Journal of Personality and Social Psychology
1985, Vol. 48, No. 2, 349-363


0022-3514/85/$00.75

# The Effects of Jury Size and Polling Method on the Process and Product of Jury Deliberation

Norbert L. Kerr and Robert J. MacCoun
Michigan State University

The U.S. Supreme Court has repeatedly assumed the functional equivalence of different sized juries (at least in the range of 6- to 12-person groups). Several formal models of jury decision making predict that larger juries should hang more often, particularly for very close cases. Failures to confirm this prediction in several previous studies were attributed to inadequate sample sizes or to insufficiently close cases. An experimental simulation study that minimized these problems was undertaken to test the models' prediction. Social decision scheme and social transition scheme analyses permitted comparisons of the decision-making processes of the different-sized mock juries. The effect of the method used to poll group members' verdict preferences was also examined. As group size increased, the observed probability of a hung jury increased significantly. No process differences between 6- and 12-person groups were detected, but 3-person groups did exhibit several process differences from the larger groups. When cases were very close, the likelihood of a hung jury for typically sized juries was found to be lower when the group was polled by secret ballot than when a show-of-hands polling method was used.

A number of Supreme Court rulings have focused the attention of social scientists and legal scholars on the issue of jury size. Although juries traditionally have 12 members, economic considerations have encouraged the use of smaller juries. In *Williams v. Florida* (1970), the Supreme Court held that there was no constitutional barrier to the use of juries with fewer than 12 members. The Court's explicit reliance on psychological reasoning has led to a number of studies to check the Court's assumptions of functional equivalence of different-sized juries (Beiser & Varrin, 1975; Bermant & Coppock, 1973; Buckhout, Weg, Reilly, & Frohboese, 1977; Davis, Kerr, Atkin, Holt, & Meek, 1975; Friedman & Shaver, 1975; Institute of Judicial Administration, 1972; Kessler, 1973; Mills, 1973; Padawer-Singer, Singer, & Singer, 1977; Roper, 1980; Saks, 1977; Valenti & Downing, 1975). Nearly all of these researchers have compared juries of 12 and 6 persons. The nearly univeral finding is one of no reliable jury size effects on jury verdicts (see Hastie, Penrod, & Pennington, 1984, for a recent comprehensive review). However, some of these studies have documented several other interesting effects that were due to jury size (e.g., in participation rates, length of deliberations, recall of evidence). In addition, several theoretical analyses based on sampling theory (e.g., Lempert, 1975) have demonstrated that a reduction of jury size has an adverse effect on jury representativeness. Such findings and the vigorous objections of legal scholars (e.g., Lempert, 1975) to jury size reduction were cited by the Court in a subsequent ruling that proscribed juries with fewer than 6 persons in criminal trials (*Ballew v. Georgia*, 1978; see Saks, 1982, for an analysis of the *Ballew* ruling).

There are, however, reasons to believe that variations in jury size should have small but reliable effects on juries' verdicts. A number of formal models of jury decision making that are intuitively plausible (e.g., Grofman,

Portions of this article were presented at the annual convention of the Law and Society Association, Toronto, Canada, June 5, 1982, and the Third International Conference on Small Group Processes, Nags Head, North Carolina, June 1983. Support for this research was provided by National Institute of Mental Health Grant MH29919-01 to the first author. We would like to thank Richard Lempert, Sam Gaertner, and two anonymous reviewers for their helpful comments on an earlier draft.

Requests for reprints should be sent to Norbert L. Kerr, Department of Psychology, Michigan State University, East Lansing, Michigan 48824-1117.

1976; Lempert, 1975; Tanford & Penrod, 1983) and have been empirically validated (e.g., Davis et al., 1975; Hastie et al., 1984) have all led to the prediction that larger juries should hang more often. With these models, one does not generally assume differences in the decision-making process for different group sizes; rather, one assumes with all models that unanimity is more likely when there is a preponderance of support in the group for one of the verdict alternatives at the beginning of deliberation, and sampling theory indicates that this is relatively more likely in smaller groups. There is some weak empirical support for the models' prediction. Zeisel (1971) reported a hung jury rate of 2.4% in a sample of 290 six-person criminal juries, slightly lower than the 5% rate observed in his national sample of 12-person juries. However, it is difficult to determine how comparable the cases and groups in these two samples really are. Padawer-Singer et al. (1977) reported that 8.7% of their 6-person mock juries hung, whereas 21.7% of their 12-person groups hung. However, these results were due entirely to the absence of any hung juries in their 6-person groups operating under a nonunanimous rule; there were no differences between the 6- and 12-person groups that were required to agree unanimously. Roper (1980) also reported a strong trend for more hung juries in 12- than 6-person mock juries, but again the effect was not significant. None of the researchers in several other studies who examined jury size reported a reliable effect of jury size on hung jury rate (see Hastie et al., 1984, for a review). In summary, there is currently no firm empirical support for the models' prediction.

Fortunately, these models also suggest reasons why the existing research may have failed to obtain the effect. First, the effect is only predicted when cases are "close" ones—that is, when the individual juror conviction rate is near 50% (see Davis et al., 1975, Figure 1, or Tanford & Penrod, 1983, Table 1). A number of the researchers testing for jury size effects have used rather lopsided cases (e.g., Davis et al., 1975; Kessler, 1973). Second, even when the cases are very close ones, the predicted differences in hung juries tend to be fairly small. Although in a few studies reasonably large samples have been used (e.g., Davis et al., 1975), the enormous cost of using 6- and 12-person groups as replicates has generally resulted in sample sizes that are too small to test the models' prediction adequately.

A second prediction made with some of these models (e.g., Davis et al., 1975; Grofman, 1976; Tanford & Penrod, 1983) is that when cases are not close, the larger the group is, the more group deliberation will tend to polarize opinion (cf. Myers & Lamm, 1976). In general, these predicted effects are even smaller than those involving hung juries, and can be further obscured by ceiling and floor effects. Thus it is not surprising that, like the prediction on hung juries, this prediction has not been empirically confirmed.

The first objective of the present study was to provide a clearer test of these formal model predictions involving jury size. The verdicts (and especially the rate of hung juries) of 12-, 6-, and 3-person mock juries were compared.[1] The stimulus cases used in this study were carefully developed to be very close ones. The overall sample size was also quite large. We further increased the power of our tests by having each group consider several different cases, thereby increasing the reliability of the estimate of each group's tendency to hang.[2] An experimental simulation methodology was used to help avoid the confounds that have plagued field and archival studies of jury size (e.g., Institute of Judicial Administration, 1972). However, it should

[1] As noted, the *Ballew v. Georgia* (1978) decision has barred reduction of jury size below 6 persons for criminal trials. However, states may still use juries with fewer than 6 persons for civil trials. Thus the study of 3-person groups is not only of theoretical interest, but also of some interest for jury applications.

[2] At each group size our design called for over 250 verdicts to be reached. With an alpha of .05, Cohen (1977, Ch. 6) suggests that a sample of this size would detect a "small" difference in proportions (viz., an effect size = $h$ = .20) with a probability of .83. For purposes of comparison, Davis, Stasser, Spitzer, & Holt's (1976) model predicts that for perfectly balanced cases, the effect size for comparing 6- and 12-person juries would be approximately .75. (A comparison of 12- and 3-person groups would produce an even larger effect). A sample of 250 observations would have power in excess of .99 to detect an effect of this magnitude. Thus the proposed design should have sufficient power to detect the predicted effect in hung jury rates, assuming the cases are very close ones.

also be noted at the outset that statistical power and experimental control were purchased at a cost of realism. We attempted to confirm a theoretical prediction under conditions that were experimentally ideal, but highly artificial.

The use of a highly controlled method was also crucial to the study's second and chief objective: to examine certain features of the group deliberation process. We were particularly interested in examining a process issue that has figured heavily in the Court's and in modelers' reasoning about the jury size issue: How does the absolute and relative size of a faction affect its likelihood of maintaining or increasing that size? In *Williams v. Florida* (1970), the Court explicitly assumed a proportionality model, as have certain modelers. For example, according to Klevorick and Rothschild's (1979) model, juries with identical ratios of opposing factions are functionally equivalent. For example, a minority of one in a 6-person group is assumed to have the same chance of yielding to the majority as a member of a 2-person minority in a 12-person group. However, conformity research (e.g., Asch, 1956) tends to contradict such a proportionality model. Asch has shown that having some social support strongly attenuates the power of a majority. Latané's social impact theory (Latané, 1981; Latané & Wolf, 1981) makes a similar prediction. Asch's findings suggest, for example, that a minority of one in a 3-person group should be more easily influenced than a member of a 2-person minority faction in a 6-person group. Although there is much social influence research bearing indirectly on such issues (cf. Allen, 1965; Tanford & Penrod, 1984), there is very little in which these questions have been examined in the context of interacting, decision-making groups of different sizes. A study by Godwin and Restle (1974) is a noteworthy exception. They found that a model that took account of both the absolute and relative size of a faction better accounted for observed shifts in position than a model in which only the faction's relative size (i.e., relative to the opposing faction) was considered. They also found that minorities of one were particularly vulnerable to the influence of a unanimous majority. However, Godwin and Restle's task (estimating which of several stimuli another group had judged as "most outstanding") is a rather poor facsimile of the jury's task or any other common group decision task (e.g., it is unclear how persuasion might occur on Godwin and Restle's task).

Our second objective, therefore, was to examine the effect of factions' relative and absolute size on social influence in different-sized mock juries. The unit of analysis in our process analyses was the distribution or "split" of verdict preferences in the group. We examined both the likelihood of movements from initial split to final verdict (i.e., we estimated the D matrices of the Social Decision Scheme model; Davis, 1973, 1980) and the probabilities of movement between possible splits during deliberation (i.e., we estimated the T matrix transition probabilities of the Social Transition Scheme model; Kerr, 1981, 1982). Besides the validity and generality of the proportionality model, these analyses also allowed us to examine in the context of an interacting group a social influence question of longstanding interest (e.g., Asch, 1951; Gerard, Wilhelmy, & Conolley, 1968; Latané & Wolf, 1981; Tanford & Penrod, 1984): the effect of the size of a unanimous majority on yielding by a minority of one. In summary, our process analyses sought to determine whether and how a group's size affects its decision-making process. Although this is an issue of particular relevance to the debate on jury size, it clearly has broader theoretical significance.

A third and final objective of this research was to examine the effect of the method a group used to poll its members on the group's final decision. Hawkins (1960) reported an association between the polling method freely chosen by his mock juries and the likelihood that the jury would hang; juries that had used a secret ballot technique were found to be more likely to hang than groups that had used public forms of polling (e.g., a show of hands). One explanation for this finding is that secret balloting made it more difficult for the majority to identify and hence pressure or persuade minority faction members. There are indications, though, that jurors are expected to state and defend their position (e.g., Lempert, 1975), particularly when the jury is operating under a unanimity requirement. Another explanation is that the relationship

is a spurious one; juries composed of or led by jurors who prefer to use a secret ballot may also act in other ways that make agreement less likely (e.g., avoiding direct and open conflict that may be necessary to achieve unanimity). This interpretation would suggest that secret polling is merely a symptom, not a cause of ineffective jury deliberation. One could even argue that open polling tends to publicly identify a juror with a particular position and to foster early and strong commitment to what may have initially been only a weak preference. This argument suggests that secret polling should lead to fewer, not more, hung juries. In any case, we would expect the effect of the method of polling to be most pronounced for larger juries deliberating close cases. In small juries considering lopsided cases, minority factions will necessarily be very small and should be easily identified, regardless of the method of polling. Because a number of close cases were considered by different-sized mock juries in this study, it was possible to test the latter prediction.

## Method

### *Subjects*

The subjects were 612 undergraduate students at Michigan State University who participated to earn extra credit in an introductory psychology course. Subjects were scheduled in same-sex groups approximately 20% larger than the scheduled jury size to allow for no-shows. If too many subjects came, the surplus subjects were excused; if too few subjects came, the group was run in the next smaller group-size condition. In addition, there were several regularly scheduled 6- and 3-person sessions.

### *Design*

The basic experimental design was a 3 × 2 (3- vs. 6- vs. 12-Person Mock Juries × Open vs. Secret Polling Method) factorial. There were approximately 15 same-sex groups in each of the experimental conditions.

### *Materials and Equipment*

Each group considered a set of nine one-page summaries of armed robbery cases, purportedly tried in San Diego, California within the last 5 years. In fact, the cases were fictional ones that had been developed to produce nearly equal rates of conviction and acquittal by individual mock jurors from this population. Prior research with these cases (e.g., Kerr, 1981, 1982) has established that the cases are indeed close ones. Each summary consisted of two paragraphs, one summarizing the evidence for the prosecution and the other summarizing the evidence for the defense. Immediately before the start of deliberations, each group member received a folder containing all nine cases. On a cover sheet there was a summary of instructions typically provided by a judge to juries considering armed-robbery cases (e.g., a definition of armed robbery, a reminder of the presumption of innocence, a definition of the reasonable-doubt concept). The order of cases in the folder was randomized separately for each group.

Subjects recorded their predeliberation verdict preferences and any change in verdict preference occuring during deliberation by pressing one of two buttons (for guilty and not guilty) on a response panel. Each subject had one panel, and the foreman had a second panel to indicate the group's verdict. The buttons on the response panels were interlocking: When a button was pressed it lit up as the light for the last response was extinguished; hence group members could always check their own last response with a glance at their panels. When a subject made a response, the response and the time it was made were automatically encoded and stored in the memory of a microcomputer. Because of some chronic hardware and software problems, these deliberation process data were not encoded for several groups (viz., for 7 of the 28 twelve-person groups, for 12 of the 31 six-person groups, and for 8 of the 30 three-person groups).

Subjects were seated around a large rectangular table. One seat was arbitrarily designated as the foreperson's, and whichever subject took this seat served as foreperson. Low partitions separated the seating positions. They permitted group members to see and to talk to one another without obstruction but prevented observation of one another's response panels.

### *Procedure*

Each group of 3, 6, or 12 subjects was asked to take on the role of a jury and deliberate each of the nine armed robbery cases. The use of the response panels and the procedure to be followed was then explained. After reading a case summary, subjects were to individually and privately decide how they would vote if they were on the actual jury that tried the case, and to indicate this preference on their response panels. When everyone had recorded their predeliberation verdicts for the case, the group was to begin deliberating. To enable each group to deliberate all nine cases, the groups had a maximum of 10 min to deliberate each case. The foreperson was to note the time at which deliberations began and ended and was to record these times on a supplied form. The group was to keep track of the deliberation time; a digital clock was visible to all members of the group. In addition, the experimenter would signal when 10 min had elapsed. If the group was unable to agree unanimously on a verdict within the time limit, the foreperson was to enter a verdict of hung and the group was to proceed to the next case.

The foreperson was advised to take a poll of the jury whenever necessary to check the group's progress toward unanimity. The manipulation of polling method was introduced at this point. In the open-polling groups, the foreperson was instructed to use only a show-of-hands method to poll the group. In the secret-polling condition, the foreperson was instructed to use secret written ballots

for every poll of the group; ballots were supplied for this purpose.

If the group reached a unanimous verdict, the foreperson was to record it on the jury's panel, and the group was to proceed to the next case. After giving the foreperson a sheet that summarized the procedure and after answering all questions, the experimenter gave subjects their case folders and told them to read the cover sheet carefully and then to begin reading the first case. The experimenter then left and remained in a nearby room until the end of the session (except when it was necessary to tell a group that its time had elapsed on a case). When either a group had completed all nine cases or the time for the 2-hour experimental session had elapsed, groups were debriefed, thanked, and excused.

## Results

### *Deliberation Products: Verdicts*

We hypothesized that the variables of jury size and polling method would exert their strongest effects when the cases were very close ones. Although the cases used here have been developed to produce individual conviction rates near 50%, in any particular experimental situation some of the cases come closer than others in achieving this (see Davis & Nagao, 1980, for an illustration of temporal drift in such parameters). In the present instance, when subjects' predeliberation verdict preferences were tallied for each case, four cases were found to have conviction rates between 35% and 65%. This set of cases were designated as the *close set.* The remaining five cases were designated as the *clear set* (although it may be noted that none of them resulted in highly lopsided conviction rates; the range in the clear set was between 32%–35% and 65%–74% guilty). In all the following product analyses, the group was the unit of analysis.

Because we were most interested in the likelihood of a hung jury, we first examined the rate at which the mock juries hung. The proportion of cases for which the group hung was computed for each mock jury for the close and clear case sets. These proportions were entered into a 3 × 2 × 2 (Group Size × Polling Condition × Case Set) least-squares repeated measures analysis of variance (ANOVA; see Woodward & Overall, 1976).[3] Two effects were significant. First, and most important, there was a strong main effect for size, $F(2, 76) = 28.94$, $p < .001$. As the groups became smaller, the proportion of hung juries declined; the means were 12-person = 45% hung, 6-person = 28% hung, and 3-person = 9% hung. Newman-Keuls post hoc comparisons indicated that all pairwise comparisons of size conditions were significant ($p < .01$). There was also a tendency for the size effect to be stronger for the close cases than for the clear cases (e.g., the difference between 3- and 12-person groups was 44% for the close cases, but only 28% for the clear cases), but the Case Set × Size interaction was not statistically significant, $F(2, 76) = 2.72$, $p < .08$. The other significant effect was the Polling × Case Set × Group Size interaction effect, $F(2, 76) = 4.36$, $p < .025$. The means for this effect are presented in Table 1. Tests of the polling simple main effects indicated that when the cases were clear, there were no significant polling simple effects at any group size. However, when the cases were close, the polling simple effect was significant ($p < .05$) at every group size. For these cases, the 3-person groups revealed an effect similar to that obtained by Hawkins (1960): Hanging was less likely when polling was public. However, when the groups were larger the direction of the polling effect was opposite to that obtained by Hawkins: Public

Table 1
*Effect of Polling × Size × Case Set on Hung Juries*

| Case set/Jury size | Polling method: Open | Polling method: Secret |
|---|---|---|
| Close | | |
| 12 | .56 | .45 |
| 6 | .42 | .17 |
| 3 | .02 | .13 |
| Clear | | |
| 12 | .40 | .42 |
| 6 | .27 | .21 |
| 3 | .13 | .08 |

*Note.* Values represent the mean proportion of deliberations resulting in a hung jury, averaged within groups.

[3] In a preliminary analysis of rates of hanging and convicting, juror sex was also included as a factor. Sex entered into no significant interactions with the primary independent variables, jury size and polling method. In light of these results and Kerr's (1981) finding of no deliberation process differences between male and female mock juries, juror sex was not examined further.

polling tended to increase rather than decrease the likelihood of a hung jury.

To examine the effect of jury size and polling method on the verdict reached by those juries that did not hang, we considered only the cases in which a group reached a verdict. For each group the conviction rates among close and among clear cases were computed. The few groups that failed to reach a verdict on any cases within a case set were dropped from this analysis. These data were analyzed in a 3 × 2 × 2 (Group Size × Polling Method × Case Set) least-squares repeated measures ANOVA. The only significant effect was the main effect for case set, $F(1, 72) = 6.7$, $p < .05$; convictions were more likely in the clear (58%) than the close set (47%). Because four of the five clear cases tended to favor conviction (i.e., individual conviction rates > 65%), this effect is neither surprising nor especially interesting. It is interesting to note, however, that mock jury size did not affect the relative likelihood of conviction versus acquittal, a pattern obtained in nearly all previous work.[4]

We also checked to see whether jury size or polling method had any effect on the degree of polarization produced by group deliberation (cf. Myers & Lamm, 1976). Again, certain formal models (e.g., Davis et al., 1975; Grofman, 1976; Tanford & Penrod, 1983) predict that the larger the group is, the more the initially preferred alternative will gain in popularity as a result of group deliberation. This means, for example, that for a case on which individual jurors tended to favor conviction, the larger the jury is, the stronger this tendency to convict should be among juries. In order to test this conjecture, group verdicts were reclassified as either the same or opposite from the modal individual verdict preference; a group verdict that agreed with the individual preference was classified as a "polarized" verdict. Within each case set and group we calculated the proportion of all reached verdicts that were polarized verdicts. This variable was entered into a 3 × 2 × 2 (Group Size × Polling Method × Case Set) least-squares repeated measures ANOVA. Unsurprisingly, the clear cases were less likely to result in a reversal of the normative preference, $F(1, 72) = 46.6$, $p < .001$. There were no other significant effects. In particular, jury size was not significantly associated with this index of group polarization: Size, $F(2, 72) = 1.2$, *ns*; Size × Case Set, $F(2, 72) = 1.9$, $p > .10$.

In summary, the results were consistent with most previous research: When the mock jury was able to reach a verdict, the group's size had no significant effect on what that verdict was, nor did the method of polling. The analyses of the verdict product data clearly confirm the prediction that jury size and polling method do not materially affect what verdict is reached but can affect the likelihood that any unanimous decision will be reached. The latter finding represents the strongest empirical evidence to date that reductions in group size will increase the likelihood that the group will reach consensus.

### *Deliberation Products: Deliberation Time*

We first computed the mean time each group spent per close case and per clear case, regardless of the initial split or jury verdict. A 3 × 2 × 2 (Group Size × Polling Method × Case Set) least-squares repeated measures ANOVA yielded three significant effects. Unsurprisingly, it took significantly more time for groups to agree on the close cases than on the clear cases, $F(1, 71) = 6.5$, $p < .001$. There was also a significant three-way interaction effect, $F(2, 71) = 4.0$, $p < .05$. The polling effect was negligible in every condition but the 6-person clear-case condition, in which secretly polled groups took considerably longer to reach their verdicts. The result of primary interest was the group size main effect, $F(2, 71) = 22.45$, $p < .001$. As the hung jury data suggests, the larger the group was, the longer the groups tended to take. Means were 3-person = 167.2 s, 6-person = 321.8 s, and 12-person = 350.3 s. When interpreting these means, one should keep in mind that they reflect all deliberations, including initially unanimous groups that should have taken

[4] In one study, Valenti and Downing (1975) reported that their 6-person groups were significantly more likely to convict than their 12-person groups when the case against the defendant was strong. However, it was later discovered (e.g., Saks, 1977) that random assignment had failed to equate conditions before deliberation, and that when allowance was made for this, the reported size effect was eliminated.

Table 2
*Predeliberation Distribution to Final Outcome Frequencies*

| % for conviction | 12-person jury | | | | 6-person jury | | | | 3-person jury | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Initial split | Outcome | | | Initial split | Outcome | | | Initial split | Outcome | | |
| | | C | A | H | | C | A | H | | C | A | H |
| 100 | 12–0 | 14 | 0 | 0 | 6–0 | 16 | 0 | 0 | 3–0 | 49 | 0 | 0 |
| 92 | 11–1 | 16 | 0 | 0 | | | | | | | | |
| 83 | 10–2 | 6 | 0 | 5 | 5–1 | 26 | 0 | 6 | | | | |
| 75 | 9–3 | 10 | 0 | 15 | | | | | | | | |
| 67 | 8–4 | 5 | 3 | 9 | 4–2 | 12 | 10 | 16 | 2–1 | 38 | 11 | 11 |
| 58 | 7–5 | 2 | 4 | 16 | | | | | | | | |
| 50 | 6–6 | 1 | 9 | 13 | 3–3 | 3 | 13 | 12 | | | | |
| 42 | 5–7 | 1 | 9 | 5 | | | | | | | | |
| 33 | 4–8 | 0 | 7 | 5 | 2–4 | 0 | 21 | 4 | 1–2 | 5 | 38 | 6 |
| 25 | 3–9 | 0 | 3 | 0 | | | | | | | | |
| 17 | 2–10 | 0 | 5 | 0 | 1–5 | 0 | 17 | 2 | | | | |
| 8 | 1–11 | 0 | 4 | 0 | | | | | | | | |
| 0 | 0–12 | — | — | — | 0–6 | 0 | 3 | 0 | 0–3 | 0 | 24 | 0 |

*Note.* C = convict, A = acquit, H = hung jury. In the "initial split" columns the first number in each pair = number of votes for conviction; second number = number of votes for acquittal.

only a few seconds to reach a verdict. If the groups that were initially unanimous are dropped from the preceding analysis, the mean deliberation times naturally increase (3-person = 238.8 s, 6-person = 357.0 s, 12-person = 366.3 s). In this subsample, only the main effects were significant. Larger (i.e., 6- and 12-person) groups deliberated longer ($p < .001$); groups took longer to deliberate the close cases ($p < .05$); and groups deliberated longer under secret polling ($p < .01$). The latter effect may just reflect the extra time required to collect secret written ballots.

### *Deliberation Process: Initial Split to Final Outcome Frequencies*

One way of concisely summarizing the deliberation process is to determine the likelihood of groups moving from each possible beginning distribution of opinion to each possible outcome (viz., convict, acquit, or hung jury). In effect, this estimates the applicable social decision scheme matrix, D; see Davis (1973, 1980) and Kerr, Stasser, and Davis (1979). These data are presented in Table 2. The one pattern that is immediately evident is that, as sampling theory requires, the smaller the group is, the greater is the proportion of groups beginning deliberation at or near unanimity. For example, there were no 12-person groups beginning deliberation with a 0-guilty–12-not-guilty split, whereas there were 4 six-person and 24 three-person groups with initial unanimity for acquittal.

Several log-linear analyses were performed on these data.[5] First, the comparable rows

[5] In this analysis the unit of analysis was the group verdict. In the following process analyses of shifts between states, the unit of analysis is a shift by a group member. Statistically, this means that we have assumed that each verdict (or shift) is an independent event. Of course, because each group considered several cases, the same group produced several verdicts. Likewise, the same kind of shift was sometimes made more than once by a single group (either on different cases or, infrequently, on the same case). The reason for making these assumptions was that parameter estimation and statistical inference would have effectively been impossible for several of the questions of interest had we followed the standard practice of using the group as the unit of analysis. If only one observation per group were permitted, it is clear that the reliable estimates of the applicable social decision schemes (Table 2) or social transition schemes (see Table 3) could not have been obtained. The only other alternative—obtaining sufficiently many observations from each group to estimate the complete matrix for each group—was equally impractical. Because of such problems, the use of the shift as the unit of analysis has been standard practice in research of this type. Statistical analyses of all previous applications of similar stochastic process models (e.g., Davis et al., 1976; Godwin & Restle, 1974) have made such assumptions, either explicitly or implicitly.

(viz. those with initial ratios of convictors to acquitters of 2:1 and 1:2) were contrasted across the three group size conditions; that is, a 2 × 3 × 3 (Ratio × Outcome × Group Size) contingency table was analyzed. There was a strong association of size with verdict, $\chi^2(4, N = 201) = 21.9$, $p < .01$, $\tau = .07$.[6] This effect was due to differences between the 3-person groups and the larger groups; analyses in which we compared the 6- and 12-person groups only produced no such effect, $\chi^2(2, N = 92) = 3.1$, *ns*. Inspection of Table 2 suggests that the 3-person groups were both less likely to hang and more likely to convict than the larger groups. Collapsing the convict and acquit categories resulted in a 2 × 2 × 3 (Ratio × Outcome × Group Size) contingency table. Log-linear analysis resulted in a significant Size × Outcome association, $\chi^2(2, N = 201) = 14.1$, $p < .01$, $\tau = .08$, confirming that the overall effect was due, at least in part, to the lower probability of a hung jury in the 3-person groups. Dropping the hung-jury outcomes from the original contingency table resulted in a 2 × 2 × 3 (Ratio × Outcome × Size) table. Analysis of this table also produced a significant Size × Outcome association, $\chi^2(2, N = 150) = 7.0$, $p < .05$, $\tau = .05$. Thus part of the overall effect was also attributable to the greater likelihood of conviction in the 3-person groups.

The near equivalence of 6- and 12-person groups' initial split to final outcome probabilities was corroborated by an analysis of the five rows for which the (nonunanimous) initial splits are in the same ratios (viz., 5:1, 4:2, 3:3, 2:4, 1:5). Although there was a tendency for the 12-person groups to be more likely to hang, the Size × Outcome association was not significant, $\chi^2(2, N = 210) = 4.99$, $p < .10$. None of the separate comparisons within each row yielded a significant size effect.

The earlier analyses of the verdict data showed a steady increase in the likelihood of a hung jury as group size increased. The present analysis suggests that this effect does not reflect process differences between the 6- and 12-person groups. Twelve- and 6-person groups beginning with comparable distributions of verdict preferences (i.e., in the same ratio) had statistically indistinguishable outcome probabilities. Thus the substantial difference in hung-jury rates for 6- and 12-person groups was due to differences in the distribution of starting splits, not to differences in their decision-making processes.

This equivalence breaks down, though, when we include 3-person groups in the analysis. Three-person groups beginning with verdict splits comparable (i.e., proportional) to the larger groups' had markedly different final verdict distributions; they were generally less likely to hang, were more likely to convict when the majority favored conviction, and were less likely to acquit when the majority favored acquittal. Altogether, these findings suggest genuine process differences between the 3-person groups and the larger groups (cf. Penrod & Hastie's 1980 analysis for nonunanimous juries). The following analyses also support this conclusion.

### *Deliberation Process: Transition Frequencies*

The frequencies with which groups moved between adjacent "states" (i.e., distributions of verdict preferences) were tallied. In effect, this is an estimate of the applicable social transition scheme matrix, T; see Kerr (1981, 1982). These data appear in Table 3. The rows of this table represent a group's verdict split before a member shifted position, and the columns represent the two possible directions of movement. So, for example, there were 34 instances in which a member of a group with a 10-guilty–2-not-guilty (10G–2NG) split of opinion changed position. In 4 of those 34 instances, the shift served to reduce the number of jurors supporting conviction (G↓); that is, one of the members of the majority defected and the group went from 10G–2NG to 9G–3NG. In the other 30 instances, the shift served to increase the number supporting conviction (G↑); that is, one of the two minority-faction members joined the majority, and the group went from 10G–2NG to 11G–1NG.

One reasonable basis for comparison of the decision-making process in the different size groups is to examine the relative "attractiveness" (i.e., power to attract converts) of

[6] See Bishop, Fienberg, & Holland (1975) for a description of the $\tau$ statistic, which indexes strength of association.

Table 3
*Observed Transition Frequencies*

| % for conviction | 12-person jury | | | 6-person jury | | | 3-person jury | | |
|---|---|---|---|---|---|---|---|---|---|
| | Pre-shift distribution | Shift | | Pre-shift distribution | Shift | | Pre-shift distribution | Shift | |
| | | G↓ | G↑ | | G↓ | G↑ | | G↓ | G↑ |
| 100 | 12–0 | 2 | — | 6–0 | 0 | — | 3–0 | 1 | — |
| 92 | 11–1 | 3 | 41 | | | | | | |
| 83 | 10–2 | 4 | 30 | 5–1 | 4 | 41 | | | |
| 75 | 9–3 | 9 | 25 | | | | | | |
| 67 | 8–4 | 15 | 16 | 4–2 | 22 | 16 | 2–1 | 11 | 43 |
| 58 | 7–5 | 21 | 15 | | | | | | |
| 50 | 6–6 | 29 | 11 | 3–3 | 37 | 8 | | | |
| 42 | 5–7 | 40 | 8 | | | | | | |
| 33 | 4–8 | 46 | 7 | 2–4 | 48 | 7 | 1–2 | 49 | 4 |
| 25 | 3–9 | 43 | 2 | | | | | | |
| 17 | 2–10 | 41 | 1 | 1–5 | 63 | 0 | | | |
| 8 | 1–11 | 43 | 0 | | | | | | |
| 0 | 0–12 | — | 0 | 0–6 | — | 1 | 0–3 | — | 0 |

*Note.* G↓ = shift of position reducing the number of guilty votes; G↑ = shift of position increasing the number of guilty votes. In the "Pre-shift distribution" columns, the first number in each pair = number of votes for conviction; the second number = number of votes for acquittal.

factions of the same relative size (i.e., relative to the opposing faction). In the present instance, there were two proportional splits that could be compared across all three group sizes, namely, the 2:1 and 1:2 ratios of guilty to not-guilty voters. The corresponding entries in Table 3 can be viewed as a Ratio (2:1 vs. 1:2) × Direction of Movement (G↑ vs. G↓) × Size (3-, 6-, and 12-person group) contingency table. Applying a log-linear analysis to this table resulted in a significant Size × Direction effect, $\chi^2(2, N = 284) = 6.7$, $p < .05$, $\tau = .08$, qualified by a significant three-way association effect, $\chi^2(2, N = 284) = 9.7$, $p < .01$, indicating that the likelihood of transitions to and from factions depends not only on their relative sizes but also on the overall group size. Inspection of Table 3 suggests that this effect was primarily attributable to the 3-person group with the 2:1 ratio. Additional log-linear analyses confirmed this. An analysis based on only the 2:1 row produced a significant Size × Direction effect, $\chi^2(2, N = 123) = 15.27$, $p < .01$, $= .12$; this effect vanished, $\chi^2(1, N = 69) = .6$, when one dropped the 3-person groups from the analysis. A similar analysis based on only the 1:2 row produced no Size × Direction effect, $\chi^2(2, N = 161) = 1.1$, *ns*.

In the 2:1 ratio row, the majority for conviction was much more successful in attracting a minority member in the 3-person groups than in either the 6- or 12-person groups. In fact, in these larger groups, the minority position was as or more likely to gain a convert as the majority position was. This pattern clearly did not hold in the 1:2 ratio row, in which regardless of group size, the majority of acquitters exhibited much greater drawing power than the minority of convictors. (This asymmetry in the drawing power of proconviction vs. proacquittal factions has been repeatedly observed in jury research. It is most clearly evident here in the 50% row of Table 3. See Stasser, Kerr, & Bray, 1982, for a review and discussion of this leniency bias.)

As the preceding analyses show, the 6- and 12-person groups appear to satisfy the proportionality rule fairly well. We further substantiated this conclusion by comparing transition frequencies for all the comparable splits in 6- and 12-person groups. There are five such ratios: 5:1, 2:1, 1:1, 1:2, and 1:5. Analysis of the 5 × 2 × 2 (Ratio × Direction of Movement × Size) contingency table resulted in no size effects: Size × Direction, $\chi^2(1, N = 446) = .83$, *ns*, and Size × Direction × Ratio

$\chi^2(4, N = 446) = 1.12$, *ns*. Altogether, these analyses indicate that a proportionality model that assumes equal attractiveness of equal ratios fits the 6- and 12-person groups, but breaks down when we include the smaller, 3-person groups, for which minorities are necessarily minorities of one.

A second reasonable basis for comparison—suggested by intuition, by Asch's conformity research, and by Godwin and Restle's (1974) study of decision-making groups—is between splits in which a minority of one is opposed by an unanimous majority. There has been some controversy about the nature of the functional relation between the size of a unanimous majority and yielding to that majority (Asch, 1951; Gerard et al., 1968; Latané & Wolf, 1981). In previous studies conformity paradigms have been used; this study permits examination of the same question in the context of an interacting, decision-making group. Therefore, the transition frequencies for the $(n - 1)$G–1NG (where $n$ = group size), and the 1G–$(n - 1)$NG splits were compared across groups sizes. Analyses of the 2 × 2 × 3 (Unanimous Majority for Conviction vs. Acquittal × Direction of Movement × 3 Group Size) contingency table resulted in a significant three-way association, $\chi^2(2, N = 302) = 13.1$, $p < .01$. As Table 3 indicates, the minority of one was more successful at attracting a convert from the unanimous majority in the 3-person group than in either of the larger groups. Again, comparisons of 6- with 12-person groups resulted in no such effect, $\chi^2(1, N = 195) = .004$, *ns*.

In summary, by comparison with similar factions in 6- and 12-person groups, a minority (necessarily of one) in a 3-person group is somewhat less able to win converts than minority factions of the same proportional size in the larger groups, but somewhat more able to win converts than minority factions of the same absolute size in the larger groups.

### *Deliberation Process: Time to Shift*

A previous analysis suggested that larger groups took more time to deliberate. The question we address here is whether this was due to process differences in the rate of movement toward consensus. But first, it is of interest to reexamine total deliberation time. We have seen that small groups are more likely to be near or at unanimity at the outset of group deliberation. This could account for the shorter deliberation times of the smaller groups. To explore this possibility, we performed a regression analysis on the juries' deliberation times. The study's two main independent variables (group size and polling method) were dummy coded as predictor variables. In addition, two other predictors variables were defined. The first indexed how sharply divided the group was at the beginning of deliberation. It was defined as EXTDEV = |(#G)/(#NG) − .5|, where #G = the number of votes for a guilty verdict before deliberations, and #NG = the number of votes for a not-guilty verdict. The second predictor was closely related but conceptually distinct. It was the minimum number of verdict changes required for the group to achieve unanimity (termed NTU, or number to unanimity). These predictor variables and all possible interaction terms were entered into a standard regression analysis in order to predict deliberation times for each jury on each case. This analysis produced only one significant predictor, EXTDEV, $F(1, 703) = 13.6$, $p < .001$. The more even the initial split in the group, the longer the group deliberated. But when one controlled for group size differences in this variable, and in the related NTU variable, group size per se did not affect deliberation time.

More direct analyses of group process were then performed. As with the previous process analyses, we first compared groups with identical ratios of convictors to acquitters. In the earlier analyses we asked whether the *direction* of vote shifts varied with group size. Here we asked whether the *speed* of shifts varied with group size. The time required for the next shift was analyzed in a 2 × 2 × 2 (Ratio, 2:1 vs. 1:2 × Size × Polling Method) least-squares ANOVA. Two effects were significant: size, $F(2, 266) = 6.5$, $p < .01$; and ratio, $F(1, 266) = 7.1$, $p < .01$. Shifts were more rapid when the groups were larger (3-person = 162.7 s, 6-person = 154.1 s, and 12-person = 89.8 s) and when the majority favored acquittal (2:1 ratio = 167.7 s, 1:2 ratio = 114.7 s). These effects were spurious, however. Kerr (1981)

showed that a group's first shift in opinion was slower than later shifts, all else being equal. In the present study, smaller groups tended to have had fewer shifts preceding the shift out of the 2:1 or 1:2 state (largely because, as we have already shown, smaller groups are more likely to *begin* deliberation with an extreme split). Likewise, because of the leniency bias in jury deliberation, groups with nearly even initial splits were more likely to move toward acquittal than conviction and, therefore, groups in a 1:2 ratio were more likely to have had a previous shift than groups in a 2:1 ratio. A simple way to remove these confounds was to examine groups' first shift only. An ANOVA like the previous one was performed on this subsample. It produced no significant effects. Thus, in terms of the time taken to produce a shift, a proportionality rule seemed to apply.

We performed a final analysis to see whether the time to shift out of the minority of one split depended on the size of the unanimous majority. A 2 × 2 × 3 (Unanimous Majority for Conviction vs. Acquittal × Polling Method × Size) ANOVA resulted in an affirmative answer, even when only first shifts were examined ($p < .01$; 11-person majority = 97.6 s, 5-person majority = 158.4 s, and 2-person majority = 175.2 s).

## Discussion

### *Products of Deliberation*

The prediction that the rate of hung juries would increase with mock jury size was confirmed. This effect was not entirely attributable to a low rate of hung juries in the 3-person groups; the 6- and 12-person groups also differed strongly and significantly. This is, to our knowledge, the first of many field and experimental simulation studies of jury size that has demonstrated a significant positive relation between jury size and the rate of hung juries.

Of course, because this was only a simulation of the jury's task, considerable caution must be exercised in generalizing these findings to actual jury behavior. Like the majority of juror/jury studies, the present simulation was artificial in many ways (see Bray & Kerr, 1982). In order to have enough data to fairly test the models' prediction and to undertake the process analyses of interest, it was essential that a very large number of deliberations be observed at each group size. This made it infeasible to present each group with a single case of realistic length and detail and to impose no time limit on group deliberation. An efficient alternative was to have each group deliberate several different cases. This required that the case materials be brief and that deliberation time be limited.

Of the several differences between actual juries and our laboratory groups, the imposition of the 10-min time limit may raise the most important ecological validity questions. Although the deliberation time of actual juries is not limitless, it is rarely as short as (and is never restricted to) the interval imposed in the present experiment. Both intuition and research (Kerr, 1981) suggest that giving juries less time to deliberate makes it more difficult for them to reach consensus. So it seems probable that the overall rate of hung juries would have been lower had there been no time limits. Time limits also seem likely to create a greater impediment to unanimous agreement in larger groups. Any time limit leaves less time per member to speak in a larger group. With less speaking time per member, larger groups may have more trouble accumulating the preponderance of verbal support that has been linked to group consensus (Hoffman, 1979). Furthermore larger groups are likely to face greater problems of coordination (Steiner, 1972), such as managing speaking orders, avoiding interruptions, and so on, especially when time is limited. Such coordination problems would tend to make it harder for larger groups to complete the task of reaching agreement. All these considerations suggest that short time limits tend to make it particularly hard for larger groups to reach agreement.

We now know that larger groups are less likely to begin deliberation at or near unanimity. Or, as Roper (1980) puts it, larger groups are relatively more likely to have "viable minorities." We also know from a great deal of research on mock juries (e.g., Davis, 1980; Stasser et al., 1982) and actual juries (Kalven & Zeisel, 1966) that the further a jury is from unanimity, the more likely it is to hang. Together, these facts strongly imply that larger juries should hang more often,

especially for very close cases. The present study confirms this prediction for mock juries with rather limited deliberation times. The unresolved question is whether this effect will also obtain in actual juries, who operate without effective time limit. (Of course, we should not forget that many if not most decision-making groups do operate under time pressures and time limits. The time limit of the present study poses less of an ecological validity concern for this large and interesting class of groups.)

The generalizability question is, of course, an empirical issue that cannot be resolved without additional research. Unfortunately, the crucial experiment is exorbitantly expensive (requiring many larger groups that deliberate without time limit). Two alternative approaches to the generalizability question may be noted. First, it should be possible to do more careful field research on the effects of jury size, particularly because the use of juries with fewer than 12 members is becomming more commonplace. Second, theory and research on the causes of jury hanging would help us to evaluate the relevance of deliberation time limits. For example, if juries hang largely out of fatigue, then very short time limits may strongly distort results. However, if hanging is largely due to the presence of intransigent, uncompromising jurors (cf. Tanford & Penrod, 1983), then larger groups should be more likely to hang whether or not deliberation time is limited.

Disregarding hung juries, group size had no effect on the relative likelihood of conviction or acquittal. It would be misleading to conclude from this finding that the use of smaller juries results in no net pro- or anti-defendant bias. A hung jury is a relatively favorable outcome for a defendant (cf. Kerr et al., 1976; Lempert, 1975). It may prompt the dropping of charges, may lead to renewed plea bargaining, and always offers the possibility of acquittal upon retrial. Thus although the use of smaller juries may result in certain economies, by reducing the odds of a hung jury it may do so at the defendant's expense.

There were two other results of interest involving size and deliberation products. Smaller juries required shorter deliberation times, even when hung juries were excluded from the analysis. This replicates a finding of several earlier studies (e.g., Friedman & Shaver, 1975; Padawer-Singer et al., 1977). However, if one controls for size effects on initial splits, size did not affect deliberation time. Second, group polarization was not detectably stronger in the larger groups, as some formal models have predicted. However, the models also predict that this effect should be very small, particularly for very close cases, such as those used here.

### *Process of Deliberation*

Examination of the deliberation process revealed no differences between the 6- and 12-person groups. The estimated decision schemes and transition schemes (Tables 2 and 3, respectively) indicated that movement between states and to a final verdict were highly similar for those two conditions. For example, factions of the same relative size were equally able to attract converts and to ultimately prevail in 6- and 12-person groups. Also, the time required for a shift was the same for comparable 6- and 12-person groups. These findings reinforce the conclusion that the lower likelihood of a hung jury in the 6-person group did not stem from deliberation process differences between 6- and 12-person groups, but rather stemmed from the greater probability that the group was at or near unanimity at the outset of deliberations for the smaller groups (cf. Grofman, 1976; Padawer-Singer et al., 1977).

But as we consider still smaller groups, certain process differences did begin to emerge from the transition data. The proportionality model broke down for 3-person groups. For example, when there was a 2:1 ratio of convictors to acquitters, the majority faction was relatively more likely to attract a convert in the 3-person groups than in the two larger group sizes. In essence, a minority for acquittal had less influence in the 3-person group than in the larger groups. This can be most plausibly attributed to the lack of social support for minority members of 3-person groups. Asch (1951) has demonstrated the importance of social support for resisting a majority in a conformity situation; in this experiment we find a similar pattern within interacting decision-making groups.

We also observed in the group decision-

making context what Asch (1956) and Gerard et al. (1968) have observed in conformity paradigms: As the size of a unanimous majority increases, the likelihood of a minority-of-one yielding increases. Just as in Asch's experiments, yielding to a majority appeared to level off at higher group sizes. In our study, however, floor and ceiling effects may have had more to do with this leveling off than the psychological equivalence of unanimous majorities of three or more. In addition, we found that shifts occurred faster as the size of the majority faced by a minority of one increased.

We might note several implications of these findings. First, quite apart from other relevant concerns about jury size (e.g., representativeness), our results tend to support the Supreme Court's assumption of process equivalence between 6- and 12-person groups, although they are at odds with the Court's assumption of outcome equivalence (if we may consider a hung jury an outcome). They also tend to support the general conclusion of *Ballew v. Georgia* (1978) that juries may not continue to shrink in size much below 6 without somewhat different group dynamics coming into play. However, whether the effects reported here enhance or corrupt the jury's functioning is ultimately not an empirical question, but a matter of weighing competing sociolegal values. Finally, these results are of particular relevance for those attempting to develop formal models of the jury decision-making process (e.g., Davis, 1973, 1980; Grofman, 1976; Klevorick & Rothschild, 1979; Penrod & Hastie, 1979). Like Godwin and Restle's, (1974), our results strongly suggest that a complete model must take both the relative and absolute size of a faction into account when attempting to estimate its power to attract converts, particularly in smaller groups. On the other hand, even though they were reliable, the size effects on deliberation process were fairly weak. This suggests that a model that assumed proportionality would not predict much more poorly than one that did not.

### *Method of Polling*

The method used to poll the jury was also found to significantly affect the likelihood of hung juries for close cases. In a correlational study, Hawkins (1960) found that 12-person mock juries that used a secret ballot tended to hang more often than those that did not. However, in the present experimental study, exactly the opposite effect was found, at least for the 6- and 12-person groups considering close cases. Under these conditions, secret balloting significantly reduced the likelihood of hanging.

Hastie et al. (1984) have recently suggested that juries adopt one of two distinct styles of deliberation. In *verdict-driven* deliberations, jurors quickly align in factions and jurors become committed advocates of their positions. In *evidence-driven* juries, jurors tend to review and evaluate the trial evidence in a more impartial and accommodative fashion. Hastie et al. suggest that early public ballots tend to encourage verdict-driven deliberations. Their data also suggested that this style was more likely to produce a hung jury. Our results for the larger groups are entirely consistent with this analysis. Being publicly identified with a position may force early commitment to that position and make it difficult to change one's position without appearing inconsistent or irresolute. The result is a verdict-driven deliberation and, commonly, a hung jury. But if one can avoid making a clear public commitment to a position, it should be easier to switch sides and hence easier for the group to reach consensus. Whenever one is likely to be part of a small minority, it may be difficult to avoid public disclosure of one's preference regardless of the polling procedure, but when the opposing factions are both fairly large, a juror may be able to avoid commitment until the stronger position becomes evident. And one would expect two large opposing factions to result most often when the group is large and the case is close. Of course, these were exactly the conditions under which secret balloting discouraged hanging. These findings suggest that the correlation reported by Hawkins (1960) may have been spurious. Juries (or, perhaps, just forepersons) predisposed to avoid the open conflict necessary to achieve consensus may have preferred the secret ballot technique. Thus in Hawkins's study, secret balloting may have been an effect rather than a cause of an inability to reach agreement.

The only other effect of polling method was an effect like Hawkins's in the 3-person groups considering close cases. We have no compelling explanation of this effect.[7] Although the polling effect in the 3-person groups is theoretically intriguing, because most juries now have a minimum size of 6, it is the larger groups' polling effect that has greater practical significance for actual juries. It seems most relevant to foreperson behavior, because she or he would probably have the greatest influence on such procedural matters as polling. The present findings suggest that under the appropriate conditions (viz., when the case is a very close one), the jury foreperson's choice of a secret balloting technique can help reduce the risk of a hung jury. Of course, the relevance of these polling effects, like the size effects, for actual jury behavior must still be demonstrated empirically in future research.

[7] There should have been little difficulty identifying a minority member in these small groups so there should have been little anonymity to be gained through secret balloting. On very close cases, every juror has only a weak preference of one alternative over the other. Perhaps under secret polling, the difficulty of choosing between the verdict alternatives, and the risk that one could become a minority of one is made more salient as one personally debates how to mark one's ballot. Such considerations might temper the majority faction's determination to advocate their verdict preference, which would improve the probability of a hung jury. But under open polling, the lack of social support may make yielding to the other, narrowly rejected position an easy choice for a minority of one. The latter reasoning would also seem to suggest very low rates of hanging in the corresponding (i.e., 3-person, open-polling) clear-case condition; but there was no polling effect in this condition (See Table 1). Perhaps with fairly clear cases one is more likely to find oneself in the minority position as a result of carelessness (e.g., missing or misinterpreting significant facts) rather than as a result of careful reasoning (which recommends the majority position for a truly clear case). When being in the minority is a result of carelessness, it may be preferable to defend one's ill-advised choice rather than to reveal that one had failed to discern the "obvious" choice.

## References


Allen, V. (1965). Situational factors in conformity. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2, pp. 133–176). New York: Academic Press.

Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgments. In H. Guetzkow (Ed.), *Groups, leadership and men.* Pittsburgh: Carnegie Press.

Asch, S. E. (1956). Studies of independence and submission to group pressure: I. On minority of one against a unanimous majority. *Psychological Monographs, 70* (9, Whole No. 417).

Ballew v. Georgia 435 U.S. 223 (1978).

Beiser, E., & Varrin, R. (1975) Six-member juries in the federal courts. *Judicature, 58,* 423–433.

Bermant, G., & Coppock, R. (1973). Outcomes of six- and twelve-member jury trials: An analysis of 128 civil cases in the State of Washington. *Washington Law Review, 48,* 593–596.

Bishop, Y., Fienberg, S., & Holland, P. (1975). *Discrete multivariate analysis: Theory and practice.* Cambridge, MA: MIT Press.

Bray, R. M., & Kerr, N. L. (1982). Methodological considerations in the study of the psychology of the courtroom. In N. Kerr & R. Bray (Eds.), *The psychology of the courtroom* (pp. 287–323). New York: Academic Press.

Buckhout, R., Weg, S., Reilly, V., and Frohboese, R. (1977). Jury verdicts: Comparison of 6- and 12-person juries and unanimous vs. majority decision rule in a murder trial. *Bulletin of the Psychonomic Society, 10,* 175–178.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (rev. ed.). New York: Academic Press.

Davis, J. H. (1973). Group decision and social interaction: A theory of social decision schemes. *Psychological Review, 80,* 97–125.

Davis, J. H. (1980). Group decision and procedural justice. In M. Fishbein (Ed.), *Progress in social psychology* (pp. 157–229). Hillsdale, NJ: Erlbaum.

Davis, J. H., Kerr, N. L., Atkin, R. S., Holt, R., & Meek, D. (1975). The decision processes of 6- and 12-person juries assigned unanimous and two-thirds majority rules. *Journal of Personality and Social Psychology, 32,* 1–14.

Davis, J. H., & Nagao, D. (1980). Some implications of temporal drift in social parameters. *Journal of Experimental Social Psychology, 16,* 479–496.

Davis, J. H., Stasser, G., Spitzer, C. E., & Holt, R. W. (1976). Changes in group members' decision preferences during discussion: An illustration with mock juries. *Journal of Personality and Social Psychology, 34,* 1177–1187.

Friedman, H., & Shaver, K. (1975). *The effect of jury deliberations and decisions in mock criminal cases of 6- and 12-member juries and of unanimous and non-unanimous verdict requirements.* Unpublished manuscript, College of William and Mary, Williamsburg, VA.

Gerard, H., Wilhelmy, R., & Conolley, E. (1968). Conformity and group size. *Journal of Personality and Social Psychology, 8,* 79–82.

Godwin, F., & Restle, F. (1974). The road to agreement: Subgroup pressures in small group consensus processes. *Journal of Personality and Social Psychology, 30,* 500–509.

Grofman, B. (1976). Not necessarily twelve and not necessarily unanimous. In G. Bermant, C. Nemeth, & N. Vidmar (Eds.), *Psychology and the law* (pp. 149–168). Lexington, MA: D. C. Heath.

Hastie, R, Penrod, S., & Pennington, N. (1984). *Inside the jury.* Cambridge, MA: Harvard University Press.

Hawkins, C. H. (1960). *Interaction and coalition realignments in consensus seeking groups: A study of experimental jury deliberations.* Unpublished doctoral dissertation, Department of Sociology, University of Chicago.

Hoffman, L. R. (1979). *The group problem solving process: Studies of a valence model.* New York: Praeger.

Institute of Judicial Administration (1972). *A comparison of six- and twelve-member juries in New Jersey Superior and County Courts.* New York: Author.

Kalven, H., & Zeisel, H. (1966). *The American jury.* Boston: Little, Brown.

Kerr, N. L. (1981). Social transition schemes: Charting the group's road to agreement. *Journal of Personality and Social Psychology, 41,* 684–702.

Kerr, N. L. (1982). Social transition schemes: Model, method, and applications. In H. Brandstatter, J. H. Davis, & G. Stocker-Kreichgauer (Eds.), *Group decision making* (pp. 59–80). London: Academic Press.

Kerr, N. L., Atkin, R., Stasser, G., Meek, D., Holt, R., & Davis, J. H. (1976). Guilt beyond a reasonable doubt: Effects of concept definition and assigned decision rule on the judgments of mock jurors. *Journal of Personality and Social Psychology, 34,* 282–294.

Kerr, N. L., Stasser, G., & Davis, J. H. (1979). Model-testing, model-fitting, and social decision schemes. *Organizational Behavior and Human Performance, 23,* 339–410.

Kessler, J. B. (1973). An empirical study of six- and twelve-member jury decision-making processes. *University of Michigan Journal of Law Reform, 6,* 712–734.

Klevorick, A. & Rothschild, M. (1979). A model of the jury decision process. *Journal of Legal Studies, 8,* 141–161.

Latané, B. (1981). Psychology of social impact. *American Psychologist, 36,* 343–356.

Latané, B., & Wolf, S. (1981). The social impact of majorities and minorities. *Psychological Review, 88,* 438–453.

Lempert, R. O. (1975). Uncovering "nondiscernible" differences: Empirical research and the jury-size cases. *Michigan Law Review, 73,* 643–708.

Mills, L. R. (1973). Six-member and twelve-member juries: An empirical study of trial results. *University of Michigan Journal of Law Reform, 6,* 671–711.

Myers, D. G., & Lamm, H. (1976). The group polarization phenomenon. *Psychological Bulletin, 83,* 602–627.

Padawer-Singer, A. M., Singer, A. N., & Singer, R. (1977). An experimental study of twelve vs. six member juries under unanimous vs. nonunanimous decisions. In B. D. Sales (Ed.), *Psychology in the legal process.* New York: Spectrum.

Penrod, S. D., & Hastie, R. (1979). Models of jury decision making: A critical review. *Psychological Bulletin, 86,* 462–492.

Penrod, S., & Hastie, R. (1980). A computer simulation of jury decison making. *Psychological Review, 87,* 133–159.

Roper, R. T. (1980). Jury size and verdict consistency: "A line has to be drawn somewhere"? *Law & Society Review, 14,* 977–999.

Saks, M. (1982). Innovation and change in the courtroom. In N. Kerr & R. Bray (Eds.), *The psychology of the courtroom* (pp. 325–352). New York: Academic Press.

Saks, M. (1977). *Jury verdicts: The role of group size and social decision rule.* Lexington, MA: Lexington Books.

Stasser, G., Kerr, N. L., & Bray, R. M. (1982). The social psychology of jury deliberations: Structure, process, and product. In N. Kerr & R. Bray (Eds.), *The psychology of the courtroom* (pp. 221–256). New York: Academic Press.

Steiner, I. D. (1972). *Group process and productivity.* New York: Academic Press.

Tanford, S., & Penrod, S. (1983). Computer modeling of influence in the jury: The role of the consistent juror. *Social Psychology Quarterly, 46,* 200–212.

Tanford, S., & Penrod, S. (1984). Social influence model: A formal integration of research on majority and minority influence processes. *Psychological Bulletin, 95,* 189–225.

Williams v. Florida 399 U.S. 78C (1970).

Valenti, A. C., & Downing, L. L. (1975). Differential effects of jury size on verdicts following deliberations as a function of the apparent guilt of a defendant. *Journal of Personality and Social Psychology, 32,* 655–663.

Woodward, J. A., & Overall, J. E. (1976). Nonorthogonal analysis of variance in repeated measures experimental designs. *Educational and Psychological Measurement, 68,* 855–859.

Zeisel, H. (1971). . . . And then there were none: The diminution of the federal jury. *University of Chicago Law Review, 38,* 710–724.

Received February 20, 1984 ■