

3. *Effect of Group Size on General Accuracy*

A final variation could consider the effect of large group size on the quality of a group's decisionmaking. It is possible that increasing a jury's size will improve the accuracy of its decisionmaking by reducing both the probability that the guilty are acquitted and that the innocent are convicted. This improvement in accuracy may occur if adding members to the jury means adding more knowledge, perception, intelligence, and deliberation, which may reduce both kinds of jury errors. Increasing the size of a group may, however, increase the irresponsibility of individuals and thus the sloppiness of the group product. The psychological data on the effect of group size on the quality of group decisionmaking is not only conflicting but also seems to apply principally to groups that decide by majority decision, or to advisory groups. If a group must decide unanimously, or nearly unanimously, as a criminal jury generally must, then the mathematics of unanimous decisionmaking that we have emphasized seems more relevant than the psychological studies determining the relation between group size and the number and kinds of errors likely to be made in other situations.[56] Nevertheless, if the psychological data could ascertain the effect of jury size on deliberation, irresponsibility, and error-making, then those findings could be incorporated into the model.[57]

VIII. CONCLUSIONS

Delivering the opinion of the Court in *Williams v. Florida*, Justice White stated, "[C]ertainly the reliability of the jury as a factfinder hardly seems likely to be a function of its size."[58] In *Apodaca v. Oregon*, speaking for the majority, he observed, "In terms of this function [of providing commonsense accuracy] we perceive no difference between juries required to act unanimously and those permitted to convict or acquit by votes of 10 to two or 11 to one."[59]

Because the empirical premises of our model have not been tested, we cannot definitively state how much effect jury size or the fraction required to convict has on the jury's reliability or accuracy. Nevertheless, Justice White seems to be wrong in saying that jury size and the fraction required to convict have no effect on reliability or accuracy. The model shows that the probability of an innocent person's being convicted increases both as the jury size decreases and as the fraction required to convict decreases, although the magnitude of these changes depends on the untested premises. These relations are not suprising, but rather are in conformity with common sense; the model simply clarifies these relations. Perhaps Justice White meant only that the increase is not great enough in his eyes to cross the threshold of unconstitutionality.

56. *See, e.g.*, GROUP DYNAMICS (D. Cartwright & A. Zander eds. 1968); I. JANIS, VICTIMS OF GROUPTHINK: A PSYCHOLOGICAL STUDY OF FOREIGN-POLICY DECISIONS AND FIASCOES (1972); George, *The Case for Multiple Advocacy in Making Foreign Policy*, 66 AM. POL. SCI. REV. 751 (1972). For a discussion of the psychological aspects of group decisionmaking as they pertain to juries, see Kessler, *An Empirical Study of Six- and Twelve-Member Jury Decision-Making Processes*, 6 U. MICH. J.L. REFORM 712 (1973); Powell, *Reducing the Size of Juries*, 5 U. MICH. J.L. REFORM 87 (1971); Rosenblatt & Rosenblatt, *Six-Member Juries in Criminal Cases: Legal and Psychological Considerations*, 47 ST. JOHN'S L. REV. 615 (1973).

57. The effect of jury size on general accuracy is unclear, especially when a unanimous rather than a majority rule is operating, but there are other institutional changes that might simultaneously reduce both kinds of error. These changes might include allowing jurors to take notes, having training sessions for jurors to improve their perceptions and memories, and providing for videotape replaying of portions of the trial. Better-qualified defense attorneys and prosecutors, as well as judges with larger operating budgets, might also reduce both kinds of error.

58. 399 U.S. 78, 100-01 (1970).

59. 406 U.S. 404, 411 (1972).

The other Justices similarly failed to indicate any substantial awareness of the probabilistic considerations involved in decreasing the size of the jury or the fraction required to convict.[60] They did, however, intuitively recognize that decreasing the unanimity rule has a greater effect on the chances of an innocent person being convicted than decreasing the jury size, at least within the factual ranges presented by a jury size of six in *Williams* and a fraction required to convict of ten out of twelve in *Apodaca*. That recognition is implicit in the number of dissenting votes in the two cases: one dissent in *Williams* versus four in *Apodaca*. That three additional judges should dissent in *Apodaca* is in conformity with a comparison of the effects of fractions required to convict in Tables 3 and 5 with the effects of jury size in Tables 2 and 4. If PIC_{12} is estimated to be 0.40, Table 3 indicates that an innocent person has a 92 percent chance of being convicted under a 10/12 rule, whereas Table 2 indicates that such a person has only a 63 percent chance of being convicted under a 6/6 rule.

Regardless of the value given to PIC_{12}, a 10/12 rule will always result in a higher probability of the innocent being convicted than a 6/6 rule. Under our independent probability model, the probability of convicting an innocent person with a six-person unanimous jury (PIC_6) is pic multiplied by itself six times. The probability of an individual juror convicting an innocent person (pic) is $(PIC_{12})^{1/12}$ since $(PIC_{12}) = (pic)^{12}$; therefore,

$$PIC_6 = [(PIC_{12})^{1/12}]^6 = (PIC_{12})^{0.5}.$$

Applying the same reasoning and the binomial probability formula, we find that the simplified $PIC_{10/12}$ formula is[61]

60. Justice Douglas, for instance, possibly exaggerates the effect of changing the fraction required to convict from 12/12 to 10/12 by saying the opinion of the court "permits prosecutors in Oregon and Louisiana to enjoy a conviction-acquittal ratio substantially greater than that ordinarily returned by unanimous juries." Johnson v. Louisiana, 406 U.S. 356, 388 (1972).

61. The $PIC_{10/12}$ equation was derived as follows, using P as a symbol for PIC_{12} and rounding 1/12 to 0.083:

1. $(P^{0.083})^{12} = (P)$

   (The probability of exactly twelve out of twelve)

2. $12(P^{0.083})^{11}(1 - P^{0.083}) = 12(P)^{0.917} - 12(P)$

   (The probability of exactly eleven out of twelve)

3. $(\frac{1}{2})(12^2 - 12)(P^{0.083})^{10}(1 - P^{0.083})^2 = 66(P)^{0.833} - 132(P)^{0.917} + 66(P)$

   (The probability of exactly ten out of twelve)

4. $66(P)^{0.833} - 120(P)^{0.917} + 55(P) = PIC_{10/12}$

   (The probability of at least ten out of twelve, which is the sum of the three probabilities above; *see* note 22 *supra*).

$$PIC_{10/12} = 55(PIC_{12}) - 120(PIC_{12})^{0.92} + 66(PIC_{12})^{0.83}.$$

The PIC_6 formula will always produce a lower probability than the $PIC_{10/12}$ formula, provided the same PIC_{12} probability is used in both formulas. As PIC_{12} approaches zero, both probabilities approach zero, and as PIC_{12} approaches 1.0, both probabilities approach 1.0. PIC_6, however, always remains lower than $PIC_{10/12}$. If PIC_{12} is between 0.10 and 0.90, as it is likely to be in reality, the difference is substantial. The direction of these differences remains unchanged when the independent probability model is converted into the combination independent and collective model.

This comparison of PIC_6 and $PIC_{10/12}$ shows that even without any numerical estimates for the parameters, PIC, PGC, #G, and W, the model presented in this Article can still be used to make comparisons between different jury decision rules. Rank-order comparisons can be developed for any two or more decision rules, reflecting their probabilities of convicting or not convicting the innocent, the guilty, or the average defendant.[62] With numerical estimates of the parameters, more precise comparisons can be made of degrees of difference rather than simple rank order. In addition, one can compare the numbers of each kind of error and thus determine an optimum jury size or fraction required to convict.

The comparison of PIC_6 and $PIC_{10/12}$ also demonstrates that intuitive notions about probability may be incorrect. Intuitively, one might believe that the probability of convicting an innocent person under a 6/6 rule is greater than under a 10/12 rule, because only six persons have to be convinced of guilt under the 6/6 rule, but ten have to be convinced under the 10/12 rule. That kind of reasonable thinking, however, does not adequately consider how much easier it is not to have to convince everybody on a jury, even though the total number of decisionmakers is somewhat larger. Probabilistic modeling can identify and correct such errors of commonsense when, as with jury size, direct experiments are difficult or impossible.

Nevertheless, if the model is wrong or totally untestable, a purely deductive approach may be just as useless as a purely empirical ap-

62. With a 10/12 or a 9/12 decision rule, a defendant may as well ask for a bench trial. If we assume a judge is like a typical juror, and assume the meaningfulness of the combined independent and collective probability model, then a defendant (whether innocent, average, or guilty) has about as much chance of being convicted with a 1/1 rule as with a 10/12 or 9/12 rule. Compare the 10/12 and 9/12 lines of Tables 5 and 3 with the 1/1 line of Tables 4 and 2.

proach. What seems to be needed is a basically deductive model, but one based on as much empirical data as can be obtained. The main empirical data we were able to use consisted of the Kalven-Zeisel findings that twelve-person unanimous juries convict 64 percent of the time in major criminal cases, and that the average juror votes to convict 67.7 percent of the time. Additional data might be obtained by interviewing legislative policymakers and appellate court judges to determine what trade-off weight they would prefer, and by interviewing judges, prosecutors, defense counsel, and experienced defendants to determine their perceptions of the number of guilty and innocent defendants per 1000 defendants.

With additional empirical data, we would have more confidence in our optimizing model and in the more precise effects indicated for alternative decision rules. As of now, we can consider the numerical data we have used in our optimizing model and our causal analysis to be illustrative data, designed mainly to demonstrate the methodology of the model and the analysis. Even without data, the model is capable of providing insights into the effects of different jury sizes and different fractions required to convict, as well as the effects of different normative and empirical premises on the optimum jury size or fraction required to convict, when the weighted sum of errors is minimized. In addition, the model can help persons recognize and understand the independent probability and collective-mind aspects of jury decisionmaking.[63]

With estimated data, the model can do all those things better. And with more accurate estimates, the model's capability for developing better causal theories and policy judgments about jury rules and jury decisionmaking would improve still further. Perhaps extensions of the jury model could be applied to other kinds of committees or small groups.[64] Certainly, there is a substantial potential here on which others may build.

---

63. Justice Powell seemed to be referring to the collective-mind or averaging perspective when in *Johnson v. Louisiana*, he stated, "the rule that juries must speak with a single voice often leads not to full agreement among the 12 but to agreement by none and compromise by all." 406 U.S. 356, 377 (1972) (concurring opinion).

64. For those who are more concerned with the optimum size and decisionmaking fraction for multimember courts rather than juries, Tables 4 and 5 may provide suggestive insights for analyzing those issues. *See* S. ULMER, COURTS AS SMALL AND NOT SO SMALL GROUPS 4-7 (1971). The model is also relevant to the more general issue of the number of individuals and the degree of unanimity that should be required to take collective action in other situations. *See generally* J. BUCHANAN & G. TULLOCK, THE CALCULUS OF CONSENT 63-84 (1962).