

Small Group Research

<http://sgr.sagepub.com>

Do Juries Deliberate? A Study of Deliberation, Individual Difference, and Group Member Satisfaction at a Municipal Courthouse

John Gastil, Stephanie Burkhalter and Laura W. Black
Small Group Research 2007; 38; 337
DOI: 10.1177/1046496407301967

The online version of this article can be found at:
<http://sgr.sagepub.com/cgi/content/abstract/38/3/337>

Published by:

 SAGE Publications

<http://www.sagepublications.com>

Additional services and information for *Small Group Research* can be found at:

Email Alerts: <http://sgr.sagepub.com/cgi/alerts>

Subscriptions: <http://sgr.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations (this article cites 23 articles hosted on the SAGE Journals Online and HighWire Press platforms):
<http://sgr.sagepub.com/cgi/content/abstract/38/3/337#BIBL>

Do Juries Deliberate?

A Study of Deliberation, Individual Difference, and Group Member Satisfaction at a Municipal Courthouse

John Gastil

University of Washington

Stephanie Burkhalter

Humboldt State University

Laura W. Black

Ohio University

Despite long-standing interest in juries and a growing body of work on public deliberation, we have a limited understanding of how often everyday juries actually engage in meaningful deliberation. This study uses deliberative theory and small group research to develop a set of research questions and hypotheses regarding how juror characteristics promote deliberation as well as how deliberation influences juror satisfaction. Examination of 267 jurors' accounts of their experiences deliberating on municipal criminal juries suggests that juries do, indeed, deliberate at a remarkably high level of competence. Results show complex relationships between juror characteristics and their levels of deliberation as well as a direct link between the quality of deliberation and juror satisfaction.

Keywords: *decision making; jury deliberation; political knowledge; political efficacy; satisfaction*

A trial jury composed of average citizens is an American institution whose roots run deep in English common law (Dwyer, 2002; Hans & Vidmar, 1986). Although sometimes burdensome for the individual, jury duty is still prized in the contemporary United States as one of the main ways that citizens participate in democracy. Jury deliberation is a unique form of communication that takes place in a public institution and has a substantial impact on the lives of others, yet the communication itself is not observable by the public.

Despite its cultural significance and tremendous impact on society, relatively few studies have examined the communicative aspects of deliberation

among real jurors serving on empanelled juries. In large part this is because of restrictions on access to juries. In all federal courts and in almost all states and municipalities, field observation of decision making in the jury room is illegal. The secrecy of jury deliberation is so important to jurists that most judges refuse to grant social scientists systematic access to jurors.

Most of the information gleaned thus far about what goes on in real juries has been gathered through postverdict interviews with jurors, self-published accounts by jurors (and these typically are highly variable, biased, and from high-profile cases or studies with small samples that permit exploratory but not generalizable research, e.g., Pettus, 1990). The vast majority of the research on jury communication relies on mock juries, which may or may not provide an accurate representation of how juries actually deliberate (e.g., Burnett & Badzinski, 2000).

The importance, then, of interviewing actual trial jurors and seeing jury deliberation "from the jurors' point of view" (Pettus, 1990, p. 96) cannot be overstated. Do jurors themselves believe that they have deliberated? What individual- and group-level factors make actual jurors more or less likely to report a deliberative experience? Is the quality of their deliberative experience related to their satisfaction with the process and outcome of their labors?

In this study we aim to address those questions. Through the generosity of Seattle municipal court officials, we received unusually extensive access to a large number of real criminal jurors whose attitudes we were able to track over time. Pretrial surveys measured juror characteristics such as motivation to serve on a jury, political knowledge, level of education, political self-confidence, and partisanship. Posttrial surveys assessed the deliberative quality of juror discussions and their satisfaction with the deliberative process and its outcome.

Authors' Note: An earlier version of this article was presented on July 28, 2006, at the first annual conference of the Interdisciplinary Network for Group Research, Pittsburgh, PA. This research is supported by the National Science Foundation's Directorate for Social, Behavioral, and Economic Sciences: Law and Social Science Program (Award No. 0318513). Assistance with the design of the study was offered by Valerie Hans, Reid Hastie, Dick Madsen, Kay Schlozman, Phil Weiser, and Cindy Simmons. For assistance with collecting and processing the data reported herein, we wish to thank Jay Leichter, Giorgia Aiello, Ray Calluori, Perry Deess, Tina Gall, Irina Gendelman, Hillary Gillings, Sophia Liu, and Rachel Nez. For editorial assistance, thanks go to Jamie Moshin. At the Municipal Court of Seattle, we received helpful cooperation and oversight from court services manager Kathleen Friedman and Judge Fred Bonner, assistance collecting survey data from jury administrators Kendra Lafayette and Juliet Dickens, and court records retrieval assistance from Sharon Petty. For additional details on this and related studies, see the Jury and Democracy Project at www.jurydemocracy.org.

Before discussing our survey methods and findings, however, we begin with an exposition of the specific research questions and hypotheses this study addresses. In reviewing previous work on jury deliberation, we ask three general questions: What is the evidence that juries engage in thorough and respectful deliberation, along the lines imagined by deliberative theory? What are the factors that facilitate such deliberation? Also, does participating in meaningful jury deliberation translate into overall satisfaction for jurors?

Do Juries Deliberate?

What counts as competent deliberation in the jury room? Studies of actual and simulated juries have observed that most jury decision making is verdict driven, meaning that most jurors enter the jury room with a preliminary decision on the verdict with the goal of quickly making a decision (Hastie, Penrod, & Pennington, 1983; Sandys & Dillehay, 1995). It has been estimated that the ideal of full consideration of facts and evidence with an open mind and respect for others' positions may occur in only 35% of cases and that only 1 in 10 cases results in a reversal of the initial majority's verdict preferences after discussion in the jury room (Hastie et al., 1983; Sandys & Dillehay, 1995). The role of deliberation in the jury room is not just to help the group come to a decision but also to invest the ultimate decision with legitimacy. As Hans and Vidmar (1986) argue in their examination of the jury system, "even if its impact on the ultimate verdict is modest, deliberation helps to assure the integrity of jury decision making" (p. 112).

Defining Deliberation

A precise communicative ideal of jury deliberation on which we could base hypotheses and measures has not been developed; however, Burkhalter, Gastil, and Kelshaw's (2002) general conception of democratic deliberation in small face-to-face groups can be applied to the jury context. Following Burkhalter et al., people deliberate when they rigorously examine a problem and arrive at a well-reasoned solution after a period of inclusive, respectful consideration of diverse points of view. This broader definition provides guidelines for what could be considered deliberative communication in the jury. This approach also underscores the connection between the conventional practice of deliberation in juries and the larger

literature on public deliberation, which has often overlooked the jury (e.g., Chambers, 2003) or at least doubted its deliberative credentials (e.g., Sanders, 1997).

In the context of a jury trial, we can say that a group has participated in deliberative discussion if its analysis of the case was rigorous, jurors weighed the evidence carefully, and they discussed the instructions given by the judge (these being analogous to a predefined set of evaluative criteria). Burkhalter et al. (2002) also stress the social dimension of deliberation, which entails adequate opportunities for each juror to speak and mutual respect. Mutual respect, in particular, is a key component of successful deliberation, and feeling respected by other group members is especially important during disagreement (Benhabib, 1992; Gutmann & Thompson, 1996; Pearce & Littlejohn, 1997).

This adaptation of Burkhalter et al.'s (2002) general definition to the context of jury service fits well with the conventional conception of deliberation advocated by the judges themselves. The instructions given to juries vary from one case to the next, but the model instructions for the circuit court in which Seattle operates is representative of the words judges give to jurors. These match up well with the conceptual definition advanced above. For instance, the "Duty to Deliberate" section of the instructions explains to jurors, "Each of you must decide the case for yourself, but you should do so only after you have considered all the evidence, discussed it fully with the other jurors, and listened to the views of your fellow jurors." Judges are encouraged to add, "Do not be afraid to change your opinion if the discussion persuades you that you should. But do not come to a decision simply because other jurors think it is right" (Office of Circuit Executive, 2003). Thus, the aim is rigorous analysis, full discussion, and listening with respect to other views.

Weighing the Evidence for and Against the Quality of Jury Deliberation

Prior research presents mixed findings regarding how deliberative jury decision making works in actual practice. On one hand, there is evidence that many juries do not engage in high-quality deliberation. As indirect evidence, many studies have shown that jurors' initial verdict preferences are a strong predictor of the jury's final verdict. In fact, about 90% of the time, the jury's final verdict is the same as the majority of jurors' predeliberation positions. This result has been found in field studies of actual juries (Kalven & Zeisel, 1966; Sandys & Dillehay, 1995) as well as mock juries (MacCoun & Kerr,

1988). This finding does not prove that juries failed to engage in rigorous face-to-face deliberation; rather, it points to the power the initial majority has over the final decision, which can be problematic for aspects of deliberation, such as thorough case analysis and careful consideration of each juror's point of view (Burkhalter et al., 2002). If the final outcomes are consistent with initial opinions, there is likely little motivation for jurors to thoroughly evaluate the facts of the case and interpretations of the evidence or the judge's instructions that differ from their own. The general concern about jury deliberation has led some legal researchers to create guidebooks to advise jurors on how to deliberate effectively (e.g., Boatright & Murphy, 1999).

On the other hand, some studies indicate that many juries do engage in thorough, respectful, and egalitarian deliberation. Rejecting the common misperception that most juries start their deliberations with a preliminary vote, Diamond and her colleagues found that preliminary votes are often suggested but typically not employed, as juries instead begin discussing evidence (Diamond, Vidmar, Rose, Ellis, & Murphy, 2003a, 2003b). Additional research using posttrial interviews (Sandys & Dillehay, 1995) and mock trials (Hastie et al., 1983) found that only 10% to 27% of the juries hold early votes. These studies do not definitively show that jurors are engaging in respectful and egalitarian deliberation. However, the absence of early votes implies that jurors may be able to consider evidence more fully than if they had voted. As Kalven and Zeisel (1966) note, early votes very often determine the trial verdict.

From another perspective, it is remarkable that 10% of trials do result in a verdict that is the opposite of the jury's initial preferences (Devine, Clayton, Dunford, Seying, & Price, 2001). One potential explanation for why some juries reverse their verdict is the deliberation style. Research has shown a major difference between juries that use a verdict-driven style of decision making and those that use an evidence-driven style (Hastie, Schkade, & Payne, 1998). Verdict-driven juries tend to start with an initial vote and then spend the bulk of their time discussing verdict options. Evidence-driven juries, on the other hand, structure their discussion around the various pieces of evidence and do not vote until a great deal of discussion has taken place. An evidence-driven style is likely to be more deliberative by promoting thorough discussion of the case and may also allow minority voices more equal time and respect in the discussion than a verdict-driven style.

Inconsistent findings regarding the quality of jury deliberation make it difficult to make a prediction. Moreover, we have chosen to work with a specific conception of deliberation (Burkhalter et al., 2002) that reaches

beyond narrower conceptions of jury deliberation employed in past research. Finally, our survey research approach orients us to jurors' experience of deliberation, a vantage point distinct from the external assessments conventional in the literature. For all of these reasons, we begin with Research Question 1: Do juries tend to perceive their jury room discussions as thorough, respectful, and/or egalitarian?

What Promotes (and Obstructs) Jury Deliberation?

In their self-reinforcing model of deliberation, Burkhalter et al. (2002) were able to draw on the larger literature on small group communication and political participation to theoretically justify a set of four general predictions about which individual characteristics promote deliberation. Adapting the general claims of Burkhalter et al. (2002), we briefly justify each of these four hypotheses in relation to jury deliberation, generally, and in the context of the particular data we have available for this study.

Facilitators of Deliberation

First, following Burkhalter et al. (2002), we predict that (Hypothesis 1a) juries will report more deliberative experiences when they have a favorable disposition toward the jury system. This adapts Burkhalter et al.'s general contention that people are more likely to deliberate when they perceive the activity as appropriate, following the basic tenets of structuration theory regarding habituation to normal social practices (Giddens, 1984). In this context, those jurors who believe in the general efficacy and value of juries are more likely to promote and safeguard the deliberative tradition of jury discourse.

Second, Burkhalter et al. (2002) expect more deliberation when there is a perception of potential common ground among the deliberators. In the context of the present study, this is approached in general terms given the considerable differences in the facts of each trial. We predict that (Hypothesis 1b) juries will report more deliberative experiences when their members have relatively similar ideological self-identifications (i.e., liberal vs. conservative). The jury is, after all, a political body, and its deliberations and judgments have been theorized as political expressions (Amar, 1995; Iontcheva, 2003). Moreover, long-standing partisan debates about the wisdom of civil jury verdicts and the leniency of criminal juries make clear the conventional framing of juries as political bodies (Hans & Vidmar, 1986).

Thus, it is likely that jurors could construe partisan diversity on their jury as a threat to the common ground they believe necessary for effective deliberation.

Third, Burkhalter et al. (2002) reason that deliberation is more likely to occur when participants possess greater cognitive abilities and discussion skills. In the jury context, we translate this into two background variables available in our data and predict that (Hypothesis 1c) jurors will report more deliberative experiences when they have higher levels of political knowledge and formal education. The claim here is not that juries lacking formal education or political awareness cannot have honest, thoughtful discussions; rather, it is that education and knowledge help jurors work through complex legal issues and dilemmas in the same way they help citizens make sense of complex political problems and disputes (Delli Carpini & Keeter, 1996).

Finally, Burkhalter et al. (2002) hypothesize that "participants are likely to deliberate only if they are sufficiently motivated to process the content of the arguments they hear, rather than simply focusing on the peripheral cues of messages" (p. 417). One would hope that all juries can meet this standard, but nonetheless, we expect that (Hypothesis 1d) jurors will report more deliberative experiences when they have higher levels of motivation. In this study, we operationalized motivation in three different ways: political self-confidence (i.e., political efficacy), willingness to serve on a jury, and interest in the case being tried. This hypothesis can be viewed as a specific articulation of the general finding that the systematic processing of potentially persuasive messages requires the motivation to do so (Petty & Cacioppo, 1986).

Unequal Distributions of Talent and Inspiration

Beyond these hypotheses, we also examine the impact of uneven distributions of these same individual characteristics in juries. It is possible that jurors who are highly educated, motivated, and favorable toward the jury system are catalysts that promote deliberation in their group regardless of the characteristics of other jurors. Sunwolf and Seibold (1998) found that jurors gain access to rules and resources about communication in many aspects of their lives, and they employ these rules and resources in jury decision making. Therefore, it seems likely that jurors who are motivated to deliberate and have higher levels of education may know of and draw on communication resources that can help them facilitate deliberative conversations in the jury room. If there is a disparity in education, confidence, and political knowledge in the group, those jurors who have higher levels of

these factors may take on leadership roles and craft the decision making to promote thoroughness, equality, and respect for all jurors.

It is also possible, however, that those same exceptional participants could actually detract from the group's deliberation. A common theme in critiques of deliberation is a concern with the detrimental effects of inequalities among group members (Benhabib, 1992; Chambers, 2003; Sanders, 1997). Given that group members' evaluations of group behavior take into account their perceptions of other members' behavior (e.g., Flanagan, Park, & Seibold, 2004; Gouran, 1973;), it is reasonable to suspect that unequal knowledge, skill, or motivation could indirectly undermine group members' assessments of jury deliberation.

For instance, it is possible that a large discrepancy in political self-confidence could undermine deliberation because jurors who feel less confident may not take or feel that they have adequate opportunities to speak. Likewise, in groups with a disparity in education or political knowledge, those with lower levels of these factors may not feel respected in the group. If jurors hold back because of feeling overshadowed by those who are more knowledgeable, confident, and educated than they are, then deliberations may be less thorough.

Thus, rather than a clear hypothesis, we are left with two rival views—one stressing the value of knowledge, skill, and motivation, whatever its distribution, and the other suspecting that unequal distribution can detract from deliberation. This leads us to investigate this issue through a second research question. Research Question 2 asks, Is the inequality of the distribution of deliberation-promoting characteristics in a jury associated with lower levels of deliberation?

Does Deliberation Promote Juror Satisfaction?

Finally, juries know that deliberation is their primary task, but do juries really find deliberation satisfying? Satisfaction is a ubiquitous variable in small group research (Keyton, 1991) and a classic concern of research on democratic versus autocratic group behavior (Lewin, Lippitt, & White, 1961). In the context of the jury, we expect that jurors' deliberation ratings will predict their satisfaction with their jury experience and the verdict reached.

Given the enduring acceptance in American civic culture of what Bormann (1996) calls the "public discussion model," jurors are more likely to be satisfied with the jury experience if they feel it is respectful, egalitarian, and thorough. Because their job is to deliberate, jurors are likely to be satisfied with a discussion in which they believe that they have appropriately

addressed the task at hand. Past group communication research has found that both the quality of information acquired by group members and the equity of group member participation were important predictors of member satisfaction (Flanagin et al., 2004). Although Flanagin's research investigates these questions in the context of computer-mediated groups, it makes sense that the attributes of group processes such as information quality and contribution equity would also influence jurors' satisfaction with their experience. More directly, Oetzel (2001) found that group members who feel that they were listened to, were respected, and had adequate opportunity to speak were more satisfied with their groups. Thus, Hypothesis 2 predicts that overall satisfaction with jury deliberation and verdict will be associated positively with (a) one's own active participation in the discussion and (b) the perception that one's jury deliberated.

Method

Participants

We collected complete jury service records and survey data for 267 Seattle residents who served on juries that deliberated at the Seattle municipal courthouse between March 9, 2004, and July 20, 2004. Fifty-two percent of jurors were female, and 84% were White, 8.2% Asian American, 3.4% African American, 1.9% Hispanic, and 1.7% Native American. The median educational level was a college degree, and the median age was 49.

These 267 jurors came from 60 different criminal trials in Seattle municipal court, only 3 of which failed to return complete verdicts. Typical jury size in this court is 6 jurors, though there were cases with 5 or 7. Nearly a third (32.2%) of those participating in this study sat on juries hearing low-level assault charges, another 16.1% heard drunk-driving cases, and the rest heard a range of minor offenses from sexual indiscretion to reckless driving. The median juror spent 2 days in the courtroom, with 91% spending 3 or fewer days there. The median juror deliberated for no more than 1 hour, with 84% deliberating for 2 hours or less.

Survey Administration

Data were collected in two waves—one immediately after orientation to jury service (but prior to being assigned to a jury) and a second after the completion of jury service. During the Wave 1 survey, researchers attempted to approach every person who signed in for jury duty to request his or her

voluntary participation in a study on community life. Seventy-nine percent of those appearing for jury service completed the Wave 1 survey.

In 50% of the trials studied, the Seattle municipal court administrators presented jurors with Wave 2 (postservice) questionnaires, and for all other trials, jurors were recontacted 2 to 7 days later and invited to complete Wave 2 on paper or online. A repeated-contact design (Dillman, 1999) was used to send postcards, up to two survey booklets, reminders, and thank-you cards to potential respondents. Response rate on the Wave 2 Survey (combining both methods of reaching jurors) was 87%, and we received three or more complete surveys in 55 of the 60 trials.

Deliberation Measures

The questions in the Wave 2 survey were designed to capture essential components of small group face-to-face deliberation presented in Burkhalter et al.'s (2002) model. In the context of a jury trial, a group is deliberative to the extent that (a) its analysis of the evidence and relevant law is thorough (i.e., it carefully weighs evidence and uses the evaluative criteria given in the judge's instructions) and (b) individual jurors respect one another and provide each other with adequate opportunities to express their views.

To measure the careful weighing and use of evaluative criteria requisite in deliberation, jurors were asked to indicate their level of agreement on a 5-point Likert-type scale (ranging from *strongly disagree* to *strongly agree*) with the following statements: "Jurors thoroughly discussed the relevant facts of the case" (scale name = Fact), and "The jury thoroughly discussed the instructions the judge provided" (Instruct). The level of respect toward jurors in the jury room was measured by the same scale in response to the statement "All of the jurors listened respectfully to each other during deliberation" (Listen). In addition, a single item asked, "How were you treated by fellow jurors?" (Treatment); responses were coded *less than satisfactory* (1), *satisfactory* (2), *good* (3), *very good* (4), and *excellent* (5). The adequacy of speaking opportunities was measured by response to the question on a 5-point Likert-type scale similar to the one described above: "The other jurors gave me enough of a chance to express my opinions about the case" (Express).

Satisfaction Measures

Two Wave 2 items assessed jurors' overall satisfaction with the jury's deliberation and its ultimate verdict. Jurors were asked how they would rate

“their satisfaction with the jury’s final verdict” and “the quality of the jury’s deliberations.” Responses were recorded on a 4-point scale from 1 (*very unsatisfied*) to 4 (*very satisfied*). Responses to the two questions were moderately correlated ($r = .55$) and produced a reliable two-item scale ($M = 3.42$, $SD = 0.67$, $\alpha = .63$).

Antecedents to Deliberation

The Wave 1 survey included items measuring jurors’ preservice attitudes and background characteristics. Item wordings and descriptive statistics are provided for each of these in the order they appear in the hypotheses.

Trust in the jury system. Three items in Wave 1 measured jurors’ trust in the jury system. Respondents were asked to rate their “confidence in the jury system” on a scale from 1 (*very low*) to 5 (*very high*). They also used a 5-point agreement scale to respond to two statements: “The criminal jury system is the fairest way to determine guilt or innocence of a person accused of a crime,” and “The civil jury system is a good way to settle many civil lawsuits.” These three items yielded a reliable scale, TrustJury, with higher scores indicating greater trust in the jury system as a means of administering justice ($M = 3.84$, $SD = 0.56$, $\alpha = .71$).

Partisanship. Surveys also included conventional National Election Study measures of party membership (“Generally speaking, do you think of yourself as a Republican, a Democrat, an independent, or what?”) and the strength of affiliation (“How strongly do you support that party?”). Responses were combined to create a 7-point partisanship scale ranging from 1 (*very strong Democrat*) to 7 (*very strong GOP*), with the midpoint reserved for the 26 self-identified independents and 6 third-party members in the sample.

Liberalism–conservatism was also measured using a conventional scale (“When it comes to politics, which of the following do you think of yourself as?”). Other than the 11 *don’t know* and 4 *decline to answer* responses, this scale arrayed respondents from 1 (*extremely liberal*) to 7 (*extremely conservative*).

The two items were highly correlated ($r = .67$) and were combined to form a 7-point Partisanship scale, with higher scores indicating a more conservative or strong Republican self-identification ($M = 2.80$, $SD = 0.125$, $\alpha = .80$). Because variance in this variable is particularly important, it is worth noting that Partisan scale scores had a bell-shaped distribution that was skewed toward the lower end of the scale (mode = 2.0, skew = 1.17, kurtosis = 1.25).

Political knowledge. Following the conventional procedures for measuring political knowledge, the Wave 1 survey asked respondents to attempt to answer correctly five multiple-choice knowledge items (e.g., “Who is currently the King County Executive?”). A summary scale, PolKnowledge, was calculated as the number of correct answers to the five items listed ($M = 3.48$, $SD = 1.30$). Because of the small number of knowledge items relative to full-length scales, scale reliability was low ($\alpha = .54$).

Education. The first survey included a straightforward measure of educational level (“What is the highest level of formal education you have completed?”), with available responses corresponding to U.S. Census categories. The Education scale ranged from 0 (*no high school degree*) to 8 (*doctorate*), with the median response being a 4 (*BA or equivalent*), $M = 4.42$, $SD = 1.64$.

Political self-confidence. To measure political self-confidence, items were taken from a pool of conventional political science measures of internal political efficacy (Craig, Niemi, & Silver, 1990). Respondents used a 5-point agreement scale to respond to three items (e.g., “I have a pretty good understanding of the important issues facing this country”). Together, the items produced a reliable scale, PolConfidence, with high scores indicating greater political self-confidence ($M = 3.82$, $SD = 0.74$, $\alpha = .86$).

Eagerness to serve on a jury. To assess participants’ interest in serving on a jury, the Wave 1 survey asked respondents to use the 5-point agreement scale to respond to this item: “I am hopeful that I will be required to serve on a jury today.” This item was labeled JuryHope ($M = 3.52$, $SD = 1.05$).

Interest in the trial. To assess interest in the trial, the Wave 2 survey asked respondents to use the 5-point agreement scale to respond to this item: “The trial was very interesting to think about.” This item was labeled TrialInterest ($M = 3.83$, $SD = 0.82$).

Level of Analysis and Data Aggregation

Small group researchers increasingly attend to concerns about identifying the appropriate level of analysis (Kenny, Manetti, Pierro, Livi, & Kashy, 2002). In this study, our hypotheses and research questions were cast at the group level, as they discuss the composition, behavior, and outcomes of whole juries. To assess the appropriateness of aggregation and group-level analysis of data measured at the individual level, intraclass correlations

(ICC) were computed for each variable. Positive correlations for all scales indicated the appropriateness of group-level analysis (e.g., for listen, $ICC = .310$, $p = .03$). Readers who might have preferred an alternative analytic approach may be reassured to know that the results shown below were approximately equivalent to those obtained by corresponding individual-level analysis.

Results

Research Question 1: Jury Deliberation

The first research question does not test relationships between variables; rather, it simply asks whether the juries studied herein appear to have deliberated based on jurors' rating of the essential components of small group face-to-face deliberation. We believe this is one of those moments in social science where simple descriptive, rather than inferential, statistics should get the chance to speak for themselves, and Tables 1 and 2 provide relevant summary data on all the measured facets of jury deliberation.

Table 1 shows group-level means and minimums for the 55 juries that had three or more completed Wave 2 surveys. On scales ranging from 1 to 5, average scores are skewed heavily toward the top end of the scale for Fact ($M = 4.51$), Instruct ($M = 4.35$), Listen ($M = 4.41$), Express ($M = 4.50$), and Treatment ($M = 4.40$). Translating these numbers back to scale labels, the first four scores fall between *agree* and *strongly agree*, and Treatment falls between *very good* and *excellent*.

The same table shows the lowest (minimum) score within each group to give a sense for the experience of the juror who experienced the least deliberation. The average minimum score across the 55 groups was closest to the fourth scale point in all cases, suggesting a "minimum experience" that was usually comparable to that of the rest of the jurors. Table 1 also shows the distribution of minimum experiences across the groups, and this analysis showed exactly how few juries included a member who gave his or her jury experience low marks. Only 2% of juries had a juror who reported a Fact score below 3 on the 5-point scale, and such low scores were equally rare for Instruct (5%), Listen (4%), Express (2%), and Treatment (7%).

Hypothesis 1: Predicting Deliberation

The overall answer to Research Question 1 is that jurors reported that their jury room discussions were deliberative. But there was nonetheless

Table 1
Group-Level Descriptive Statistics for Deliberation Measures in Municipal Criminal Juries

Measure	<i>M</i>	<i>SD</i>	Average Minimum Score	Percentage of Groups With Minimum Response Level		
				Agree or Strongly Agree	Neutral	Disagree or Strongly Disagree
Fact (Jurors thoroughly discussed the relevant facts of the case.)	4.52	0.33	4.04	89	9	2
Instruct (The jury thoroughly discussed the instructions the judge provided.)	4.35	0.42	3.73	67	27	5
Listen (All of the jurors listened respectfully to each other during deliberation.)	4.41	0.36	3.91	89	7	4
Express (The other jurors gave me enough of a chance to express my opinions about the case.)	4.50	0.31	4.02	95	4	2
				Very Good or Excellent	Good	Satisfactory or Less Than Satisfactory
Treatment (How were you treated by fellow jurors?)	4.40	0.46	3.68	71	24	7

Note: $N = 55$ groups for all rows. All variables are scored on scales ranging from 1 to 5.

variation in the degree to which a jury was deliberative, and Hypothesis 1 predicted that juries would report the highest levels of deliberation when their members had (a) a favorable disposition toward the jury system; (b) greater potential for common ground; (c) relatively more political knowledge and formal education; and (d) greater political self-confidence, willingness to serve on a jury, and interest in the case being tried.

Correlations tested each of these hypotheses, and Table 2 shows the results. There were no significant effects for the knowledge–skill measures (PolKnowledge, Education), but each other group had one or more significant associations. Scores on TrustJury had a positive relationship to Treatment

Table 2
Correlations of Antecedents With Mean Group
Scores on Deliberation Measures

Category and Measure(s)	Fact	Instruct	Listen	Express	Treatment
Appropriateness					
TrustJury	.10	.04	.03	-.03	.23**
Common ground					
PartisanSD ^a	.10	.27**	-.05	.04	-.03
Cognitive/Communication skill					
PolKnowledge	.14	-.05	-.09	.13	-.11
Education	.05	-.01	-.08	.04	-.05
Motivation					
PolConfidence	.23**	.17	-.12	.23**	.05
JuryHope	.07	.09	.10	-.03	.04
TrialInterest	.38***	.11	.04	.32***	.29**

Note: $N = 55$ for all rows except PartisanSD, where $N = 41$.

a. PartisanSD measures the variance in partisanship within juries. These are partial correlations that control for the mean level of Partisan in each group.

** $p < .05$. *** $p < .01$.

mean scores ($r = .23$). The variance in juror partisanship (PartisanSD, as in “standard deviation in partisanship”) was associated with Instruct scores ($r = .27$). As for the motivation measures, TrialInterest had moderate correlations with mean scores for Fact ($r = .38$), Express ($r = .32$), and Treatment ($r = .29$). PolConfidence also had significant effects for Fact ($r = .23$) and Express ($r = .23$). In sum, these data showed many—but not all—of the linkages predicted in Hypothesis 1. Exceptions were the absence of effects flowing from PolKnowledge and Education and toward Listen.

Research Question 2: Antecedent Distribution Effects

The latter finding logically leads to the question of whether the hypothesized antecedents can have a backlash effect when unevenly distributed. Research Question 2 asks whether high scores on the very same measures that we posit as individual characteristics likely to promote deliberation, such as PolKnowledge, can actually detract from a jury’s average deliberation scores when paired up with persons scoring low on those same factors. Addressing this question required running partial correlations between group-level standard deviations in antecedents and group-level means in deliberation measures, controlling for both the number of completed surveys in each jury and the mean score on the corresponding antecedent variable. For

Table 3
Partial Correlations of Within-Group Variance of Antecedents
with Mean Deliberation Scores

Category and Measure(s)	Fact	Instruct	Listen	Express	Treatment
Appropriateness					
TrustJurySD	.07	.09	.19*	.13	.12
Cognitive/Communication skill					
PolKnowledgeSD	-.21**	-.31**	-.21*	-.23*	-.38***
EducationSD	.06	.01	.30**	.10	.14
Motivation					
PolConfidenceSD	-.38***	-.36***	-.05	-.26**	-.14
JuryHopeSD	.11	.11	-.10	.09	.09
TrialInterestSD	.22*	.15	.07	.17	-.06

Note: These are partial correlations that control for the number of surveys in each group and the mean score of the corresponding antecedent variable (e.g., the correlations for TrustJurySD control for the average group score on TrustJury). $N = 52$ for all rows except JuryHopeSD, where $N = 51$.

* $p < .10$. ** $p < .05$. *** $p < .01$.

example, a predictor in this instance is the within-group variation in political knowledge, which we label PolKnowledgeSD.

Using this approach, Table 3 shows that the results of the tests are two clear patterns of near-significant and significant effects: (a) PolKnowledgeSD reduces scores on the deliberation measures—Fact, Instruct, Listen, Express, and Treatment; and (b) PolConfidenceSD reduces scores on some deliberation measures (Fact, Instruct, Expression). In other words, the more political knowledge and self-confidence variation there is within a jury, the lower the jurors rate their deliberation. Moreover, the nature of these partial correlations demonstrates that these associations are not because of the difference in average antecedent score between low- and high-variance juries.

Hypothesis 3: Satisfaction

The final hypothesis predicted that jurors would be more satisfied with the jury deliberations and verdict when they perceived the jury as deliberative. Satisfaction with deliberation was correlated with Fact ($r = .25$), Listen ($r = .31$), and Express ($r = .26$), and it had a near-significant correlation with Treatment ($r = .22$, $p = .058$). Verdict satisfaction was correlated with Listen ($r = .29$) and Treatment ($r = .28$). In sum, all but the Instruct deliberation measure correlated substantially with jurors' satisfaction with their jury's deliberation and/or verdict.

Conclusion

Juries and Public Deliberation

That juries really do deliberate is now clearer, not only because of the use of actual juries in this study but also owing to the more detailed conceptualization of what deliberative jury discussion entails. The results provide clear evidence that jurors do understand their experience as being a largely *deliberative* one in the strongest sense of that word. When broken down into the components described in Burkhalter et al.'s (2002) self-reinforcing model, almost all of the juries studied cleared a high bar. For instance, looking at only those jurors completing the survey, fewer than 10% of juries had a single member who disagreed with any of a range of statements descriptive of deliberative rigor, such as listening, equality, and respect. Even though participation rates varied considerably, a finding that has previously raised red flags for critics of deliberation (e.g., Sanders, 1997), fully 98% of participants maintained that they had adequate opportunities to express their views of the case.

Though we found it useful as a theoretical guide, this study provides mixed evidence in support of Burkhalter et al.'s (2002) self-reinforcing model of deliberation. The significant association between discrete deliberative measures and jurors' overall satisfaction with deliberation suggests that the theoretical conception of deliberation advocated by Burkhalter et al. taps into lay jurors' sense of what juries are supposed to be doing.

Whereas each of the deliberation measures was associated with deliberation satisfaction, it is noteworthy that the only measures associated with verdict satisfaction were jurors' sense of how they were treated by one another and how well jurors listened to one another. This finding foregrounds the importance of listening and respect in deliberative theory, as these aspects of deliberation translated into jurors' satisfaction not only with the process but also with their ultimate decision. Deliberative theory is often concerned with bolstering the legitimacy of public institutions (Habermas, 1975), and this result suggests that it may be the relational—rather than the rational—element of deliberation that most readily leads participants to view deliberative decisions as legitimate.

In addition, support for the Burkhalter et al. (2002) model comes from finding some of the expected effects of four antecedents on deliberation: To the extent that jurors had faith in the jury process, they were more likely to experience mutual respect; to the extent that jurors were self-confident and interested in the trial, they were more likely to engage in careful analysis, perceive adequate speaking opportunities, and experience mutual respect.

At the same time, some of the deliberative antecedents hypothesized in Burkhalter et al. (2002) were found to have negative effects on deliberation when they were unevenly distributed within a jury. Even after controlling for differences in average group levels of political knowledge and self-confidence, increasing inequality in the distribution of these individual traits was consistently negatively correlated with average group scores on the deliberation measures. Knowledge, self-confidence, and civic engagement go hand in hand (Verba, Schlozman, & Brady, 1995), with the strongest speaking the loudest and with the most force. Moreover, citizens readily pick up on social cues that help them identify those around them who have real political acumen (Huckfeldt, 2001). Within a jury, the net result of having significantly unequal knowledge and confidence levels may be a shared perception that the experts will (and do) run the show. This finding suggests that a rising tide of political skill does not lift all boats, at least when it comes to jury deliberation. The potentially disruptive role of experts among novices is one reason deliberation practitioners routinely stress the role of effective facilitators in issue-focused deliberative forums, such as citizen juries (Crosby & Nethercutt, 2005).

Reflecting on Anomalies

Keeping in mind the limits of statistical power, some nonfindings were noteworthy. The ideological diversity of juries was unrelated to deliberative experience, except that the more diverse juries were actually more likely to scrutinize the judge's instructions. We take this to indicate that the political diversity on these juries was constructive, likely prompting jurors to consider different interpretations of the law. This suggests that ideological diversity, when present in an appropriately structured setting, such as a jury room, stimulates discussion but does not preclude the perception that common ground can be found. Such results should be reassuring to deliberative theorists who emphasize the need for diversity in groups to facilitate exposure to new ideas and opinions (Barabas, 2004).

Another finding contrary to hypotheses was the absence of a significant correlation between expertise and the various measures of deliberation. Higher levels of expertise, as measured by higher levels of political knowledge and education, were associated with higher juror participation in discussion but not deliberation *per se*. This finding should be encouraging to practitioners developing deliberation programs in communities with little formal education or political experience.

Limitations and Future Directions

Our findings are limited to some extent by the study design and data we were able to collect. Because court officials restricted the questions we were able to ask jurors, a few of our measures are based on a small number of general questionnaire items. Similarly, like most jury studies, we were not able to have access to the jurors' closed deliberations. Rather, our data are drawn from jurors' questionnaire responses. It is possible that jurors' self-reports reflect a positive-response bias, thus providing a more positive image of their deliberative process than actual observations of their deliberation would have (Diamond et al., 2003b).

Self-reported experiences are, in and of themselves, important to measure. In the case of public deliberation, it is particularly important how citizens recall—and choose to describe—their experiences in public service. After all, one of the important aims of public deliberation is to ensure the legitimacy of government institutions (Chambers, 2003), and in the case of closed-door juries, this legitimacy flows from citizens' perceptions of deliberative quality.

Nonetheless, external observation is important for improving our understanding of the correspondence between recalled and directly observed deliberative behaviors, which do not always perfectly correspond. If researchers are able to gain access to videotaped jury deliberations, future research should investigate the extent to which our findings hold up with observation data.

In addition, the findings in this study should be qualified by the fact that these data refer to a specific size and level of trial—smaller (5- to 7-person) municipal criminal trials. Such trials make up a large proportion of jury trials, very few of which grab headlines with brutal murders or high-stakes litigation. If another jurisdiction will permit researchers to examine 12-person juries and higher stakes trials, it will be useful to compare those results with the ones found here. To the extent that group size remains a predictable influence on group member performance and satisfaction, it is likely that the larger groups will have more difficulty deliberating than did these relatively compact juries.

Implications for Juries and Civic Engagement

Even with these limitations, this study contributes to our understanding of how juries deliberate, and it has important implications for a new line of research investigating the civic impact of jury service. Legal scholar Vikram

Amar argues that jury service should be thought of as a political act akin to voting and not simply as a legal-administrative tool. "After all," he writes, "jurors *vote* to decide the winners and losers in cases" (Amar, 1995, p. 205). Not only are juries part of public-political life, the experience of deliberating on a jury may be a vital civic educational experience that inspires many Americans to heighten their sense of civic commitment and do things such as vote, join local boards, and so on. Alexis de Tocqueville (1835/1961) advocated precisely this view in his famous treatise, *Democracy in America*.

New research suggests that jury service does have this kind of civic effect on participants. Gastil, Deess, Weiser, and Larner (2006) conducted a national study of jurors in counties across the United States and found that jurors who deliberated, compared to those who were seated in the jury box but never got to deliberate (e.g., because of a mistrial, etc.), increased their rate of participation in subsequent elections. Using a large data set with thousands of jurors, it was possible to determine that the critical distinction was between those who deliberated and those who did not. The data set was also large enough to permit breaking down participants into two subgroups, and this analysis found that the increased voting effects were apparent only for previously infrequent voters (voting less than 50% of the time) who served on criminal trials. In contrast, jurors characterized as frequent voters and who served on civil juries did not have a significant increase in voting after jury service.

The present study aids in the interpretation of the findings presented in Gastil et al. (2006). In particular, this study found that the vernacular category of jury deliberation contains within it experiences that fit squarely into the more rarefied conception of public deliberation advocated by Burkhalter et al. (2002). Future research may find that civil juries (rather than the criminal juries used here) are less deliberative, which would help explain the absence of a voting effect for civil jurors in Gastil et al.

Our results also suggest that one does not need political knowledge or high levels of formal education to experience effective deliberation. Although statistical power limitations make it impossible to rule out small effect sizes, our data do clearly show that those with less knowledge and experience (such as the infrequent voters in Gastil et al., 2006) can have a deliberative experience on juries. In other words, deliberation is an experience equally available to both political novices and experts. This is important because, as Gastil et al. (2006) found, novices appear to have the most to gain from deliberation. If, in the end, the experience of jury deliberation is a civic educational experience as much as an administrative tool, it is comforting to know that the deliberative elixir works at least as well for those with the most to gain by its consumption.

References

- Amar, V. D. (1995). Jury service as political participation akin to voting. *Cornell Law Review*, 80, 203-259.
- Barabas, J. (2004). How deliberation affects policy opinions. *American Political Science Review*, 98, 687-701.
- Benhabib, S. (1992). *Situating the self: Gender, community, and postmodernism in contemporary ethics*. New York: Routledge.
- Boatright, R. G., & Murphy, B. (1999). Behind closed doors: Assisting jurors with their deliberations. *Judicature*, 83, 52-58.
- Bormann, E. G. (1996). Symbolic convergence theory and communication in group decision making. In R. Y. Hirokawa & M. S. Poole (Eds.), *Communication and group decision-making* (2nd ed., pp. 81-113). Thousand Oaks, CA: Sage.
- Burkhalter, S., Gastil, J., & Kelshaw, T. (2002). A conceptual definition and theoretical model of public deliberation in small face-to-face groups. *Communication Theory*, 12, 398-422.
- Burnett, A., & Badzinski, D. M. (2000). An exploratory study of argument in the jury decision-making process. *Communication Quarterly*, 48, 380-396.
- Chambers, S. (2003). Deliberative democratic theory. *Annual Review of Political Science*, 6, 307-326.
- Craig, S. C., Niemi, R. G., & Silver, G. E. (1990). Political efficacy and trust: A report on the NES pilot study items. *Political Behavior*, 12, 289-314.
- Crosby, N., & Nethercutt, D. (2005). Citizens juries: Creating a trustworthy voice of the people. In J. Gastil & P. Levine (Eds.), *The deliberative democracy handbook* (pp. 111-119). San Francisco: Jossey-Bass.
- Delli Carpini, M. X., & Keeter, S. (1996). *What Americans know and why it matters*. New Haven, CT: Yale University Press.
- Devine, D. D., Clayton, L. D., Dunford, B. B., Seying, R., & Price, J. (2001). Jury decision making: 45 years of empirical research on deliberating groups. *Psychology, Public Policy and Law*, 7, 622-727.
- Diamond, S. S., Vidmar, N., Rose, M., Ellis, L., & Murphy, B. (2003a). Inside the jury room: Evaluating juror discussions during trial. *Judicature*, 87, 54-58.
- Diamond, S. S., Vidmar, N., Rose, M., Ellis, L., & Murphy, B. (2003b). Jury discussions during civil trials: Studying an Arizona innovation. *Arizona Law Review*, 45, 1-81.
- Dillman, D. A. (1999). *Mail and internet surveys: The tailored design method*. New York: Wiley.
- Dwyer, W. L. (2002). *In the hands of the people*. New York: St. Martin's.
- Flanagin, A. J., Park, H. S., & Seibold, D. R. (2004). Group performance and collaborative technology: A longitudinal and multilevel analysis of information quality, contribution equity, and members' satisfaction in computer-mediated groups. *Communication Monographs*, 71, 352-372.
- Gastil, J., Deess, E. P., Weiser, P., & Lerner, J. (2006). *Jury service and electoral participation: A strong test of the participation hypothesis*. Paper presented at the annual meeting of the International Communication Association, Dresden, Germany. Available from <http://www.jurydemocracy.org>
- Giddens, A. (1984). *The constitution of society*. Berkeley: University of California Press.
- Gouran, D. (1973). Correlates of member satisfaction in group decision-making discussions. *Central States Speech Journal*, 24, 91-96.

- Gutmann, A., & Thompson, D. (1996). *Democracy and disagreement*. Cambridge, MA: Harvard University Press.
- Habermas, J. (1975). *Legitimation crisis*. Boston: Beacon.
- Hans, V. P., & Vidmar, N. (1986). *Judging the jury*. New York: Plenum.
- Hastie, R., Penrod, S. D., & Pennington, N. (1983). *Inside the jury*. Cambridge, MA: Harvard University Press.
- Hastie, R., Schkade, D. A., & Payne, J. W. (1998). A study of juror and jury judgments in civil cases: Deciding liability for punitive damages. *Law and Human Behavior*, 22, 287-314.
- Huckfeldt, R. (2001). The social communication of political expertise. *American Journal of Political Science*, 45, 425-438.
- Iontcheva, J. (2003). Jury sentencing as democratic practice. *Virginia Law Review*, 89, 311-383.
- Kalven, H., & Zeisel, H. (1966). *The American jury*. Chicago: University of Chicago Press.
- Kenny, D. A., Manetti, L., Pierro, A., Livi, S., & Kashy, D. A. (2002). The statistical analysis of data from small groups. *Journal of Personality and Social Psychology*, 83, 126-137.
- Keyton, J. (1991). Evaluating individual group member satisfaction as a situational variable. *Small Group Research*, 22, 200-219.
- Lewin, K., Lippitt, R., & White, R. K. (1961). *Autocracy and democracy*. New York: Harper & Row.
- MacCoun, R. J., & Kerr, N. L. (1988). Asymmetric influence in mock jury deliberation: Jurors' bias for leniency. *Journal of Personality and Social Psychology*, 54, 21-33.
- Oetzel, J. G. (2001). Self-construals, communication processes, and group outcomes in homogeneous and heterogeneous groups. *Small Group Research*, 32, 19-54.
- Office of Circuit Executive. (2003). *Manual of model criminal jury instructions*. San Francisco: United States Courts for the Ninth Circuit. Retrieved October 9, 2006, from <http://www.ce9.uscourts.gov/web/sdocuments.nsf/crim>
- Pearce, W. B., & Littlejohn, S. (1997). *Moral conflict: When social worlds collide*. Thousand Oaks, CA: Sage.
- Pettus, A. B. (1990). The verdict is in: A study of jury decision making factors, moment of personal decision, and jury deliberations—from the juror's point of view. *Communication Quarterly*, 38, 83-97.
- Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 16, pp. 123-205). New York: Academic Press.
- Sanders, L. M. (1997). Against deliberation. *Political Theory*, 25, 347-376.
- Sandys, M., & Dillehay, R. C. (1995). First-ballot votes, pre-deliberation dispositions, and final verdicts in jury trials. *Law and Human Behavior*, 19, 175-195.
- Sunwolf & Seibold, D. R. (1998). Jurors' intuitive rules for deliberation: A structural approach to communication in jury decision making. *Communication Monographs*, 65, 282-307.
- Tocqueville, A. de. (1961). *Democracy in America*. New York: Schocken. (Original work published 1835)
- Verba, S., Scholzman, K. L., & Brady, H. E. (1995). *Voice and equality: Civic voluntarism in American politics*. Cambridge, MA: Harvard University Press.

John Gastil is a professor of communication at the University of Washington. He is coeditor, with Peter Levine, of *The Deliberative Democracy Handbook: Strategies for Effective Civic Engagement in the Twenty-First Century* and the author of *Democracy in Small Groups*, By

Popular Demand, and *Political Communication and Deliberation* (forthcoming in 2007). His research on juries is available at www.jurydemocracy.org.

Stephanie Burkhalter is an assistant professor of political science at Humboldt State University. Her doctoral dissertation assesses the prevalence of message discipline and deliberation in the U.S. Congress.

Laura W. Black is an assistant professor of communication studies at Ohio University. Her research interests include group deliberation and dialogue in public and organizational settings. Her current research examines how telling and responding to personal narratives during a disagreement influences deliberative group members' sense of collective or civic identity.