

THE SLIPPERY SLOPE

Jury Size and Jury Verdict Requirements—Legal and Social Science Approaches

BERNARD GROFMAN

University of California, Irvine

We review six recent U.S. Supreme Court cases dealing with the constitutionality of decision-making by juries with fewer than twelve members or operating under verdict requirements less stringent than unanimity. For the issues raised in each of these cases, we contrast the views of Supreme Court justices with data drawn from social science research and from statistical models of group decision processes.

REVIEW OF RECENT SUPREME COURT CASES ON JURY SIZE AND JURY VERDICT REQUIREMENTS

In the past decade there have been six Supreme Court cases which have dealt with the constitutionality of jury decisions reached by juries with fewer than twelve members and/or by juries permitting less than unanimous verdicts. In the first of these cases, *Williams v. Florida* (1970: 102), the Court held that

the fact that the jury at Common Law was composed of precisely 12 is a *historical accident*, unnecessary to effect the purposes of the jury system and wholly without significance 'except to mystics' [italics added].

AUTHOR'S NOTE: *This research was partially supported by NSF Grant SOC 77-24702, Law and Social Sciences Program. An earlier version of this article was delivered as a paper at a graduate colloquium of the Department of Political Science, State University of New York at Stony Brook in September 1979. I would like to acknowledge the assistance of my secretary, Sue Pursche, and the Word Processing Center of the School of Social Sciences, University of California, Irvine, in translating my handwritten scribbles into finished copy.*

LAW & POLICY QUARTERLY, Vol. 2 No. 3, July 1980 285-304
© 1980 Sage Publications, Inc.

In that case, in a seven-to-one decision (with Justice Marshall the lone dissenter) the Court held that six-member juries were constitutional in noncapital trials in state courts. In addition to its rejection of any historical requirement for twelve-member juries, the decision rested largely on the assertion (*Williams v. Florida*, 1970: 100) that

the essential feature of a jury obviously lies in the interposition between the accused and his accuser of the *commonsense judgment* of a group of laymen, and in the community participation and shared responsibility that results from the group's determination of guilt or innocence. *The performance of this role is not a function of the particular number of the body that makes up the jury.* [italics added].

The court did, however, also assert (*Williams v. Florida*, 1970: 101) that “what few experiments have occurred—usually in the civil area—indicate that *there is no discernible difference* between the results reached by the different-sized juries” (italics added), i.e., six vs. twelve.

In the next two cases, which were concurrently considered by the Court—*Johnson v. Louisiana* (1972) and *Apodaca v. Oregon* (1972)—the Court upheld the constitutionality of a nine to three verdict in a Louisiana felony trial involving a mandatory sentence of hard labor upon conviction and a ten to two verdict in a noncapital criminal trial in Oregon. Both these cases were decided by a five to four margin, with Justice White (as in *Williams*) delivering the opinion of the Court, and Justice Marshall now joined in dissent by Justices Brennan, Douglas, and Stewart. The dissenting justices (in a total of four separate but partly shared opinions) argued that a unanimous jury was embedded in legal history as a constitutional standard in criminal cases, and that it was basic to the accusational system and necessary to properly effectuate the fundamental constitutional standard of guilt beyond a reasonable doubt. The Court majority repeated its assertion in *Williams* that “the essential feature of a jury obviously lies in the interposition between the accused and the accuser of the commonsense judgment of a group of laymen” and

then went on to claim (*Apodaca v. Oregon*, 1972: 410-411) that

a requirement of unanimity . . . does not materially contribute to the exercise of this commonsense judgment. As we said in *Williams*, a jury will come to such a judgment as long as it consists of a group of laymen representing a cross section of the community who have the duty and the opportunity to deliberate, free from outside attempts at intimidation, on the question of a defendant's guilt. In terms of this function, we see no difference between jurors required to act unanimously and those permitted to convict or acquit by votes of 10 to two or 11 to one.

Earlier (*Williams v. Florida*, 1970: 361-362) the Court majority stated that

we have no grounds for believing that majority jurors, aware of their responsibility and power over the liberty of the defendant would simply refuse to listen to arguments presented to them in favor of acquittal, terminate discussion, and render a verdict. . . . Appellant offers no evidence that majority jurors simply ignore the reasonable doubts of their colleagues or otherwise act irresponsibly in casting their votes in favor of conviction.

In the next case, *Colegrove v. Battin* (1973), again decided by a five to four lineup but with Justice Powell now in the minority and Justice Brennan now in the majority, the Court ruled that six-member federal civil juries met the Seventh Amendment requirement of trial by jury. As in *Williams*, the argument rested primarily on a reading of the historical evidence that twelve-member juries were not constitutionally mandated. The Court also made the quite striking claim (*Colegrove v. Battins*, 1973: n.15; italics added) that "four very recent studies have provided *convincing empirical evidence* of the *Williams* conclusion that 'there is no discernible difference'" in verdicts between six-member and twelve-member juries.

These court decisions met with a mixed reaction. While some judges hailed them (Bloom, 1973), most legal scholars condemned them as a threat to the integrity of the "guilt beyond a reasonable doubt" standard (Zeisel, 1971; *New York Times*, 1972), repudiated the accuracy of the historical arguments on

which they were based (Rosenblatt and Rosenblatt, 1973), strongly disputed the “no discernible difference” claim (Lempert, 1975), and criticized as methodologically flawed the social science studies which purportedly supported this claim (Zeisel and Diamond, 1974; Diamond, 1974; Saks, 1977). Nonetheless, these court decisions triggered a reduction in jury size or jury verdict requirements in a number of states (Delsner, 1975) and, for civil cases, in the federal courts as well (Sperlich, 1979: 201, n.43).

The fifth Supreme Court case involving jury decision-making was one in which the Court confronted the constitutionality of Georgia’s five-member criminal juries, *Ballew v. Georgia* (1978). Justice Blackmun posed the issue in *Ballew* as follows:

When the Court in *Williams* permitted the reduction in jury size . . . it expressly reserved ruling on the issue whether a number smaller than six passed constitutional scrutiny. . . . The Court refused to speculate when this so-called ‘*slippery slope*’ would become too steep. We face, now, however, the two-fold question whether a further reduction in the size of the state criminal jury does make the grade too dangerous, that is, whether it inhibits the functioning of the jury to a significant degree, and, if so, whether any state interest counterbalances and justifies the disruption so as to preserve its constitutionality [*Ballew v. Georgia*, 1978: 230-231; italics added].

In *Ballew*, the court drew a line as to the minimum jury size constitutionally permitted by requiring juries in state court criminal trials to consist of at least six members. Georgia’s five-member jury was rejected as threatening Sixth and Fourteenth Amendment guarantees.¹ In *Williams*, the court had declined to judge what minimum number can still constitute a jury but asserted that “we do not doubt that six is above that minimum” (*Williams v. Florida*, 1970: 92, n.28).

Justice Blackmun announced the judgment of the Court, which was unanimous in rejecting five-member juries. However, there was little agreement on the court as to the reasons underpinning that judgment. There were four separate opinions in the case. Only Justice Stevens joined in the Blackmun opinion, whose line of reasoning was savagely attacked by Justice Powell (joined by

Justice Burger and Justice Rehnquist). Blackmun's opinion made extensive use of social science studies of verdict outcomes and of the nature of the deliberation process in mock juries, cited approvingly a statistical model of jury judgmental accuracy developed by two social scientists (Nagel and Neef, 1975), and reviewed other studies of small group decision processes as well. In his opinion in *Ballew*, Blackmun asserted that smaller juries have been shown to be less representative, less reliable, and less accurate than larger juries. One leading scholar has called Blackmun's opinion in *Ballew* the most extensive use ever made of social science by the court—one in which, for the first time, "social science moved out of the footnotes and into the body of the text" (Zeisel, 1978). Justices Powell, Rehnquist, and Burger, however, expressed "reservation as to the wisdom—as well as the necessity of Mr. Justice Blackmun's heavy *reliance on numerology* derived from statistical studies" (*Ballew v. Georgia*, 1978: 246, italics added). Powell's opinion in *Ballew* goes on to say that

neither the validity nor the methodology employed by the studies cited was subject to the traditional testing mechanisms of the adversary process. The studies relied on merely represent *unexamined findings* of persons interested in the jury system [*Ballew v. Georgia*, 1978: 246, italics added].

The sixth and most recent court case, *Burch v. Louisiana* (1979), appears to have been written so as to be able to support the claim that it does not require either empirical social science data or statistical models for constitutional scholars, such as Supreme Court justices, to decide what jury verdict requirements violate constitutional guarantees of the right to trial by jury. In *Burch*, there was unanimity as to the principal holding that a five of six verdict for a state criminal trial was unconstitutional. The opinion of the court in *Burch*, written by Justice Rehnquist, cited no social science studies at all. It supported its conclusion as to the unconstitutionality of nonunanimous verdicts in six-member juries by recourse to two arguments. First, Justice Rehnquist pointed out that "of those states that utilize six-member juries in trials of nonpetty offenses, only two, including Louisiana, also allow nonunanimous verdicts" (*Burch v. Louisiana*, 1979: 1628).

Then Rehnquist went on to assert that “we think that this near-uniform judgment of the Nation provides a useful guide in delimiting the line between those jury practices that are constitutionally permissible and those that are not” (*Burch v. Louisiana*, 1979: 1628). Second, Rehnquist asserted that

we think when a State has reduced the size of its juries to the minimum number of jurors permitted by the Constitution, the additional authorization of nonunanimous verdicts by such juries sufficiently threatens the constitutional principle that led to the establishment of the size threshold that any countervailing interest of the State should yield [*Burch v. Louisiana*, 1979: 1628].

Of course, if the first argument is accepted, then perhaps *Johnson* and *Apodaca* should have been decided differently, since nonunanimous verdicts in state felony trials were the exception rather than the rule when these cases were decided (*Institute for Judicial Administration*, 1971a, 1971b). As to the second argument, it is simply unsupported assertion. Rehnquist provides no evidence that a five of six (83%) rule is more likely to threaten “constitutional principles” than the nine of twelve (75%) rule previously found to be constitutional in *Johnson*. It is interesting to compare Justice Rehnquist’s standard of proof in *Burch* with Justice White’s statement in *Johnson* (*Johnson v. Louisiana*, 1972: 362) that “before we alter our own longstanding perceptions about jury behavior and overturn considered legislative judgment that unanimity is not essential to reasoned jury verdicts, we must have some basis for doing so other than unsupported assumptions.”

With *Ballew*, the court set the minimum jury size at six. With *Burch*, the court rejected nonunanimous verdicts for six-member juries. In principle, still open is the question of where the court will ultimately draw the line on nonunanimous verdicts. With nine of twelve verdicts permitted by *Johnson*, what about eight or twelve or even seven of twelve? There is little in Rehnquist’s opinion in *Burch* or in White’s opinion in *Johnson* or *Apodaca* to aid us.

Rehnquist's opinion in *Burch* cites a dictum in an earlier case (*Duncan v. Louisiana*, 1968: 161) about the process of drawing lines which "although essential, cannot be wholly satisfactory, for it requires attaching different consequences to events which, when they lie near the line, actually differ very little" (*Duncan v. Louisiana*, 1968: 161). Justice Powell said in *Ballew* that "a line has to be drawn somewhere" and of course, this is true. It does not follow, however, that the line has to be drawn arbitrarily and without good reason. As Justice Blackmun says in *Ballew* in justification for his heavy reliance on social science studies and analytic models of the jury decision process:

We have considered them [these studies] carefully because they provide the only basis *besides judicial hunch*, for a decision about whether smaller and smaller juries will be able to fulfill the purpose and functions of the Sixth Amendment. Without an examination about how juries and small groups actually work, we would not understand the basis for the conclusion of Mr. Justice Powell that 'a line has to be drawn somewhere' [*Ballew v. Georgia*, 1978: n.10; italics added].

Having briefly reviewed recent Supreme Court decisions in the jury area, let us now turn to the data and models in the social science literature which bear on the issue of drawing the line between "permissible" and "impermissible" jury size/decision requirements. In doing so, we shall address both the question of how social science findings were actually used by the court and the question of how social science findings might (ideally) have been used by the court.

ASSESSING THE IMPACT OF CHANGES IN JURY SIZE AND/OR JURY VERDICT REQUIREMENTS

We shall review the social science literature on jury decision, focusing on modeling the implications of changes in jury decision rule and/or jury size for (a) verdict outcomes, (b) the nature of the

jury deliberation process, (c) jury representativeness, and (d) the achievement of substantive justice. For each of these issues, we shall compare social science findings with the assertions made by Supreme Court justices.

Verdict Outcomes

To model verdict outcomes as a function of jury size/verdict rule, we require (1) a population pool with some posited distribution of characteristics related to juror verdict preferences, (2) a sampling mechanism, and (3) a model of the process through which individual juror predeliberation verdict preferences are transformed into a final jury verdict.

The two most common models of underlying juror sampling characteristics are the one-parameter model, in which p is the probability that a randomly chosen juror can be expected to vote to convict (Walbert, 1971; Saks and Ostrom, 1975; Grofman, 1976); and the two-parameter model, in which p_{GG} is the probability that a randomly chosen juror will vote to convict a guilty defendant and p_{GI} is the probability that a randomly chosen jury will vote to convict an innocent defendant (Grofman, 1974, 1980c; Gelfand and Solomon, 1973, 1974, 1975, 1977).

Three approaches have commonly been taken to modeling the jury deliberation process. The simplest is to assume that the jury deliberation process has no effect and that jury outcomes can be treated as if perfectly predicted by juror predeliberation preferences (Saks and Ostrom, 1975). A second, somewhat more sophisticated view, is to assume that the de facto quorum rule and the de jure quorum rule need not coincide, but that both can be described as a K/N decision process—e.g., juries ostensibly requiring unanimity will reach unanimous accord once, say, eight of twelve are in agreement as to verdict (Gelfand and Solomon, 1973; Grofman, 1974, 1976). The third and most realistic approach is to assume that the jury deliberation process can be represented by a $(N + 1) \times 3$ matrix whose cell matrices provide the likelihood that any given lineup of predeliberation preferences will eventuate in final verdicts of acquittal, conviction, or in

a hung jury (Davis, 1973; Davis et al., 1975; Gelfand and Solomon, 1977; Grofman, 1980a).²

Space does not permit us to review these models in any further detail (see Penrod and Hastie, 1979; Grofman, 1980a). Suffice it to note that the most sophisticated modeling efforts now available predict aggregate conviction differences between twelve-member and six-member juries of only one to three percentage points (Gelfand and Solomon, 1977; Grofman, 1980a);³ and for juries of a given size, predict little impact of reductions in jury unanimity requirements as long as at least a two-thirds majority is required (Grofman, 1976, 1979a, 1980b).

In the University of Chicago jury project's study of 225 criminal cases in Chicago and Brooklyn courts, 92% of the verdicts in the twelve-member juries accorded with the views of the initial majority, 5% of the juries remained hung, and in only 3% of the cases did the minority persuade the majority (Kalven and Zeisel, 1966: ch. 37). A reanalysis of the Kalven and Zeisel (1966) data showed (Grofman, 1980b) that even if the 225 cases examined by Kalven and Zeisel had been decided by simple majority, there would have been little difference in aggregate verdict outcomes. Under simple majority verdict, there would have been 65% convictions and 30% acquittals and (assuming no further revotes for juries split six to six) 4% hung juries compared to the 63% convictions, 32% acquittals, and 5% hung juries obtained under a unanimous verdict requirement. There does, however, appear to be some limited evidence (see, e.g., Nemeth, 1977) that minorities for acquittal are more resistant to majority persuasion than proconviction minorities.

Let us now turn to what various justices have had to say about the verdict consequences of changes in jury decision requirements and changes in jury size.

In *Johnson v. Louisiana* (1972: 391), Justice Douglas claimed that "the use of the nonunanimous jury stacks the truthdetermining process gainst the accused." Elsewhere (Grofman, 1980b), we have shown that the conclusion rests on a misreading of the data in Kalven and Zeisel (1966). When Douglas asserts (*Johnson v. Louisiana*, 1972: 389) that "initial

majorities normally prevail in the end but about a tenth of the time the rough and tumble of the jury room operates to *reverse completely* their preliminary perception of guilt or innocence," he is off by a factor of three. "A correct reading of the data in Kalven and Zeisel (1966) leads to the conclusion that verdict reversals occur not the nearly 10% of the time that Douglas claimed, but rather, roughly 3% of the time" (Grofman, 1980b, and see discussion above).

In *Williams* (*Williams v. Florida*, 1970: 101), Justice White asserted that lowering the sizes of juries from twelve to six (while preserving the unanimity requirement) would have no "discernible difference" on verdict outcomes, a view which he reasserted in *Colegrove* (*Colegrove v. Battin*, 1973: 159). We believe that Justice White was correct in this view only to the extent that (a) it is only aggregate percentages of convictions and acquittals which are ever discernible, not verdict differences in particular trials (see n.3); and (b) as previously noted, the expected differences in aggregate conviction rates that various scholars have predicted, using the most sophisticated models now available, are on the order of only a few percentage points. However, while we agree with Justice White's conclusion, we find the reasoning which led him to make it to have been erroneous on both occasions.

In *Williams*, Justice White cited six studies in support of his claim that what few "experiments have occurred" show that lowering the size of juries from twelve to six would have no "discernible difference" on verdict outcomes. However, as Saks trenchantly (and quite correctly) puts it, of the six "experiments"

the first was a mere assertion with no evidence; the next three were casual observations of the verdicts rendered by smaller juries, two of those reports being of the same set of juries; the fifth was merely a report that a smaller jury had been used; and the sixth was a discussion of economic advantage, irrelevant to the question [Saks, 1977: 10].

In *Colegrove*, Justice White (*Colegrove v. Battin*, 1973: 150) asserted that since 1970 "much has been written about the

sixmember jury but nothing that persuades us to depart from the conclusion reached in Williams.” He then goes on to assert that “four very recent studies have provided convincing empirical evidence of the correctness of the Williams conclusion” of no “discernible difference” (Colegrove v. Battin, 1973: 158-159, n. 13). These four studies have been reviewed and devastatingly critiqued by Zeisel and Diamond (1974) and Saks (1977: 37-49), and we shall not bother to repeat their lengthy criticisms here (see also Grofman, 1980b). Suffice it to state that the four studies were sufficiently marred by *ceteris paribus* problems or other methodological flaws as to be useless for proving anything one way or the other about probable verdict differences between six-member and twelve-member juries. Convincing empirical evidence they were not—a point fully acknowledged by Justice Blackmun in his opinion in Ballew (Ballew v. Georgia, 1978: n. 30).

Following Saks (1977: 86-87), Justice Blackmun in Ballew makes the claim that smaller juries are less consistent in their verdicts than larger ones. By this he means that if we imagined the same case tried repeatedly before a number of different juries drawn from the same jury pool (as has been done in mock jury research, such as that of Saks), then the larger juries will reach agreement on the same verdict a higher proportion of time than will the smaller. In a study cited in the Blackmun opinion in Ballew, Saks (1977) finds, for example, in one study that the preponderant verdict was reached by 83% of twelve-member juries but by only 69% of the (unanimous) six-member juries. While one can construct hypothetical social decision schema for six-member and twelve-member juries for which the six-member juries will be *more* consistent than the twelve-member juries, based on the available empirical evidence on the actual social decision scheme likely to be operative in juries, we concur with Blackmun’s judgment that the larger juries will be the more consistent.⁴ However, with sufficiently large *N*s, it is likely that the differences in the percentage of juries reaching the preponderant verdict between six-member and twelve-member juries will be considerably smaller than the fourteen percentage points found in the Saks (1977) study, which Blackmun cites.

Jury Deliberation

Justice White, discussing nonunanimous verdicts, asserted in *Johnson* (*Johnson v. Louisiana*, 1972: 361) that “we have no grounds for believing that majority jurors . . . would simply refuse to listen to arguments presented to them in favor of acquittal, terminate discussion and render a verdict.” Rebutting this sanguine view, Justice Douglas, speaking for himself and two of the three other members of the minority in *Johnson* (*Johnson v. Louisiana*, 1972: 388), asserted that “as soon as the requisite majority is obtained further consideration is not required . . . even though the dissident jurors might, if given the chance, be able to convince the majority.”

We have no grounds for believing either Justice White or Justice Douglas to have been correct.

The best available evidence for juries which do not require unanimous verdicts is that the presence of minority jurors after the necessary votes are identified

causes continued deliberation and occasional further votes. But what happens in the epilog interval has no effect on the outcome of the trial. The minimum vote decision is as psychologically binding on the nonunanimous jury as the unanimous consensus is for a unanimous jury [Saks, 1977: 94; italics added; see also Nemeth, 1977).⁵

As for the relationship between jury size and jury deliberation, we have already noted that, according to Justice White, there were no “reason(s)” to think the requirement stated in *Williams* that “the number (of jurors) should . . . be large enough to promote group deliberation, free from outside attempts to intimidation” was “less likely to be achieved when the jury numbers six than when it numbers 12”—particularly if the requirement of unanimity is retained (*Williams v. Florida*, 1970: 100-101).⁶ Of course, as noted previously, Justice White provides no indication as to what these reasons might be.

Taking a quite different view, in *Ballew*, Justice Blackmun asserts that “recent empirical data suggest that progressively

smaller juries are less likely to foster effective group deliberation,” noting in particular that “as juries decrease in size . . . they are less likely to have members who remember each of the important pieces of evidence or argument” (*Ballew v. Georgia*, 1978: 232, and footnote citations therein). These conclusions seem not unreasonable ones, but rest too heavily on what is *very* limited empirical evidence (notably Saks, 1977). Furthermore, it is not clear why jury deliberations are relevant *except* insofar as they relate to jury fact-finding accuracy. In his discussion of jury deliberation, Blackmun does cite studies dealing with group accuracy and also with the issue of community representativeness. These topics we shall, however, discuss separately below.

Jury Representativeness

According to Justice White in *Williams*,

While in theory the number of viewpoints represented on a randomly selected jury ought to increase as the size of the jury increases, *in practice the difference between the 12-man and the six-man jury in terms of the cross-section of the community represented seems likely to be negligible* [*Williams v. Florida*, 1970: 102, italics added].

As before, Justice White offers not one scintilla of evidence for his claim.

Justice Blackmun in *Ballew* relying on Lempert (1975)—who uses a straightforward binomial model to generate probability values—sets forth the following results:

If a minority viewpoint is shared by 10% of the community, 28.2% of 12-member juries may be expected to have no minority representation, but 53.1% of six-member juries would have none. Thirty-four percent of 12-member panels could be expected to have two minority members while only 11% of six-member panels would have two. As the number diminishes below six even fewer panels would have one member with the minority viewpoint and still fewer would have two [*Ballew v. Georgia*, 1978: 237].

Differences which appeared negligible to Justice White's intuition appeared to Justice Blackmun to be quite substantial in light of Lempert's (1975) exact calculations, a conclusion which we share (see also Grofman, 1980a).

Jury Accuracy

Justice White, in *Williams*, asserts that "the reliability of the jury as a fact-finder hardly seems likely to be a function of its size" (*Williams v. Florida*, 1970: 100). Justice Blackmun in *Ballew*, citing several studies by social psychologists, takes a diametrically opposed point of view.

The smaller the group, the less likely is it to overcome the biases of its members to obtain an accurate result. When individual and group decision making were compared, it was seen that groups permitted better decisions because prejudices of individuals were frequently counterbalanced, and objectivity resulted. Groups also exhibited increased motivation and self-criticism. All these advantages, except perhaps self-motivation tend to diminish with group size [*Ballew v. Georgia*, 1978: 233, footnotes deleted].

The issue of jury judgmental accuracy is an extremely complex one, and we shall do no more than touch upon it here (see Grofman, 1974, 1979a, 1980a, 1980c). The most sophisticated work on the judgmental accuracy of juries is that of Gelfand and Solomon (1973, 1974, 1975, 1977) who make use of the two-parameter model of the jury decision process and an adaptation of the Davis (1973) social decision scheme approach. Modeling data from Kalven and Zeisel (1966), they find (Gelfand and Solomon, 1977) that for twelve-member juries the probability of convicting an innocent person is .0221 and the probability of acquitting a guilty person is .0615. For six-member juries, these errors are increased by more than 50%—the error probabilities became .0325 and .1395 respectively. From these results, they conclude that twelve-member juries are superior to six-member juries. Gelfand and Solomon (1977) also look at judgmental accuracy of six-member and twelve-member juries. Here they

find, quite counterintuitively, that majority verdict juries are *less* likely to make errors of either the Type I or the Type II kind than unanimous juries of the same size—although twelve-member juries are still more accurate than six-member ones for both sorts of errors. However, a six-member majority verdict jury is actually considerably *less* likely to convict an innocent defendant than is a twelve-member jury operating under unanimity (.006 v. .022). For a discussion of these findings and why they must be interpreted with considerable caution, see Penrod and Hastie (1979) and Grofman (1980a).

In *Ballew*, a model of jury judgmental accuracy proposed by Nagel and Neef (1975) which is conceptually quite similar to those offered by Gelfand and Solomon and by Grofman is given considerable prominence in Blackmun's opinion.⁷

Statistical studies suggest that the risk of convicting an innocent person (Type I error) rises as the size of the jury diminishes. Because the risk of not convicting a guilty person (Type II error) increases with the size of the panel, an optimal size can be selected as a function of the interaction between the two risks. Nagel and Neef (1975) concluded that the optimal size, for the purpose of minimizing errors should vary with the importance attached to the two types of mistakes. After weighting Type I error as 10 times more significant than Type II, perhaps not an unreasonable assumption, they *concluded that the optimal jury size was between six and eight*. As the size diminished to five and below, the weighted sum of errors increased because of the enlarged risk of the conviction of innocent defendants [*Ballew v. Georgia*, 1978: 233; italics added, footnotes deleted].

While in general we have high praise for Blackmun's use of social science in his *Ballew* opinion, in summarizing the Nagel and Neef (1975) study he neglected the fact that it rested not just on one assumption (a tradeoff ratio of ten between Type I and Type II errors), but on several. The supposed optimality of juries between six and eight rests on an assumption made as to the number of defendants who are truly guilty. If that number is varied even slightly, a quite different optimal jury size is reached by Nagel and Neef (1975: 967-968), a fact which they, as

conscientious social scientists, are careful to point out. In particular, the Nagel and Neef (1975) study cannot be used as Blackmun used it, to justify rejecting five-member juries while accepting juries with six members.

CONCLUSIONS

We share the view of Peter Sperlich (1979: 208) that, in the jury cases, with the notable exception of Justice Blackmun's opinion in *Ballew*, "the Court's use of empirical evidence is uniformly dreadful." Even Justice Blackmun, however, is not without fault. Although his opinion is admirable in the range of social science research which it summarizes, and in its handling of complex issues and sophisticated statistical concepts, it is flawed "because . . . the Court's desire to preserve *Williams* required the misreading of texts and the skirting of inferences clearly demanded by the "date" (Sperlich, 1979: 224).

First, the studies cited by Blackmun which deal with jury or mock jury decision making all refer exclusively to comparisons of *six*-member and twelve-member juries. Thus, if these jury studies are to be used to conclude that smaller juries are constitutionally impermissible, it should be the case that it is *six*-member juries which are to be held impermissible. Second, the methodological studies which Blackmun approvingly cites (e.g., Zeisel and Diamond, 1974; Diamond, 1974) are ones which undermine totally the reliability of the studies cited in *Colegrove* to support the claim that six-member juries are not discernibly different from twelve-member juries. Hence, if those methodological antiques are to be taken seriously, as Blackmun apparently does, then the major empirical props underpinning the legitimacy of six-member juries have been knocked down. Third and finally, the single jury study which Blackmun cites to sustain the decision to draw the line at five but not at six does not justify such a conclusion [Grofman, 1980b; some renumbering].

With the considerable body of social science mock-jury data and modeling efforts now available to draw upon (much but not all of which was cited by Justice Blackmun in *Ballew*), a strong

case for the superiority of twelve-member juries over six-member juries can be made. Such a case is unlikely to be made by the present court. Judging from the mishandling of social science evidence in *Williams* and *Colegrove*, the complete absence of social science evidence from Justice Rehnquist's opinion in *Burch* and Justice White's opinion in *Johnson*, Justices Powell, Burger, and Rehnquist's views on social science as numerology, and the fact that only Justice Stevens joined Justice Blackmun's opinion in *Ballew*, law and social science continue as at best uneasy bedfellows in the Burger Court as they have been in previous courts. At issue here is the practical incompatibility between social science standards of proof and judicial needs to reach reasonable decisions on specific cases.

NOTES

1. If asked to guess, we would predict that the court will hold the line at nine of twelve. Our reason for guessing so is that one justice in the *Johnson* majority, Justice Blackmun, is on record as expressing dislike for rules which do not require a "substantial majority of the jury to be convinced" (*Johnson v. Louisiana*, 1972: 366). Also, nine minus three is six, and White's dictum in *Williams* (and, the court's unanimous judgment in *Ballew*) suggests that six is the "magic" number. A nonunanimous verdict of say, eight to four leaves a preponderance of only four jurors on the majority side. Of course, given such reasoning, an eight of ten (80%) verdict requirement would be permitted even though a six of seven (86%) verdict requirement would not be.

2. Still, a fourth approach is to represent deliberation as a sequential process and to model it either in terms of simulation (Penrod and Hastie, 1979) or in terms of a stochastic process such as a Markov chain (Klevorick and Rothschild, 1979).

3. That some *particular* trials may be decided differently under juries of different sizes we have no doubt, although in comparing the proportion of such trials as less than 10% (see Grofman, 1980b; Lempert, 1975). On the other hand, we fully agree with Blackmun's observation in *Ballew* that "nationwide . . . small percentages will represent a large number of cases. And it is with respect to these cases that the jury trial might have its greatest value." In civil cases, smaller juries can be expected to have a higher variance in damage awards (Zeisel, 1971; Lempert, 1975).

4. Zeisel's (1971) analysis suggests that in criminal cases, the percentage of hung six-member juries will be half that for twelve-member juries, 2.4% vs. 5%.

5. In fairness to Justice Douglas, he points out that nonunanimous verdicts may also eliminate the circumstances under which a minority "while unable to persuade the majority to acquit, nonetheless could have convinced them to convict only on a lesser-included offense" (*Johnson v. Louisiana*, 1972: 188).

6. When juries are allowed to reach nonunanimous verdicts, the probability that the jurors will already have achieved sufficient consensus for a verdict before they begin deliberation is extremely high in smaller juries (see Grofman, 1976).

7. Neither the work of Gelfand and Solomon (1973, 1974, 1975, 1977) nor the very similar work of Grofman (1974, 1980a, 1980c), has ever been referenced in opinions by the justices. Although the work of the latter was called to the court's attention in the defendant's oral argument in *Burch v. Louisiana* (1979), as noted above that was a case in which no social science studies were cited in the opinion. For a straightforward introduction to (and critique of) the Gelfand and Solomon, Grofman, and Nagel and Neef approaches to jury accuracy, see Penrod and Hastie (1979).

CASES

APODACA v. OREGON (1972) 406 U.S. 404.
BALLEW v. GEORGIA (1978) 435 U.S. 223.
BURCH v. LOUISIANA (1979) 99 S. Ct. 1623.
COLEGROVE v. BATTIN (1973) 413 U.S. 149.
DUNCAN v. LOUISIANA (1968) 391 U.S. 145.
JOHNSON v. LOUISIANA (1972) 406 U.S. 356.
WILLIAMS v. FLORIDA (1970) 399 U.S. 78.

REFERENCES

- BLOOM, M. (1973) "Here come the six-man jury." *Reader's Digest* (September): 126-129.
- DAVIS, J. H. (1973) "Group decision and and social interaction: a theory of social decision schemes." *Psych. Rev.* 80, 2 (March): 97-125.
- R. M. BRAY, and R. HOLT (1977) "The empirical study of decision processes in juries: a critical review," in J. L. Tapp and F. J. Levine (eds.) *Law, Justice and the Individual in Society: Psychological and Legal Issues*. New York: Holt, Rinehart & Winston.
- DAVIS, J. H., H. L. KERR, R. S. ATKIN, R. HOLT, and D. MEEK (1975) "The decision processes of 6- and 12-person mock juries assigned unanimous and 2/3 majority rules." *J. of Personality and Social Psychology*.
- DELSNER, L. (1975) "Smaller juries increase: divided verdicts allowed." *New York Times* (July 20).
- DIAMOND, S. S. (1974) "A jury experiment reanalyzed." *Michigan J. of Law Reform* (Spring): 520-531.
- FRIEDMAN, H. (1972) "Trial by jury: criteria for conviction, jury size, and Type I and Type II errors." *Amer. Statistician* 26 (April): 21-23.
- GELFAND, A. A. and H. SOLOMON (1977) "An argument in favor of twelve-member juries," pp. 205-220 in S. Nagel (ed.) *Modeling the Criminal Justice System*. Beverly Hills: Sage.

- (1974) "Modelling jury verdicts in the American legal system." *J. of Amer. Stat. Assn.* 69 (March): 32-37.
- (1973) "A study of Poisson's models for jury verdicts in criminal and civil trials." *J. of Amer. Stat. Assn.* 68 (June): 271-278.
- GROFMAN, B. (1980a) "Mathematical modeling of jury/juror decision making," in B. D. Sales (ed.) *Perspectives in Law and Psychology, Volume II: The Jury, Judicial and Trial Processes*. New York: Plenum.
- (1980b) "Jury decision-making models and the Supreme Court: the jury cases from *Williams v. Florida* to *Ballew v. Georgia*." *Policy Studies J.*
- (1980c) "Preliminary models of jury decision-making," in G. Tullock (ed.) *Frontiers of Economics* 3.
- (1979a) "The case for majority jury verdicts." *Trial Magazine* (December).
- (1979b) "Research note: a pilot study of individual behavior as mediated by the group context: three- and five-member mock juries." *Experimental Study of Politics*: 41-54.
- (1977) "Jury decision-making models," pp. 191-203, in S. Nagel (ed.) *Modeling the Criminal Justice System*. Beverly Hills: Sage.
- (1976) "Not necessarily twelve and not necessarily unanimous: evaluating the impact of *Williams v. Florida* and *Johnson v. Louisiana*," pp. 149-168 in G. Bermant et al. (eds.) *Psychology and the Law: Research Frontiers*. Lexington, MA: D. C. Heath.
- (1974) "Mathematics and politics: mathematical reasoning and optimal jury decision process" pp. 331-337 and 544-547 in M. Black (ed.) *Problems of Choice and Decision: Proceedings of a Colloquium Held in Aspen Colorado, June 24-July 6, 1974*. Cornell University Program on Science, Technology and Society and Aspen Institute for Humanistic Studies.
- Institute for Judicial Administration (1971a) "Juries of less than twelve and verdicts of less than unanimous." New York University, May. (mimeo)
- (1971b) "States in which criminal cases may be tried by fewer than twelve jurors." New York University, June 15. (mimeo)
- KALVEN, H. and H. ZEISEL (1966) *The American Jury*. Boston: Little, Brown.
- KERR, N., R. ATKIN, G. STRASSER, D. MEEK, R. HOLT, and J. DAVIS (1976) "Guilt beyond a reasonable doubt: effects of concept definition and assigned decision rule on the judgments of mock jurors." *J. of Personality and Social Behavior* 34: 282-294.
- KLEVORICK, A. K. and M. ROTHSCHILD (1979) "A model of the jury decision process." *J. of Legal Studies* (January): 141-164.
- LEMPERT, R. O. (1975) "Uncovering nondiscernible differences: empirical research and the jury-size cases." *Michigan Law Rev.* 73 (March): 644-707.
- NAGEL, S. S. and M. NEEF (1975) "Deductive modeling to determine an optimum jury size and fraction required to convict." *Washington University Law Rev.*: 933-978.
- NEMETH, C. (1977) "Interactions between jurors as a function of majority vs. unanimity decision rules." *J. of Applied Social Psychology* 7: 35-86.
- New York Times (1972) "Editorial: retreat on rights." (May 25): 44.
- PENROD, S. and R. HASTIE (1979) "Models of jury decision-making: a critical review." *Psych. Bull.* 86: 462-492.
- ROSENBLATT, A. M. and J. C. ROSENBLATT (1973) "Six member juries in criminal cases: legal and psychological considerations." *St. John's Law Rev.* 47 (May): 615-633.
- SAKS, M. J. (1977) *Jury Verdicts: The Role of Group Size and Social Decision Rule*. Lexington, MA: D. C. Heath.

- and T. M. OSTROM (1975) "Jury size and consensus requirements: the laws of probability v. the laws of the land." *J. of Contemporary Law* 1 (Spring): 163-173.
- SIMON, R. J. (1967) *The Jury and the Defense of Insanity*. Boston: Little, Brown.
- SPERLICH, P. (1979) "Trial by jury: it may have a future," pp. 191-224, in P. Kurland and G. Casper (eds.) *Supreme Court Review*. Chicago: Univ. of Chicago Press.
- WALBERT, T. (1971) "Note: effect of jury size on probability of conviction—an evaluation of *Williams v. Florida*." *Case Western Reserve Law Rev.* 22: 529-555.
- ZEISEL, H. (1978) Personal communication.
- (1971) "And then there were none: the diminution of the federal jury." *University of Chicago Law Rev.* 3: 710-724.
- and S. DIAMOND (1974) "Convincing empirical evidence on the six-member jury." *University of Chicago Law Rev.* 41.

BERNARD GROFMAN is Associate Professor of Social Psychology and Political Science in the School of Social Sciences, University of California, Irvine. He is the author of articles focusing on models of collective decision-making, which have appeared in scholarly journals such as *Behavioral Science*, the *American Political Science Review*, the *Journal of Mathematical Sociology*, and *Public Choice*.