

Deliberative Preferences in the Presidential Nomination Campaign:
Evidence from an Online Deliberative Poll^{*}

Shanto Iyengar, Stanford University

Robert C. Luskin, University of Texas at Austin

James S. Fishkin, Stanford University

November 29, 2005

Media coverage typically highlights the horserace and strategic aspects of political campaigns, at the expense of the candidates' policy positions and records (for a review of the evidence, see Iyengar, Norpoth, and Hahn 2004). These tendencies are especially pronounced during the early stages of nomination campaigns. With a relatively large field of sometimes unfamiliar candidates, the task of delineating platforms and records is even harder than in November. Reporters find it far easier to use poll numbers, fund-raising data, and attendance at campaign events to handicap the contest (Zaller 2004; Polsby 1983; Amundson, Lichter, and Noyes 1988). In contrast to the candidates' policy pronouncements, which are important but repetitive, the state of the horse race is *new* news.

This media focus makes the Iowa caucuses and New Hampshire primary, decided by small fractions of the voters in small, unrepresentative states, disproportionately influential. As the first real tests of public support, they generate a torrent of reports on the "winners," "losers," and trajectory of the race (Robinson and Sheehan 1983, Adams 1987, Busch and Mayer 2003). Donors, campaign workers, and voters respond. The winners surge; the losers fade. The 2004 Democratic primaries provided textbook examples. Over the space of just a few weeks, beginning with the Iowa caucuses, John Kerry rose from a self-financed, under-achieving candidate to the presumptive nominee, while Howard Dean fell from front-runner to hopeless also-ran.

For their part, the media are partly following voters, who, whatever they *say*, exhibit a strong behavioral preference for reading and hearing about the horse race, rather than policy issues (Iyengar, Norpoth, and Hahn 2004). The horse race is what attracts the most viewers and sells the most newspapers and magazines. But that is partly because most voters know so little about policy issues and the candidates' positions on them (Delli Carpini and Keeter 1996, Kinder

1998, Price 1999). Which is in turn a function of media behavior. Given what they see, hear, and read, all but the most determined policy information seekers must rely heavily on the candidates' sheer visibility and appearance of electability (Bartels 1988, Brady and Johnston 1987). It is, in short, a vicious circle.

The relative invisibility of issues in primary campaigns may also be attributed to the candidates, who often prefer to deemphasize their policy positions, which may not differ much from those of their co-partisan opponents (Aldrich 1980). Instead, they often campaign on the basis of background and personality, seeking to present themselves as likeable, competent, and the best bet to win in November. But in doing so they are partly responding to incentives set by media judgments of newsworthiness. It is another vicious circle.

These unfortunate tendencies have lately been exacerbated by “frontloading”: the concentration of primaries during the early stages of the campaign (Busch and Mayer 2003). The rapid-fire succession of contests leaves less time for reporters to attend to anything beyond the horse race implications of the most recent results, and the abbreviated length of the whole sequence leaves less time for candidates to repeat their policy positions often enough to reach many voters or for voters to absorb them.

The dearth and inconspicuousness of policy issues in primary campaigns raises serious normative questions. Primaries were adopted as a means of strengthening the responsiveness of the eventual nominee to the policy preferences of rank and file partisans (Polsby 1983). In practice, however, primary voting shows little policy basis. It is dominated instead by perceptions of the candidates' personal qualities and vote-getting prowess. Evaluations of the candidates tend to be “primitive”—lightly informed, undifferentiated, and therefore unstable and

relatively disconnected from more basic predispositions (Gelman and King 1993, Mayer 1996, Bartels 1988). As Brady and Johnston (1987, p. 184) put it,

Citizens ... learn too slowly about every aspect of the candidates except their viability. And, one of the major reasons that citizens learn quickly about viability is the enormous emphasis placed on the horse race by the media, especially right after the Iowa caucuses and the New Hampshire primary.

The implications for the outcome are similarly troubling. Either voters attach themselves to a front-runner whose early successes force competitors out of the race and who wins the nomination with little serious scrutiny, or, as in 2004, they gravitate to a relatively unknown candidate who does surprisingly well in the early contests. Neither scenario entails widespread comparisons of the candidates on policy issues or credentials.

One useful way of evaluating this state of affairs is to consider the impact of a counter-factually enhanced information environment. How might the candidates' standings in the polls and chances of achieving the nomination be altered if voters were given greater opportunity and incentive to learn, think about, and discuss the candidates' positions on the issues during the period of the Iowa caucuses and New Hampshire primary?

This paper addresses this question, using the first Deliberative Poll (DP) about a presidential nomination campaign to simulate a more informed and thoughtful public.¹ This particular DP was conducted online during the first few weeks of the campaign, with the deliberations running from mid-January to mid-February. The participants, meeting online, discussed the candidates' policy positions. They were given access to carefully balanced information, as well as the opportunity to seek clarification or explanation from a panel of non-partisan issue experts. This was quite different from the experience of voters exposed only to the real-world campaign. Policy information was more readily available, and discussions of the campaign were more frequent, focused more on policy information, were more balanced, and

involved more heterogeneous interlocutors. We estimate the effects of this experience on candidate preferences and the evaluative processes behind them.

Our results suggest that deliberating about the candidates' policy positions increases the weights voters give policy issues and thus affects their evaluations of the candidates. This in turn suggests that more substantive media coverage would produce more substantively oriented decision-making by voters—and, at least sometimes, depending on exactly how the relevant substantive considerations play out, produce different outcomes.

Design

We attempted to emulate face-to-face Deliberative Polling as closely as possible. Face-to-face DPs assemble random samples of citizens in a single location to deliberate over policy issues or electoral choices. The participants are sent balanced briefing materials laying out the principal arguments for and against the principal options. Then, on site, they deliberate in randomly assigned small groups led by trained moderators and put questions formed in the small group discussions to panels of experts, policy-makers, or politicians in plenary sessions. There have been two national-level face-to-face DPs in the U.S., five in the U.K., two in Australia, and one each in Denmark Bulgaria, and Hungary (see Fishkin and Luskin 1998 and Luskin, Fishkin, and Jowell 2002 for two reports and Fishkin and Luskin 2005 for an overview).

While face-to-face deliberation is fuller-channel and more vivid, online deliberation has some significant practical advantages. It costs substantially less, in man-hours and dollars. A random sample can be assembled online for a small fraction of the cost of assembling it in the flesh. From a participant's perspective, online deliberation reduces the inconvenience and opportunity costs of participation. Our online participants did not have to travel and did not have to give over whole weekends to the process. (For more on the advantages and disadvantages of

online versus face-to-face Deliberative Polling, see Luskin, Iyengar, and Fishkin 2003.) As a platform for administering surveys, moreover, the Internet has significant advantages over the telephone, including reduced random measurement error, increased differentiation of responses, and reduced social desirability response bias (Chang and Krosnick 2003).

Sampling and Randomization

The chief obstacle to replicating the face-to-face design online is that large numbers of Americans still lack home access to the Internet (U.S. Department of Commerce 2002). We overcome this “digital divide” by first drawing a random sample of voting-age citizens, not just Internet users, then providing those who are “offline” with free Internet access.

More precisely, we invited a random sample of 755 adult Americans from the 40,000-strong Knowledge Networks (KN) national panel to participate. The KN panel, recruited by telephone and offered free Web TV service as an incentive for regular participation in market research, is equivalent to a standard RDD sample (Chang and Krosnick 2003).² We also invited 247 veterans of a previous national online DP, also initially a random sample of the KN panel, who had spent four weeks (from mid-December 2002 through mid-January 2003) discussing American foreign policy, with no reference to the presidential nomination campaign or election (for further details, see Luskin, Iyengar, and Fishkin 2003). In all, 385 invitees could and agreed to participate, of whom 317 completed the pre-deliberation survey, and 284 the post-deliberation survey. This paper focuses on the Democrats and independents—roughly, the potential electorate for the Democratic nomination campaign—who were 180 (75 veterans and 105 novices) of the 284.

We included the veterans because we were interested in seeing whether the prior foreign policy deliberation had any cumulative or conditioning effect on the subsequent deliberation

about the presidential candidates. As it turned out, the effects we examined were insignificantly different for veterans and novices—with the sole exception that the veterans, who began what was for them this second DP with somewhat higher knowledge levels, thanks to their participation in the earlier one, also, by one measure, gained somewhat more. Apart from detailing this one difference, therefore, we examine the pooled treatment group.

We also had a pre-post “quasi control group,” separated from the treatment group by something close, though not quite equivalent, to random assignment. We randomly selected 762 KN panel members to complete the pre- and post-deliberation questionnaires, at the same times as the treatment group, without participating in the deliberations. Of these, 546 completed the pre-deliberation questionnaire, and 460 also completed the post-deliberation questionnaire. This would yield a true control group if everyone drawn into either the treatment or control group sample had in fact participated and completed both questionnaires, since two independent random samples are equivalent to a single random sample, of the same joint size, randomly divided in two. As it is, our treatment group is a nonrandom subsample of a random sample, since not everybody invited could or did participate or complete both questionnaires, and the same is true, in lesser degree, of the control group, since not everybody invited completed both questionnaires.

The roughly twenty percent of the treatment group lacking Internet access were offered a free computer and a two-month trial ISP membership in exchange for their participation. The remaining treatment group members were offered a financial incentive of \$200. Control group members, whose obligations were limited to completing the pre- and post-deliberation questionnaires, were offered \$40.

Briefing Materials

Treatment group members who agreed to participate were sent two sorts of briefing materials.³ The first was a document prepared by MacNeil/Lehrer Productions, our media partner, describing the policy options and the arguments for and against them in the areas of national security and international trade.⁴ Among other things, the document examined the questions of multilateral and unilateral military intervention generally, the invasion of Iraq in particular, and protectionism versus free trade.

The second briefing material was a multimedia CD mailed to every treatment group member. The CD provided biographical sketches of the ten declared presidential candidates (President Bush and nine Democrats) and outlined their positions on national security, trade, healthcare, taxes, and the economy. The candidates' positions, taken from their stump speeches, televised interviews, or debates, were presented as comparably as possible. In the case of healthcare and the economy, for instance, all nine Democrats responded to the identical questions posed by the AFL-CIO (e.g., "What steps will you take to create and retain good jobs?"). We approximated Bush's "answers" by using excerpts from speeches, interviews, and press releases.

In contrast to the news media, the CD provided roughly equal coverage of each candidate and each of the five issues.⁵ The sole instance of unequal candidate treatment concerned TV ads—those candidates with no advertisements were disadvantaged. Using the average number of "candidate pages" as a rough gauge of coverage, Al Sharpton, Carol Mosely-Braun and Dennis Kucinich (all with no advertisements) were allocated fifteen pages each, while the rest of the field (all with ads) was allocated an average of twenty pages. Thus, the CD provided somewhat greater exposure for "serious" candidates.

Small Group Discussions

Beginning January 19 (coincidentally, the day of the Iowa caucuses), participants attended five weekly hour-long online meetings. They selected a discussion group from a list of sixteen (defined by meeting times) and remained with the same group for the duration of the study. This was a departure from the practice in face-to-face Deliberative Polling, where the small groups are randomly assigned. The small groups varied in size between 12 and 21, averaging 14.

The weekly discussion agendas corresponded to the content of the briefing materials. Week 1 focused on the Iraq War and the question of unilateral versus multilateral military intervention. In Week 2, the focus shifted to the candidates' positions on Iraq and the use of military force. Week 3 dealt with international trade, Week 4 with taxes, the economy, and health care. Week 5, the last session, provided the participants an opportunity to reflect on and synthesize what they had learned about the issues and candidates.

The discussions were voice- rather than text-based. The Lotus Sametime software permitted the moderator to regulate the order of speaking, thus precluding simultaneous comments by different participants. In most instances, participants formed a "speaking queue," and the microphone was passed accordingly. When the discussion faltered, the moderator would pose questions or otherwise attempt to stimulate contributions. As in other DPs, the moderators also attempted to prevent anyone from dominating the discussion.⁶

This design replicates the small group discussions in face-to-face Deliberative Polling quite closely. The interactions are oral, not written. The biggest differences are the absence of the paralinguistic and other nonverbal communication in face-to-face interactions and the greater dispersion of the discussions over time. In face-to-face DPs, the small group deliberations

normally total five or six hours, spread over roughly two days. In this online version, they consumed roughly the same amount of time, but spread over five weeks. Thus our online participants had both more time to think and more time to forget. They also had more opportunity, between sessions, to discuss the topics they were deliberating with everyday discussion partners.

As in face-to-face Deliberative Polling, each small group was encouraged, during the last fifteen minutes of each session, to formulate questions to put to the expert panelists. The moderators submitted these questions electronically to the *Online Newshour with Jim Lehrer*, which in turn relayed them to policy experts representing both Democratic and Republican perspectives. The experts' answers were then posted on the small group's website in advance of the next online session.

Measurement

As indicated, both treatment and control group members completed a web-administered pretest questionnaire in advance of the first online meeting (on a median date of January 5) and a posttest questionnaire, also web-administered, immediately following the conclusion of the fifth discussion session. The posttest contained the same questions as the pretest, plus a few new ones.

Results

The well-known findings of widespread voter ignorance (reviewed by Delli Carpini and Keeter 1996, Price 1999) come mostly from surveys conducted shortly before or after elections, after the campaigns have had the bulk of their educative effect. That voters still know so little at that point says something about voters, but also about the campaigns. The results of this DP provide a window on the educative effects of a counterfactually policy-focused campaign, as

well as on the effects that the heightened policy focus and attendant learning and thinking have on candidate preferences and the criteria underlying them. We take these effects one by one, confining the analysis to Democrats and independents, who are, by and large, the relevant electorate for the Democratic nomination campaign.

Knowledge

Deliberation's effects presumably stem from the learning and thinking involved. Thinking is hard to measure but highly correlated with learning, and we do have two sorts of items gauging knowledge. Seven are factual, asking about economic conditions (the impact of Bush's tax policy on investment income and changes in the rate of unemployment since Bush took office), trade issues (the reason for the termination of US steel tariffs and the Bush Administration's position on the FTAA), the previous experience of Gephardt and Clark, and the conflict in Iraq (the number of Americans killed since the invasion). The index is the proportion answered correctly.⁷

We also have the respondent's placements of Bush and five of the Democratic candidates (Clark, Dean, Edwards, Kerry, Sharpton) on four NES-type seven-point policy scales. The scales were defined by their endpoints of free trade (1) versus protecting US industries (7), intervening militarily on our own (1) versus only with international approval and cooperation (7), increasing defense spending at the expense of domestic spending (1) versus the reverse (7), and cutting services to lower taxes (1) versus increasing services at the cost of higher taxes (7).⁸ We take the candidates' positions on these scales to be somewhere on the lower taxes, increased defense spending, unilateral military intervention, and free trade sides (points 1-3) for Bush and somewhere on the opposite sides (points 5-7) for all five Democrats. That leads to 24

placement-based knowledge items (six candidates times four issues). Again the index is the proportion answered correctly.⁹

By both measures, deliberation produced learning. To be sure, the control group, exposed just to the campaign, also learned. As Table 4 shows, their percentages of correct answers increased by 2.1 points on the factual measure and by 8.7 points on the placement

(Table 4 about here)

measure. But the participants learned much more. Their percentages of correct answers increased by 8.4 points (four times as much) on the factual measure and by 16.9 points (almost twice as much) on the placement measure. For both measures, the treatment-control group difference is highly significant.

The only significant differences between novice and veteran participants also lay here. The veterans entered these deliberations with elevated knowledge levels, thanks to their prior Deliberative Polling experience. They averaged 57.5% percent correct on the factual items and 32.8% correct on the placement items, compared to 42.2% and 24.3% for the novices. Both groups increased their factual knowledge scores by 8.4 points, but the veterans increased their placement knowledge scores by 20.7 points, compared to the novices' 14.2, result hinting at a cumulative, indeed greater-than- linear effect of repeated deliberation on learning.

Candidate Preference

From previous DPs, we know that deliberation stimulates learning and thought and thus often produces significant opinion change (Fishkin and Luskin 1999; Luskin, Fishkin, and Jowell 2002).¹⁰ Here too the opportunity to read balanced, substantive presentations about the candidates, to participate in small group discussions, and to seek clarification from Democratic

and Republican experts should create more informed and thoughtful citizens. In the nature of things, this may benefit some candidates and penalize others.

As indicators of candidate support, we focus on the NES “feeling thermometers,” whose treatment and control group means are shown in Table 1.¹¹ Both the treatment and the control

(Table 1 about here)

group grow significantly warmer or cooler toward most of the candidates, and in generally similar directions. Predictably, these Democrats and independents grew significantly cooler toward Bush and significantly warmer toward the leading Democratic contenders, Kerry and Edwards. They also grew significantly cooler toward the fading Democrats, Dean and Clark (counting $p \leq .06$ in the case of the control group’s rating of Clark as significant). In every case, the warming toward the Democratic candidate or cooling toward the Republican candidate was greater in the treatment than in the control group, although the difference was significant only for Kerry and Edwards.

For these latter two, the control group ratings increased roughly 6 or 8 points, reflecting the effect of the real-world campaign, while the treatment group ratings leapt up by roughly 15 points. Edwards outgained Kerry by roughly 2 points, although this difference is insignificant.¹² In a campaign counterfactually dominated by policy issues (at least by the same policy issues as were salient in the actual campaign), therefore, the race would still likely have narrowed to Kerry and Edwards.

We hesitate to make too much of the treatment-control group contrasts for Dean and Clark, since neither is statistically significant, but the results suggest some modest insulation from real-world momentum. In the more issue-oriented, less horse-race-obsessed environment

of the DP, they faded somewhat less. In a campaign counterfactually dominated by policy issues, they might have been able to survive a bit longer than they did.

But let us focus more closely on the contest between the two “finalists,” Kerry and Edwards. Which candidate, in a more deliberative universe, would have attracted more votes? We can predict a Kerry vote (1) if the Kerry thermometer score is strictly greater than the Edwards thermometer score, predict an Edwards vote (-1) if vice versa, and consider the voter undecided if the two thermometer scores are equal.¹³ Table 2 presents the means of this predicted vote variable (the percentage difference between Kerry and Edwards, with positive numbers representing a Kerry edge), as well as of the difference between the Kerry and Edwards

(Table 2 about here)

feeling thermometers, in both the treatment and control groups, both pre- and post- deliberation. At the end of the day, our participants exhibited a slight, and slightly increased, preference for Kerry. So, however, did the control group, and in greater degree. Deliberation thus seems to have slightly reduced Kerry’s already thin edge. But all these differences are small and statistically insignificant.

Does this mean that deliberation had little effect? Only in the aggregate, and only in a sense. As we shall see, deliberation moved some participants distinctly toward Kerry, others distinctly toward Edwards. In the aggregate, these effects cancelled out, but they need not always. Even in the aggregate, moreover, deliberation seems to have affected the *criteria* by which the candidates were evaluated.

Criteria

To see what criteria received what weight, we estimate a more explicit multivariate model, with deliberation playing a conditioning role. Since we have already-cast or intended

vote choices for only a small minority of the control and treatment samples, we approximate vote choice by the difference between the Kerry and Edwards thermometer scores. The correlation between the thermometer difference and vote choice, for respondents on whom we have the latter, is .60.

On the explanatory side, we might expect the degree of preference for Kerry or Edwards to depend on:

- *The difference in the distances between the voters' policy positions and their perceptions of Kerry's and Edwards'.* This is based on the respondents' placements of themselves and the candidates on the four NES-type seven-point policy scales, concerning protectionism, multilateralism, spending priorities, and taxes. We reckon the distances by the usual Euclidean formula, then subtract the distance from Kerry from the distance from Edwards (so that voters who are closer to Kerry receive positive scores and voters who are closer to Edwards receive negative ones, consistent with the scorings of the dependent variables and other regressors). If politics is, in Lasswell's phrase, "who gets what," this is what voters should be giving greatest weight.
- *The difference in the candidates' perceived electability.* Media accounts suggested that many Democratic primary and caucus voters were deciding on the basis of which candidate they thought would stand the best chance of defeating Bush in November. Ideally, therefore, we would have asked for ratings of each candidate's prospects of prevailing in a contest against Bush. We did ask "Who do you think will be elected President in November?" We use the responses to construct a variable scored -1 for answers of Edwards, +1 for answers of Kerry,

and 0 for all other answers. This too is a sensible criterion, at least for Democrats, the great majority of whom would have preferred any Democratic candidate to Bush. It would do little good for one's most preferred candidate to get the nomination if he could not beat Bush in November.

- *The difference in their perceived personalities.* Empirically, candidate personalities seem to be the most powerful influence, in this age of media politics, on vote choice, although normatively they are among the least valuable of these voting criteria (Luskin and Globetti 2002). Our questionnaire asked respondents to say whether the phrases "intelligent," "sincere," and "thinks like me" described each candidate "extremely well," "pretty well," "not too well," or "not well at all." We average the linearly scored responses, then subtract the Edwards score from the Kerry score.
- *Whether the voter resides in the South.* One might expect Edwards' Southern origins to help him with Southern voters.
- *Whether the voter is a Democrat versus an independent.* According to plausible media accounts, Kerry, the more liberal and familiar of the two Democratic "finalists," had relatively greater appeal to Democrats, while Edwards had relatively greater appeal to independents. (Republicans, remember, are excluded.)
- *Experimental condition.* This is a dummy variable distinguishing the treatment and control groups. Given the results in Tables 1-2, we do not expect deliberation to move the mass of voters greatly toward or away from either Kerry or Edwards, but we do expect it to heighten the effect of policy issues on their preference for Kerry versus Edwards (and thus to move some in one direction, and others in the

other). It would also be normatively pleasing if deliberation reduced the role of candidate personality in vote choice, although the evidence from general election analyses conditioning on information is unencouraging (Rahn, Aldrich, Borgida, and Sullivan 1990; Pierce 1993; Miller, Wattenberg, and Malanchuk 1986; Glass 1985). By and large, these studies suggest that image-related factors dominate all others—for the well and poorly informed alike. Some even suggest that the better informed weigh the candidates' personalities *more* heavily.¹⁴

Consistent with these hypotheses (and allowing deliberation to condition the effects of personality and electability as well as policy considerations), we estimate the equation:

$$KE = \gamma_0 + \gamma_1 TREAT + \gamma_2 DEM + \gamma_3 SOU + \gamma_4 TRAIT + \gamma_5 ELECT + \gamma_6 POLICY + \gamma_7 TRAIT * TREAT + \gamma_8 ELECT * TREAT + \gamma_9 POLICY * TREAT + u,$$

where KE is the Kerry thermometer rating minus the Edwards feeling thermometer rating, TREAT is a dummy variable scored 1 for members of the treatment group and 0 for members of the control group, DEM is a dummy variable scored 1 for Democrats and 0 for independents, SOU is a dummy variable scored 1 for residents of the South and 0 otherwise, TRAIT is the difference between the respondent's mean rating of Kerry on the three personality traits and his or her mean rating of Edwards on the same traits (linearly translated to the -1 to +1 interval), POLICY is the Euclidean distance between the respondent's own positions and his or her perceptions of Edwards' positions minus the Euclidean distance between his or her own positions and his or her perceptions of Kerry's positions (also linearly translated to the -1 to +1 interval), ELECT is the electability variable (already running from -1 to +1), the γ 's are the coefficients, and u is the usual disturbance or error term.¹⁵

The coefficients for DEM, SOU, TRAIT, ELECT, and POLICY ($\gamma_2, \gamma_3, \dots, \gamma_6$) are those variables' effects in the control group. In the treatment group, we assume DEM's and SOU's effects to be unchanged (still γ_2 and γ_3), as preliminary estimations suggested, while TRAIT's, ELECT's, and POLICY's effects are $\gamma_4 + \gamma_7, \gamma_5 + \gamma_8,$ and $\gamma_6 + \gamma_9,$ respectively. The differences that being in the treatment group makes to TRAIT's, ELECT's, and POLICY's effects are thus $\gamma_7, \gamma_8,$ and $\gamma_9.$ The treatment effect, similarly contingent, is $\gamma_1 + \gamma_7\text{TRAIT} + \gamma_8\text{ELECT} + \gamma_9\text{POLICY}.$ To the extent that $\gamma_7, \gamma_8,$ and γ_9 are nonzero, therefore, deliberation's effect depends on the voter's perceptions of the candidates' personalities, chances of winning in November, and policy positions (in relation to his or her own).

All these effects, excepting SOU's, should be positive. Southerners should be less inclined to prefer Kerry to Edwards. In the control group, based on what we know about primary voting under normal circumstances, electability and image should dominate policy voting (Bartels 1988; Williams et al. 1976; Gopoian 1982; for a review, see Kinder 1998). But we expect policy to become important in the treatment condition.

The ordinary least squares (OLS) estimates are in Table 3. SOU and DEM, whose effects are unconditioned by treatment, both carry significant coefficient estimates. The effect of

(Table 3 about here)

being a Southerner is negative, as expected, the effect of being a Democrat positive, also as expected. The former decreases Kerry's edge by about four degrees; the latter increases it by a similar amount. More interesting, however, are the estimated effects of ELECT, TRAIT, and POLICY, all allowed to vary by condition. In the control group, only ELECT and TRAIT have statistically significant effects. POLICY does not. In the treatment group, however, the first two effects remain essentially the same, while the POLICY effect becomes significantly greater,

indeed becomes significant. Thinking better of Kerry's personal qualities and thinking him more electable increase his edge in both the treatment and control group, and to roughly the same degree in both. The control and treatment group coefficient estimates are 69.5 and 71.5 ($= 69.5 + 1.98$) for TRAIT and 6.41 and 8.14 ($= 6.41 + 1.73$) for ELECT, all significant. Agreeing more with Kerry's policy positions, however, increases his edge only in the treatment group, with an estimated coefficient of 40.4 ($= -10.1 + 50.5$), also significant.

To provide some perspective on these estimates, Table 3 also reports the "first differences" (in the expected value of the Kerry - Edwards thermometer difference KE). The column headed E_t bases the first differences on each variable's theoretical range—from its minimum possible value to its maximum possible value, i.e. from -1 to +1 for ELECT, TRAIT, and POLICY and from 0 to 1 for DEM, SOU, and TREAT. In the treatment group, then, a change in POLICY from -1 (maximally distant from Kerry and at the same point on every issue as Edwards) to +1 (the reverse) could be expected to increase KE by 80 degrees of its 200 degree range. A change in TRAIT from -1 (Edwards rated as positively as possible and Kerry as negatively as possible on all traits) to +1 (the reverse) could be expected to increase KE in the treatment group by 139 degrees.

To be sure, the TRAIT and POLICY effects are not quite as large as they look, because neither variable actually occupies much of its theoretical range. ELECT can only be -1, 0, or +1 and is frequently either -1 or +1, but POLICY and TRAIT are another story. POLICY never comes close to either -1 or +1. Its observed minimum is actually only -.26, its observed maximum only .22. TRAIT does occasionally hit -1 or +1, but only occasionally. For these two regressors, the column headed E_e therefore adds the less clear-cut but more revealing first differences based on their "effective" or "greatest likely" ranges, defined for present purposes as

two standard deviations below to two standard deviations above the mean.¹⁶ This measure yields slimmer but still muscular treatment-group first differences of 21.5 and 62.8 for POLICY and TRAIT, respectively.

In all, these results show the patterns we expect. In the control group, TRAIT's effect towers over any other, followed at some considerable distance by ELECT's. In the treatment group, POLICY acquires a major effect, roughly the equal of ELECT's, though still much smaller than TRAIT's. Even our deliberators weigh the candidates' personalities most heavily and give some weight to electability. Unlike the treatment group, however, they also weigh policy.

As for the treatment effect, the first difference, too long an expression to fit in its table cell, is $0.85 + 1.98*\text{TRAIT} + 1.73*\text{ELECT} + 50.5*\text{POLICY}$. Its size and sign are thus mostly governed by the size and sign of POLICY. For participants equidistant between Kerry and Edwards on policy (POLICY = 0) but revering Kerry while reviling Edwards (TRAIT = 1) and thinking Kerry more electable (ELECT = 1), deliberating only increases the preference for Kerry by about 4.5 degrees. But for participants seeing Kerry and Edwards as equally appealing (TRAIT = 0) and equally electable (ELECT = 0) but agreeing decidedly more with Kerry on policy (POLICY = say 0.2, not far from its largest observed value, though well short of its theoretical maximum), deliberating increases Kerry's edge by roughly 11 degrees. Symmetrically, for participants agreeing equally decidedly more with Edwards than Kerry (POLICY = -0.2), while still finding them equally appealing and electable (TRAIT = ELECT = 0), deliberating *decreases* Kerry's edge by a similar amount (slightly smaller, given the positive intercept). So the deliberation had an appreciable effect on the preferences of voters who saw themselves as sufficiently closer to Kerry than to Edwards, or vice versa, in the policy space.

At the same time, this same formula suggests the reason deliberation's effect on the *aggregate* preference for Kerry versus Edwards was so modest: *on average*, the voters saw them as roughly equidistant from themselves. The mean estimated treatment effect, across the sample, is $0.85 + 1.98 * \bar{T} + 1.73 * \bar{E} + 50.5 * \bar{P}$, where \bar{T} , \bar{E} , and \bar{P} denote the post-deliberation sample means of TRAIT, ELECT, and POLICY. The mean effect thus depends largely on \bar{P} , which was only 0.03 (on its -1 to +1 scale). Of course another nomination contest could narrow to a pair of candidates the voters saw, on average, as more unequally distant from themselves. In that case, a similar deliberation could be expected to have a greater effect on the aggregate distribution of preferences.

Conclusion

The original idea of Deliberative Polling was to provide a tonic for the presidential nomination process (Fishkin 1988, 1991). As matters stand, early contests in two small, atypical states generate the “momentum” that may carry candidates to the nomination precisely when—and partly because—public familiarity with the candidates and their positions on policy issues is meager. This initial input is neither very representative nor very well considered. A national DP preceding the Iowa caucuses and New Hampshire primary, if it received similar attention, could provide a more representative and better considered one. With different “initial conditions,” the process might play out differently.

The first national American DP, the 1996 National Issues Convention, included appearances by several Republican presidential candidates and Vice President Gore but focused on policy issues, not candidate preference (Fishkin and Luskin 1999). While some DPs in other countries have focused on electoral choice by the mass public in either general elections (Luskin, Fishkin, Jowell, and Park 1999) or national referenda (Luskin et al. 2000, Hansen 2004), this is

the first application of Deliberative Polling to the arena for which it was originally proposed—presidential nomination campaigns.

Our results show that, in this connection too, deliberation matters. After only five hours of discussion, spread over five weeks, the treatment group became significantly more knowledgeable about the candidates and evaluated them to a significantly greater degree on the basis of policy issues. A longer or more enveloping deliberation could be expected to give policy still greater weight. More speculatively, the deliberations may also have created a context where “momentum” worked differently. They may have given the two leading candidates an extra boost, perhaps because our participants thought them likeliest to prevail in November, while also providing the lagging candidates with something of a cushion, presumably because the participants acquired a broader portfolio of criteria than just their primary-losing ways with which to assess them.

At least three notable differences between this process and the original proposal remain. The first is the most obvious but least consequential. The original idea envisioned face-to-face deliberations, whereas ours here were online. Research on tandem face-to-face and online DPs suggests that the two modes have broadly similar effects, with the online effects running slightly weaker (Luskin, Iyengar, and Fishkin 2003). So our present results may be slightly conservative. The second difference is that our deliberations concluded well after the Iowa caucuses and New Hampshire primary—not until just before Super Tuesday, by which point all the candidates but Kerry and Edwards had officially or effectively dropped out. The third difference concerns the level of media coverage. The original aspiration was for coverage rivaling that of the Iowa caucuses and New Hampshire primary. Even face-to-face DPs, which are generally large-scale media events, have never attained that level of coverage. This online DP provided a good basis

for examining the effects of deliberation on candidate preferences but was not designed to be a media event.

Thus the original proposal—a widely heavily covered national DP about candidate preferences (and policy issues) at the very beginning of the primary season—still remains to be realized. Such an event could provide the media, politicians, and the voters a valuable cue at a moment when it could make an important difference. Indeed it would be nice to see an early national DP institutionalized as part of the nomination process. The prospects, since the frequency of face-to-face Deliberative Polling has always been held down by its cost, may be enhanced by the continued development of the online version (which promises to become still cheaper over time, as the percentage of the population needing to be provided with computer access—the biggest expense in this design—continues to dwindle).

References

- Adams, William C. 1987. As New Hampshire goes . . ., in Gary R. Orren and Nelson W. Polsby eds., *Media and Momentum*. Chatham, NJ: Chatham House.
- Aldrich, John H. 1980. *Before the Convention: Strategies and Choices in Presidential Nomination Campaigns*. Chicago: University of Chicago Press.
- Amundson, Daniel, Lichter, S. Robert, and Richard Noyes. 1988. *The Video Campaign: Network Coverage of the 1988 Primaries*. Washington DC: American Enterprise Institute Press.
- Bartels, Larry M. 1988. *Presidential Primaries and the Dynamics of Public Choice*. Princeton: Princeton University Press.
- Brady, Henry E., and Richard Johnston. 1987. What's the primary message? Horse race or issue journalism? in Gary R. Orren and Nelson W. Polsby eds, *Media and Momentum*. Chathan, NJ: Chatham House.
- Busch, Andrew E., and William G. Mayer. 2003. The front-loading problem, in Andrew E. Busch and William G. Mayer eds., *The Front-Loading Problem in Presidential Nominations*. Washigton, DC: Brookings Institution Press.
- Chang, Lin Chiat, and Jon A. Krosnick. 2003. "RDD Telephone vs. Internet Survey Methodology: Comparing Sample Representativeness and Response Quality." Unpublished Paper, Ohio State University.
- Delli Carpini, Michael X., and Scott Keeter. 1996. *What Americans Know About Politics and Why It Matters*. New Haven: Yale University Press
- Fishkin, James. 1988. "The Case for a National Caucus: Taking Democracy Seriously." *The Atlantic*, August, 16-18.

- Fishkin, James. 1991. *Democracy and Deliberation: New Directions for Democratic Reform*. New Haven, CT: Yale University Press.
- Fishkin, James S. and Robert C. Luskin. 1999. Bringing deliberation to the democratic dialogue, in Maxwell McCombs and Amy Reynolds eds., *The Poll with a Human Face: The National Issues Convention Experiment in Political Communication*. Mahwah, NJ: Erlbaum.
- James S. Fishkin and Robert C. Luskin. 2005. "Experimenting with a Democratic Ideal: Deliberative Polling and Public Opinion," *Acta Politica* 40 (September): 284-298.
- Gelman, Andrew, and Gary King. 1993. "Why are American Presidential Election Campaign Polls so Variable When Votes are so Predictable?" *British Journal of Political Science*, 23, 409-51.
- Glass, David P. 1985. "Evaluating Presidential Candidates: Who Focuses on Their Personal Attributes?" *Public Opinion Quarterly*, 49:517-34.
- Gopoiian, David J. 1982. "Issue Preferences and Candidate Choice in Presidential Primaries." *American Journal of Political Science*, 26, 523-46.
- Hansen, Kasper M. 2004. *Deliberative Democracy and Opinion Formation*. Odense, Denmark: University of Southern Denmark.
- Iyengar, Shanto, Norpoth, Helmut, and Kyu Hahn. 2004. "Consumer Demand for Election News: The Horse Race Sells." *Journal of Politics*, 66, 157-175.
- Kinder, Donald R. 1998. Opinion and action in the realm of politics, in Daniel T. Gilbert, Susan T. Fiske, and Gardner Lindzey eds., *Handbook of Social Psychology*. New York: McGraw-Hill.
- Luskin, Robert C. 1991. "Abusus Non Tollit Usum: Standardized Coefficients, Correlations and R^2 s." *American Journal of Political Science*, 35, 1032-46.

- Luskin, Robert C., Fishkin, James S., and Roger Jowell. 2002. "Considered Opinions: Deliberative Polling in Britain." *British Journal of Political Science*, 32, 455-87.
- 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," Presented at the Annual Meeting of the American Political Science Association, Washington, DC, August 31-September 3.
- Luskin, Robert C., James S. Fishkin, and Dennis L. Plane. 1999. "Deliberative Polling and Policy Outcomes: Electric Utility Issues in Texas," Presented at the Annual Meeting of the Association for Public Policy Analysis and Management, Washington, DC, November 4-6.
- Luskin, Robert C., James S. Fishkin, Roger Jowell, and Alison Park. 1999. "Learning and Voting in Britain: Insights from the Deliberative Poll." Presented at the Annual Meeting of the American Political Science Association, Atlanta, GA, September 2-5.
- Luskin, Robert C. and Suzanne Globetti. 2002. Candidate versus Policy Considerations in the Voting Decision: The Role of Political Sophistication Ms., Dept. of Government, University of Texas at Austin.
- Luskin, Robert C., Shanto Iyengar, and James S. Fishkin. 2003. Considered Opinions on U.S. Foreign Policy: Face-to-Face versus Online Deliberative Polling. Presented at the annual meeting of the American Political Science Association, Philadelphia, PA, August 27-31, 2003.
- Mayer, William G. 1996. Forecasting presidential nominations, in Mayer ed., *In Pursuit of the White House: How We Choose our Presidential Nominees*. Chatham, NJ: Chatham House.

- Miller, Arthur H., Wattenberg, Martin P., and Oksana Malanchuk. 1986. "Schematic Assessments of Presidential Candidates." *American Political Science Review*, 80, 521-40.
- Pierce, Patrick A. 1993. "Political Sophistication and the Use of Candidate Traits in Candidate Evaluation." *Political Psychology*, 14, 21-35.
- Polsby, Nelson W. 1983. *Consequences of Party Reform*. New York: Oxford University Press.
- Rahn, Wendy M., Aldrich, John H., Borgida, Eugene, and John L. Sullivan. 1990. "A Social-Cognitive Model of Candidate Appraisal." In *Information and Democratic Processes*, ed. John A. Ferejohn and James H. Kuklinski. Urbana: University of Illinois Press.
- Robinson, Michael J., and Margaret A. Sheehan. 1983. *Over the Wire and on TV: CBS and UPI in Campaign '80*. New York: Russell Sage.
- United States Department of Commerce. 2002. *A Nation Online: How Americans Are Expanding Their Use of the Internet*. Washington, D.C.: Telecommunications and Information Administration.
- Williams, Daniel C., Weber, Stephen J., Haaland, Gordon A., Mueller, Ronald H., and Robert E. Craig. 1976. "Voter Decision-making in a Primary Election: An Evaluation of Three Models of Choice." *American Journal of Political Science*, 20, 37-50.
- Zaller, John. 2004. *A Theory of Media Politics*. Chicago: University of Chicago Press, forthcoming.
- _____. 1992. *The Nature and Origins of Mass Opinion*. New York: Cambridge University Press.

Table 1
Thermometer Ratings

	Treatment Group (<i>n</i> =180)				Control Group (<i>n</i> =278)				Difference	
	T1	T2	T2-T1 (s.e.)	<i>p</i>	T1	T2	T2-T1 (s.e.)	<i>p</i>	Mean diff. (s.e.)	<i>p</i>
George Bush	40.7	35.1	-5.58 (1.58)	.00	42.8	38.4	-4.37 (1.19)	.00	-1.21 (1.95)	.54
John Kerry	50.0	65.3	15.3 (1.75)	.00	50.4	58.8	8.41 (1.28)	.00	6.87 (2.12)	.00
John Edwards	49.3	63.8	14.6 (1.60)	.00	49.7	56.0	6.37 (1.28)	.00	8.18 (2.05)	.00
Howard Dean	54.3	50.9	-3.56 (1.73)	.05	52.2	45.8	-6.32 (1.37)	.00	2.87 (2.20)	.19
Wesley Clark	53.2	50.3	-2.91 (1.53)	.06	53.3	48.9	-4.37 (1.12)	.00	1.46 (1.86)	.43
Al Sharpton	33.8	36.4	2.58* (1.66)	.12	38.2	36.9	-1.37 (1.24)	.27	3.95 (2.04)	.06

NOTE: This and all succeeding tables are confined to Democrats and independents. The *p*-values are two-tailed.

Table 2
Kerry versus Edwards: Feeling Thermometer Differences and Predicted Votes

	Treatment Group (<i>n</i>=180)				Control Group (<i>n</i>=278)				Difference	
	T1	T2	T2-T1 (s.e.)	<i>p</i>	T1	T2	T2-T1 (s.e.)	<i>p</i>	Mean (s.e.)	<i>p</i>
Kerry – Edwards Difference										
Feeling Thermometer	0.78	1.50	0.72 (1.98)	.72	.766	2.80	2.03 (1.49)	.17	-1.31 (2.44)	.24
Predicted Vote	0.04	0.08	0.04 (0.07)	.59	.047	.122	0.08 (0.05)	.12	-0.04 (0.09)	.68

NOTE: The *p*-values are two-tailed.

Table 3
Explaining the Kerry-Edwards Thermometer Difference

	Coeff.	s.e.	<i>p</i>	<i>E_t</i>	<i>E_e</i>
Constant	-3.12	2.33	.18	--	--
TREAT	0.85	2.71	.76	See text	--
DEM	4.62	2.16	.02	4.62	--
SOU	-3.95	2.17	.04	-3.95	--
ELECT	6.41	2.45	.01	12.8/16.3	--
TRAIT	69.5	7.24	.00	139/135	61.6/62.8
POLICY	-10.1	12.9	.67	-20.2/80.8	-4.7/21.5
TREAT*ELECT	1.73	3.41	.61	--	--
TREAT*TRAIT	1.98	10.5	.85	--	--
TREAT*POLICY	50.5	17.9	.01	--	--
<i>R</i> ²	.556				
Adj. <i>R</i> ²	.542				
<i>F</i>	38.7				
<i>p</i>	.00				
<i>n</i>	288				

The p-values for P-val's are 2-tailed. The entries separated by slash marks in the *E_t* and *E_e* columns are the first differences in the control group to the left of the slash mark and in the treatment group to the right of it.

N = 340, Adj. *R*² = .508, *F* = 39.852, *p* = .000.

**p* < .05, one-tailed.

NOTE: The entries separated by slash marks are the first differences in the control group to the left of the slash mark and in the treatment group to the right of it.

Table 4
Knowledge Gains

	Treatment (n = 180)				Control (n =278)				Difference	
	T1	T2	T2-T1 (s.e.)	p	T1	T2	T2-T1 (s.e.)	p	Mean (s.e.)	p
Factual index	.486	.570	.084 (.014)	.00	.441	.462	.021 (.011)	.03	.063 (.018)	.00
Placement index	.278	.448	.169 (.017)	.00	.230	.316	.087 (.012)	.00	.083 (.021)	.00

NOTE: *p*-values are one-tailed.

*Earlier versions of this paper were presented at the conference “Voice and Citizenship: Re-thinking Theory and Practice in Political Communication,” University of Washington, April 23-24, 2004, the annual meeting of the American Political Science Association, Chicago, IL, September 2-5, 2004, the annual meeting of the International Communications Association, New York, NY, May 26-29, 2005, the annual meeting of the American Political Science Association, Washington, DC, September 1-4, 2005, and the biennial General Conference of the European Consortium for Political Research, Budapest, Hungary, September 8-10, 2005. We are grateful to Larry Bartels, Todd Belt, John Gastil, William Jacoby, and Christopher Wlezien for comments and to Jennifer McGrady, Alice Siu, and Kyu Hahn for research assistance.

¹The “National Issues Convention” Deliberative Poll of 1996 involved appearances by several Republican presidential candidates and then-Vice President Al Gore but was not *about* the nomination campaign.

²The KN panel covers slightly less than the 95 percent of the population with a working telephone, since the design excludes residents of areas not serviced by a Web TV ISP. Almost two-thirds of those invited to join the panel do so (see Chang and Krosnick 2003).

³Of course, “offline” participants were sent their computers considerably in advance (no later than two weeks before the first online session). We also sent them a checklist of items to complete in preparation for the first session as well as a schedule of “practice” online sessions.

⁴A team at the Woodrow Wilson School at Princeton University prepared the non-partisan briefing document on international security and economics for *By the People*, an initiative of MacNeil/Lehrer Productions. It can be found on the PBS web site at <http://www.by-the-people.org>.

⁵The coverage ranged from 40 pages (trade) to 48 (national security).

⁶Our moderators, all Stanford University graduate students, were all experienced with Sametime, having moderated in the previous online Deliberative Poll on U.S. foreign policy.

⁷ Cronbach's alpha (post-test) is .72.

⁸More precisely the endpoints were “the federal government should provide fewer services such as health and education in order to lower taxes” versus the federal government [should] provide more services even if it means higher taxes”; “we should spend much less money for defense and focus much more on solving domestic problems” versus “defense spending should be greatly increased even if it means reduced spending on domestic problems”; “the United States should intervene militarily when its interests are threatened without working through international organizations such as the U.N and NATO” versus “we should not intervene without first obtaining international approval and cooperation”; and “we should pursue increased free trade even if we may lose some existing jobs” versus “we should subsidize American industries even if that means we lose markets for our goods abroad.”

⁹Cronbach's alpha (post-test), using each candidate as the individual indicator, was .91.

¹⁰See also Barabas (2004), who uses matching techniques to approximate random assignment in the America Speaks data (a very nonrandom and very unrepresentative sample, engaged in less careful deliberation).

¹¹The item ran as follows: “Next, we’d like to get your feelings toward the presidential candidates using what is sometimes called a ‘feeling thermometer.’ Ratings between 50 degrees and 100 degrees mean that you feel favorable and warm toward the candidates. Ratings between 0 degrees and 50 degrees mean that you don’t feel favorable that you don’t care too much for that candidate. You would rate the candidate at the 50-degree mark if you don’t feel particularly warm or cold. If you don’t know enough to rate a given candidate, just check ‘can’t say.’” In Table 1, “Can’t say’s” are assigned scores of 50.

¹²There is also a nearly significant treatment-control group difference in the change in Sharpton’s ratings (insignificantly positive in the treatment group, insignificantly negative in the control group). Sharpton remained, however, the least popular candidate in both groups.

¹³If Kerry but not Edwards was rated, the vote was allocated to Kerry if his rating exceeded 50, was allocated to Edwards if it was less than 50, and was counted as undecided if it equaled 50. Symmetrically, if Edwards but not Kerry was rated, the vote was allocated to Edwards if his rating exceeded 50, was allocated to Kerry if it was less than 50, and was counted as undecided if it equaled 50. If neither candidate was rated, the vote was counted as undecided.

¹⁴These same studies also fail to show the effects of policy as increasing with information, although Luskin and Globetti (2002) argue that the failure is artifactual and provide results showing the theoretically expected interaction.

¹⁵The Euclidean distance, for concreteness' sake from Kerry, is $[\sum(K_{ij} - R_{ij})^2]^{1/2}$, where R_{ij} and K_{ij} are the i th respondent's own position and placement of Kerry on the j th issue, the summation is over all four issues, and we take $K_{ij} - R_{ij}$ to be zero for those issues on which the respondent either takes no position or ascribes none to Kerry. A reasonable alternative would be to define the distance as $[\sum(K_{ij} - R_{ij})^2/n_{iK}]^{1/2}$, where the summation is only over the n_{iK} (≤ 4) issues on which the respondent both takes a position and ascribes one to Kerry. This yields essentially the same results.

¹⁶More generally, one could choose any number c of standard deviations above and below the mean ($c > 0$). In a linear, additive model, E_e , so defined, divided by the difference between c standard deviations above and below the mean on the dependent variable works out to the standardized regression coefficient (no matter what the c). Here the interactions break up that equivalence (Luskin 1991).