AN EMPIRICAL STUDY OF SIX-AND TWELVE-MEMBER JURY
DECISION-MAKING PROCESSES

There is a marked recent trend in the American judicial system to employ juries of six rather than twelve members. The United States Supreme Court has specifically upheld the constitutionality of the six-member jury in criminal cases and is considering during its current term the validity of this smaller-sized jury in civil cases as well. The use of the six-member jury appears to be based on the proposition that its results will not differ significantly from those produced by the traditional jury of twelve and hence will not result in an infringement of rights guaranteed by the Constitution.

The finding of no significant differences in results between the two different-sized panels should, as a statement of fact, be empirically tested. Moreover, it is a proposition the validity of which should be ascertained if the current trend of reducing the size of

3 Colgrove v. Battin, 456 F.2d 1379 (9th Cir. 1972), cert. granted, 409 U.S. 841 (1972).
5 The issue is whether the right to trial by jury as expressed in the sixth and seventh amendments requires a jury of twelve. The answer turns in part on the intent of the framers of the constitution. The Williams court resolved the problem by finding no implicit requirement of twelve-member juries.

We do not pretend to be able to divine precisely what the word “jury” imported to the Framers, the First Congress, or the States in 1789. It may well be that the usual expectation was that the jury would consist of twelve . . . . But there is absolutely no indication in “the intent of the Framers” of an explicit decision to equate the constitutional and common law characteristics of the jury. Nothing in this history suggests, then, that we do violence to the letter of the Constitution by turning to other than purely historical considerations to determine which features of the jury system, as it existed at common law, were preserved in the Constitution. The relevant inquiry . . . must be the function that the particular feature performs and its relation to the purposes of the jury trial. Measured by this standard, the 12-man requirement cannot be regarded as an indispensable component of the Sixth Amendment.

Williams v. Florida, 399 U.S. at 98-100 (footnote omitted). But cf. Capital Traction Co. v. Hof, 174 U.S. 1, 6-16 (1899), in which the Court, relying upon the framers’ intent to incorporate into the seventh amendment the common law of England, ruled that a trial by jury in a civil action required a twelve-member panel.
juries is to proceed free from suspicion that it is restricting sixth and seventh amendment rights. There are at least two levels at which the proposition can be empirically tested: an analysis of the literal finding that there are no significant differences in results,\(^6\) and an examination of whether the process by which those results are achieved differs between the different-sized panels. The latter is the purpose of this article: to seek to discover by the use of empirical research the degree to which the deliberative processes in civil cases of six- and twelve-member juries differ.

Legal analysis alone—that is, interpretation of cases and statutes—would not appear sufficient to make this determination. Because a jury verdict is the product of group decision-making, the skills of the social scientist, one who has expertise in examining the mechanics of group interaction, are required to undertake an analysis of deliberations. Moreover, when legal scholars have sought to study the jury as a decision-maker, they have in fact relied quite heavily on the social sciences.\(^7\) Past social scientific research has investigated, among other aspects of the deliberative process, participation in deliberation as a function of sex,\(^8\) variations in the personal characteristics of the victim and the defendant in criminal cases,\(^9\) and the persuasiveness of courtroom arguments with respect to the sequence in which they were offered.\(^{10}\) There is also a body of knowledge which has been obtained from the study of communication in small groups, including those groups whose responsibility it is to make a decision.\(^{11}\) Therefore, because the social sciences have both the expertise and methodology for the study of the jury, it is appropriate

---


\(^{7}\) See, e.g., H. KALVEN & H. ZEISEL, THE AMERICAN JURY (1966), which was a pioneer empirical study conducted at the University of Chicago Law School, undertaken to examine the functions of the jury in criminal and civil trials. The wide-ranging study was thoroughly interdisciplinary in character. The thrust of the program was to bring together into a working partnership the lawyer and the social scientist: the hope was to marry the research skills and fresh perspectives of the one to the socially significant problems of the other, and in the end to produce a new scholarship and literature for both. ... It was the intention that this be field research that we do more than collate existing social science writings with a series of legal problems, that new data be collected about old legal institutions.

\(^{8}\) Id. at v.

\(^{9}\) James, Status and Competence of Jurors, 64 Am. J. Sociology 563 (1959).


\(^{12}\) See, e.g., D. CARTWRIGHT & A. ZANDER, GROUP DYNAMICS RESEARCH AND THEORY (3d ed. 1968).
that an investigation into the effects of jury size be undertaken jointly by lawyers and social scientists.

This article employs the techniques of the social sciences in testing a legal proposition. After setting forth the hypotheses and methodology utilized by the experiment discussed herein, it presents the results obtained by examining the deliberations of different-sized juries concerning the same civil litigation. This article does not purport to be definitive; it does, however, attempt to indicate one methodology of interdisciplinary research which can be undertaken and the utility of this research to both the social sciences and the legal profession.12

I. DEVELOPMENT OF HYPOTHESES

A. Differences Between Verdicts

In June, 1970, the Supreme Court concluded in Williams v. Florida that six-member juries were constitutional in criminal cases. The basis for this decision, wrote Justice White, was that “[w]hat few experiments have occurred—usually in the civil area—indicate that there is no discernible difference between the results reached by the two different-sized juries.”13 The “experiments” referred to here are six articles14 which rely on opinions only and are not the product of disciplined, empirical research.15 One cited article merely predicts that a six-member jury would deliberate equally as well as a twelve-member jury.16 Other sources were the views of a judge,17 clerks of courts,18 and lawyers.19 Because this evidence is not the product of rigorous research, it should be viewed with some skepticism,20 especially

12 Fred Strodtbeck, a prominent social psychologist who devised the methodology employed in The American Jury, supra note 7, has called for a stronger bond between legal and sociological research, writing, “The frontiers of collaboration can accommodate legal doctrine and sociological theory without resulting in an interpenetration that will erase the unique prerogatives and responsibilities of the respective disciplines.” Strodtbeck, Social Process, the Law and Jury Functioning, in LAW AND SOCIOLOGY 164 (W. Evan ed. 1962).
14 Id. at 101, n.43.
19 Cronin, supra note 18, at 28-29; Six-Member Juries Tried in Massachusetts District Court, supra note 18, at 136.
20 Tapp, Reflections, 27 J. SOCIAL ISSUES, No. 2, 1971, at 14: “[L]awyers—socialized in a scientific milieu—may be too uncritical in their acceptance of ‘research’ findings.”
when the evidence is used to support a proposition of great importance to trial by jury.

There is additional evidence bearing on the question of similarity of verdicts, but it is equally inconclusive. Basing his conclusions on the opinions of judges and some judicial research into six- and twelve-member juries, Chief Judge Augelli of the United States District Court for the District of New Jersey has written:

While there is not enough empirical data available on this subject, experiments with six-member civil juries have shown that the verdicts reached by smaller juries are as carefully considered as, and comparable to, those reached in similar cases by twelve-member juries.21

Some differences have, however, been detected. A 1972 New Jersey study states that nonunanimous verdicts (five-sixths majority) were found in 20.2 percent of six-member juries and 45.0 percent of twelve-member juries.22 The experimenters caution, however, that because the option of using a six- or twelve-member jury was available, the twelve-member juries were requested in more difficult cases than were six-member juries.23

Further research is needed to control for the differences in cases tried by six- and twelve-member juries in order to give a more accurate evaluation of any six- and twelve-member differences in deliberation and verdicts. One study examined the question of jury size by having groups of four, eight, and twelve read the same sections from a preparation text for a law school aptitude examination. The study found no significant variations among the verdicts of the different-sized juries.24 Another researcher presented a filmed dramatization of a trial to three juries each of six, nine, and twelve students. He found that there was little difference among the results reached by the three different-sized juries, except that six-member juries had more trouble achieving unanimity.25 The researcher used a very small sample (N = 3), which weakened the impact of his findings. Addi-

---

22 INSTITUTE OF JUDICIAL ADMINISTRATION, A COMPARISON OF SIX- AND TWELVE-MEMBER CIVIL JURIES IN NEW JERSEY SUPERIOR AND COUNTY COURTS 7 (1972) [hereinafter cited as IJA Study].
23 Id.
25 R. Gordon, A Study in Forensic Psychology: Petit Jury Verdicts as a Function of the Number of Jury Members (1968) (unpublished doctoral dissertation at the University of Oklahoma).
tionally, he evaluated neither juror satisfaction nor the deliberative process itself.

In spite of the fact that the evidence regarding similarity of verdicts appears to be inconclusive, an increasing number of courts is utilizing six-member juries. 26 This pronounced trend necessitates assurance that six-member juries do in fact reach verdicts similar to those rendered by twelve-member juries. Thus, the first research hypothesis of the study reported in this article is: There will be no significant difference between the verdicts of six-member juries and the verdicts of twelve-member juries.

B. Differences Between Deliberation Times

One of the major reasons for the introduction of the smaller jury was to accelerate trials by jury. 27 Fewer jurors would presumably take less time to impanel, to poll, and to pass exhibits. Nevertheless, a rigorous comparison of deliberation times of six- and twelve-member juries is needed to determine if significant time savings can be expected in that area of jury activity. Previous research in this area has produced conflicting opinions. 28 The New Jersey study stated that “The average time for six-member deliberations was 1.2 hours; for twelve-member deliberations it was 1.8 hours . . . .” 29 This study, as mentioned previously, presents a distorted view because the twelve-member juries were requested for more complicated cases which presumably would take longer to deliberate under any circumstances.

It has been argued that in order to compare accurately the deliberation time of the different-sized juries both groups would need to see the same trial. Yet,

[t]his of course, would be virtually impossible, because precisely the same case would be unlikely to arise in two jurisdictions—one of which used twelve jurors and the other using

26 See note 1 supra. See also Augelli, supra note 21, at 281–82.
27 See, e.g., Devitt, supra note 1, at 275–76.
28 One researcher found that when the task is one of modifying opinion, in the absence of clear objective criteria for judgment, groups of six took longer than groups of three: “On abstract tasks, groups of six were faster than groups of three . . . .” A. HARE, HANDBOOK OF SMALL GROUP RESEARCH 235 (1962). Other research indicates that on a concept formation task groups of two obtained the answer in shorter time, used fewer questions, and failed less often at the task than did individuals. Groups of four failed least often, but were not otherwise superior to groups of two. The number of man minutes required per problem increases with size.
29 IJA Study, supra note 22, at 28–29.
fewer. Even if it did, the participants would be different, and therefore no accurate comparison could be made.\textsuperscript{30}

Chief Judge Augelli reiterates the vagueness of findings in this area:

[It] has been suggested that the time it would take a six-member jury to reach a verdict would be less than the time required by the traditional twelve-member jury. No statistical evidence on the subject is yet available, though judges who have worked with six-member juries have reported such findings.\textsuperscript{31}

In view of the demonstrated need for comparisons of deliberation time, the present study hypothesizes that: \textit{There will be no significant differences in the deliberation times of six- and twelve-member juries deliberating in the same case.}

\textbf{C. Differences Between Factfinding Abilities}

The jury is a factfinding body, and Williams expressed the Court’s view that juries of fewer than twelve members would be equally able to perform this responsibility.\textsuperscript{32} One means of testing the ability of the factfinder is to analyze and compare the number of relevant issues actually discussed during the deliberations of each of the different-sized panels.\textsuperscript{33} Therefore, a further research hypothesis of the present study is that: \textit{There will be no significant difference between the number of issues discussed by six-member juries and that discussed by twelve-member juries.}

\textbf{D. Differences in Group Participation}

In the area of group participation several tentative generalizations may be posited. One summary of the literature on size and participation demonstrates that there is less cohesion in larger groups, and that the larger the size of the group, the more likely it is that an individual member will become peripheral to the

\textsuperscript{30} Note, \textit{Reducing the Size of Juries}, 5 U. Mich. J.L. Ref. 87, 100 n.80 (1971). See also id. at 101 & n.83, indicating that in a study of groups composed of from five to twelve members confronted with the same problem, “larger groups consistently required more time to reach a decision.”

\textsuperscript{31} Augelli, supra note 21, at 292 (footnote omitted).

\textsuperscript{32} “Certainly, the reliability of the jury as a factfinder hardly seems likely to be a function of its size.” 399 U.S. at 100-01. See also Augelli, supra note 21, at 292-93.

\textsuperscript{33} The addition of members has the following impact on group task ability:

Many abilities or resources needed in task performance tend to have an additive character. They may include such things as the number of items of
group. Also, as groups increase in size a smaller percentage of members takes the opportunity to speak. Studies of organizational structure reinforce some of these findings. Increased size yielded poorer cohesion and poorer communication among members. Another investigation into the impact of group size on organizational units examined twenty-three studies of sets of comparable organizational units. The experimenter found that in twenty-one studies there was "a consistent significant finding that larger organizations have lower member-participation rates." Another study reports that members of twelve-member groups are less likely to speak than members of five-member groups. A 1964 survey of the literature of laboratory and field studies found a negative correlation between group size and participation. In regard to jury deliberations, smaller groups might give jurors a better chance to be heard and interact. There might also be fewer opinions on each issue to discuss. This suggests another research hypothesis: There will be a significantly greater number of individual contributions in the six-member jury than in the twelve-member jury.

E. Differences in Juror Satisfaction

A final consideration in the jury size question is that of the satisfaction each juror receives from his participation in the deliberative process. Juror satisfaction may well be a very important consideration in the community's acceptance of the judicial system. In general, group dynamics literature indicates that the smaller the group, the more likely it is that the individual will be satisfied with the discussion. Professor Strubbeck found that the

---

information which can be absorbed and recalled, the number of critical judgments available to correct errors of information and inference, the number of ideas or suggestions available for solution of problems . . . . However, the familiar phenomenon of "diminishing returns" tends to set in at some point. The addition of a person to a group of a given size may not represent a unit addition to task ability.

A. Hare, supra note 28, at 233.


25 Id. at 375.

26 See, e.g., Porter & Lawler, Properties of Organizational Structure in Relation to Job Attitudes and Job Behavior, 64 Psychology Bull. 23 (1964).

27 Id. at 40.

28 Indik, Organization Size and Member Participation, 18 Human Relations 339 (1965).

29 A. Hare, supra note 28, at 231.


31 Note. supra note 30, at 103 n.93.

32 Id. at 104-05 & n.100.
level of juror satisfaction was positively correlated with the level of each juror’s participation. Another study, examining work produced by large organizations, found less satisfaction with work as the size of the organization increased. A third experiment revealed a similar negative correlation between size and total group satisfaction. Still another found that dissatisfaction increased as size increased from two to seven members in problem-solving, production, and discussion groups. A final research hypothesis arises here: There will be greater juror satisfaction among the members of the six-member jury than among members of the twelve-member jury.

II. RESEARCH METHODOLOGY

The study reported here consisted of showing a videotaped trial to eight juries of six members and eight juries of twelve members. On March 24, 1972, this researcher produced and directed the videotaping of a mock trial involving an actual automobile negligence case which had been settled out of court. The actual plaintiff and defendant involved in the 1971 accident agreed to portray themselves. A resident in orthopedics from the University of Michigan Hospital agreed to testify as an expert witness for the plaintiff. An experienced trial attorney served as judge, and two third-year law students who had excelled in a trial practice course acted as the attorneys.

The facts of the case, as they actually occurred and were presented to the mock court, indicated that while plaintiff was driving south along a main thoroughfare within the city limits of Ann Arbor, Michigan, defendant pulled onto that thoroughfare from a private driveway, crossing two lanes of traffic in order to move into the southbound lanes. When plaintiff moved to the right to avoid defendant, plaintiff lost control of his car and struck a tree. Plaintiff suffered injuries to his knees, incurred substantial medical bills, and missed several weeks of work. The two cars did not collide, and defendant was unhurt.

The trial proceeded according to the procedures employed in

44 See generally Indik, supra note 38.
45 See generally Porter & Lawler, supra note 36.
46 Hackman & Vidmar, Effects of Size and Task Type on Group Performance and Member Reactions, 33 SOCIOMETRY 37–54 (1970).
47 The author acknowledges the assistance of the University of Michigan Journal of Law Reform and Edward Stein, Visiting Lecturer in Law at the University of Michigan Law School and member of the Michigan Bar.
Michigan trial courts. Counsel for plaintiff and defendant each presented a brief opening argument. Plaintiff’s first witness was the plaintiff himself, who testified as to the accident and the resulting injuries. Defendant’s counsel cross-examined. Plaintiff’s second witness was the expert witness who testified as to the nature and effect of plaintiff’s injuries, relying on the actual medical records. Defendant’s counsel followed with cross-examination. The defendant testified as the sole witness for the defense. Plaintiff’s counsel then cross-examined. Finally, closing arguments were presented, with defendant’s argument preceding that of plaintiff. Instructions by the judge concluded the trial.\footnote{Michigan Supreme Court Committee on Standard Jury Instructions, Michigan Standard Jury Instructions—Civil §§ 10.00–11.05 (Institute of Continuing Legal Education, 1970).}

Plaintiff sought to prove that defendant had not exercised due care and was therefore negligent when he moved onto the street and had thereby caused plaintiff to veer off the roadway. Defendant contended that because plaintiff had been driving at an excessive speed, plaintiff also was not exercising due care and was therefore contributorily negligent. A defense of contributory negligence, if believed by the jury, prevents a verdict for plaintiff in any amount, even if both parties were negligent.\footnote{W. Prosser, Handbook of the Law of Torts § 65, at 425 (4th ed. 1971).}

The trial was recorded on a one-inch Ampex videotape machine and lasted for two and one-half hours. Time cues were given periodically to the lawyers and judge to keep the trial relatively brief. Four unidirectional microphones were used: one in front of the judge, one at the witness stand, and one at each attorney’s table. An engineer from the University of Michigan Television Center maintained the audio levels of the various participants, thus creating a uniform audio level. Since the realism of the proceedings was of utmost concern, the camera was placed in the jury box, thus giving the student jurors who ultimately viewed the tape the perspective of an actual juror. As a result of a questionnaire prepared by the researcher and administered to nonjuror viewers of the live mock trial and its videotape, the trial was found to be perceived as realistic and equally weighted for both sides of the suit.\footnote{Fifty nonjuror subjects were asked if the trial seemed to them to be a real trial.}

<table>
<thead>
<tr>
<th></th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very real</td>
<td>11</td>
<td>22</td>
</tr>
<tr>
<td>Real</td>
<td>32</td>
<td>64</td>
</tr>
<tr>
<td>Unreal</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td>Very unreal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>50</td>
<td>100</td>
</tr>
</tbody>
</table>
The subjects of the experiment were 144 student jurors randomly assigned from undergraduate speech classes at the University of Michigan. They viewed the videotape of the trial between April and October of 1972 in either six- or twelve-member groups. The jury deliberations were recorded on audiotape in a laboratory equipped with a one-way mirror through which visual observations of the deliberations were recorded. In accordance with Michigan law, a verdict was reached when five-sixths of the jurors agreed.\textsuperscript{51}

After the subjects viewed the trial, they were given a preliminary questionnaire to ascertain their immediate individual verdict. After the deliberations in the laboratory resulted in a group decision, a second individual questionnaire was administered for comparison of pre- and postdeliberation opinions. The latter questionnaire also evaluated the juror's individual satisfaction with the deliberations. In addition to the above information, segments of each deliberation were coded by three observers in order to analyze and compare various group communication aspects of the two different-sized juries.\textsuperscript{52} The entire deliberation of each panel was also coded in order to ascertain the number of issues discussed and the relevance of those issues to the deliberation.

The two basic problems most often mentioned by jury researchers are the method of selection of juror subjects and the method of presenting the trial information. One experimenter used jurs from the county courts of Minneapolis and Chicago in analyzing various methods of jury research.\textsuperscript{53} In subsequent studies this same researcher has used student jurors, and he concludes that there are no major differences in employing actual or student jurors for purposes of experimentation. Moreover, the cost and problems involved with actual jurors seem to outweigh any marginal benefit of using them in experimentation.\textsuperscript{54}

\begin{table}[h]
\centering
\begin{tabular}{ccc}
\textbf{Equally strong on both sides} & \textbf{Number} & \textbf{Percent} \\
\hline
& 25 & 50 \\
\hline
Slightly one sided & 21 & 42 \\
\hline
Basically one sided & 4 & 8 \\
\hline
Total & 50 & 100 \\
\end{tabular}
\end{table}


\textsuperscript{52} This data is not discussed here but will appear in the author's dissertation. The data is on file with the \textit{University of Michigan Journal of Law Reform}.


\textsuperscript{54} Forston compared eight groups of real jurors and eight groups of student jurors and
The problem of motivation of subjects was resolved in this study by giving the students the impression that they were viewing a real trial and indicating that the results of their deliberations were critical to this study. Each of the sixteen groups took its job quite seriously. The juries were thoughtful in their analysis of the issues, and even when there was initial agreement within the group, they carefully reviewed the issues to be certain of their decisions.

One experimenter utilized three methods of case presentation in his study of the differences in jury experimentation. He employed live mock trials, audiotaped trials, and the reading of edited trial transcripts. He concluded that although there were some differences among the methods and problems within each method, all three were viable alternatives in jury research. That researcher is currently attempting to use videotaped trials in jury research, for this method seems to be more realistic than any of those previously used, except for the costly actual trial simulation which cannot be replicated.

III. Results

A. Verdicts

Although the case used in this study was actually settled out of court in the plaintiff’s favor, all sixteen juries either found for the defendant or were unable to arrive at a verdict. There were two six-member juries which were hung, that is, unable to decide by a
five-sixths margin; an additional six-member jury rendered a verdict for the defendant, although two jurors privately, as indicated by their postdeliberation questionnaires, were still for the plaintiff. There was one hung jury among the twelve-member panels, seven jurors in favor of the plaintiff and five for the defendant. This was the only jury with a majority of members in favor of the plaintiff.

Several factors may have been the cause of this discrepancy between the actual out-of-court settlement and the verdicts of the student juries. Review of the audiotaped jury deliberations revealed that the defendant’s demeanor on the witness stand appeared to be more informal and therefore more appealing to student jurors than the plaintiff’s, and that the defendant’s counsel was a more eloquent speaker than was plaintiff’s counsel. A further explanation might be that defendant’s insurance company might have been willing to settle the actual case rather than risk losing in court. In the negotiations preceding the settlement, plaintiff’s actual lawyer argued that the defendant in fact had the last clear chance to look and see the plaintiff and thereby avoid the accident. This argument was not used by the plaintiff’s lawyer in the mock trial, and this omission might explain the discrepancy. This article is not, however, concerned with the discrepancy between the actual settlement and the verdicts. Rather what is significant is that the facts of the case were sufficiently uncertain to provide the researcher the opportunity to observe and compare meaningful deliberations and not merely to compare the verdict rendered.

Table 1 displays the results reached by the different-sized juries. Six of the eight six-member juries and seven of the eight twelve-member juries found for the defendant. Hung juries occurred in the other three deliberations. Table 1 also indicates that

between six- and twelve-member juries is predicted the hypothesis may be statistically tested directly. Even where differences between the two different-sized juries are initially predicted, only the statement of no difference between the subject groups, the null hypothesis, can be statistically tested. Once the null hypothesis has been tested, the results of the test are evaluated in terms of the predicted hypothesis.

To test the null hypothesis, the data are collected. Differences may then be observed between the different-sized groups. Statistical testing is then applied to determine whether the observed differences reflect real differences or merely chance occurrence.

The probability level set as a criterion for rejection of a null hypothesis in this study is .05. This criterion means that if the probability of chance occurrence is greater than five in one hundred, the differences observed cannot be conclusively said to be true differences, and that therefore the null hypothesis cannot be rejected.

The fact that the level for statistical significance may not be reached does not mean that the finding is not important and that the tendency expressed by the findings cannot be considered as an interesting finding. For a discussion of statistical significance, see F. Williams, Reasoning with Statistics: Simplified Examples in Communication Research 55–72 (1968). See also W. Dixon & F. Massey, Introduction to Statistical Analysis 88–91 (1957).
TABLE 1

<table>
<thead>
<tr>
<th>Verdict</th>
<th>Six-Member Jury</th>
<th>Twelve-Member Jury</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>Percent</td>
</tr>
<tr>
<td>Plaintiff</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>Defendant</td>
<td>6</td>
<td>75.0</td>
</tr>
<tr>
<td>Hung</td>
<td>2</td>
<td>25.0</td>
</tr>
</tbody>
</table>

\[ \chi^2 = .41, \, df = 1, \, p < .52 \]

where: \( \chi^2 \) is the chi-square statistic;
\( df \) is the number of degrees of freedom;
and \( p \) is the probability that the observed differences would occur by chance.

there was a very high probability that the differences which did occur arose by chance. The value of the probability associated with the chi-square statistic (\( \chi^2 \)),\(^{59}\) the device used here to test for significant differences, is so high that the hypothesis of no difference cannot be rejected. The six-member juries in this study as in another,\(^{60}\) however, had more difficulty in reaching consensus.

Because there were no verdicts in plaintiff’s favor, it was impossible to compare damage awards between the six- and twelve-member juries. Ten juries did, however, mention the issue of damages during deliberation.

B. Time

The data presented in Table 2 indicate that the difference in the time taken by the different-sized juries was not statistically significant. The Mann-Whitney U statistic, another test statistic, was used to analyze the data, as the data were not normally distributed and the sample size (\( N \)) was small.\(^{61}\) With an \( N \) of 8 in each group compared and a Mann-Whitney U statistic of 29.5, the probability

---

\(^{59}\) The chi-square statistic may be used when the data is categorical in nature, and there is a counting of the number of cases falling into each category. To employ the chi-square statistic there should be an assumption of the independence of groups and the random sampling of subjects. The tossing of a coin and counting the number of heads is an example of counting the cases in a category. The chi-square statistic compares observed frequencies in each category in relationship to the theoretical frequencies expected. The statistic, by determining acceptance or rejection of the hypotheses of no difference, indicates whether the two samples are significantly different. See, e.g., W. Dixon & F. Massey, supra note 58, at 221–43; H. Blalock, Social Statistics 212–21 (1960).

\(^{60}\) Gordon, supra note 25.

\(^{61}\) The data of deliberation time was tested with the Mann-Whitney U statistic. The Mann-Whitney U statistic is a nonparametric test. A nonparametric test may be used when "the sample is small and normality cannot be assumed." See H. Blalock, supra note 59, at 188, 197–203.
of differences between the deliberation time of the six- and twelve-member juries occurring by chance is approximately .4. Therefore, the hypothesis of no difference between the two sized juries cannot be rejected. Although presented here for purely descriptive purposes, the mean deliberation time of the six-member juries, 22.22 minutes, however, was somewhat higher than the mean deliberation time of the twelve-member juries, 15.27 minutes. This is an indication that the six-member panels might have had more controversy to resolve within their deliberations. This supposition, however, runs counter to the literature which suggests that larger groups generally take longer to reach decisions than smaller groups. Further findings on the participation of jurors discussed below which suggest increased member participation in the smaller juries may provide a basis for the slight increase in time needed by six-member juries.

C. Issues Discussed

The results presented in Table 3 were obtained by a procedure of content analysis. In order to analyze the content of the deliberations, three coders listened to recordings of the entirety of each deliberation. At any point at which a new speaker entered the discussion or at which a new issue was raised, the tape was stopped to allow the coders to designate the issue that had been

---

62 The means (X̄) presented in Table 2 are purely descriptive. The mean is not a valid tool for measurement because the distribution of scores did not form a normal distribution. Each score greatly deviates from the mean score, thus invalidating the use of the mean.
63 See note 28 and accompanying text supra.
64 See Part III D infra.
discussed. The following are examples of discussions of single issues.

Example # 1:

Juror A: My personal view is that I believe the man was exceeding the speed limit, um, because of the injuries that he he sustained.

Juror B: Um, I'm not crystal clear on it. There're still doubts, um, as to as to whether he was. um, going fifty miles over the speed limit. . . .

Juror C: But I think he went, um, a considerable number of miles over the speed limit. 65

Example # 2:

Juror A: [He] should'a had insurance.

Juror B: He probably did. . . .

Juror C: Yeah.

Juror D: I'm sure they had insurance . . . cause like you can't . . . well like the guy . . . . They don't want to say that.

Juror C: No they can't say that. . . . If they'd said well this guy had to pay a thousand dollars in hospital bills, but the insurance company paid it, and you'd think, well, you know, so he gets twenty-five thousand dollars free, you know.

Juror E: Aw, but his premiums go up. 66

Example # 3:

Juror A: Why don't we decide right now if we have a verdict?

Juror B: Yeah.

Juror C: Let's.

Juror D: Why don't we take a preliminary vote.

65 Quoted from twelve-member jury deliberations. Sept. 27, 1972.
Juror B: Okay.

Juror D: All those that. ah. feel in favor of the defendant, raise their right hand. All those in favor of the plaintiff in the case raise their right hand.\textsuperscript{67}

The content analysis instrument used for this part of the study\textsuperscript{68} provided a list of all issues discussed in all of the deliberations. After the coders indicated the issues heard, lawyers and law students were asked to designate the legally irrelevant issues, that is, issues which could not properly be considered by the jury. Example \# 2 above, the discussion of insurance, was one irrelevant issue. Most other irrelevant issues pertained to personal experiences, unrelated to the solution of the case, or referred to the experiment itself. Issues were further divided as to issues of evidence, e.g. example \# 1 above, referring to speed, and issues of verdict, e.g. example \# 3 above, referring to voting.\textsuperscript{69}

A content analysis of each deliberation found no statistically significant difference between the number of relevant and irrelevant issues discussed by the different-sized juries.\textsuperscript{70} The means $\bar{X}$ are presented in Table 3 for purely descriptive reasons, as they are not meaningful as predictive measures in this case. To interpret properly the data obtained, it was necessary to determine whether the differences noted were significant. Once again, the researcher checked a table of Mann-Whitney U values. The eight six-member juries and eight twelve-member juries were compared with regard to the number of irrelevant issues discussed. With an $N$ of 8 juries in each condition compared and a Mann-Whitney U statistic of 31, the probability that the differences among the number of irrelevant issues discussed by the six- and the twelve-member juries would occur by chance is approximately .4. Therefore, the hypothesis of no differences between the two different-sized juries cannot be rejected.

\textsuperscript{67}Quoted from twelve-member jury deliberations, Oct. 2, 1972.
\textsuperscript{68}A copy of the form used is on file with the \textit{University of Michigan Journal of Law Reform}.
\textsuperscript{69}The average coder reliability for this section of the study was 84.6 percent, determined according to the following formula:
\textsuperscript{70}Content analysis of small group communication is an accepted research tool. See, e.g., R. Bales, \textit{Interaction Process Analysis: A Method for the Study of Small Groups} (1950).
\[ \frac{2M}{N_1 + N_2} \]
TABLE 3

<table>
<thead>
<tr>
<th></th>
<th>Irrelevant Issues</th>
<th>Relevant Issues</th>
<th>Total Issues</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Six-Member</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Juries</strong></td>
<td>9.33</td>
<td>42.33</td>
<td>51.66</td>
</tr>
<tr>
<td></td>
<td>1.33</td>
<td>19.33</td>
<td>20.66</td>
</tr>
<tr>
<td></td>
<td>3.66</td>
<td>21.00</td>
<td>24.66</td>
</tr>
<tr>
<td></td>
<td>6.66</td>
<td>40.34</td>
<td>47.00</td>
</tr>
<tr>
<td></td>
<td>9.66</td>
<td>36.00</td>
<td>45.66</td>
</tr>
<tr>
<td></td>
<td>3.00</td>
<td>13.00</td>
<td>16.00</td>
</tr>
<tr>
<td></td>
<td>6.00</td>
<td>14.00</td>
<td>20.00</td>
</tr>
<tr>
<td></td>
<td>7.00</td>
<td>40.33</td>
<td>47.33</td>
</tr>
<tr>
<td><strong>X̄ = 5.83</strong></td>
<td><strong>X̄ = 28.29</strong></td>
<td><strong>X̄ = 34.12</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Twelve-Member</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Juries</strong></td>
<td>7.00</td>
<td>18.00</td>
<td>25.00</td>
</tr>
<tr>
<td></td>
<td>3.00</td>
<td>20.66</td>
<td>23.66</td>
</tr>
<tr>
<td></td>
<td>15.33</td>
<td>38.67</td>
<td>54.00</td>
</tr>
<tr>
<td></td>
<td>1.33</td>
<td>19.67</td>
<td>21.00</td>
</tr>
<tr>
<td></td>
<td>3.00</td>
<td>14.00</td>
<td>17.00</td>
</tr>
<tr>
<td></td>
<td>6.00</td>
<td>30.66</td>
<td>36.66</td>
</tr>
<tr>
<td></td>
<td>3.33</td>
<td>27.33</td>
<td>30.66</td>
</tr>
<tr>
<td></td>
<td>7.66</td>
<td>23.34</td>
<td>31.00</td>
</tr>
<tr>
<td><strong>X̄ = 5.83</strong></td>
<td><strong>X̄ = 24.04</strong></td>
<td><strong>X̄ = 29.87</strong></td>
<td></td>
</tr>
</tbody>
</table>

Mann-Whitney U Statistic = 31, p < .4

*Based on the average of the content analysis of three coders

D. Juror Participation

One statistically significant finding was obtained with regard to juror participation or nonparticipation. In arriving at these results, the question arose whether to use the group as the unit to be evaluated or whether to use each individual as the unit for analysis. Arguably, exposure to group interaction affects the behavior of each individual member, and, therefore, the behavior of the group should be viewed as the unit for analysis rather than the individual. In analyzing the participation-nonparticipation question, three methods of analysis were used in order to evaluate the data fully.

First, the individual was employed as the unit for analysis. This analysis indicated that, to a statistically significant extent, jurors in twelve-member groups were far less likely to speak at all, while six-member jurors almost always contributed to the discussion. Table 4A indicates that 95 percent of the six-member jurors contributed to the deliberation, while only 75 percent of the twelve-member jurors spoke. On the basis of the data presented, the probability of differences between the six- and twelve-member juries occurring by chance is .002. This is far below the .05 level set in this study, and thus the null hypothesis of no differences
between the participation of members in the two different-sized groups can be rejected in favor of the original prediction of the six-member jurors' participating more than twelve-member jurors. This finding is quite consistent with group literature and common sense. A shy or hesitant person will usually not contribute when in a larger group; in a smaller group his contributions are more necessary to, and encouraged by, the group. Experts in the field of group communication suggest that a very large group may inhibit participation:

A committee of thirty people, for purely physical reasons, may make it difficult to get into the conversation. The presence of a few monopolizers may leave little chance for others who are less aggressive and unwilling to interrupt.

Further, in examining the reasons given by the nonparticipating jurors for their lack of participation, the most common response was, "I usually don't say much in a group." This study seems to indicate that a juror is less likely to participate in the discussion when serving on a larger jury.

A second method used to analyze this data compared for juries of each size the number of groups which had 100 percent participation with the number of groups having less than 100 percent participation. Table 4B indicates that in six of the eight six-member juries all members contributed, and that in only one of the eight twelve-member juries was there full participation. The findings show a statistically significant difference, with
Table 4B

| Participation: 100% Participation vs. Less Than 100% Participation |
|-----------------------------------------------------|-----------------------|
| 100% Participation (all members contributed) | Less Than 100% Participation (one or more silent members) |
| **Six-Member Jury** | 6 juries | 2 juries |
| **Twelve-Member Jury** | 1 jury | 7 juries |

Fisher’s Exact Probability for $2 \times 2$ Test, $p < .02$

six-member juries rather than twelve-member juries much more likely to have 100 percent participation.

A third method for analyzing these data compared the percent of participation in six-member juries with the percent of participation in twelve-member juries. The mean percentage of participation for each sized jury is presented for descriptive reasons only (Table 4C). In order to find the significance level achieved, the Mann-Whitney U statistic was employed. Thus with eight juries in each condition, the probability of differences between the two sized groups occurring by chance is .00035, far below the .05 level set in this study. Therefore, the null hypothesis can be rejected in favor of the prediction that six-member jurors participate more than do twelve-member jurors. These findings indicate that the six-member jury might afford the potential nonparticipant a better opportunity to be heard.

Additionally, a comparison of participation of minority jurors indicated no statistically significant differences between the different-sized juries. Table 5 indicates that the average percent of minority member participation in six-member juries was approximately 20 percent of all comments made, while the parallel figure for twelve-member juries was approximately 13 percent. On the basis of these data, the probability of differences between

developed by R. A. Fisher gives a more exact probability for rejection of the null hypothesis than does the chi-square approximate probability. H. Blalock, *supra* note 59, at 221–25.

55 A minority juror is a juror who on the pre-deliberation questionnaire administered by the researcher took a position contrary to that of the majority of those on his panel.

56 To test the significance of these findings, the $t$-test was employed. Student's $t$-test is a test of the differences between the means ($X$) of the two samples, controlling for the differences of each score from the sample's mean (standard deviation). The test assumes a normal distribution of scores and random sampling of subjects. See, e.g., H. Blalock, *supra* note 59, at 144–53.
### Table 4C

**Participation: Percent of Participation for Each Jury**

<table>
<thead>
<tr>
<th>Six-Member Juries</th>
<th>Twelve-Member Juries</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of</td>
<td>Number of</td>
</tr>
<tr>
<td>Participators</td>
<td>Participators</td>
</tr>
<tr>
<td>Percent of</td>
<td>Percent of</td>
</tr>
<tr>
<td>Participation</td>
<td>Participation</td>
</tr>
<tr>
<td>5</td>
<td>86.66</td>
</tr>
<tr>
<td>6</td>
<td>100.00</td>
</tr>
<tr>
<td>5</td>
<td>86.66</td>
</tr>
<tr>
<td>6</td>
<td>100.00</td>
</tr>
<tr>
<td>6</td>
<td>100.00</td>
</tr>
<tr>
<td>6</td>
<td>100.00</td>
</tr>
<tr>
<td>6</td>
<td>100.00</td>
</tr>
<tr>
<td>6</td>
<td>100.00</td>
</tr>
</tbody>
</table>

\[ \bar{X} = 96.67\% \]  
\[ \bar{X} = 77.08\% \]

Mann-Whitney U statistic = 5.00, p < .00035 (one-tailed)

*Percent of jury members who contributed in each jury; i.e., 86.66% of the members of jury #1 in the six-member condition spoke at all.

### Table 5

**Percent of Minority Juror Participation**

<table>
<thead>
<tr>
<th>Six-Member Minority Jurors (N = 11)</th>
<th>Twelve-Member Minority Jurors (N = 18)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.49</td>
<td>0.00</td>
</tr>
<tr>
<td>38.43</td>
<td>21.95</td>
</tr>
<tr>
<td>29.51</td>
<td>9.76</td>
</tr>
<tr>
<td>1.15</td>
<td>18.99</td>
</tr>
<tr>
<td>11.83</td>
<td>6.49</td>
</tr>
<tr>
<td>26.72</td>
<td>1.68</td>
</tr>
<tr>
<td>6.93</td>
<td>23.56</td>
</tr>
<tr>
<td>52.81</td>
<td>0.00</td>
</tr>
<tr>
<td>23.40</td>
<td>20.97</td>
</tr>
<tr>
<td>14.04</td>
<td>44.44</td>
</tr>
<tr>
<td>18.84</td>
<td>7.01</td>
</tr>
</tbody>
</table>

\[ \bar{X} = 20.83 \]

<table>
<thead>
<tr>
<th></th>
<th>Twelve-Member Minority Jurors (N = 18)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>14.65</td>
<td></td>
</tr>
<tr>
<td>0.64</td>
<td></td>
</tr>
<tr>
<td>8.28</td>
<td></td>
</tr>
<tr>
<td>21.49</td>
<td></td>
</tr>
<tr>
<td>23.33</td>
<td></td>
</tr>
<tr>
<td>17.50</td>
<td></td>
</tr>
</tbody>
</table>

\[ \bar{X} = 13.37 \]

*t-test = 1.46, df = 27, p < .075 (one-tailed)

*Percent = minority jurors' number of comments divided by total comments made by all members of that jury.
the six- and twelve-member juries occurring by chance is .075 and only approaches the .05 level set for rejection of the null hypothesis. Thus, under this mode of analysis the null hypothesis cannot be rejected in favor of the prediction that there will be greater participation in the six-member juries.

The finding that twelve-member jurors, when in the minority, contributed slightly less than six-member minority jurors is inconclusive. Nevertheless it may indicate a tendency for twelve-member minority jurors to yield to the majority and be overwhelmed by the majority to a greater extent than are minority members of six-member juries. Comments such as, "I can see your point of view,"77 from a minority member of a jury rendering an 11–1 verdict, and "Doesn't anyone else agree with me? ... [Silence],"78 from a minority member of a jury rendering a 10–2 verdict illustrate the lack of reinforcement encountered by the minority jurors. The group communication literature suggests that minority members of a group can be swayed, and perhaps a larger group offers more pressure toward conformity.79 Although no definite conclusions can be drawn, there is the suggestion that the six-member jury offers the potential nonparticipator and minority decision jurors a better chance to be heard.

E. Juror Satisfaction

Table 6 relates juror participation to juror satisfaction without regard to the size of the jury on which each juror served. For example, the 67 jurors who were "very satisfied" with the deliberative process had an average percent of participation of approximately 10 percent. Contrary to most group literature on participation and satisfaction, comparing the mean participation per-

77 Quoted from twelve-member jury deliberations, Sept. 23, 1972.
78 Quoted from twelve-member jury deliberations, Sept. 25, 1972.
79 One researcher investigated the impact of group pressure on the perception of a minority group member:

The autokinetic situation requires the individual to judge the movement of a light that is really stationary, although it appears to move; hence, the individual is faced with a very ambiguous situation and he is very uncertain about the actual amount of perceived movement. Under these conditions, about eight of every ten subjects yield to the unanimous group pressure (Sherif & Sherif, 1956). [In another study] ... the subject is asked to judge which of three vertical lines is the same length as a standard line which is presented at the same time. Typically, there is little doubt about the correct answer; when individuals respond in isolation, few errors of judgment are made. Nevertheless, about one-third of all subjects show some degree of conformity to unanimous group pressure in the form of a false norm (Asch, 1951).

Table 6

<table>
<thead>
<tr>
<th>Satisfaction Levels</th>
<th>Number</th>
<th>Mean Percent of Participation*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-Very Satisfied</td>
<td>67</td>
<td>9.64</td>
</tr>
<tr>
<td>2-Satisfied</td>
<td>52</td>
<td>11.79</td>
</tr>
<tr>
<td>3-Unsatisfied</td>
<td>19</td>
<td>13.97</td>
</tr>
<tr>
<td>4-Very Unsatisfied</td>
<td>6</td>
<td>12.54</td>
</tr>
</tbody>
</table>

Analysis of Variance $F = .97, df = 9/186, p < .41$

*Percent of participation was found by dividing number of observed comments of member by total comments of the group. The mean was then calculated for each satisfaction level.

Percentage at each of the four levels of satisfaction designated by the researcher, there was no statistical significance found; in fact, the means were quite similar. As noted in Table 6, the probability that the differences between the six- and twelve-member juries would occur by chance is .41, and thus the null hypothesis cannot be rejected in favor of the prediction of more satisfaction of participants in six-member rather than twelve-member juries.

Table 7 relates juror satisfaction to the sized jury in which the jurors served. For example, 33 percent of six-member jurors were "very satisfied" with the deliberative process, while 53 percent of twelve-member jurors were similarly satisfied. The table as a whole shows that twelve-member jurors tended to be more satisfied with the group product. This may be because less controversy occurred than in the six-member condition, where more people participated and the diverse ideas of the minority were more clearly drawn. The statistical comparison of size and satisfaction, however, yielded no significant finding. Given the data of Table 7, the probability that differences between the six- and twelve-member juries would occur by chance is .15. Thus, the null hypothesis cannot be rejected in favor of the original prediction of increased juror satisfaction with the six-member jurors.

The added controversy in the six-member juries seems to have led to more difficulty in reaching consensus, and consequently, to less satisfaction with the group product. When a group is unable to resolve conflict, the members will naturally be less satisfied with the group’s efforts. Perhaps, in a real trial situation the

---

80 The significance test employed here was an analysis of variance. H. Blalock, supra note 59, at 242-55. The analysis of variance is similar to the t-test, and is used when more than two groups are to be compared. See note 76 supra.

81 The chi-square statistic is used here for frequency data to test the independence of categories.
TABLE 7

Satisfaction and Size

<table>
<thead>
<tr>
<th>Satisfaction Levels</th>
<th>Six-Member Jury</th>
<th>Twelve-Member Jury</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>Percent</td>
</tr>
<tr>
<td>1-Very Satisfied</td>
<td>16</td>
<td>33.3</td>
</tr>
<tr>
<td>2-Satisfied</td>
<td>21</td>
<td>43.7</td>
</tr>
<tr>
<td>3- Unsatisfied</td>
<td>8</td>
<td>16.7</td>
</tr>
<tr>
<td>4-Very Satisfied</td>
<td>3</td>
<td>6.2</td>
</tr>
</tbody>
</table>

\[ \chi^2 = 5.27, \, df = 3, \, p < .15 \]

six-member jurors would take more time to resolve the conflict and ultimately reach a decision. Presumably the jurors would then be more satisfied in a smaller group, in which more people would be heard if consensus could eventually be reached.

IV. CONCLUSION

Because of the small number of student juries analyzed in this study, the conclusions can hardly be applied generally. The use of only one case at trial also limits the ability to generalize these findings. On the basis of statistical comparisons, the study found: that there were no significant differences between the verdicts, times of deliberation, and numbers of issues discussed in the two different-sized panels; that six-member jurors participate rather than remain silent significantly more often than do twelve-member jurors; and that six-member jurors are not significantly more satisfied with the deliberative process than are twelve-member jurors. While not significant, there is a tendency for six-member minority jurors to participate more than twelve-member minority jurors. From a small group communication viewpoint, the six-member jury may be superior to the larger group, as the smaller size may encourage greater overall juror participation. Although not conclusive, these findings are certainly relevant to an examination of the present trend toward the use of smaller juries.

—Joan B. Kessler

*Mrs. Kessler is a doctoral candidate in Speech Communication and Theater at the University of Michigan. This project is part of Mrs. Kessler’s doctoral dissertation, in progress, under the direction of Dr. William C. Donaghy. The project has been funded by research funds of the University of Michigan Law School.