

Experimenting with Deliberative Democracy:
Effects on Policy Preferences and Social Choice

Cynthia Farrar, Yale University

James S. Fishkin, Stanford University

Donald P. Green, Yale University

Christian List, London School of Economics

Robert C. Luskin, University of Texas at Austin

Elizabeth Levy Paluck, Yale University

For presentation at the ECPR Conference, Marburg, Germany, September, 18-21, 2003

Abstract

Using data from the first fully randomized field experiment within a Deliberative Poll, we examine the effects of the Deliberative Poll's formal on-site deliberations on both policy preferences and "preference structuration" or single-peakedness (operationally, the proportion of individuals whose preferences are aligned along the same shared dimension). The issues were airport expansion and revenue sharing in New Haven, Connecticut and its surrounding towns. Half the participants deliberated revenue sharing, then the airport, the other half the reverse. This split-half design enables us to distinguish the effect of deliberation from those of other aspects of larger Deliberative Polling treatment. We find that deliberation significantly altered aggregate policy preferences and increased the degree of single-peakedness on revenue sharing, though not airport expansion. These results both confirm the promise of civic deliberation as a means of transforming citizen preferences and raise the question of how deliberation's effects may depend on the kind of issue being deliberated.

Recent years have witnessed the emergence of the new democratic institution of Deliberative Polling. Ordinary polls gauge opinions without regard for whether the respondents possess much information or have thought much about the topics in question, not to mention whether they have considered competing perspectives on them. In contrast, Deliberative Polls invite randomly selected participants to discuss the issues with one another and to question policy experts about them. Beginning in 1994, there have been twenty Deliberative Polls, eleven in the U.S. (two national and nine regional) and nine abroad, in Britain, Denmark, Australia, and Bulgaria (all national). Fueled by normative and empirical concerns about the quality of popular decision-making, a growing body of research is using the data from Deliberative Polls to gauge how deliberation shapes people's views (as in, for example, Fishkin 1997, Fishkin and Luskin 1999, Luskin, Fishkin, and Jowell 2002).¹

The study described here addresses two of the most central hypotheses on which Deliberative Polling data shed light. The first is that deliberation frequently alters policy preferences, both individually and in the aggregate. The second is that deliberation tends to increase *preference structuration*, meaning the systematic alignment of the preferences of all individuals along the same shared set of underlying dimensions.² We expand on this latter concept and its importance below.

Previous evidence from Deliberative Polling supports both hypotheses. The participants do frequently change their views, both individually and in the aggregate, and their preferences do generally become more structured. Both effects, moreover, seem related to learning. The participants typically learn a great deal (Luskin, Fishkin, and Jowell 2002; Luskin, Fishkin, Jowell, and Park 1999; Luskin, Fishkin, and Plane 1999; Fishkin and Luskin 1999; Luskin, Fishkin, MacAllister, Higley, and Ryan 2000), and those who learn the most tend to change their preferences the most (Luskin, Fishkin, and Jowell 2002, Luskin, Fishkin, Mcallister, Higley, and Ryan 2000) and show the greatest increase in preference structuration.³

But questions about the sources of these before-after changes remain. One question is the extent to which they actually result from the Deliberative Polling experience. The whole public, after all, could be changing at the same time, in the same ways, and to the same degree. That may be generally unlikely

but is at least possible for issues sufficiently in the headlines and on people's lips. Several previous Deliberative Polls have therefore compared the participants to "quasi control groups" consisting of either reinterviewed "nonparticipants" (members of the initial random sample who declined to participate) or an independent random sample interviewed at roughly the time the Deliberative Poll was ending. But while lending some considerable assurance that the before-after changes do indeed result from something in the Deliberative Polling experience (see, e.g., Luskin and Fishkin 1999, Luskin, Fishkin, MacAllister, Higley, and Ryan 2000) these comparisons still lack the full authority of comparisons to true control groups separated from the participants by random assignment.

A second, so far less explored question is the extent to which the before-after changes stem from deliberation, as distinct from other aspects of the Deliberative Polling experience. Previous Deliberative Polls have involved one, grand, undifferentiated "treatment," consisting of everything from the invitation to participate, to the provision of briefing materials laying out competing arguments, to a weekend's worth of formal, balanced deliberation, to the casual, self-started, anticipatory deliberation between the initial interview and the deliberative weekend, to the informal, conversational spillover in corridors and dining rooms during the deliberative weekend itself. The treatment thus involves more than just discussion, and the discussion is not all of a kind, nor confined to the weekend. Yet the heart of the intervention is the formal deliberation during the weekend, which constitutes the greatest departure from the participants' everyday experience and the closest approximation to what theorists of "deliberative democracy" have in mind. The specific contribution of this formal on-site deliberation has never before been estimated.

This study reports on the first Deliberative Poll enfolded a randomized experiment. The experiment is designed to address both of the lingering questions just posed—both to buttress the inference that it is the Deliberative Polling experience that produces the observed before-after policy attitude changes and preference structuration increases and to provide evidence that it is the deliberation within the Deliberative Polling experience that is producing them. Specifically, the participants are first randomly assigned to deliberate one of two distinct policy issues, then answer the same questions as when

first interviewed and recruited, then deliberate the other of the two issues, then again answer the same questions. On each issue, to the extent that deliberating it matters, the midterm questionnaire should reveal differences between the preferences and preference structuration among those who have just finished deliberating it and those who have not yet done so. The second half of the proceedings, during which the two treatment groups switch issues, should then re-close the gap, since everyone will at that point have deliberated both. After showing that one but only one of our two issues follows this pattern, we offer some thoughts about how deliberation's effects may vary with the issue and the context.

Deliberative Democracy and Deliberative Polling

Democracy, according to deliberative democrats, should not just aggregate preferences but help shape them. Votes and opinions should emerge from processes of discussion and reflection (Elster 1998). But the attainability of this ideal is unclear. The public may never achieve much more than the modest and imperfect deliberation that already occurs. Competing demands for time and attention, coupled with typically minuscule probabilities of actually affecting the outcome, may make ignorance too rational (Downs 1957). But while it may not be possible to get everyone to discuss and reflect seriously on policy issues, it *is* possible to get random samples numbering in the hundreds to do so. The resulting distributions of policy and electoral preferences provide glimpses of what the whole public would think if it deliberated to the same degree. That, in short, is the strategy of Deliberative Polling.

The basics of the design are these: A random sample is drawn, interviewed, and invited to attend a weekend of deliberations at a common site. Those agreeing to attend are sent carefully balanced briefing materials laying out the major arguments for and against the major policy proposals. On-site, they discuss the issues in randomly assigned small groups led by trained moderators and question balanced panels of competing experts or policy-makers in plenary sessions. Then, at the end, they answer the same questions as at the beginning. They receive a financial inducement for participating, and the deliberations are generally televised.

The intent is to approach a counterfactual ideal in which deliberation is not only more pervasive than in the real world but in several important senses “better”—more substantive, better informed, more balanced, more deeply reflective, and more inclusive in the sense of involving more socio-demographically and attitudinally diverse discussants. Thus the briefing materials are provided to jump start the participants’ learning and thinking about the issues. Thus the briefing materials and expert panels are carefully balanced, and the small group moderators strive to ensure that all the arguments in the briefing materials get considered. The participants read, hear, and voice arguments and counterarguments. Thus the moderators keep the discussions on-topic and civil. Thus the combination of random sampling and random assignment maximizes the heterogeneity of both the attitudes expressed and the people expressing them. These features make the formal on-site deliberations very different from naturally occurring discussion in the real world—but much closer to what those writing of “deliberative democracy” have in mind. The participants talk with people very unlike themselves, expressing views very unlike their own, and in circumstances in which it is difficult not to give them and their views serious attention. We suspect that these artificial—more ideal—features give the deliberation in Deliberative Polling much of its effect.

Deliberation, Preferences, and Preference Structuration

But let us say a bit more about the effects we are looking for, why we expect them, and why they matter. Our first hypothesis, again, is that deliberation frequently changes preferences, both individually and in the aggregate. The degree and direction of change will naturally depend on the nature of the issue and the circumstances of the day. (For some speculations see Luskin 2003.) One possible mechanism is that the participants come to draw truer, tauter connections between their preferences and their own more fundamental values and interests. Another is that they come to redefine their interests or re-weight their values. In particular, we suspect that they often gravitate toward thinking in terms of a wider public interest. To the extent that the first mechanism predominates, we should expect to see preferences change so as to increase their predictability from sociodemographic variables proxying interests; to the extent that

the second predominates, they should change so as to decrease it. In either case, we should expect change (Luskin, Fishkin, and Jowell 2002, Luskin 2003). Of course the individual-level changes may largely cancel out, with some participants moving one way, and others moving equally the other way. There could be much gross but little net change. But absent any reason to expect such balancing to be the rule—and if anything there is reason to suspect the contrary, given some correlation between interests and initial thought and information—we may expect net change to be quite common (as the evidence from previous Deliberative Polls, e.g. in Fishkin and Luskin 1999, Luskin, Fishkin, and Jowell 2002, confirms).

Our second hypothesis—that deliberation tends to increase preference structuration—may need more explanation.⁶ Following List, Fishkin, Luskin, and McLean (2000), we use “structuration” to mean the systematic alignment of the preferences of all individuals along the same shared set of underlying dimensions, also called the “issue-space.” An illustrative issue-space might be two-dimensional, with one dimension being left-right, and the other religious-secular or urban-rural. Here, however, we focus on the special case of one-dimensional preference structuration, with the issue-space consisting of a single left-right dimension. This special case, of particular interest in the field of social choice, is called *single-peakedness*. Informally, preferences are single-peaked if the alternatives can be aligned from left-most to right-most such that each individual has a most preferred position and gives alternatives decreasing preference as they get more distant from it. A simple measure of single-peakedness is m/n , the proportion of individuals whose preferences are single-peaked with respect to the same dimension (Niemi 1969, List et al. 2002). This refers not to the cognitive organization of individuals’ preferences but to the cohesion or aggregate patterning of preferences *across* individuals.

Deliberation may increase preference structuration by any or all of at least three mechanisms. As people talk, learn, and think about the relationships among the alternatives and the criteria for choosing among them, they may simply adopt an ordering they come to recognize as conventional among political elites. Or they may influence each other’s thinking, acquiring more of a shared understanding of what the relevant issue-space is and how the alternatives are positioned within it. Or they may independently

excogitate a natural ordering urged if not quite compelled by logic. By whatever mix of such mechanisms, deliberation should tend to make preferences more single-peaked on a shared evaluative dimension (Miller 1992, Knight and Johnson 1995, Dryzek and List 2002, List et al. 2002). This need not mean that deliberation will tend to yield agreement on how to rank the alternatives (what List (2002) calls “agreement at a substantive level”). That may or may not occur. Our contention, rather, is that the post-deliberation rankings tend to share more of the same issue-space (what List (2002) calls “agreement at a meta-level”).

The importance of net preference change is obvious—majorities supporting given policies, parties, or candidates may form or disappear. A more deliberative democracy might bring different governments, enacting different policies. Preference structuration is important because it inversely affects the probability of cycles in majority preferences (Black 1948; Niemi 1969; Holmes and List forthcoming).⁷ Condorcet’s paradox famously shows that pairwise majority voting can lead to cyclical majority preferences. Several generalizations of the mechanism underlying Condorcet’s paradox have been found, among them Arrow’s impossibility theorem (1953) and the McKelvey-Schofield chaos results (McKelvey 1979). The threat of cycling and associated general impossibility results strike at the very heart of democracy’s rationale, as Riker (1982) has forcefully argued. What is the value of majority rule, if there is no meaningful majority preference?⁸

One possible escape from such cycling was provided by Black’s median voter theorem (1948), which demonstrated that if the preferences of *all* individuals are single-peaked with respect to the same dimension (i.e., in cases of *full* single-peakedness), the threat of cycling can be avoided, and there will be a Condorcet winner. More recent work has shown that even *partial* single-peakedness can help and that the *degree* of single-peakedness bears a strong positive relationship to the probability that a *Condorcet winner* (an alternative that ties or beats all others in pairwise majority voting) exists and a strongly negative relationship to the probability of cycles (Niemi 1969, Holmes and List 2002).⁹

The Split-Half Deliberative Poll: Design and Measurement

To isolate the effects of the formal on-site deliberations, we have built a fully randomized field experiment into a Deliberative Poll. The sample was drawn from the fifteen towns surrounding New Haven, Connecticut, and the issues were two: the level of service to be provided by the local airport and what if any sharing there should be of property-tax revenues from new commercial development. A random sample of 1,032 local residents was interviewed and invited to attend the on-site deliberations, held from Friday evening, March 1, through midday Sunday, March 3, 2002. A total of 133 showed up. The Friday evening session, at which participants dined with members of their randomly assigned small groups, was designed to orient the participants and acquaint them with one another. The actual deliberations began Saturday morning. Those interviewees who said they would attend were sent the briefing materials, and those who did attend were paid \$200 on completing the final questionnaire. As Table 1 shows, the participants were somewhat more educated than the nonparticipants but comparable to them in terms of income, gender, racial composition, and voter registration.¹⁰

(Table 1 about here)

At the beginning of the weekend, each of the sixteen randomly assigned small groups was itself randomly assigned to one of the two possible orders in which the two issues could be deliberated. Eight groups (containing 64 participants) deliberated the airport Saturday morning and the taxes Saturday afternoon, the other eight (containing 68 participants) the reverse. We refer to these two treatment groups as “airport-first” and “revenue-sharing-first.”¹¹

The formal on-site deliberations consisted of three “deliberative sessions,” each involving both small-group discussions and plenary questions-and-answers with panels of policy experts and advocates. The first two sessions, occupying the whole of Saturday, concentrated on one issue apiece, with the first confined to the airport for the airport-first treatment group and to revenue-sharing for the revenue-sharing-first treatment group, and the second to revenue-sharing for the airport-first group and to the airport for the revenue-sharing-first group. The third, on Sunday morning, was more synoptic, with all

the participants revisiting both issues in their small groups and then questioning a panel of local and state officials about both.

There were three waves of measurement: the initial telephone interview (T1), a written version of the same questionnaire after the first deliberative session (T2), and the same written version (plus a few additional questions) again at the end of the weekend (T3). The T1-T2 interval brackets both the first deliberative session on-site and the casual, anticipatory learning and deliberation occurring between the first interview at T1 and the beginning of the weekend. The T2-T3 interval brackets the second and third deliberative sessions on-site.

From the standpoint of the randomized experiment, the T2 measurement is particularly revealing. At that point, one treatment group had deliberated the airport but not revenue-sharing, while the other had deliberated revenue-sharing but not the airport. Both groups had had the experience of deliberating in a casual, imbalanced way, with relatively homogeneous interlocutors, in the period between the initial interview and their arrival on site. Both had also had the experience of the more formal and balanced deliberation with more heterogeneous interlocutors on site, but on different issues. The randomization provides assurance that they differed minimally in other ways. On each issue, therefore, the contrast between their T2 attitudes should reflect the effect of the on-site deliberations on that issue.

The T3 measurement can be used similarly, if less certainly, to assess the effects of the second deliberative session, in which the small groups beginning with revenue-sharing switched to the airport, and vice versa. But this comparison is harder to interpret, since by T2 each group had already deliberated its first issue, and the T2-T3 interval bracketed not only second deliberative session but the third, during which both issues were deliberated. Perhaps discussing revenue-sharing is better preparation for discussing the airport than the reverse; or perhaps considering them simultaneously at the end alters the effect of the earlier sequencing. How far the T3-T2 comparison should be expected to mimic the T2-T1 comparison is therefore uncertain.

The two questions used to measure opinions toward revenue sharing read as follows:

My town should maintain local control over all of its tax revenues from new businesses and industries [scored 1 if “agree strongly” and 0 otherwise]

Now that we’ve asked you about several options for dealing with property tax revenues from new business and industries, which would you most like to see? [Scored 1 if “My town should maintain local control over all of its tax revenues” and 0 otherwise].¹²

These two scores were summed to create a scale ranging from 0 to 2, with higher scores indicating greater opposition to revenue sharing.

The questions measuring opinions toward airport expansion are similar:

Commercial passenger service should be expanded to provide more flights to more places.
[scored 1 if “agree strongly” and 0 otherwise]

Now that we’ve talked about some options for the future of Tweed airport, which of the following would you most like to see? [scored 1 if “Expanding passenger service, providing more flights to more places” and 0 otherwise]¹³

Again these two scores were summed to create a scale ranging from 0 to 2, with higher scores in this case indicating greater support for airport expansion. As we shall see below, both scales divide the sample roughly in thirds at T1.

One reason for selecting this pair of issues is that they were likely to be seen as separate and distinct concerns. As expected, opinions on them are nearly orthogonal. Table 2 presents the Pearson correlations between revenue sharing and airport opinions over the three measurements. The correlations are all small, and none is significant at the .01 level using a two-tailed test. This weak connection suggests little possibility that deliberating one issue changes opinions on the other.

(Table 2 about here)

Results

Policy Preferences

Our first hypothesis is that deliberation often changes preferences, at both the individual and aggregate levels. Previous Deliberative Polls have typically seen notable opinion change but have not isolated the effects of the formal on-site deliberations from those of other aspects of the Deliberative Polling experience. The split-half design enables us to examine this question.

Consider first the issue of revenue-sharing. The timewise distributions of preferences are given in Table 3. At T1, the two treatment groups differ insignificantly (chi-square, 2 df. = .10, two-tailed $p = .95$), with roughly 45% of the sample giving the most pro-sharing response. Over the whole weekend, the two groups also changed equally, ending as well as starting with views similar to each other's, but ending with views dramatically different from those with which they began. At T3, 74% of the whole sample took the most pro-sharing position, an increase of 29 points from T1.

(Table 3 about here)

Much of this change was clearly attributable to the formal on-site deliberations. From T1 to T2, the percentage taking the most pro-sharing position increased by 25 points (from 44 to 69%) among the revenue-sharing-first group, then deliberating the issue, but by only 13 points (from 45 to 58%) among the airport-first group, then deliberating the airport instead. From T2 to T3, similarly, it increased by 15 points (from 58 to 73%) among the airport-first group, but by only 6 points (from 69 to 75%) among the revenue-sharing-first group. These contrasts suggest that somewhere between 9 (= 15 – 6) and 12 (= 25 – 13) points of the 29 point change between T1 and T3 can be directly attributed to the formal on-site deliberation. This contrast sharpens, moreover, when we control for T1 opinion. Estimating an ordered logit model with T1 opinion as a covariate yields an estimated treatment effect of .441 with a standard error of .227 and a two-tailed p of .05.

Most of the remaining 17 to 20 points of the 29-point change from T1 to T3 change is presumably due to the anticipatory learning and deliberation during the portion of the T1-T2 interval preceding the weekend (when even the airport-first group increased its support for the most pro-sharing position by 13%), and much of the rest presumably to the casual, spillover discussions and the synoptic concluding session between T2 and T3. Of course, it is also possible that some of the T2-T3 change in the revenue-sharing-first group represented a delayed effect of their on-site revenue-sharing deliberation, in which case the latter's effect may exceed our 9-12% estimate.

On airport expansion, the effect of the formal on-site deliberation is less clear. Table 4 presents the timewise distributions of preferences on this issue. Although the airport-first treatment group begins, at T1, with somewhat more extreme views (in both directions), the difference does not approach statistical significance (chi-square with 2 df. = 2.6, two-tailed $p = .27$). From T1 to T2, both groups grew more enthusiastic about expansion, with some opponents moving to the middle of the scale. The airport-first group, which deliberated the issue, moved slightly more but not enough to produce a statistically significant difference between the two groups at T2 (chi-square with 2 df. = 1.6, two-tailed $p = .45$).

(Table 4 about here)

The second deliberative session, on the switched topics, seemed to have quite a different effect. From T2 to T3, the revenue-sharing-first group, now deliberating the airport, turned *less* favorable toward airport expansion, with the percentage strongly opposing it rising from 32% to 43%. Among the airport-first group, now deliberating revenue-sharing, however, both the percentage strongly opposing expansion *and* the percentage strongly supporting it increased (from 25% to 30% and from 38% to 45%). Thus the T3 difference between treatment groups comes close to statistical significance. Among the airport-first group, 45% strongly supported expansion and 30% strongly opposed it, while among the revenue-sharing-first group, an almost mirror-image 28% strongly supported it and 45% strongly opposed it (chi-square with 2 df. = 4.5, two-tailed $p = .11$).¹⁴

To confirm these treatment effects, we pool the observations from both the T1-T2 and T2-T3 intervals and model the respondent's T2 or T3 preference as a function of his or her preceding preference

(at T1 or T2, respectively) and whether his or her small group deliberated the issue over the preceding interval (from T1 to T2 or from T2 to T3, respectively). We adopt an ordered logit model. The results, in Table 5, indicate that, controlling for prior attitude, deliberation affected attitudes on revenue sharing, with those who deliberated the issue becoming significantly more favorable toward sharing. The estimated treatment coefficient of .645 (with a standard error of .298) may be interpreted to mean that individuals who would otherwise have a 50% chance of favoring the most anti-sharing position have just a 34% chance of doing so in the wake of deliberation. On the airport, deliberation does not seem to have any effect over the whole T1-T3 interval, but this null finding conceals a nearly significant effect between T2 and T3, visible in Table 4. The airport deliberations in the second deliberative session—but not the first—seem to have decreased support for expansion. This suggests a contingency to deliberation’s effects that wants exploring.

(Table 5 about here)

Preference Structuration

Our second hypothesis is that deliberation may increase preference structuration. More precisely, the hypothesis is that deliberation may increase one-dimensional preference structuration, in the form of single-peakedness.¹⁵ We begin by briefly describing the method of measuring the degree of single-peakedness within the sample at each point in time, following List et al. (2002).

The participants were asked to rank four revenue sharing alternatives and three airport alternatives in order of preference. On the revenue-sharing question, the alternatives were:

1. My town should maintain local control over all of its tax revenues from new businesses and industries.
2. My town should try for a voluntary agreement with other towns in the region to share some of the tax revenues from new businesses and industries.

3. The state should provide incentives for towns in the region to share some tax revenues from new businesses and industries.
4. The state should require towns in the region to share some tax revenues from new businesses and industries.

On the airport, the alternatives were:

1. Maintaining commercial passenger service to nearby cities but not expanding service to serve a larger market
2. Expanding commercial passenger service, providing more flights to more places
3. Ending commercial passenger service and providing service only for private airplanes.

We have preference orderings over these alternatives at T1, T2, and T3.

Table 6 shows the results for the entire sample and for the two treatment groups separately at all three measurements. Each cell contains four entries: (i) the structuring dimension along which the greatest number of respondents displayed single-peaked preferences, (ii) the index of structuration (for our purposes, single-peakedness) for that dimension (m/n , where n is the number of participants in the relevant treatment group or the whole sample, as the case may be), (iii) the bootstrap standard error of the single-peakedness index,¹⁶ and (iv) the corresponding Condorcet winner.

Consider first the top row, displaying the results for the whole sample. It is entirely possible for an issue showing no net preference change to show some change (presumably an increase) in structuration to the extent that the gross preference change is such as to increase (or, less likely, decrease) single-peakedness. But here too the airport issue shows no appreciable change. The structuration index increases modestly from T1 to T2 (from .77 to .84) but then recedes from T2 to T3 (from .84 back to .81). The structuring dimension is the same throughout, ordering the alternatives enumerated above as [2 1 3],

and the Condorcet winner, throughout, is 2 (“expanding commercial passenger service, providing more flights to more places”).

Revenue sharing, however, is again another story. Structuration increases dramatically both from T1 to T2 and from T2 to T3, from .52 to .70 and then to .80. On the constant structuring dimension ordering the alternatives above as [1 2 3 4], the participants’ preferences become more and more single-peaked. The Condorcet winner changes from T2 to T3—from 3 (“the state should provide incentives for towns in the region to share some tax revenues from new businesses and industries”) to 2 (“my town should try for a voluntary agreement with other towns in the region to share some of the tax revenues from new businesses and industries”)—but the structuration index continues to rise.

Comparing the treatment groups again speaks to deliberation’s effect. Not surprisingly, since there was not much to account for, deliberation seems to have had no real effect on the structuration of airport preferences. Among the airport-first group, the structuration index increased only from .80 to .81 from T1 to T2, when they deliberated the airport, while among the revenue-sharing-first group, it decreased slightly, from .88 to .84, from T2 to T3, when they deliberated it.

Deliberation had a profound effect, however, on the structuration of revenue sharing preferences. The structuration index surged from .47 to .78 from T1 to T2 among the revenue-sharing-first group, when they deliberated the issue, and from .63 to .84 among the airport-first group, when they in turn did so. By contrast, it increased only mildly, from .56 to .63, among the airport-first group from T1 to T2 and decreased trivially, from .78 to .76, among the revenue-sharing-first group from T2 to T3, when they were instead deliberating the airport. The structuring dimension remains [1 2 3 4] for both groups throughout, and the Condorcet winner changes in both groups from 3 to 2, in each case over the interval during which the group deliberates the issue—from T1 to T2 in the revenue-sharing-first group and from T2 to T3 in the airport-first group. With respect to preference structuration, in sum, as with respect to the preferences themselves, deliberation seems to have large but contingent effects.¹⁷

(Table 6 about here)

Discussion

These results demonstrate experimentally—for the first time—that the formal on-site deliberations account for a substantial portion of the changes in both preferences and preference structuration over the course of a Deliberative Poll. Although there is not much change of either sort on the airport issue, the formal deliberations seem to account for roughly 13 points of the 29 point before-after increase in support for revenue-sharing. This greatly strengthens the inference, based heretofore only on statistical associations with information gains, that these before-after changes are substantially driven by deliberation. It also, incidentally, strengthens the logically prior inference, based heretofore on contrasts with quasi control groups and the usual absence of anything noticeable that could account for parallel, contemporaneous changes in the wider public, that the before-after changes are real. At least on some issues, deliberation—at least serious, balanced deliberation with diverse conversational partners—does seem to produce both net preference change and increased preference structuration.

Some portion of these effects would seem to rest on the absorption of factual information. At all three points of measurement we asked the participants to say (1) whether the region's population was closest to 250,000, 350,000, 550,000, or 750,000; (2) whether its rate of job growth during the 1990s was more than, about the same as, or less than in the rest of the United States; (3) whether the major source of revenue for most of the region's town governments is sales taxes, property taxes, direct state subsidies, or direct federal subsidies; (4) whether New Haven's population increased, decreased, or did not change during the 1990's; (5) whether state law allows communities to share property tax revenues; (6) whether those communities with the most valuable property tend to have the lowest, average, or the highest property tax rates; (7) whether the Federal Aviation Authority classifies the regional airport as a major hub, a medium hub, a minor hub, or not a hub; and (8) whether maintaining the regional airport at its current level of service would require any significant investment. The correct answers are (1) 555,000, (2) less, (3) property taxes, (4) decreased, (5) yes, (6) the lowest, (7) a non-hub, and (8) yes. Items (7) and (8) are specifically relevant to airport expansion, items (5) and (6) specifically relevant to revenue sharing, and items (1) through (4) generally relevant to the politics of the region.

Table 7 shows the percentages of the whole sample and of the two treatment groups answering the two airport items, the three revenue sharing items, the three general items, and all eight items correctly at T1, T2, and T3. As in previous Deliberative Polls, the participants absorbed a great deal of factual information. For the whole sample across all eight items, the average percentage answering correctly increased by 22.3%, from 36.8% to 59.1%. The largest gains on this overall measure occur from T1 to T2, doubtless because that is the much longer interval and because it brackets not only the first on-site deliberative session but the distribution of the briefing materials and the anticipatory, off-site deliberations with friends, family, and coworkers.

We present the issue-specific results with some diffidence, since the airport and revenue-sharing indices consist of only two items apiece, and any comparison must rest on a rather shaky assumption of equal average difficulty. Still, the differences are generally sizable, and the pattern extremely satisfying. For the whole sample, revenue sharing information increased by 34.5%, general information by 19.5%, and airport information by only 15.5%. From T1 to T2, both treatment groups gained more information on the issue they deliberated then but also gained on the other issue, undoubtedly because the briefing materials and anticipatory, off-site deliberations covered both issues. The airport-first group gained 15.6% on the airport and 6.3% on revenue sharing, while the revenue-sharing-first group gained 40.4% on revenue-sharing and 8.8% on the airport. Both groups also gained, and to similar degree (19.1% and 19.9%), on regional politics more generally. From T2 to T3, however, each treatment group showed significant information gains *only* on the issue it deliberated during that interval. The airport-first group gained 18.8% on revenue sharing, and the revenue-sharing-first group 5.9% on the airport. Between T1 and T2, in short, the participants learn a good deal (presumably before arriving on-site) about all the topics of deliberation and a good deal more (presumably after arriving) about the topic they deliberate during the first deliberative session. Between T2 and T3 they learn still more but only about the topic they deliberate during the second deliberative session. The on-site learning does seem to be deliberation-based.

(Insert Table 7 about here)

These information gains line up nicely with the changes in net preference and preference structuration on revenue sharing. From T1 to T2, the airport-first group learned a good deal about revenue sharing and changed its views a good deal. It also showed a modest increase in structuration. The revenue-sharing-first group learned still more, changed its views still more, and showed a much greater increase in structuration. From T2 to T3, the airport-first group, then deliberating revenue-sharing, learned a good deal on top of what it had learned from T1 to T2, changed its views a good deal further, and added greatly to its T1-T2 increase in structuration. The revenue-sharing-first group learned only a little, changed its views only a little, and showed no increase in structuration. (Compare Tables 7 with Tables 3 and 6.)

But what of the airport issue, where there was no great change in either aggregate preference or preference structuration? The result is hardly at odds with those of previous Deliberative Polls, where most but not all issues show much change. But why the difference between these particular two? The possible reasons are many, but the most obvious and perhaps likeliest lies in the issues' pre-deliberation salience. Previous Deliberative Polls have tended to show smaller net changes of preference and smaller increases in single-peakedness when the issue is more salient, and the participants consequently begin with more firmly rooted and more highly structured preferences.¹⁸ In the months preceding the present event, airport expansion was far more salient than revenue sharing. The region's most widely circulated daily newspaper mentioned revenue sharing only seven times but ran 74 articles on airport expansion during the year preceding the Deliberative Poll.¹⁹ Thirteen of the articles on airport expansion appeared during the two months immediately preceding the Poll (not counting the coverage of the Poll itself). These articles, along with editorials and letters from citizens, covered the full range of commonly held views about airport expansion and maintenance. To a much higher degree than usual, therefore, the public may have already deliberated about the issue, considering a balance mix of arguments and acquiring enough basic information to come to a considered judgment.

This explanation receives additional support from the issue-specific information measures in Table 7. Again acknowledging that each measure rests on only two items, we note that the participants

seem to have been much better informed about the airport than about revenue-sharing at T1, before they had ever heard of the Deliberative Poll, answering 25.4% of the revenue-sharing information items but 38.6% of the airport information items correctly. The revenue sharing information items do not appear simply to have been intrinsically harder, given that by T3 the participants actually fared slightly better on them (59.9% correct, versus 54.2% on the airport items).²⁰ Rather, the participants would appear to have entered the Deliberative Poll with better informed—and hence less easily changed and more highly structured—preferences about the airport than about revenue sharing.

Another question invited by these results is what accounts for the remaining 16 points of the increased support for revenue sharing. It is probably not mere stimulation by the first two waves of measurement, since the preferences regarding the airport, measured at exactly the same points, do not exhibit much net change or increased structuration. Rather, the sources must be other aspects of the Deliberative Polling experience, notably including the information in the briefing materials; the anticipatory, casual, and typically imbalanced deliberation in which many participants engage between the initial interview and their arrival on-site; and the informal conversations on-site, over meals and in the corridors.

A companion question is how—in two distinct senses—to apportion the 9-12 points that *are* attributable to the formal deliberations on-site. The first sense is more operational, disaggregating the Deliberative Polling experience. Recall that each deliberative session consisted of both small-group discussion and plenary questions and answers with panels of policy experts or policy makers. To what extent did the changes result from the one versus the other? Also recall that the T2-T3 interval encompassed two deliberative sessions: the Saturday afternoon session on the second issue (revenue sharing for the airport-first group, the airport for the revenue-sharing-first group) and the synoptic Sunday morning session on both issues. To what extent did the T2-T3 changes result from the deliberations on the second issue versus the ensuing deliberations on both issues? Even seemingly small features of the design may have effects worth trying to isolate. The requirement that each small group agree on one or more questions to pose to the expert panels may increase structuration, for example.

The second sense is more theoretical, distinguishing the social and psychological mechanisms by which deliberation may affect preferences and preference structuration. How far do the changes stem from the sheer quantity of information acquired, from the degree to which it is balanced, from the social properties or intellectual content of small group discussions, from the attitudinal or socio-demographic heterogeneity of one's fellow discussants? How far do they depend on reading versus listening? On solitary versus social processes more generally? On the adoption of "empirical premises," which are debatable, versus the absorption of "facts," which are not (Luskin, Fishkin, and Jowell 2002)? On learning versus thinking? On empathy or identification versus argument? It is already clear, from both earlier statistical analyses and the experimental results above, that the quantity of factual information absorbed plays a major intervening role, but much more remains to be explored.²¹

The answers to these empirical questions bear on more normative ones. To what extent is deliberation subject to the "pathologies" concerning critics of deliberative democracy (Stokes 1998, Mendelberg 2001)? To what extent do the substantive focus, balance, and heterogeneity built into the deliberations in Deliberative Polling protect against them?

All these questions need further exploration. We do not wish to anoint any single approach, but further randomized field experiments will help. The beauty of the split-half design is that it folds a randomized experiment into a Deliberative Poll without sacrificing the latter's public mission or recommending force. The key is that everyone eventually deliberates both issues. The only variation is in the timing. An experimental manipulation that gave different subsets of participants more substantively different experiences—as of discussing different issues or discussing given issues differently—would be exposing only some of the sample to the optimal treatment (whichever it was). Assuming only two treatment groups of equal size, a pooled n of say 300 would become an n of only 150 for the purpose of estimating deliberative public opinion.

The importance of the present findings goes without saying. The American Founders believed that only an elected elite would be capable of informed deliberation about how best to reconcile particular interests with the welfare of the nation as a whole. Thus U.S. Senators and the Electors in the Electoral

College were once mainly elected by state legislatures. But such Madisonian institutions of deliberative “filtration” have gradually given way to formal and informal inputs of “unfiltered” public opinion (Fishkin 1997). The vestigial state-by-state aggregation in the Electoral College aside, elections are now mainly direct. Ballot initiatives and referenda have grown increasingly common. Nominations are generally decided in direct primary elections or by open, primary-like caucuses. With the advent of modern polling, elected representatives are deliberating less and following the polls more. The same trends are evident around the democratic world.

But ordinary polls tend to register top-of-the-head, even phantom opinions, and the level of thought and information underlying many votes in elections and referenda is scarcely greater. Whether from Downsian “rational ignorance” (Downs 1957) or sheer uncalculated obliviousness, not many people know, think, or talk about politics very much (Delli Carpini and Keeter 1996, Kinder 1998, Price 1988, Luskin 2002). When they do talk about politics, people tend to choose conversational partners, topics, and specific sources of information to minimize disagreement. Mostly, therefore, they talk with others very like themselves, expressing opinions very like their own (Kinder 1998). And what they *hear*, given “confirmatory” and similar psychological biases (Higgins and Bargh 1987), is still more like what they already think.

Can ordinary citizens do better? The results here confirm that opportunities for serious, balanced discussion with heterogeneous fellow citizens can permit people to learn about the issues, weigh competing arguments, reflecting on their values in relation to given policies, reconsider their views, and arrive at a shared understanding of the criteria for preferring one alternative to another. The resulting preferences, much more than those in conventional polls or elections, command respect. They are the opinions of people who, much more than usual, know what they are supporting or opposing and why. They are deliberative opinions, filtered not by elected elites but by randomly selected citizens themselves.

Our results on the airport issue may seem to suggest that everyday deliberation, reflected in conventional polls, may sometimes suffice. But it is worth bearing in mind that roughly two-thirds of all the policy attitude items in all the Deliberative Polls to date have showed statistically significant net

change. The airport issue in the split-half Deliberative Poll was unusually salient, well covered by the media, and on the public's mind. But sustained, substantive, and balanced coverage of policy issues in widely consumed media is a rarity, and even then many readers and viewers may pay little attention or absorb mainly material they already agree with. In the far commoner case in which most people have little information and have never weighed competing arguments, deliberative public opinion can only be glimpsed through something like Deliberative Polling.

The effects we have been documenting should hearten reformers interested in designing fully democratic, participatory, and effective processes of civic filtration (Crosby 1995, Crosby, Kelly and Shaefer 1986, Fishkin 1997, Gastil 2000, and Leib 2002). The results demonstrate the power of even relatively small doses of serious, balanced deliberation with a wide assortment of fellow citizens. More sustained deliberation of this kind can be expected to do still more.

References

- Arrow, Kenneth. 1953. *Social Choice and Individual Values*. New York: Wiley, second edition 1963.
- Black, Duncan. 1948. "On the Rationale of Group Decision-Making," *Journal of Political Economy* 56: 23-34.
- Brams Steven J., Michael A. Jones, D. Marc Kilgour. 2002. "Single-Peakedness and Disconnected Coalitions", *Journal of Theoretical Politics*, 14(3): 359-383.
- Converse, Phillip. 1964. "The Nature of Belief Systems in Mass Publics." In *Ideology and Discontent*, ed. David E. Apter. New York: Free Press.
- Crosby, Ned. 1995. "Citizen Juries: One Solution for Difficult Environmental Questions." *In Fairness and Competence In Citizen Participation: Evaluating Models for Environmental Discourse*, ed. Ortwin Renn, Thomas Webler, and Peter Wiedemann.
- Crosby, Ned, Janey M. Kelly, and Paul Schaefer. 1986. "Citizen Panels: A New Approach to Citizen Participation." *Public Administration Review* (46): 170.
- Delli Carpini, Michael X., and Keeter, Scott. 1996. *What Americans Know About Politics and Why it Matters*. New Haven: Yale University Press.
- Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper.
- Dryzek, John and Christian List. 2003. "Social Choice Theory and Deliberative Democracy: A Reconciliation", *British Journal of Political Science*: forthcoming.
- Elster, Jon. 1998. "Introduction", in *Deliberative Democracy*, ed Jon Elster. New York: Cambridge University Press, 1-18.
- Fishkin, James S. 1997. *The Voice of the People: Public Opinion and Democracy*. New Haven and London: Yale University Press, expanded paperback edition.
- Fishkin, James S., and Robert C. Luskin. 1999. "Bringing Deliberation to the Democratic Dialogue: The NIC and Beyond" in *A Poll with a Human Face: The National Issues Convention Experiment in Political Communication*, ed Maxwell McCombs and Amy Reynolds. Mahwah, NJ: Lawrence Erlbaum.

- Gaertner, Samuel L., John F. Dovidio, Mary C. Rust, Jason A. Nier, Brenda S. Banker, Christine M. Ward, Gary Mottola, and Missy Houlette. 1999. "Reducing Intergroup Bias: Elements of Intergroup Cooperation." *Journal of Personality and Social Psychology* 76(March): 388-402.
- Gastil, John. 2000. *By Popular Demand: Revitalizing Representative Democracy Through Deliberative Elections*. Berkeley, Los Angeles, and London: University of California Press.
- Gilens, Martin. 2001. "Political Ignorance and Collective Policy Preferences." *American Political Science Review* 95(June): 379-396.
- Higgins, E. Tori, and John Bargh. 1987. "Social cognition and social perception." In *Annual review of psychology*, vol 38, ed. Mark Rosenzweig and Lyman W. Porter: 369-425.
- Holmes, Susan and Christian List. 2002. "Partial Single-Peakedness and the Probability of Cycles: A Correlation." Working paper.
- Kinder, Donald R. 1998. "Opinion and Action in the Realm of Politics." In *The Handbook of Social Psychology*, vol 4, ed. Daniel Gilbert, Susan T. Fiske, and Gardner Lindzey. New York: McGraw-Hill.
- Knight, Jack and James Johnson. 1994. "Aggregation and Deliberation: On the Possibility of Democratic Legitimacy", *Political Theory* 22: 277-296.
- Leib, Ethan J. 2002. "Towards a Practice of Deliberative Democracy: A Proposal for a Popular Branch," *Rutgers Law Journal* 33: 359-455.
- List, Christian. 2002. "Two Concepts of Agreement", *PEGS: The Good Society* 11(1): 72-79.
- List, Christian, James Fishkin, Robert Luskin and Iain McLean. 2000. "Can Deliberation Increase Preference Structuration?" Paper presented at the annual meeting of the American Political Science Association, Washington DC.
- Luskin, Robert C. 2003. "The Heavenly Public: What Would the Ideal Citizenry Be Like?" In Michael B. MacKuen and George Rabinowitz (eds.), *Electoral Democracy*. Ann Arbor: University of Michigan Press, forthcoming.

- Luskin, Robert C. 2002. "From Denial to Extenuation (and Finally Beyond): Political Sophistication and Citizen Performance." In James H. Kuklinski (ed.), *Thinking about Political Psychology*. New York: Cambridge University Press.
- Luskin, Robert C. and James S. Fishkin. 1998. "Deliberative Polling, Public Opinion and Democracy: the Case of the National Issues Convention" Paper presented at the annual meeting of the American Association for Public Opinion Research, St Louis.
- Luskin, Robert C., James S. Fishkin, Roger Jowell, and Allison Park. 1999. "Learning and Voting in Britain: Insights from the Deliberative Poll" Paper presented at the annual meeting of the American Political Science Association, Atlanta.
- Luskin, Robert C., James S. Fishkin, Dennis L. Plane. 1999. "Deliberative Polling and Policy Outcomes: Electric Utility Issues in Texas." Paper presented at the annual meeting of the Association for Public Policy Analysis and Management, Washington DC.
- Luskin, Robert C., James S. Fishkin, Ian McAllister, John Higley, and Pamela Ryan. 2000. "Information Effects in Referendum Voting: Evidence from the Australian Deliberative Poll." Paper presented at the annual meeting of the American Political Science Association, Washington DC.
- Luskin, Robert C., James S. Fishkin, and Roger Jowell. 2002. "Considered Opinions: Deliberative Polling in Britain." *British Journal of Political Science* 32(July): 455-487.
- Manin, Bernard. 1997. *The Principles of Representative Government*. Cambridge: Cambridge University Press.
- McKelvey, Richard. 1979. "General Conditions for Global Intransitivities in Formal Voting Models", *Econometrica* 47: 1085-111.
- Mendleberg, Tali. 2001. "The Deliberative Citizen: Theory and Evidence." Forthcoming in *Research in Micropolitics, Volume 6*, edited by Michael Delli Carpini, Leonie Huddy, and Robert Y. Shapiro, Elsevier Press.
- Miller, David. 1992. "Deliberative Democracy and Social Choice", *Political Studies* 40 (special issue): 54-67.

- Niemi, Richard G. 1969. "Majority Decision-Making with Partial Unidimensionality," *American Political Science Review* 63: 488-497.
- Orbell, John M., Alphons van de Kragt, and Robyn M. Dawes. 1988. "Explaining Discussion-Induced Cooperation." *Journal of Personality and Social Psychology* 54: 811-819.
- Price, Vincent. 1988. "On the Public Aspects of Opinion - Linking Levels of Analysis in Public-Opinion Research." *Communication Research* 15(December):659-679.
- Price, Vincent, and Neijens, Peter. 1997. "Opinion quality in public opinion research." *International Journal of Public Opinion Research* 9(Winter):336-360.
- Price, Vincent, and Neijens, Peter. 1998. "Deliberative Polls: Toward Improved Measures of "Informed" Public Opinion?" *International Journal of Public Opinion Research* 10(Summer):145-176.
- Riker, William H. 1982. *Liberalism against Populism*. San Francisco (W. H. Freeman).
- Sally, D. 1995. "Conversation and Cooperation in Social Dilemmas: A Meta-Analysis of Experiments from 1958 to 1992." *Rationality and Society*, 7(January): 58-92.
- Schofield, Norman. 1976. "Instability of Simple Dynamic Games", *Review of Economic Studies* 45: 575-594.
- Stokes, Susan. 1998. "Pathologies of Deliberation." In Jon Elster, ed. *Deliberative Democracy*, Cambridge UP: 123-139
- Van Mill, David. 1996. 'The Possibility of Rational Outcomes from Democratic Discourse and Procedures', *Journal of Politics*, 58: 734-52.

Table 1
Demographic Comparisons of Participants, Nonparticipants, and Voting Population

| | Participants (n = 132) | Nonparticipants (n = 1024) | Voting Population |
|--------------------------------|---------------------------|-------------------------------|--------------------|
| Age | 50.1 yrs | 50.1 yrs | 47 yrs |
| Registered to Vote | 90.2% | 88.6% | 78.7% |
| Marital Status | | | |
| Single | 30.3% | 29.3% | 29% |
| Married | 49.2% | 52.4% | 52.5% |
| Divorced/Separated /Widowed | 19.7% | 17.3% | 18.5% |
| Education* | | | |
| Less than/some high school | 2.3% | 4.2% | 17% ^b |
| High school graduate | 9.1% | 20.2% | 30.8% |
| Some college | 25% | 21.1% | 18.2% |
| College graduate | 28% | 29.1% | 21.7% |
| Trade/Technical | 3% | 2.5% | --- |
| Graduate school | 32.6% | 22.2% | 12.4% |
| Income | \$61-70,000 | \$61-70,000 | \$64,018 |
| Race | | | |
| African American | 12.1% | 7.2% | 8.3% ^c |
| Caucasian | 72.7% | 75.3% | 78.1% |
| Hispanic or Latino/a | 3% | 4.4% | 9.6% |
| Other | 9.1% | 9.1% | 4% |
| Gender | | | |
| Male | 50.8% | 47.2% | 46.8% ^c |
| Female | 49.25 | 52.8% | 53.2% |

NOTE: percentages may not add up to 100 in nonparticipant blocks due to “refusal” category.

* Education levels in Participants and Nonparticipants differed significantly, $\chi^2 = 15.05$, $df = 5$, $p < .01$.

^a For population 15 and over.

^b For population 25 and over.

^c For population 18 and over.

Table 2
Correlations between Opinions Toward Tax Sharing and Airport Expansion

| | | | | | | |
|----------------------|------|------|------|------|------|------|
| Revenue-Sharing T1 | 1.00 | | | | | |
| Revenue-Sharing T2 | .48* | 1.00 | | | | |
| Revenue- Sharing T3 | .36* | .61* | 1.00 | | | |
| Airport Expansion T1 | .01 | .06 | .11 | 1.00 | | |
| Airport Expansion T2 | -.00 | -.08 | -.01 | .44* | 1.00 | |
| Airport Expansion T3 | -.15 | -.20 | -.07 | .41* | .59* | 1.00 |

NOTES: High scores on the tax sharing measure indicate a preference for maintaining local control over revenues. High scores on the airport measure indicate support for expanding passenger service. $n = 132$.

*Significant at the .01 level, by a two-tailed test.

Table 3
Support for Revenue-Sharing, by Experimental Treatment

Airport-First (*n* = 64)

| | Strongly Supports Sharing | In-Between | Strongly Opposes Sharing |
|----------------|------------------------------|------------|-----------------------------|
| Pre-Test (T1) | 45% | 33% | 22% |
| Midpoint (T2) | 58 | 13 | 30 |
| Post-test (T3) | 73 | 17 | 9 |

Revenue-Sharing-First (*n* = 68)

| | Strongly Supports Sharing | In-Between | Strongly Opposes Sharing |
|----------------|------------------------------|------------|-----------------------------|
| Pre-Test (T1) | 44% | 35% | 21% |
| Midpoint (T2) | 69 | 18 | 13 |
| Post-test (T3) | 75 | 21 | 4 |

Whole Sample (*n* = 132)

| | Strongly Supports Sharing | In-Between | Strongly Opposes Sharing |
|----------------|------------------------------|------------|-----------------------------|
| Pre-Test (T1) | 45% | 34% | 21% |
| Midpoint (T2) | 64 | 15 | 21 |
| Post-test (T3) | 74 | 19 | 7 |

Note: Rows may not sum to 100% due to rounding error.

Table 4
Support for Airport Expansion, by Experimental Treatment

Airport-First (*n* = 64)

| | Strongly Supports Expansion | In-Between | Strongly Opposes Expansion |
|----------------|--------------------------------|------------|-------------------------------|
| Pre-Test (T1) | 38% | 19% | 44% |
| Midpoint (T2) | 38 | 38 | 25 |
| Post-test (T3) | 45 | 25 | 30 |

Revenue-Sharing-First (*n* = 68)

| | Strongly Supports Expansion | In-Between | Strongly Opposes Expansion |
|----------------|--------------------------------|------------|-------------------------------|
| Pre-Test (T1) | 32% | 31% | 37% |
| Midpoint (T2) | 28 | 40 | 32 |
| Post-test (T3) | 28 | 29 | 43 |

Whole Sample (*n* = 132)

| | Strongly Supports Expansion | In-Between | Strongly Opposes Expansion |
|----------------|--------------------------------|------------|-------------------------------|
| Pre-Test (T1) | 35% | 25% | 40% |
| Midpoint (T2) | 33 | 39 | 29 |
| Post-test (T3) | 36 | 27 | 36 |

Note: Rows may not sum to 100% due to rounding error.

Table 5
Logistic Regression of Post-treatment Attitude on Pre-treatment Attitude and Deliberation, by Issue

Revenue Sharing

| Variables | B | Wald Test (z-ratio) | 95% CI | |
|---------------------|----------------|------------------------|-----------|-------|
| | | | upper | lower |
| Prior attitude | 1.54 (.19) | 65.49 | 1.17 | 1.91 |
| Deliberated Revenue | .645 (.298) | 4.68 | .006 | 1.23 |

Model Fit: $\chi^2 = 79.27, p < .0001.$

Airport

| Variables | B | Wald Test (z-ratio) | 95% CI | |
|---------------------|----------------|------------------------|-----------|-------|
| | | | upper | lower |
| Prior attitude | 1.37 (.16) | 70.13 | 1.05 | 1.69 |
| Deliberated Airport | .003 (.241) | .013 | -.445 | .501 |

Model Fit: $\chi^2 = 79.23, p < .0001.$

Table 6
Structuration (Single-Peakedness), by Experimental Treatment

| | | Airport | | | Revenue Sharing | | |
|---------------------------------------------------|------------------------------|----------------|----------------|----------------|------------------------|------------------|------------------|
| | | T1 | T2 | T3 | T1 | T2 | T3 |
| Whole Sample <i>n</i> = 132 | Dimension | [2 1 3] | [2 1 3] | [2 1 3] | [1 2 3 4] | [1 2 3 4] | [1 2 3 4] |
| | Structuration (<i>m/n</i>) | 0.77 | 0.84 | 0.81 | 0.52 | 0.70 | 0.80 |
| | standard error | (.033) | (.031) | (.032) | (.042) | (.037) | (.037) |
| | Condorcet Winner | 2 | 2 | 2 | 3 | 3 | 2 |
| Airport-First <i>n</i> = 64 | Dimension | [2 1 3] | [2 1 3] | [2 1 3] | [1 2 3 4] | [1 2 3 4] | [1 2 3 4] |
| | Structuration (<i>m/n</i>) | 0.80 | 0.81 | 0.86 | 0.56 | 0.63 | 0.84 |
| | standard error | (.043) | (.049) | (.049) | (.060) | (.062) | (.044) |
| | Condorcet Winner | 2 | 2 | 2 | 3 | 3 | 2 |
| Revenue-Sharing-First <i>n</i> = 68 | Dimension | [2 1 3] | [2 1 3] | [1 2 3] | [1 2 3 4] | [1 2 3 4] | [1 2 3 4] |
| | Structuration (<i>m/n</i>) | 0.82 | 0.88 | 0.84 | 0.47 | 0.78 | 0.76 |
| | standard error | (.045) | (.037) | (.036) | (.057) | (.052) | (.051) |
| | Condorcet Winner | 2 | 2 | 2 | 3 | 2 | 2 |

Notes: Boldfaced cells show structuration before and after the relevant treatment group deliberated the issue. The top entry in each cell is the proportion in each group holding single-peaked preferences on the most frequently occurring preference dimension. The bottom entry is the Condorcet winner. On the airport, Alternative 2, the Condorcet winner throughout, is “expanding commercial passenger service, providing more flights to more places.” On revenue sharing, Alternatives 2 and 3 are “my town should try for a voluntary agreement with other towns in the region to share some of the tax revenues from new businesses and industries” and “the state should provide incentives for towns in the region to share some tax revenues from new businesses and industries,” respectively.

Table7
Information Gains, by Experimental Treatment

| | Means (Percentages of Correct Answers) | | | Differences of Means | | |
|----------------------------------------|----------------------------------------|------|------|----------------------|---------|---------|
| | T1 | T2 | T3 | T2 – T1 | T3 - T2 | T3 - T1 |
| Airport-First (<i>n</i> = 64) | | | | | | |
| Airport | 39.8 | 55.5 | 56.3 | 15.6*** | 0.8 | 16.4*** |
| Revenue Sharing | 27.3 | 33.6 | 52.3 | 6.3* | 18.8*** | 25.0*** |
| General | 39.8 | 60.1 | 59.0 | 20.3*** | -1.2 | 19.1*** |
| Overall | 36.7 | 52.3 | 56.6 | 15.6*** | 4.3* | 19.9*** |
| Revenue-Sharing-First (<i>n</i> = 68) | | | | | | |
| Airport | 37.5 | 46.3 | 52.2 | 8.8** | 5.9* | 14.7*** |
| Revenue Sharing | 23.5 | 64.0 | 66.9 | 40.4*** | 2.9 | 43.4*** |
| General | 43.4 | 61.0 | 63.2 | 17.6*** | 2.2 | 19.9*** |
| Overall | 37.0 | 58.1 | 61.4 | 21.1*** | 3.3* | 24.5* |
| Whole Sample (<i>n</i> = 132) | | | | | | |
| Airport | 38.6 | 50.8 | 54.2 | 12.1*** | 3.4 | 15.5*** |
| Revenue Sharing | 25.4 | 49.2 | 59.9 | 23.9*** | 10.6*** | 34.5*** |
| General | 41.7 | 60.6 | 61.2 | 18.9*** | 0.6 | 19.5*** |
| Overall | 36.8 | 55.3 | 59.1 | 18.5*** | 3.8** | 22.3*** |

Note: Standard errors in parentheses. *P*-values based on one-tailed tests.

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

NOTES

NOTES

*An earlier version of this paper was presented at the annual meeting of the International Society of Political Psychology, Boston, MA, July 6-9, 2003. We are grateful to Ethan Leib, the League of Women Voters of Connecticut Education Fund Inc.; The Guild Group, and Regional Plan Association for assistance in organizing and carrying out the split-half Deliberative Poll and to Carnegie Corporation, the Ford Foundation, the Renee B. Fisher Foundation, Fannie Mae, William C. Graustein, New Haven Savings Bank, the South Central Connecticut Regional Water Authority, the United Way of Greater New Haven, and United Illuminating Company for making the project possible. The Center for Deliberative Polling and the Public Policy Clinic at the University of Texas at Austin, the Community Foundation for Greater New Haven, and Yale University's Institution for Social and Policy Studies also provided financial and in-kind support. The research and much of the writing were conducted while Fishkin and Luskin were Fellows at the Center for Advanced Study in the Behavioral Sciences, supported by grants from the William and Flora Hewlett foundation (Grant #2000-5633), the Center General Fund, and the University Research Institute of the University of Texas.

¹Topics have included the choices in a parliamentary election (in Britain), in referenda (in Australia and Denmark), and on such policy issues as the future of the American family, how best to meet regional electricity needs, how to deal with crime, the future of Britain's National Health Service, and Britain's role in Europe.

²For theoretical discussions of this hypothesis, see Miller (1992), Knight and Johnson (1994), and Dryzek and List (2003).

³These findings also support the widely accepted propositions that many of the opinions in conventional polls are "top-of-the-head" (Converse 1964) and that political knowledge affects policy and electoral preferences (Bartels 1996, Delli Carpini and Keeter 1996, Althaus 1997, Luskin and Globetti 1997, Gilens 2001).

⁴ To illustrate, suppose there are three individuals and three alternatives x , y , and z . Let the first individual prefer x to y to z , the second prefer y to z to x , and the third prefer z to x to y . Then two of the three individuals prefer x to y , prefer y to z , and prefer z to x . The alternatives are in a cycle, and there does not seem to be a normatively attractive winner. In particular, there is no Condorcet winner—no alternative that beats or at least ties with all other alternatives in pairwise majority voting. Thus perfectly consistent individual preferences may lead to inconsistent majority preferences.

⁵ More recently, some social choice theorists have argued that because the discussion characteristic of democratic deliberation is relatively free-wheeling and open-ended, it may exacerbate rather than reduce the threats of impossibility and instability in popular decision-making. (Van Mill 1996, Knight and Johnson 1994) This conclusion has been challenged by Dryzek and List (2003), who argue that “there are ‘structuration’ processes endogenous to deliberation that can speak to the problems highlighted by social choice theory.”

⁶When not referring specifically to preference *structuration*, we use the term “preference” loosely and broadly, intending it as rough synonym for “attitude.” When it comes to preference structuration, however, we confine the term to orderings over sets of alternatives—captured by rankings.

⁷To illustrate, suppose there are three individuals and three alternatives x , y , and z . Let the first individual prefer x to y to z , the second prefer y to z to x , and the third prefer z to x to y . Then two of the three individuals prefer x to y , prefer y to z , and prefer z to x . The alternatives are in a cycle, and there does not seem to be a normatively attractive winner. In particular, there is no Condorcet winner—no alternative that beats or at least ties with all other alternatives in pairwise majority voting. Thus perfectly consistent individual preferences may lead to inconsistent majority preferences.

⁸Ironically, given his generally despairing view, it was Riker (1982) who first noted that discussion might create “a common view of the political dimension” at issue, thereby preventing cycles. “If by reason of discussion, debate, civic education and political socialization, voters have a common view of the political

dimension (as evidenced by single-peakedness),” he remarked, “then a transitive outcome is guaranteed” (Riker 1982: 128).

⁹More recently, some social choice theorists have argued that because the discussion characteristic of democratic deliberation is relatively free-wheeling and open-ended, it may exacerbate rather than reduce the threats of impossibility and instability in popular decision-making. (Van Mill 1996, Knight and Johnson 1994) This conclusion has been challenged by Dryzek and List (2003), who argue that “there are ‘structuration’ processes endogenous to deliberation that can speak to the problems highlighted by social choice theory.”

¹⁰A slightly higher fraction of the participants came from New Haven, where the Deliberative Poll took place, probably a natural consequence of the longer commute from suburbs and outlying towns. This geographic bias does not affect the character of the results. The views of those residing in New Haven moved in the same direction and to the same extent as the views of those residing in the surrounding towns.

¹¹We shall thus be referring to two sorts of groups—the small groups, within which the issues are discussed, and the treatment groups, each consisting of eight small groups, which tackle the issues in different sequences.

¹²These cut-points were chosen to give the resulting scale a roughly symmetrical distribution at T1.

¹³These cut-points, too, were chosen to give the resulting scale a roughly symmetrical distribution at T1.

¹⁴A slight difference in the T1-T2 versus T2-T3 airport treatment could be responsible for this pattern, as the expert panelists were not entirely the same. But they did represent the same mix of views, and the evidence from other Deliberative Polls suggests that the results are surprisingly insensitive to this sort of variation (Luskin, Fishkin, and Plane 1999).

¹⁵Let N be the set of n individuals, and X the set of k alternatives. To each individual $i \in N$, there corresponds a preference ordering R_i over the alternatives in X which is reflexive, transitive and connected (i.e. it allows strict preferences as well as ties). We write $xP_i y$ as an abbreviation for $[xR_i y$ and

not $yR_i x$]. The profile of individual preference orderings $\{R_i\}_{i \in N}$ satisfies single-peakedness if there exists a bijection $\square: X \rightarrow \{1, 2, \dots, k\}$ (representing the “structuring dimension”) such that, for every triple of alternatives $x, y, z \in X$ and every individual $i \in N$, if $\square(x) < \square(y) < \square(z)$ or $(\square(z) < \square(y) < \square(x))$, then $xR_i y$ implies $xP_i z$. This is the notion of single-peakedness used by Black (1948) and Arrow (1951). It is now sometimes called “ordinal” single-peakedness, as distinct from the slightly more demanding notion of “spatial” or “cardinal” single-peakedness. On the distinction between the two notions of single-peakedness, see Brams, Jones and Kilgour (2002).

¹⁶Standard errors of the m/n statistic were calculated using bootstrap procedures, in which 1,000 random samples of 132 subjects were drawn with replacement. The standard deviation of the resampled statistic provides an estimate of the standard error.

¹⁷Because the sampling distribution of m/n —not a straightforward proportion because the identity of the dimension on which the numerator is premised may change— is unknown, we refrain from assertions of statistical “significance or “insignificance.” It should also be noted that Table 6 shows the bootstrapped standard errors at given measurements, not for the differences between adjacent measurements, which may be somewhat larger or (less likely) smaller, depending on the sign and magnitude of the covariance. Still, the estimated standard errors make it hard to imagine, on the one hand, that the increases in the structuration of revenue sharing preferences among the revenue-sharing-first group from T1 to T2 and among the airport-first group from T2 to T3 are not significant or, on the other hand, that any of the other differences between adjacent measurements is.

¹⁸Deliberative Polls sometimes show significant change on highly salient issues, but often where there is a side to the issue that has not received intensive exposure in the real world. The overcrowding of prisons in the case of crime is one example; the Lib-Dems’ policy positions in the case of the 1997 British general election is another.

¹⁹Property taxes for homeowners were at issue in the region, and a great deal of discussion was devoted to property tax relief for the elderly, disabled, and low income families. But regional tax sharing was only

briefly mentioned in one editorial on the future of the region and as a low-profile issue in the platform of an unsuccessful local mayoral candidate.

²⁰This does not necessarily mean that there was more learning during the Deliberative Poll about revenue sharing than about airport expansion. Information indices like these are subject to ceiling effects—the participants answering both airport items correctly at T1 could not *show* any learning, but since great deal of literature in both psychology and communications research argues and finds that the information-rich tend to get information-richer, they were very likely learning a lot, unobservably. Accordingly, there even may have been *more* learning about the airport than about revenue sharing during the Deliberative Poll. We just can't see as much of it. For more extended discussion of this general issue, see Luskin, Fishkin, and Jowell (2002) and Luskin (2002).

²¹See Dryzek and List (2003) and Luskin (2003) for further discussion of possible mechanisms.