

Deliberative Polling, Public Opinion, and Democracy:
The Case of the National Issues Convention*

Robert C. Luskin
James S. Fishkin

University of Texas at Austin

January, 2005

Slightly revised from a paper presented at the annual meeting of the American
Political Science Association, Boston, MA, September 2-6, 1998.

For nearly a decade now, we have been conducting a new species of poll, a *Deliberative Poll* (Fishkin xxxx, xxxx; Luskin, Fishkin, and Jowell 1997; Fishkin and Luskin 1998). Reduced to essentials, the idea is to gauge the opinions of a random sample of citizens after they have had the opportunity of learning about, reflecting on, and discussing a set of policy issues. They are sent briefing materials, brought together in one place for a weekend, and then asked their opinions. There have now been nine regional Deliberative Polls, all in the U.S., and nine national ones, in the U.S., the U.K., Australia, and Denmark. The first was in the U.K., in April, 1994 (Luskin, Fishkin, and Jowell 1997), but easily the most prominent Stateside, and in ways the most ambitious, was the National Issues Convention of January, 1996, on which we focus here.¹

The project is partly and in origin an innovation in democratic practice, intended to address concerns in democratic theory. The value of democracy hinges jointly, we contend, on representativeness and deliberation. The democratic ideal is to aggregate the well-formed preferences of either the whole public or some much smaller set of people representing them by some form of majority rule, where by “well-formed” we mean something like “full-information.” We see deliberation as an indispensable means of informing preferences (Fishkin xxxx, xxxx). But how far and in what directions do well-formed preferences differ from those people actually hold? The rationale of Deliberative Polling presumes the existence of some sizable gap, but is there one, really? Perhaps the distribution of well-formed opinion can already be read from ordinary polls.

Such questions also connect Deliberative Polling with the empirical literatures on public opinion and voting, where it has for some time been the strong consensus of the relevant academic community, if perhaps not so much of commercial pollsters, who have some interest in thinking otherwise, that most people know very little about politics (Kinder and Sears, 1985; Luskin, 1987; Smith, 1989; Delli Carpini and Keeter, xxxx; Bennett, xxxx, Price 1998). Unfathomably, this proposition was once controversial, even in academia. For roughly a decade-and-a-half, from the late 1960s through perhaps the mid-1980s, the public's level of political information was much debated. Otherwise sensible political scientists could be found arguing in print that most people either had always been or had recently become quite well informed. Much of the evidence for such claims derived from careless knockoffs of *The American Voter's* famed “levels of conceptualization,” itself a relatively crude measure based on people's use of “ideological” terms like “liberal” and “conservative.” Other evidence came from correlations, factor analyses, and other devices based on the covariances of individuals' positions on policy issues. We needn't dwell here on why such measures mislead (Luskin, 1987; 2002a; 2002b); it should be obvious that the more straightforward measurement afforded by more or less factual items is more credible.

But the testimony of these latter is unequivocal: most people know very little about politics, and several decades of rising education levels, whitening collars, and lightning-rod events have altered that scarcely at all (Kinder and Sears, 1985; Luskin, 1987; Smith, 1989; Delli Carpini and Keeter, xxxx; Bennett, xxxx; Price 1998; and—notably, in view of Nie, Verba, and Petrocik xxxx—Nie, Junn, and Stehlik-Barry 1996). In Delli Carpini and Keeter's (1996) canvass of more than 2,000 survey questions tapping factual knowledge of politics, most items tapping knowledge of parties or prominent politicians' stands on policy issues, probably the most

important sort of information for citizens wishing to vote in effective pursuit of their values or interests, drew only 25-50% correct responses (this without corrections for guessing). Even very salient information, of a sort that would surely be known by *everyone*, is not: in 1986 roughly a quarter of a national random sample could not say who George Bush was or what position he held, even though he had at that point been Vice-President for six years. Generally speaking, large fractions of the public, often majorities, cannot correctly answer items that anyone reading this paper could correctly answer in his or her sleep.

Yet the Panglossian impulse to celebrate the political wisdom of the common man and woman has not receded. It has “merely shifted from denial to extenuation” (Luskin 2002a). The extenuations come in stronger (individual-level) and weaker (aggregate) varieties. The stronger sort runs: yes, people are remarkably ignorant of politics, but they nonetheless manage to grope their way to accurate perceptions and appropriate preferences—the same perceptions and preferences they would have if fully informed (Popkin, xxxx; Lupia xxxx, Lupia and McCubbins xxxx). The weaker sort runs: yes, people are remarkably ignorant of politics, and, yes, they individually make lots of errors, but since the errors tend to cancel out, the public as a whole still reaches the same distribution of perceptions and preferences as it would if everyone were fully informed (Page and Shapiro, xxxx; and, less explicitly, Stimson, xxxx).

Already, these claims are disintegrating. Luskin (2002a) shows how Brady and Sniderman’s (xxxx) contention that most people manage to locate liberals and conservatives correctly on policy dimensions is simply wrong. The mean locations look plausible enough, but at the individual level people get these locations right at rates barely different from chance. Bartels (19xx) and Luskin and Globetti (1997) show that cross-sectional variations in information affect how people vote in presidential elections—both at the individual level and in the aggregate. Delli Carpini and Keeter (1996) and Althaus (xxxx) show the same for people’s policy preferences.²

All these analyses, however, share the well-known weakness of cross-sectional, nonexperimental designs with respect to internal validity (Campbell and Stanley 1963). They establish only that high and low information respondents tend to have different perceptions and preferences, not that actually boosting *given* respondents’ information would alter their perceptions or preferences. Cross-sectional differences in perceptions or preferences could always have stemmed from differences not in information but in other variables happening to be correlated with information, and in that case giving people more information would not really make any difference, *ceteris paribus*. The best way of gauging information’s effects is to make people better informed and observe the consequences.

That is of course far more easily said than done. Raising political consciousness to any meaningful degree would require lifting people out of the everyday lives in which they find it rational to be ignorant. But that is just what Deliberative Polling does. In addition to being an innovative tool of democratic practice, it provides a unique, quasi-experimental window on this central question: to what extent and in what directions are distributions of opinions and votes being tilted by the considerable ignorance in which ordinary citizens generally operate? To what extent and in what ways, accordingly, would a substantially better informed and more thoughtful

public look different? This paper focuses on the answers suggested by the National Issues Convention.³

Deliberative Polling and the National Issues Convention

“Deliberative Polling” involves both polling and deliberation. The process begins with a brief interview, followed by a self-administered questionnaire, completed while the interviewer is still present, and then, finally, an invitation to attend the deliberations. (The interviewing in every national Deliberative Poll to date has been face-to-face.)⁴ In the interval between interview and deliberations, those who agree to attend are provided with carefully balanced briefing materials, which outline the issues, provide purely factual background, and describe some of the major policy proposals being mooted by politicians and the press, along with the empirical propositions underlying them.

Once arrived on site, participants spend the largest chunk of their working time discussing the issues in moderated, randomly assigned small groups of fifteen to twenty, although smaller clusters commonly continue these discussions in the corridors, over lunch and dinner, and the like. The moderators, in many cases focus group leaders by profession, receive special training for the Deliberative Poll. The other large chunk of working time is devoted to plenary sessions in which participants put questions to policy experts (sometime in mid-proceedings) and politicians (just one small-group wrap-up short of the end).⁵

The Deliberative Poll that concerns here is the largest, longest-deliberating, and generally most ambitious of the five national Deliberative Polls so far, the National Issues Convention (NIC) of January 18-21, 1996. The NIC is also the only national Deliberative Poll to have been held so far in the U.S. The other four national Deliberative Polls, all in Britain, have had roughly 300 participants meeting for something less than two days (Friday evening through Sunday noon). The NIC involved 466 participants meeting for roughly a day longer (Thursday evening through Sunday noon). For the first time, too, the design was fortified by interviews with a couple of quasi-control groups. A more detailed description of how the deliberations were structured can be found in Fishkin and Luskin (1998).

The NIC as Quasi-Experiment

Depending on further details, above all concerning control groups, the Deliberative Polling design falls somewhere between pre- and quasi- verging on true experimentation, with the one grand manipulation consisting of everything the participants experience from initial interview to post-deliberation questionnaire. In this section, we sketch some features of the NIC’s design and instrumentation.

Participant and Interview Samples

For external validity, Deliberative Polling has the great strength, compared to most experiments, of starting with a probability sample. This is also important, from the standpoint of democratic theory, to the goal of representation. By the end of the weekend, of course, the participants should be quite *unrepresentative* of the public in its natural state (as the treatment group at the end of any successful experiment will be). But it is nevertheless important that they be as representative as possible at the outset. The point of Deliberative Polling is to estimate the *considered* opinions of the *actual* public—with its actual distribution of traits and initial attitudes—not some biased portion thereof.

One major qualification, and concern, stems from attrition. By the time the data are in, no sample of free human beings is ever truly random. All are eroded, nonrandomly, by inaccessibility and self-selection. Not every designated respondent can be contacted; not everyone contacted agrees to be interviewed. And those who do end up being interviewed are different, on average, from those who do not—more affluent, better educated, more socially central, and more interested in and knowledgeable about the subject of the interview (Brehm xxxx).

This sort of erosion varies greatly, however, by survey house, survey design, and individual survey. For this among other quality-related reasons, we have been fortunate to work with Social and Community Planning Research (SCPR) in the Britain and the National Opinion Research Center of the University of Chicago (NORC) in the U.S. All six national Deliberative Polls in the U.S. and the U.K. have rested on stratified area (cluster) sampling, and all have involved the face-to-face interviewing that area sampling makes possible.⁶ The initial response rates, or *interview rates* as we may term them in the present context, have ranged roughly from 61.9% to 74.2%.

In contrast to ordinary surveys, however, Deliberative Polls suffer a second stage of attrition. Not everyone interviewed makes it to the on-site deliberations, and again those who do may be different from those who do not. A perennial and reasonable worry is the fraction of interviewees who actually wind up deliberating, what may be termed the *participation rate*. It is not necessarily easy to get people to spend a weekend away from home, possibly missing work or family obligations, to discuss issues they may never previously have thought much about. And since the overall *response rate* is the product of the interview and participation rates, a low participation rate implies a still lower response rate. In the national Deliberative Polls to date, the participation rates have mostly been in the thirties, with a couple down in the twenties, and one—for the NIC—over 50%. The result has been overall participation rates mainly in the twenties, with one—again the NIC’s—well up in the thirties.

It is worth placing these figures in perspective. Compared to the 60 to 70-odd percent response rates commonly achieved by the most careful academic and quasi-academic survey houses, like SCPR and NORC, in face-to-face interviewing, percentages in the thirties are very low. But: (1) Deliberative Polling bears some considerable resemblance in this respect to panel surveys, which also involve multi-stage losses from both self-selection and inaccessibility. The 1972-74-76 NES panel, for example, began with 2,705 respondents in the 1972 pre-election

wave. By the 1976 post-election wave that original n was down to 1,015, for a retention rate of 37.5%. Since the 1972 pre-election response rate was 75.1%, the response rate for the whole panel was $37.5\% * 75.1\% = 28.2\%$. (2) Media polls, what with telephone interviewing and far less determined pursuit of hard-to-reach respondents, have properly calculated response rates of only 30-50%—and that’s the best of them (Brady and Orren 1992). The response rates of lesser or overnight polls run still lower, generally *much* lower. By this standard, percentages in the twenties or thirties look very respectable. (3) Experimental and quasi-experimental designs—and it must be borne in mind that the Deliberative Poll is more quasi-experiment than conventional survey—seldom have probability samples or meaningful response rates at all. (4) The real question, in any case, is not response rate per se but bias. We examine this question empirically below.

For the NIC, our sample was drawn, interviewed, and recruited to Austin by NORC. The initial survey went into the field November 4, 1995 and right up to January 18, 1996, the start of the Convention. The interviewing was face-to-face, and the sampling followed a standard area (cluster) sample design, with Kish table selection at the doorstep. Of the 1,260 households drawn, the designated respondent was successfully interviewed in 911, for an interview rate of 72.4%. Of these 911, 466 eventually showed up in Austin, for a participation rate of 51.1%.⁷ The overall response rate was thus $466/1260 = 37.0\%$.⁸

Control Groups

As previously mentioned, the NIC had two quasi-control groups. The first consisted of a random subsample of the initial interviewees who did not attend the deliberations, reinterviewed by telephone at roughly the same time as the deliberations were ending. NORC randomly selected 301 of the 452 nonparticipants, and succeed in interviewing 172 of them, for a conditional response rate of 57.1%. We shall refer to these respondents, for short, as *nonparticipants*. The second quasi-control group consisted of an independent national RDD sample interviewed by telephone, also at roughly the same time as the deliberations were ending. There were 431 respondents, corresponding to a response rate of 48%.

We call these “quasi-control groups” because true control groups would be separated from the participant sample only by random assignment. Here, the nonparticipant sample is instead separated from the participant sample by self-selection. The nonparticipant data do have the advantage, however, of repeated measurement, affording the same t_1 - t_2 contrasts as the participant data.

The independent RDD sample, though observed only at t_2 , comes much closer to being a true control group. The way to think about its status is to note that drawing two independent random samples is equivalent to drawing a single random sample and randomly dividing it in two. Now, the participant and RDD samples *are* independent; the hitch is that neither is fully random, and particularly that one (the participant sample) is much less random than the other. While both involve self-selection, the participant sample involves a second layer of it.⁹

Between them, these two quasi-control groups, the one with its t_1 - t_2 measurement, the other with its reasonably close approximation of random assignment, afford a reasonable way of assuring ourselves that any changes of attitude or knowledge we see are not something that would have occurred anyway, participation or no. They bring the NIC fairly close to true, as opposed to merely quasi-, experimentation.

The Questionnaire

The interview and written questionnaire items can be divided into several stacks. There are measures of sociodemographic variables like age, sex, and race; general political orientations and behaviors, including party and ideological identifications (as Republican or Democrat, liberal or conservative) and 1992 presidential vote choice; prior political participation; civic attitudes, including political interest and political efficacy; values like respect for authority, economic equality, future as opposed to present well-being, and pursuing the national self-interest in foreign affairs; religious preferences, practices, and beliefs; policy attitudes, concerning the degree to which the government should do or refrain from doing particular things within our three policy domains; factual knowledge of politics, especially in those domains; positions on what we call *empirical premises*, like facts but more debatable (for example, the position that “the economy can only run if businesses make good profits” or that the “biggest problem for the American family” is “economic pressure,” as opposed to “the breakdown of traditional American values”); and, post-deliberation only, the respondent’s experience of the process. It should be noted that the foreign affairs items tend to be more general and abstract, and thus closer to goals or values, than those in the other two domains.

In all, we have 49 policy items, including 16 about the family, 22 about foreign affairs, and 11 about the economy; 16 items measuring political media consumption, political interest, internal and external political efficacy, and voting and other participatory behaviors; 3 on general political orientations and behaviors; 10 gauging empirical premises; 11 on sociodemographic variables; 5 on religious preferences, practices, and beliefs; 11 measuring relevant information; and 18 on the respondent’s experience of the process. Though too numerous to describe in detail here, the majority of these items are either familiar or reasonably self-explanatory. A fair sense of them may be gleaned from the labels and ranges in Tables 1 and 2. Note that the number of categories on the policy items varies from two to five.

Results

There are three basic empirical questions to be asked of any Deliberative Polling data. The first has to do with representation, the second with policy attitude change, and the third with the reasons for any such change, particularly the extent to which it seems to be driven by the learning and reflection the process is designed to promote.

Representativeness

By relevant standards, the NIC sample was highly representative, despite some mostly quite modest but statistically significant skews. The bias in our participant sample, as compared with a good academic-standard survey sample, may be reckoned by comparing our participants (Tables 1 and 2 about here)

with our nonparticipants. Between Tables 1 and 2, showing difference-of-means tests for binary, interval, and numerically scored ordinal items and difference-of-distribution (chi-squared) tests for frankly categorical ones, statistically significant differences emerge on seven of eleven sociodemographic characteristics; eight of twelve aspects of political involvement; one of two political orientations; two of six measures of political efficacy; one of six items having to do with religion; four of eight basic values; five of eleven pieces of more or less factual knowledge; and three of eleven empirical premises. In all, the participants and nonparticipants show statistically significant differences on 42 of 114 items.

In particular, participants are significantly younger, better educated, better off financially, apter to feel that their family's economic fortunes are improving, likelier never to have been married and unlikelier to be widowed, likelier to be working and unlikelier to be retired. They watch TV public affairs shows (but not the nightly TV news), read the news, and discuss politics more often; they have more frequently given money to a political party and seen or talked with politicians; more of them voted for president in 1992; they are more interested in politics. As a matter of self-identification, they are more liberal (though not more Democratic). They are likelier to agree that they have worthwhile political opinions and to disagree that politics is too complicated to understand. They are less likely to regard the Bible as the literal word of God. They place less emphasis on putting America's interests above those of other countries and on the rights to use one's property as one wishes, regardless of community views, and greater emphasis on sacrificing for the future. They are also more inclined to disagree that a "husband's job is to earn money, [while] a wife's job is to look after the home and family." They are likelier to have an accurate idea of the percentages of American children born out of wedlock and living in single-parent households; to know that the U.S. spends more on the national defense than on foreign aid, agricultural subsidies, or the space program; to have an accurate idea of the average unemployment rate in the US. over the past ten years; and to know that the Democratic and Republican parties are left and right of center, respectively. They are likelier to disagree that welfare is the source of most problems with the American family, that businesses need to make good profits, and that lowered trade barriers cost jobs at home.

The differences are many fewer, though still well above chance levels, when it comes to policy attitudes. Eleven of 47 policy attitudes, heavily concentrated in the family domain, show significant differences. Participants are less inclined to favor strengthening the American family by emphasizing "moral education in the schools," "more family planning services," "limiting sex and violence in movies and television," and "allowing prayer in school"; to agree that "one parent should stay at home while the children are growing up" and that this country would be better off if we just stayed at home and did not concern ourselves with problems in other parts of the world"; to emphasize "insuring adequate energy supplies for the U.S." as a foreign policy goal; to say we are spending too little on Medicare and Medicaid; and to favor keeping the current tax system basically the same.

Though statistically significant, most of these differences are relatively narrow. The widest, no doubt, is for age. Participants average six-and-a-half years younger than nonparticipants. They also, not coincidentally, average roughly two-thirds of a point higher on an eight-point education scale. More typically, however, the differences are on the order of the following examples: The participants average between one-sixth and one-half a point higher on four-point scales gauging political interest, discussion, and campaign activity; a quarter of a point more agreement on the four-point scale that politics is too complicated to understand; and a quarter of a point more liberal (though still to the conservative side of the midpoint) on the seven-point liberal-conservative scale.

These differences are more numerous and larger than in our other national Deliberative Polls. In the first-ever Deliberative Poll in Manchester, England in 1994, for example, participants and nonparticipants differed significantly on only four of 17 sociodemographic items and only three of 48 policy attitudes. The only domain in which there were differences approaching those here was information. On four of eight information items (two of four concerning crime policy and two of four concerning British politics generally), participants were more knowledgeable than nonparticipants. On the whole, however, the sample was much more perfectly representative. Across the 102 items in the data, only 14 showed statistically significant differences (Luskin, Fishkin, and Jowell 2002).

To some degree, these greater than usual differences between participants and nonparticipants may be put down to meteorological bad luck. The weekend of the NIC, much of the Western U.S. was blanketed by a major snow storm, which seems to have deterred some respondents, especially elderly ones, from attending. In all, 39 respondents held airline tickets but did not come, mostly on account of the weather.

We note these differences in the interest of candor, not to concede fault. The statistically significant ones are more numerous than would be expected by chance and than in other Deliberative Polls, but almost two-thirds remain statistically *insignificant*. The proportions of men and women, of Blacks and Hispanics, of Democrats and Republicans, of Protestants, Catholics, and Jews, differ only within the range expectable by chance. In the policy domain, the same is true of the proportion agreeing that “government has become too involved with the family through social welfare and family planning programs,” of the mean support for promoting democracy and protecting weaker nations as foreign policy goals, and of the mean preference for spending more, about the same, or less on eight of nine policy areas (all but Medicare/Medicaid). These are only a few examples.

More importantly, the Deliberative Poll is a quasi-experiment, not simply a survey, and the participant sample, with all its imperfections, is incomparably more representative than the typical experimental or quasi-experimental sample, notoriously consisting of university students selecting themselves either into the experiment itself or into the courses from which they are conscripted. As we have put it elsewhere, “almost certainly the NIC sample was the best microcosm of the American public ever gathered to a single place” (Fishkin and Luskin 1998).

Finally, many of the differences we do find should act to minimize the impact of participation. If our participants arrive somewhat better educated and more knowledgeable than the general population, that leaves less room for improvement. Both knowledge gains and attendant policy attitude changes should be smaller than if the participant sample were perfectly representative. Our results, therefore, should be conservative.

Policy Attitude Change

There are many ways of slicing the notion of attitude change, differing in their notion of what constitutes change. Is it *signed* or *absolute*? (In the aggregate, signed change is *net*, absolute change *gross*.) Does it refer to *position* or merely to *side*? Are we talking about the *degree* of change or merely its *existence*?

We may begin with some remarks about gross change, which may at first blush seem the most revealing sort. Signed changes, after all may cancel out. Some people may come to support a given policy proposal more at the same time as others are coming to oppose it more. The result will be minor net change, even though large numbers of participants have in fact changed. The problem is that there is no necessary correlation between learning and gross change, given the co-occurrence of learning with the sort of Brownian motion caused by “nonattitudes.” It is possible to derive conditions, under plausible if stylized assumptions, under which gross change will be either positively or negatively (!) correlated with learning (Luskin 1997).

Mainly, we wish to discuss the amount of gross change because it is the crux of a mischaracterization of what happened at the NIC that has received some currency. A number of secondhand accounts, notably including those of Kohut (1996) and Mitofsky (1996), have somehow concluded that not many participants changed policy views. Nondeliberative polls, on this account, were vindicated. The NIC, at great expense, produced little attitude change. Fortunately for the NIC, however, this is simply untrue.

How untrue depends on the species of gross change considered. Gross change of *position*—counting all movement between scale points (as from strong agreement to agreement or agreement to neutrality)—is very common. Depending on the item, the percentage changing position ranges mainly between 20% and 50%. A couple of items show percentages under 20%; a half-dozen come in over 50%. The lowest is 14.3%, the highest 53.4%. The grand mean is 36.2%. To save space, we simply summarize these results, referring interested readers to Fishkin and Luskin (1998) for the item-by-item account.

Of course, these figures may be partly nonattitudinal noise. Perhaps the movements between adjoining categories on the same side of the scale (as between agreement and strong agreement) do not represent real change. In that case, we might consider gross change of *side*, rather than position: change among agreement, disagreement, and neutrality (if an option). Now the percentages vary to some noticeable degree with the number of categories. For items with even numbers of categories (two or four), lacking a middle or neutral category, the percentages average 25.4%. For items with odd numbers of categories (three or five), on the other hand, where traffic in and out of neutrality is possible, the percentages run higher, dwelling mainly in the 30-40% range and averaging 34.9%. The grand mean over all policy items is 30.4% changing side. Though necessarily less common than changes of position, changes of side are still very common.¹⁰

The most draconian move would be to count only *complete change of side* (from agreement to disagreement or vice versa), disregarding any movement in or out of neutrality,

which seems to be what Kohut (1996) and Mitofsky (1996) are thinking of. For items with even numbers of categories, change of side and complete change are the same, so that the mean percentage changing sides completely remains 25.4%. For the items with odd numbers of categories, the percentage changing sides completely is markedly lower, averaging only 3.8%.

These last percentages may indeed seem low. But: (1) these are only roughly half the policy items, and even on these the percentages changing opinion run in the thirties or higher under any but this strictest of definitions. (2) Anyone who prefers this definition, which effectively reduces all items to dichotomies, ought for consistency to reject all multicategory attitude items, even trichotomies, as wasted effort. But of course the native dichotomies show much greater change, even under this definition. It is only by providing but then ignoring a middle category that these low numbers can be obtained. (3) Because this definition ignores all movement in and out of neutrality for items with odd numbers of categories, the percentage who could possibly have changed opinions in this sense must be under 100%, provided the neutral category is nonvacant at either t_1 or t_2 . A recalculation of the proportions changing side completely as a fraction of those who could possibly have done so (neither starting nor ending in the middle category) yields somewhat higher percentages, averaging roughly 6%. And (4) With just two categories as possibilities, most respondents will already be in the category representing their full-information preference at t_1 (50% can be expected to be there strictly by chance, even absent any t_1 information), which severely limits the amount of change that can be expected. It is hard to find plausible assumptions under which the gross complete change of side could much exceed the percentages obtained for the two- and four-category items, and in this light even the percentages obtained for the three- and five category items look very respectable (see Luskin and Globetti 1997, Luskin 2002a).

Next we consider the politically more interesting question of net changes of position. How much, that is, did the NIC participants move in one direction versus the other on our policies and empirical premises. How much did the balance of opinion shift, and in which direction?

By this criterion, our participants show statistically significant change (at the .05 level, by a two-tailed test) on 24, or roughly half, of our 47 policy items and four of ten of our empirical premises. Given that night-and-day net changes of position—with nearly everybody moving from one end of the scale to the other, and nearly everybody in the same direction—are exceedingly unlikely, these changes are also “substantively significant,” amounting to nontrivial fractions of the items’ ranges. For example, participants gave increasing support to the proposition that the U.S. should be willing to use militarily means like airstrikes or blockades “to solve problems in other countries,” averaging .24 higher on a five point scale (with, i.e., a four point range). Roughly ten percent more of them came to favor devolving the responsibility for safety net programs to the states and to agree that “the government in Washington” should just let each person get ahead on their [sic] own,” rather than “see[ing] to it that every person has a job and a good standard of living.” Fourteen percent fewer of them ended up favoring replacing the existing income tax structure with a flat tax.

There remains, however, the question of what portions of these changes are actually due to the participants’ experience of the NIC. The NIC data permit comparisons to the two quasi-

control groups described above. The first consists of those nonparticipants who were reinterviewed contemporaneously with the end of the NIC, the second of the independent RDD sample interviewed by telephone at the same time. If we denote the participant means by P , the nonparticipant means by N , and the RDD sample means by R and distinguish pre- and post-deliberation observations by numerical subscripts, we have so far been describing $P_2 - P_1$. But we should also consider the two discounted measures these quasi-control groups permit: $(P_2 - P_1) - (N_2 - N_1)$, the extent to which the participants changed more than the nonparticipants, and $P_2 - R_2$, the extent to which the participants, after participating, departed from a random sample never invited to do so.

The first comparison controls both for what Campbell and Stanley (1963) call *history*, meaning events that affect the whole population, including our participants, and for what we shall call *stimulation*, meaning the tendency for being interviewed and invited to participate to motivate people to learn and think more about the issues and thus perhaps to change their views, whether they actually participate or not. (The invitation to participate makes this a slightly larger possible confound than Campbell and Stanley's "testing.") The extreme case we worry about is the one in which $P_2 - P_1 = N_2 - N_1$, indicating that nonparticipants changed to the same degree and in the same direction as participants. To the extent that $P_2 - P_1$ is large but $(P_2 - P_1) - (N_2 - N_1)$ small, we suspect that either history or stimulation is responsible for the NIC's apparent effect. Unfortunately, it is difficult from this comparison to say which.

The second comparison, to $P_2 - R_2$, controls only for history, since the RDD sample members have not been previously interviewed (or therefore invited to participate). Here the extreme case we worry about is the one in which $P_2 = R_2$, indicating that the participants arrived at the same point, at t_2 , as a separate random sample (who had presumably also changed, in the same degree and direction, from t_1). To the extent that $P_2 - P_1$ is large but $P_2 - R_2$ small, we suspect that history alone is responsible for the NIC's apparent effect.

Two other cases are especially worth noting. One is the configuration most suggestive of an effect of deliberation proper. Here $P_2 - P_1$, $(P_2 - P_1) - (N_2 - N_1)$, and $P_2 - R_2$ are all large, indicating substantial change that cannot be ascribed simply to history or stimulation. The other is the case in which $P_2 - P_1$ is large, $(P_2 - P_1) - (N_2 - N_1)$ small, and $P_2 - R_2$ large. Here the participants have changed and look quite different from the separate random sample at t_2 , but the nonparticipants have changed similarly. The inference in this case is that the experiment had an effect but that the bulk of it owed to stimulation by the initial interview and invitation.

As can be seen in Table 3, the statistically significant policy attitude changes largely survive these comparisons. In the family policy domain, the participants change significantly ($P_2 - P_1$) on six of twelve items, the nonparticipants ($N_2 - N_1$) on only one of the six (plus one other). The participants change significantly more than the nonparticipants ($(P_2 - P_1) - (N_2 - N_1)$) on (Table 3 about here) three of the six (although the nonparticipants also change significantly more than the participants on the one item on which they do change significantly while the participants do not) and differ significantly from the separate random sample ($P_2 - R_2$) on five of the six (plus one other).

In the foreign policy domain, the participants change significantly on seven of fifteen items, the nonparticipants on none of those seven (but one other). The participants change significantly more than the nonparticipants on four of the seven, and differ significantly from the separate random sample on four of the seven (plus two others).

In the economic policy domain, the participants change significantly on ten of twenty items, the nonparticipants on none of those ten (but two others). The participants change significantly more than the nonparticipants on three of the ten (although the nonparticipants also change significantly more than the participants on the two items on which they do change significantly while the participants do not) and differ significantly from the separate random sample on eight of the ten (plus two others).

In sum, then, these results show considerable change among our participants. Some of the change seems to have resulted from the mere fact of having been interviewed and invited to attend, which by itself seems to have been enough to get them thinking, reading, listening, watching, and talking more about politics. Some but not much of it seems to have stemmed from the parade of great-world events to which they would have been exposed NIC or no. The greatest share, however, seems to have stemmed from the anticipation and experience of participation.

Information Gains

The third question is why the participants' policy attitudes changed as they did. It would be truly extraordinary for there not to be any net change, not to mention any gross change, anywhere in the data. Opinion formation—some respondents' transition from non- to genuine attitudes—will produce some change. Random variation in the translation from thoughts and feelings to response categories will produce some more. The question is whether there is anything more, and what accounts for it, if so. One stratum of answer is suggested by the previous section: We have already seen that the observed net changes seem to have been the result partly of stimulation by the initial interview and invitation and partly of deliberating and preparing to do so (but not particularly of the mere passage of events to which everybody, the NIC participants included). But what was it in the stimulation or participation that mattered? Was it because the participants (and in some degree the nonparticipants) gained information? Because they acquired particular pieces or kinds of information? Because they thought more seriously about the information they already had? Because they acquired new perspectives on the information they already had, perhaps from others in their small groups? Because they were swayed by the balance of opinion in their small groups?

We are still in the process of sorting these matters out, but the variables that lie closest to the heart of the Deliberative Poll's rationale have to do with cognitive involvement, with thought and information. Some nice distinctions separate *information* from *good information* or *knowledge* and knowledge from *understanding* or *sophistication*, which in practice are all highly correlated (Luskin 2002a). For present purposes, we consider only factual knowledge.

The NIC questionnaire had nine purely factual items. Respondents were asked to say roughly what percentages of American children are “born out of wedlock” and of American families are receiving AFDC; to say whether the U.S. “conduct[s] the most international trade” with Canada, Japan, Germany, or Mexico”; to say whether the U.S. had sent ground troops in the past decade to Bosnia, Somalia, Rwanda, and Iraq; to say whether the U.S. spends most on the space program, foreign aid, the national defense, or agricultural subsidies; and to estimate “the average unemployment rate in the United states over the past ten years. There was also a question asking respondents to place the two major political parties on a seven-point liberal conservative scale, from which two further items may be derived.

For the open-ended items like the percentage of children born out of wedlock, it would be foolish to insist on accuracy to the nearest percent (at which level even expert estimates may vary). It is enough to have a rough idea. Thus we count respondents as having the right answer if they peg the percentage of American children born out of wedlock between 25 and 40 percent, the percentage of American families receiving AFDC between 1 and 10 percent, and the average unemployment rate in the US over the past ten years between 5 and 10 percent. For the other items, the right answers are clearer-cut. The nation of those listed with which the US conducts the most international trade is Canada; the US had sent ground troops during the preceding ten years to Somalia and Iraq but not Rwanda; and the category of those listed on which the US spends the most is the national defense. Only somewhat more debatably, the Republicans are somewhere to the right of the midpoint, the Democrats somewhere to the left of it. The only further nicety is that respondents who said in the initial interview that the US had sent ground troops to Bosnia were counted as correct if interviewed after December 15, 1995, when the UN Security Council voted to send troops and NATO began deploying advance units, but incorrect if interviewed before then.

On the whole, our participants gained information. On eight of these eleven items, they gained significantly, the only exceptions being the percentages of American children born out of wedlock and of American families receiving AFDC and the identity of the U.S.'s highest volume trading partner. The remaining eight items averaged an 11.4% increase in the percentage of

(Table 4 about here)

correct answers. The smallest significant increase, in the percentage knowing roughly what the unemployment rate in the U.S. had averaged over the preceding decade, was approximately four percent. Two items, the budget category the U.S. spends most on and whether U.S. ground forces had been sent to Bosnia showed roughly 20% gains. Table 4 gives the details.

Though not the largest of these gains, those of roughly eight percent apiece in knowledge of where the Democratic and Republican parties stand on the liberal-conservative dimension are especially striking, because nothing in the briefing materials alluded to such terms, nor did the moderators use in them in framing the discussion. Yet, somehow, either from the small group discussions, or from conversation and reading in the period preceding the on-site deliberations, some notable fraction of the participants who had previously not had much notion of how such abstractions applied to the parties acquired some. Quite arguably, this sort of knowledge, of where the parties stand on policy or ideological dimensions, is the most important for citizens to have in performing what is for most them their main civic task, voting intelligently.

It is also worth noting that the right answers to several of the factual questions received could not have been gleaned from the briefing materials, which made no mention of the percentage of families receiving AFDC, the average unemployment rate, or which country is our main trading partner; said nothing about liberalism or conservatism or the parties' positions on the liberal-conservative dimension; and while mentioning Iraq, Somalia, and Bosnia as "trouble spots" did not discuss American military involvement there. Though possibly unfortunate, this lack of coordination between the briefing materials and questionnaire makes the information gains in Table 4 all the more impressive.

The nonparticipants, however, also learned, registering significant information gains on six of the same eleven items (four of the eight on which the participants gained significantly plus two others). The $(P_2 - P_1) - (N_2 - N_1)$ comparison, consequently, is significant for only three items (the percentage of children born out of wedlock, the average unemployment rate in the U.S. over the past ten years, and the category on which the U.S. spends the most), and on one of those three (the percentage of children born out of wedlock) it was the nonparticipants who gained more.

What is happening here? People may be expected to pick up some information during election campaigns, but the NIC was held in the early dawn of the presidential nomination campaign, before the Iowa caucuses and New Hampshire primary. During the interval from initial interview to deliberation, the U.S. did reach the decision to send troops to Bosnia, and so there may well have been some learning on that score, but generally speaking there is not much reason to suspect history as the source of the nonparticipants' (and participants') information gains. Much more likely, the nonparticipants were stimulated by the initial interview and invitation to attend the NIC to think, talk, and learn more about politics. Quite plausibly, they learned, thought, and deliberated off-site. Comparing the participants to the separate RDD sample buttresses this assessment. Five of the eleven $P_2 - R_2$ differences are statistically significant (as compared to eight of the $P_2 - P_1$ differences). In short, the participants generally learned something, only a little more than the nonparticipants, but greatly more than the general public.

But did the nonparticipants really learn nearly as much as the participants? That is very hard to believe. The widespread and sizable differences between participants and nonparticipants in their changes of policy views is telling. Clearly the participants are getting a larger dose of something. Part of that something may of course be adventitious, having nothing to do with learning in any meaningful sense. But almost certainly much of it does involve learning in a broad sense. Some of the learning—for example, the increased appreciation of other people's interests alluded to at the end—may be of a sort our information items are simply not geared to capture. Some, too, may be of the sort our information items are geared to capture but elude them. We do suspect that our measures, inevitably based on a minuscule fraction of all relevant information, are not picking up all of what the participants learned but the nonparticipants did not.

Setting the Record Straight

From its inception, the Deliberative Poll has been the target of frequently overheated criticism from portions of the commercial polling community and a handful of academics closely tied to it (Ladd xxxx; Traugott xxxx, Mitofsky xxxx, Worcester xxxx). Even the more measured and thoughtful assessment by Merkle (1996) is visibly flavored by the author's position in the commercial polling community.

The reaction is understandable. Elements of the rationale of the Deliberative Poll, if accepted, place important boundaries on the utility of ordinary polls. If major fractions of the average sample have non- or minimal attitudes on most policy issues, we must accept the very real possibility that their views would change with greater thought or information and must correspondingly regard the distributions of opinion in ordinary polls in a much more qualified light. They measure public opinion as it is, and in many cases quite well, but to what extent and in what sense should that carry weight as the "voice of the people"?

This observation also accords with the very different reception the project has had within the academic survey community, where it has excited little hostility and much interest and support.¹¹ Within academia, almost everyone studying public opinion now accepts that most people know and have thought extremely little about politics and that most attitudes on most issues are correspondingly wispy. For academics qua academics, with no interest in selling polls, there is nothing inconvenient about this, although it may disquiet them as democrats. Indeed, the individual variation in political information, attitude development, and the like become interesting conditioning variables in explanations of policy and candidate preferences, and the Deliberative Poll an interesting way of tackling this conditioning quasi-experimentally.

Let us briefly address a few points in the composite critique. A particularly obtuse but persistent criticism is effectively that the Deliberative Poll is not a (nondeliberative) poll (Traugott 1992, Mitofsky 1996b). Its participants, by the end, are no longer a representative sample. Of course they aren't; the whole point is that they shouldn't be. Similarly, Mitofsky (1996b) complained that the NIC could never capture the actual nomination or election campaign. It wasn't meant to. It was meant to influence the campaign, not anticipate it.

Another persistent criticism has been the possibility of a Hawthorne effect (Ladd xxxx, Adair 1996). As we said while the NIC was still only gestating,

We should be so lucky. The classic Hawthorne experiments involved workers who became more efficient or productive workers, as a result merely of the fact that they knew they were being observed. In the same way, our participants, to the extent they are keenly aware of being studied, should learn all the more about and become all the more interested in politics (Fishkin and Luskin 1996).

Merkle (1996) segues from Hawthorne effects, which he rightly dismisses, to demand characteristics, which, following Adair (1996), he thinks problematic. He quotes the second sentence from a passage in which we say,

In the typical social psychology experiment, the true purpose is hidden. Subjects, sensing that they have not been told the true purpose, try to divine it and to behave accordingly. In contrast, we have no secrets about our purposes. We want the participants to learn, think, and talk about politics. They know that. If they try to oblige us, so much the better (Fishkin and Luskin 1996).

Then, missing the point, he adds, “A number of ways have been suggested to control demand characteristics ..., but making the purpose of the study clear to the participants is not one of them” (Merkle 1996: 609). Demand characteristics would indeed be a problem if the purpose we were making clear were to induce attitude change. It isn’t—as Merkle’s own language seems to recognize. Instead, he describes the purpose—correctly as far as he goes—as being “to determine *if* the NIC experience would lead to opinion change” (1996: 610, italics ours). But that is not the whole or most visible purpose. The more basic and more visible purpose, which he misses, is to expose the participants to relevant information and provide them with opportunities and incentives to think through and discuss the issues. This is the most evident “demand,” the only one stressed in our dealings with participants. It is a “demand” we hope they do recognize and fulfill.

Merkle’s (1996) critique, as previously noted, is far the most careful and judicious of the relatively critical reviews the project has received. It contains some nice insights. Yet it still betrays some jaundice. While correctly noting that Deliberative Polling is best viewed as quasi-experiment, Merkle continues to compare NIC sample to survey rather than experimental samples, making much of participant-nonparticipant differences that in the experimental context are utterly trivial. He allows as how the NIC’s 72% interview rate is “respectable,” while never mentioning that even the highest quality commercial polls seem not to get anywhere near that.

In the end, Merkle concedes that “upon close inspection, the methodology used by the NIC is comparable in many ways to the methodologies used by social science researchers every day.” But then he adds, straightfacedly, that “because the NIC claims a higher purpose than traditional polling, it is appropriate to hold it to a different standard” (Merkle 1996: 616). In other words, Deliberative Polling, held to the same standards as ordinary commercial polling, looks pretty good, so, well, obviously, we’d better hold it to higher ones. Please. Science is science, and ours should be held to the same standard as anyone else’s. We have acknowledged the design’s vulnerabilities as well as its strengths, but, on the whole, as we have argued here, it comes out very well.

Conclusions

We have tried in these pages to sketch some of the main features of the Deliberative Polling process and results, with particular reference to the one national Deliberative Poll so far in the U.S., the National Issues Convention of January, 1996. The analyses here do not by any means exhaust the riches of these data. We have said nothing, for one thing, of the third-wave measurement of the same respondents some ten months later, just after the presidential election. We leave this aside for now as less central. It is the t_1 - t_2 comparison that speaks to the NIC's *effect*. The t_2 - t_3 comparison has to do with the effect's *persistence*, another matter. To preview these results, the effect decays. On average, the participants revert back—though not all the way back—to t_1 positions. This is hardly surprising. Ordinary citizens, back home from the special experience and incentives of the NIC, return, in some measure, to being ordinary citizens. What matters for the NIC is the glimpse of their opinions when they are led, even if only transitorily, to become more involved in and informed about politics than before (or, in lesser degree, after).

Most importantly, we shall need, perhaps with the aid of more elaborate designs or at least more direct measurement, to sift the effects of deliberation—entailing discussion, rumination, and learning—from one another. One reason this task looms as important is that we can see the glintings in our results so far of two opposing mechanisms driving the policy attitude changes among the participants. One is increased efficiency in pursuit of unchanging interests, unchangingly viewed. To the extent that this first mechanism prevails, participation should strengthen the interest-attitude relationship. The other mechanism involves changes in—broadenings of—the interests to which participants attend. They may come to empathize with groups to which they do not belong or by much the same token to think more about the interests of more inclusive groups to which they do belong, like the community or the nation. To the extent that this second mechanism prevails, participation should weaken the interest-attitude relationship. Our hunch is that rumination and learning by themselves particularly feed the first mechanism, while deliberation (though almost inevitably accompanied by rumination and learning) feeds the second.

Table 1
Differences between Participants and Nonparticipants

Variable	Min	Max	Mean P1	Mean N1	P1-N1	Prob
Socio-Demographic Background						
Age	19	94	42.40	48.89	-6.69	
Sex	1	2	1.53	1.56	-0.03	0.3095
Number of children	0	6	0.94	0.73	0.21	0.0109
Education	1	13	7.51	6.20	1.31	0.0001
R Hispanic or Latino	0	1	0.07	0.04	0.03	0.0501
Total family income, all sources	1	22	15.36	14.14	1.22	0.0006
Finances increased or decreased	1	3	2.15	2.13	0.02	0.7517
Political Participation						
Frequency R watches news	1	7	5.95	5.96	-0.01	0.8937
Frequency R watches pub affairs	1	7	4.79	4.49	0.30	0.0119
Frequency R reads news	1	7	5.42	5.06	0.36	0.0072
R's interest in politics	1	4	3.08	2.75	0.33	0.0001
Frequency R discusses politics	1	4	2.85	2.33	0.52	0.0001
Worked on political campaigns	1	4	1.40	1.23	0.17	0.0002
Seen/talked with politician	0	1	0.57	0.36	0.21	0.0001
Given money to party	0	1	0.31	0.17	0.14	0.0001
Did R vote for president in 1992	0	1	0.78	0.67	0.11	0.0001
Political Orientations						
Liberal/conservative ID	1	7	4.21	4.45	-0.24	0.0196
Party identification	1	7	2.28	2.49	-0.21	0.0536

Table 1 Cont'd

Variable	Min	Max	Mean P1	Mean N1	P1-N1	Prob
Civic Attitudes						
I have no say in government	1	4	2.37	2.45	-0.08	0.2617
Officials care what I think	1	4	2.28	2.25	0.03	0.6010
Politics too complicated	1	4	2.55	2.81	-0.26	0.0001
Politicians out of touch	1	4	2.99	2.97	0.02	0.6859
I have worthy political opinions	1	4	3.26	3.03	0.23	0.0001
Express opinion vs. keep quiet	1	3	1.19	1.24	-0.05	0.1151
Values						
Religion vs. humanism	0	1	0.38	0.40	-0.02	0.6074
America's interest first	1	2	1.41	1.33	0.08	0.0133
Respect authority	1	2	1.27	1.23	0.04	0.1962
Property rights	1	2	1.53	1.41	0.12	0.0004
Economic equality	1	2	1.83	1.86	-0.03	0.1610
Sacrifice for future	1	2	1.42	1.50	-0.08	0.0168
Morality in foreign affairs	1	2	1.92	1.92	0.00	0.8191
Husband works/wife stays home	1	4	2.02	2.30	-0.28	0.0002
One parent stay at home w/ children	1	4	2.86	2.99	-0.13	0.0397
Religion						
Attendance at religious services	0	8	3.91	3.88	0.03	0.8868
Frequency of prayer	0	6	3.92	3.98	-0.06	0.6292
Member religious right	0	1	0.09	0.13	-0.04	0.1231

Table 1 Cont'd

Variable	Min	Max	Mean P1	Mean N1	P1-N1	Prob
Policy Attitudes						
Family						
Govt too involved with family	1	2	1.53	1.55	-0.02	0.5451
Make divorce harder	1	3	1.51	1.60	-0.09	0.0758
Pressure fathers for child support	1	3	2.66	2.62	0.04	0.3084
Moral education in schools	1	3	2.54	2.64	-0.10	0.0143
More family planning services	1	3	2.33	2.44	-0.11	0.0284
Govt help with childcare	1	3	2.29	2.35	-0.06	0.2182
Govt assistance to poor families	1	3	2.26	2.32	-0.06	0.1915
Limit sex and violence on TV&movies	1	3	2.35	2.46	-0.11	0.0386
Allow school prayer	1	3	2.17	2.35	-0.18	0.0016
Provide sex education in schools	1	3	2.39	2.35	0.04	0.3256
Employers facilitate time w/ children	1	4	2.93	2.93	0.00	0.9994
Foreign Policy						
U.S. send troops	1	5	2.71	2.70	0.01	0.7762
U.S. take other military action	1	5	2.75	2.72	0.03	0.5681
U.S. cooperate militarily	1	4	2.90	2.95	-0.05	0.3587
U.S. better off staying home	1	4	1.99	2.19	-0.20	0.0033
Prevent spread of weapons	1	3	2.67	2.62	0.05	0.2697
Improve global environment	1	3	2.55	2.53	0.02	0.5107
Insure adequate energy supplies	1	3	2.50	2.58	-0.08	0.0396
Promote democracy	1	3	1.93	1.91	0.02	0.6574
Aid U.S. business interests	1	3	1.95	2.03	-0.08	0.0688
Famine relief and medical help	1	3	2.16	2.19	-0.03	0.4995

Table 1 Cont'd

Variable	Min	Max	Mean P1	Mean N1	P1-N1	Prob
Promote human rights	1	3	2.12	2.05	0.07	0.1390
Protect weaker nations	1	3	2.03	2.03	0.00	0.9613
Promote economic development	1	3	2.08	2.05	0.03	0.4523
Most imp. shift resources back home	0	1	0.78	0.77	0.01	0.8127
Most imp. bring order to world	0	1	0.08	0.09	-0.01	0.8859
Most imp. promote democracy	0	1	0.14	0.14	0.00	0.8653
Use foreign aid for what	1	3	2.39	2.37	0.02	0.6315
Economic Policy						
Spending on environment	1	3	2.42	2.34	0.08	0.1034
Spending on Medicare/Medicaid	1	3	2.19	2.31	-0.12	0.0290
Spending on law enforcement	1	3	2.45	2.48	-0.03	0.4633
Spending on drug rehab	1	3	2.34	2.36	-0.02	0.7581
Spending on education	1	3	2.71	2.68	0.03	0.3057
Spending on national defense	1	3	1.72	1.79	-0.07	0.1463
Spending on foreign aid	1	3	1.40	1.39	0.01	0.8063
Spending on welfare	1	3	1.53	1.56	-0.03	0.5650
Spending on Social Security	1	3	2.38	2.44	-0.06	0.2458
Importance of reducing deficit	1	4	3.51	3.41	0.10	0.0375
Raise taxes vs. cut spending	1	3	2.28	2.24	0.04	0.2993
Retraining responsibility of govt	0	1	0.12	0.14	-0.02	0.2972
Retraining responsibility of employers	0	1	0.40	0.40	0.00	0.9712
Retraining responsibility of individual	0	1	0.49	0.46	0.03	0.5040
Federal vs. state welfare&healthcare	1	2	1.56	1.62	-0.06	0.0808
Govt guarantee jobs	1	2	1.35	1.39	-0.04	0.1942

Table 1 Cont'd

Variable	Min	Max	Mean P1	Mean N1	P1-N1	Prob
Graduation of taxes	1	3	2.08	2.00	0.08	0.1848
National sales tax	0	1	0.35	0.42	-0.07	0.1023
Tax reduction for savings	0	1	0.78	0.79	-0.01	0.8119
Flat tax	0	1	0.51	0.56	-0.05	0.1871
Keep current tax system	0	1	0.54	0.62	-0.08	0.0261
NAFTA was a mistake	1	4	2.48	2.60	-0.12	0.1252
Empirical Premises						
Breakdown in traditional family	1	2	1.29	1.26	0.03	0.2558
Econ pressure vs. value breakdown	1	2	1.62	1.56	0.06	0.0621
Welfare root of family problems	1	4	2.56	2.74	-0.18	0.0068
Improving economy helps families	1	4	3.25	3.32	-0.07	0.1935
Working mom does as well w/ kids	1	4	3.04	3.08	-0.04	0.5392
Economy needs good profits	1	4	3.09	3.24	-0.15	0.0078
Avg worker doesn't get fair pay	1	4	2.83	2.88	-0.05	0.5208
Achievement depends on abilities	1	4	3.16	3.19	-0.03	0.5723
Economy needs foreign trade	1	4	3.08	3.04	0.04	0.5580
Low trade barriers lose jobs	1	4	2.72	2.85	-0.13	0.0428
Free trade good for everyone	1	4	2.84	2.94	-0.10	0.1044
Knowledge						
Fact: Wedlock	0	1	0.33	0.24	0.09	0.0034
Fact: AFDC	0	1	0.09	0.04	0.05	0.0033
Fact: Trade	0	1	0.16	0.17	-0.01	0.6902
Fact: Iraq	0	1	0.81	0.80	0.01	0.6695
Fact: Somalia	0	1	0.71	0.66	0.05	0.0925

Table 1 Cont'd

Variable	Min	Max	Mean P1	Mean N1	P1-N1	Prob
Fact: Rwanda	0	1	0.25	0.22	0.03	0.3507
Fact: Bosnia	0	1	0.71	0.69	0.02	0.5471
Fact: Which U.S. spends most on	0	1	0.38	0.26	0.12	0.0001
Fact: Unemployment rate	0	1	0.46	0.39	0.07	0.0378
Rep Lib/Cons placement	0	1	0.62	0.46	0.16	0.0001
Dem Lib/Cons placement	0	1	0.59	0.44	0.15	0.0001

* P1 and N1 denote means for participants and nonparticipants based on initial interview; Prob is the two-tailed p-value for P1-N1 under the null hypothesis of no difference. All variables are scored so that higher numbers indicate greater agree, extent, or quantity.

Table 2
Differences Between Participants and Nonparticipants, Nominal Items

Variable	# Categories	Chi-Square Statistic	Prob
Marital Status	6	19.073	0.002
Employment Status	8	52.017	0.001
Type of Job Held	17	8.860	0.919
Race	6	7.988	0.157
Source R Gets News From	4	7.710	0.052
1992 Vote	4	6.622	0.085
Religious Preference	6	2.389	0.793
Interpretation of Bible	3	16.190	0.001

Table 3
Net Changes of Position in Policy Preference

Variable	Min	Max	P1	P2	P3	P2-P1	P3-P2	P3-P1	N'1	N'2	N'2-N'1	(P2-P1)- (N'2-N'1)	R2	P2-R2
Family Policies														
Govt too involved w/ family	1	2	1.53	1.54	1.49	0.01	-0.05	-0.04	1.55	1.50	-0.05	0.06	0.03	0.03
Make divorce harder	1	3	1.51	1.76	1.67	0.25**	-0.09**	0.16**	1.60	1.60	0.00	0.25**	0.22**	0.22**
Pressure fathers for child support	1	3	2.66	2.85	2.67	0.19**	-0.18**		2.62	2.75	0.13**	0.06	0.16**	0.16**
Moral educ in the schools	1	3	2.54	2.59		0.05			2.64	2.59	-0.05	0.10	0.03	0.03
More family planning services	1	3	2.33	2.52		0.19**			2.44	2.38	-0.06	0.25**	0.13**	0.13**
Govt help with childcare	1	3	2.29	2.51	2.36	0.22**	-0.15**	0.07**	2.35	2.42	0.07**	0.15**	2.40	0.11*
Govt assistance to poor families	1	3	2.26	2.34	2.28	0.08*	-0.06		2.32	2.31	-0.01	0.09	-0.01	-0.01
Limit sex&violence on TV&movies	1	3	2.35	2.32	2.24	-0.03	-0.08*	-0.11**	2.46	2.45	-0.01	-0.02	-0.06	-0.06
Allow school prayer	1	3	2.17	2.14	2.15	-0.03	-0.02		2.35	2.30	-0.05	0.02	-0.09	-0.09
Employers facilitate time w/ kids	1	4	2.93	3.24	3.12	0.31**	-0.12**	0.19**	2.93	2.97	0.04	0.27	0.23**	0.23**
Foreign Policies														
U.S. send troops	1	5	2.71	2.86	2.95	0.15**	0.09**	0.24**	2.70	2.80	0.10	0.05	0.10*	0.10*
U.S. take other military action	1	5	2.75	2.98	3.00	0.23**		0.25**	2.72	2.78	0.06	0.17	2.90	0.08
U.S. cooperate militarily	1	4	2.90	3.18	3.18	0.28**	0.00	0.28**	2.95	3.00	0.05	0.23*	0.09	0.09
U.S. better off staying home	1	4	1.99	1.75	1.73	-0.24**	-0.02	-0.26**	2.19	2.07	-0.12	-0.12*	-0.21**	-0.21**
Prevent spread of weapons	1	3	2.67	2.67	2.65	0.00	-0.02	-0.02	2.62	2.50	-0.12	0.12	-0.01	-0.01
Improve the global environment	1	3	2.55	2.45		-0.10**			2.53	2.55	0.02	-0.12**	-0.14**	-0.14**
Insure adequate energy supplies	1	3	2.50	2.52	2.53	0.02			2.58	2.63	0.05	-0.03	-0.12**	-0.12**
Promote democracy	1	3	1.93	1.95	1.99	0.02*			1.91	1.94	0.03*	-0.01	2.00	-0.05
Aid U.S. business interests	1	3	1.95	2.16	1.96	0.21**	-0.20**		2.03	2.01	-0.02	0.23**	2.00	0.16**
Famine relief and medical help	1	3	2.16	2.12	2.22	-0.04	0.10**		2.19	2.15	-0.04	0.00	-0.13**	-0.13**

Table 3 Cont'd

Variable	Min	Max	P1	P2	P3	P2-P1	P3-P2	P3-P1	N'1	N'2	N'2-N'1	(P2-P1)-		
												N'2-N'1	R2	P2-R2
Promote human rights	1	3	2.12	2.12	2.19	0.00	0.07*	0.07*	2.05	2.07		-0.02	-0.01	
Protect weaker nations	1	3	2.03	2.07	2.16	0.04	0.09**	0.13**	2.03	2.05		0.02	-0.04	
Promote economic development	1	3	2.08	2.10		0.02			2.05	2.06		0.01	-0.05	
Most imp. shift resources home	0	1	0.78	0.76		-0.02			0.77	0.75	-0.02	0.00	0.01	
Most imp. bring order to world	0	1	0.08	0.11		0.03			0.09	0.08	-0.01	0.04	0.04	
Most imp. promote democracy	0	1	0.14	0.12		-0.02			0.14	0.17	0.03	-0.05	-0.06*	
Use foreign aid for what	1	3	2.39	2.57	2.58	0.18**	0.01	0.19**	2.37	2.40		0.15	0.12**	
Economic Policies														
Spending on environment	1	3	2.42	2.44	2.34	0.02	-0.10	-0.08	2.34	2.37		-0.01	0.08	
Spending on Medicare/Medicaid	1	3	2.19	2.16	2.13	-0.03	-0.06		2.31	2.31	0.00	-0.03	0.00	
Spending on law enforcement	1	3	2.45	2.46	2.40	0.01	-0.06*	-0.05	2.48	2.45	-0.03	0.04	-0.06	
Spending on drug rehab	1	3	2.34	2.33	2.35	-0.01			2.36	2.32	-0.04	0.03	-0.03	
Spending on education	1	3	2.71	2.85	2.70	0.14**	-0.15**	-0.01	2.68	2.70		0.12**	0.12**	
Spending on national defense	1	3	1.72	1.67	1.85	-0.05	0.18**	0.13*	1.79	1.78	-0.01	-0.04	-0.06	
Spending on foreign aid	1	3	1.40	1.59	1.61	0.19**	0.21**		1.39	1.37	-0.02	0.21**	0.16**	
Spending on welfare	1	3	1.53	1.52	1.62	-0.01	0.10**	0.09**	1.56	1.62		-0.07	-0.11*	
Spending on Social Security	1	3	2.38	2.27	2.34	-0.11**	0.07*	-0.04	2.44	2.43	-0.01	-0.10	-0.10*	
Importance of reducing deficit	1	4	3.51	3.58	3.49	0.07**	-0.09*	-0.02	3.41	3.46		0.02	0.15**	
Raise taxes vs. cut spending	1	3	2.28	2.30	2.38	0.02*	0.08**	0.10**	2.24	2.39	0.15**	-0.13**	2.40	
Retraining responsibility govt	0	1	0.12	0.21		0.09**			0.14	0.12	-0.02	0.11	0.13	
Retraining responsibility of employers	0	1	0.40	0.41		0.01			0.40	0.35	-0.05	0.06	0.42	
Retraining responsibility of individual	0	1	0.49	0.39		-0.10**			0.46	0.53	0.07	-0.17*	0.45	
Federal vs. state welfare&healthcare	1	2	1.56	1.66	1.67	0.10**	0.11**		1.62	1.75		-0.03	0.01	
Government guarantee jobs	1	2	1.35	1.46	1.38	0.11**	-0.08		1.39	1.36	-0.03	0.14	0.07*	

Table 3 Cont'd

Variable	Min	Max	P1	P2	P3	P2-P1	P3-P2	P3-P1	N'1	N'2	N'2-N'1	(P2-P1)-	
												N'1	N'2-N'1
Graduation of taxes	1	3	2.08	2.11	1.94	0.03*	-0.17**	-0.14	2.00	1.88	-0.12**	0.15**	0.06
National sales tax	0	1	0.35	0.29	0.37	-0.06*	0.08*		0.42	0.40	-0.02	-0.04	-0.07
Tax reduction for savings	0	1	0.78	0.91	0.80	0.13**	-0.11**		0.79	0.76	-0.03	0.16	0.17**
Flat tax	0	1	0.51	0.37	0.41	-0.14**	-0.10**		0.56	0.48	-0.08	-0.06	-0.13**
Keep current tax system	0	1	0.54	0.63	0.62	0.09**	-0.01	0.08**	0.62	0.60	-0.02	0.11	0.02
NAFTA was a mistake	1	4	2.48	2.55	2.62	0.07	0.14*		2.60	2.67		0.00	0.04
Empirical Premises													
Breakdown in traditional family	1	2	1.29	1.28	1.31	-0.01			1.26	1.29		-0.04	0.03
Econ pressure vs. value breakdown	1	2	1.62	1.48	1.54	-0.14**		-0.08**	1.56	1.50	-0.06	-0.08	-0.13**
Welfare root of family problems	1	4	2.56	2.52	2.49	-0.04	-0.03	-0.07	2.74	2.59	-0.15	0.11	0.07
Improving economy helps families	1	4	3.25	3.41	3.32	0.16**	-	0.07*	3.32	3.50	0.18*	-0.02	0.07
Economy needs good profits	1	4	3.09	3.19	3.25	0.10*		0.16**	3.24	3.23	-0.01	0.11	-0.05
Avg worker doesn't get fair pay	1	4	2.83	3.12	2.90	0.29**	-0.22**	0.07*	2.88	2.78	-0.10	0.39**	0.37**
Achievement depends on abilities	1	4	3.16	3.17	3.23	0.01		0.07*	3.19	3.22		-0.02	-0.05
Economy needs foreign trade	1	4	3.08	3.10	3.17	0.02	0.07*	0.09*	3.04	2.96	-0.08	0.10	-0.09
Low trade barriers lose jobs	1	4	2.72	2.92	2.74	0.20**	-0.18**		2.85	2.75	-0.10	0.30*	0.15**
Free trade good for everyone	1	4	2.84	2.95	2.90	0.11*	-0.05		2.94	2.90	-0.04	0.15	-0.03

* Statistically significant at the .05 level

** Statistically significant at the .01 level

Note: P1 and P2 denote the participants' means before and after the NIC; P3 denotes the participants' means at Time 3. N'1 and N'2 denote the reinterviewed nonparticipants' means before and after the NIC; R2 denote the RDD means after the NIC. All variables are scored so that higher numbers indicate greater agreement, extent, or quantity.

Table 4
Changes in Knowledge

Variable	P1	P2	P3	N'1	N'2	P2-P1	P3-P2	P3-P1	N'2-N'1	(P2-P1)-		
										N'2-N'1	R2	
Proportion Correct												
Individual Items												
Fact: Wedlock	0.33**	0.40**	0.40**	0.23**	0.43**	0.07*	-0.00	0.07	0.20**	-0.13*	0.38**	0.02
Fact: AFDC	0.09**	0.15**	0.14**	0.05**	0.09**	0.06**	-0.01	0.05**	0.04	0.02	0.12**	0.03
Fact: Trade	0.16**	0.16**	0.24**	0.16**	0.19**	-0.00	0.08**	0.08**	0.03	-0.03	0.20**	-0.04
Fact: Iraq	0.81**	0.82**	0.92**	0.81**	0.87**	0.01	0.10**	0.11**	0.05	-0.04	0.83**	-0.01
Fact: Somalia	0.71**	0.77**	0.90**	0.70**	0.83**	0.06*	0.13**	0.19**	0.13**	-0.07	0.77**	-0.00
Fact: Rwanda	0.25**	0.39**	0.36**	0.21**	0.33**	0.14**	-0.03	0.11**	0.12**	0.02	0.41**	-0.02
Fact: Bosnia	0.71**	0.91**	0.95**	0.72**	0.95**	0.21**	0.04*	0.24**	0.24**	-0.03	0.97**	-0.06**
Fact: Which U.S. spends most on	0.38**	0.58**	0.51**	0.31**	0.36**	0.20**	-0.07**	0.13**	0.05	0.15**	0.43**	0.15**
Fact: Unemployment rate	0.46**	0.50**	0.54**	0.40**	0.16**	0.04	0.04	0.08*	-0.24**	0.28**	0.49**	0.01
Rep lib/cons placement	0.62**	0.69**	0.74**	0.52**	0.58**	0.07**	0.05	0.12**	0.06	0.01	0.62**	0.07*
Dem lib/cons placement	0.59**	0.79**	0.64**	0.48**	0.65**	0.20**	-0.15**	0.05	0.17**	0.03	0.69**	0.10**
Summary Indices												
Lib/Cons Placement	0.61**	0.74**	0.69**	0.50**	0.62**	0.13**	-0.05**	0.08*	0.12**	0.01	0.65**	0.09**
Family Issue Facts	0.21**	0.28**	0.27**	0.14**	0.26**	0.06**	-0.01	0.06**	0.12**	-0.06	0.25**	0.03
Foreign Issue Facts	0.53**	0.61**	0.67**	0.52**	0.63**	0.08**	0.06**	0.14**	0.11**	-0.03	0.64**	-0.03*
Economic Issue Facts	0.42**	0.54**	0.53**	0.36**	0.26**	0.12**	-0.01	0.11**	-0.10**	0.22**	0.46**	0.08**
Information Summary	0.43**	0.52**	0.55**	0.40**	0.47**	0.09**	0.03**	0.12**	0.07**	0.02	0.51**	0.01

* Statistically significant at the .05 level

** Statistically significant at the .01 level

Note: P1 and P2 denote the proportions of participants answering correctly before and after the NIC. N'1 and N'2 denote proportions of those nonparticipants reinterviewed after the NIC answering correctly before and after. R2 denotes the proportion of the separate RDD sample members answering correctly after the NIC.

 NOTES

*This paper was also previously presented at the annual meeting of the American Association for Public Opinion Research, Saint Louis, MO, May 14-17, 1998. Among others too numerous to cite individually, we especially wish to thank William Galston, Keith Melville, Robert Kingston, Karl Rove, Charls Walker, and Noelle McAfee for their help in making the National Issues Convention possible and Christopher Bratcher, Han Dorussen, Dennis Plane, Nedim Ogelman, and Jeffrey Ladewig for research assistance. We are also grateful to the members of the Review and Technical Review Committees, the former co-chaired by Bill Frenzl and the late Barbara Jordan, the latter chaired by Philip Converse. Deliberative Polling is a registered trademark, and the Center for at the University of Texas at Austin, of which Fishkin is Director and Luskin Research Director, is receiving fees from the trademark to fund research.

¹For further details of the National Issue Convention, see Fishkin and Luskin (1998).

²For further discussion, see Luskin (2002a).

³To be sure, the manipulation in a Deliberative Poll considered as quasi-experiment involves more than just exposure and incentives to pay attention to information. Most notably, it also involves deliberation—discussion with others, of possibly very different backgrounds and views—so that participants may be getting not only more information but a wider range of it than in their everyday lives. They may acquire new perspectives as well as new information narrowly defined. We return to these complications below.

⁴In the regional ones, it has been by telephone.

⁵In the British events, the rooms in which the policy experts were scheduled were too small to permit a single plenary session, so the participants were divided into three large groups, all of which had parallel question-and-answer sessions with the same panelists. There was just one, fully plenary session with the politicians, however.

⁶The three most recent national deliberative Polls, two in Australia and one in Denmark, have resorted to telephone interviewing for reasons of cost.

⁷Behind this surprisingly high figure lay some extraordinary efforts. In one case, NORC sent someone to milk a rural respondent's cow so that she could get away from her farm for the weekend.

⁸Though lower than we had had reason to expect (Fishkin and Luskin, 1996), at least partly because some of those who had said they were coming backed out at the last, in many cases deterred by winter storms that were complicating travel, this figure was still notably higher than in some dismissive forecasts (Mitofsky, 1996).

⁹Of course this is different only in degree from the case of the many “true” experiments whose experimental or control groups suffer attrition. Once self-selection or any other sort of nonrandom culling intrudes, there is effectively no longer fully random assignment. True control groups in the strictest sense are therefore rare.

¹⁰For di- and trichotomies, changes of side are the same as changes of position. For other items, the percentage changing sides must be less than or equal to the percentage changing positions.

¹¹The list of institutions more or less closely associated with the project includes SCPR (now NCSR) and NORC, as previously noted. The list of individuals, apart from the present authors, includes Norman Bradburn, Roger Jowell, Kenneth Rasinski, John Curtice, James A. Davis, Philip Converse, Henry Brady, Larry Bartels, Stanley Kelley, and the late Warren Miller.