

## **Considered Opinions: Deliberative Polling in Britain**

ROBERT C. LUSKIN, JAMES S. FISHKIN AND ROGER JOWELL\*

This article presents the results of the first Deliberative Poll, in which a national British sample discussed the issue of rising crime and what to do about it. We describe Deliberative Polling and its rationale, the representativeness of the deliberative sample, the extent to which the participants acquired factual information about the issue and about politics generally, and how much and how they changed their views. We also weigh the extent to which such changes of view hinge on small group influences versus information gains.

On Friday 15 April 1994 in Manchester, England, a national probability sample of 301 British subjects met for the world's first Deliberative Poll.<sup>1</sup> The topic was crime, and for the next two days the participants considered and discussed a wide range of policy proposals. They thought aloud in small groups and assembled in larger ones to hear and put questions to panels of experts and representatives of the Conservative, Labour and Liberal Democratic parties. At the end (as at the beginning), they gave their opinions. The proceedings were taped and edited down to a two-hour television programme later shown on Channel 4. This article describes the event, its rationale and some of the results. We particularly

\* Luskín and Fishkin: Department of Government, University of Texas, Austin; Jowell: National Centre for Social Research, London. An earlier version of this article was presented at the annual meeting of the International Society of Political Psychology, Vancouver, Canada, 1996. The 'Power and the People' Deliberative Poll on rising crime was co-sponsored by Channel 4 Television and *The Independent* newspaper. Among others too numerous to cite individually, we especially wish to thank David Lloyd, Andreas Whittam Smith, Fiona Chesterton, Julie Hall, Robert Kingston and Colin Hughes for their roles in making the Deliberative Poll possible. Rebecca Gray, then of SCPR, played an invaluable role in drafting the questionnaire, administering the sampling and recruitment, and gathering and preparing the data. We are also indebted to Loïc Blondiaux, Norman Bradburn, Henry Brady, Danielle Bütschi, Daniel Gaxie, Sybille Hardmeier, Richard Johnston, Stanley Kelley, Hanspeter Kriesi, Nonna Mayer, Russell Neuman and Jean-Luc Parodi for comments and to Christopher Bratcher, Han Dorussen and Dennis Plane for research assistance. Deliberative Polling is a registered trademark, and the Center for Deliberative Polling at the University of Texas at Austin, of which Fishkin is Director and Luskín Research Director, receives fees from the trademark to fund research. (The trademarking has been necessary to ensure that everything described as a Deliberative Poll meets certain standards.) We are grateful to the Public Policy Clinic, also of the University of Texas at Austin, for support. Much of the analysis and writing were done while Luskín was being supported by the University Research Institute of the University of Texas at Austin and housed as Chercheur Associé by the Centre d'Etude de la Vie Politique Française of the Fondation Nationale des Sciences Politiques, Paris.

<sup>1</sup> See James S. Fishkin, *Democracy and Deliberation: New Directions for Democratic Reform* (New Haven, Conn.: Yale University Press, 1991); James S. Fishkin, *The Voice of the People: Public Opinion and Democracy*, 2nd edn (New Haven, Conn.: Yale University Press, 1997).

examine the changes in the participants' policy preferences and the extent to which they seem to have resulted from policy-relevant learning.

#### DELIBERATIVE POLLING

When George Gallup first developed public opinion polling as we know it, he suggested that it would restore the democracy of the New England town meeting to the large nation-state. At least figuratively, he argued, citizens could now gather together, make and hear the arguments for and against various policy proposals, and vote them up or down. Radio, magazines and newspapers were already carrying proposals and arguments from political leaders to the public, and public opinion polls would complete the conversation by carrying the public's views back to political leaders. It would be, he said, as if everybody were 'in one great room'.<sup>2</sup>

Now, several decades' experience the wiser, we know that public opinion polls, while undoubtedly useful to the campaigns, businesses, government agencies and media outlets that commission them, have not been the great boon to democracy that Gallup envisioned. The problem is not mainly that the technology is sometimes abused (as in 'push polls') or shoddily implemented, although of course it sometimes is.<sup>3</sup> Nor is it just that the advent of polling as scorekeeping has been part of the long descent of campaign coverage towards sports coverage.<sup>4</sup> The most fundamental problem is that not many of the respondents answering any given question have very well considered or informed opinions about the issue.<sup>5</sup> The least considered and informed responses may represent what Converse famously called *non-attitudes*,<sup>6</sup> although minimal attitudes, involving some but meagre thought and information, are probably more common.<sup>7</sup>

One sign of such rickety opinions is the marked instability that frequently characterizes responses given by the same individuals to the same questions asked at intervals sufficient to minimize the trivial stability lent by memory (coupled with the desire to avoid being or appearing inconsistent). Also

<sup>2</sup> George Gallup, 'Public Opinion in a Democracy', *The Stafford Little Lectures* (Princeton: Princeton University Extension Service, 1939).

<sup>3</sup> 'Push polls' are persuasive communications disguised as polls, asking for example whether the respondent would be more or less likely to vote for the opposing candidate if he or she knew X (something bad, and often false) about him or her. The tip-off that the object is not research but persuasion is the sample size, frequently in the tens of thousands.

<sup>4</sup> See Thomas E. Patterson, *Out of Order* (New York: Alfred A. Knopf, 1993).

<sup>5</sup> 'Don't-know screens' – explicit invitations to admit to not having any real opinion – reduce but do not eliminate the problem. Indeed, really successful 'don't-know screens' would leave the question of what importance to attach to a relatively accurate picture of the well-formed opinions of a small fraction of the public.

<sup>6</sup> Philip E. Converse, 'The Nature of Belief Systems in Mass Publics', in David E. Apter, ed., *Ideology and Discontent* (New York: Free Press, 1964); Philip E. Converse, 'Attitudes and Nonattitudes: Continuation of a Dialogue', in Edward R. Tufte, ed., *The Quantitative Analysis of Social Problems* (Reading, Mass.: Addison-Wesley, 1970).

<sup>7</sup> Robert C. Luskin, 'Measuring Political Sophistication', *American Journal of Political Science*, 31 (1987), 856–99.

famously, Converse noted that the correlations between responses to the same question put at intervals of two years hewed rather closely to patterns generated by a 'black-white' model in which some small proportion of the sample responds unvaryingly, while a much larger one responds at random.<sup>8</sup> There was for a time some argument that this instability owed less to the insubstantiality of the opinions being expressed than to the intrinsic difficulty of measuring even the best formed opinions.<sup>9</sup> But that controversy, like the parallel one over the public's grasp of 'ideological' terms like 'liberal' and 'conservative', seems to have been doused by the more direct and sobering evidence of just how little information underlies most responses.<sup>10</sup>

By now, the fact is too well known to need much exemplifying, but we do wish to emphasize that 'major chunks of information critical to comprehending the stakes and arguments in policy debates are mostly absent'.<sup>11</sup> We suggest cruising Delli Carpini and Keeter<sup>12</sup> or Price<sup>13</sup> for macabre examples of information not widely known, but perhaps the most telling evidence lies in many respondents' inability to say correctly where the major parties stand on the issues of the day. In the United States, roughly half the sample, depending on the issue and the wording of the question, generally admits to having no idea where the Republican and Democratic parties stand or places them incorrectly, with the Republicans to the left of the Democrats.<sup>14</sup>

According to one recent stream of research, this widespread ignorance does not actually make much difference. Two claims are made, by different authors.

<sup>8</sup> Converse, 'The Nature of Belief Systems'; Converse, 'Attitudes and Nonattitudes'.

<sup>9</sup> Christopher H. Achen, 'Mass Political Attitudes and the Survey Response', *American Political Science Review*, 69 (1975), 1218–31; Charles M. Judd and Michael A. Milburn, 'The Structure of Attitude Systems in the General Public', *American Sociological Review*, 45 (1980), 627–43. See also, in rebuttal, see Philip E. Converse, 'Public Opinion and Voting Behavior', in Fred I. Greenstein and Nelson W. Polsby, eds, *The Handbook of Political Science* (Reading, Mass.: Addison-Wesley, 1975); and, in the middle, Christopher H. Achen, 'Toward Theories of Data: The State of Political Methodology', in Ada W. Finifter, ed., *Political Science: The State of the Discipline* (Washington, DC: American Political Science Association, 1983) and Stanley Feldman, 'Measuring Issue Preferences: The Problem of Response Instability', *Political Analysis*, 1 (1990), 25–60.

<sup>10</sup> Donald R. Kinder and David O. Sears, 'Public Opinion and Political Action', in Gardner Lindzey and Elliot Aronson, eds, *Handbook of Social Psychology*, Vol. 2, 3rd edn (New York: Random House, 1985); Luskin, 'Measuring Political Sophistication'; Stephen Earl Bennett, 'Trends in Americans' Political Information', *American Politics Quarterly*, 17 (1989), 422–35; Michael X. Delli Carpini and Scott Keeter, 'Stability and Change in the U.S. Public's Knowledge of Politics', *Public Opinion Quarterly*, 55 (1991), 583–612; Michael X. Delli Carpini and Scott Keeter, *What Americans Know About Politics and Why It Matters* (New Haven, Conn.: Yale University Press, 1996); Vincent Price, 'Political Information', in John P. Robinson, Phillip R. Shaver and Lawrence S. Wrightsman, eds, *Measures of Political Attitudes* (New York: Academic Press, 1999).

<sup>11</sup> James S. Fishkin and Robert C. Luskin, 'The Deliberative Poll: A Reply to Our Critics', *Public Perspective*, 7 (1996), 45–9.

<sup>12</sup> Delli Carpini and Keeter, 'Stability and Change in the U.S. Public's Knowledge of Politics'.

<sup>13</sup> Price, 'Political Information'.

<sup>14</sup> See Luskin, 'Measuring Political Sophistication', and Robert C. Luskin, 'Political Psychology, Political Behavior, and Politics: Questions of Aggregation, Causal Distance, and Taste', in James H. Kuklinski, ed., *Thinking about Political Psychology* (New York: Cambridge University Press, 2002).

The stronger is that the great majority of people somehow manage by cognitive heuristics or other shortcuts to grope their way to the same policy preferences they would hold if they were well informed.<sup>15</sup> The weaker claim is that while many *individuals* do not hold the same policy preferences as they would do if they were well informed, their errors cancel out in the aggregate, so that the *distribution* of policy preferences is essentially the same as it would be if everyone were well informed.<sup>16</sup> As the reader may have gathered, we do not agree, and a number of studies have already begun to show yet anew that, and how, information does matter, both at the individual level and in the aggregate.<sup>17</sup>

The resulting difference between well-informed and actual public opinion brings us to Deliberative Polling, whose goal is to narrow the gap. An ordinary poll is designed to show what the public actually thinks about some set of issues, however little, irreflective, and changeable that may be, and generally is. A Deliberative Poll is designed to show what the public *would* think about the issues, if it thought more earnestly and had more information about them. It is an attempt to provide some glimpse of a hypothetical public, one much more engaged with and better informed about politics than citizens in their natural surroundings actually are.<sup>18</sup>

The basic formula is as follows. Select a national probability sample of the citizen voting age population and question them about some policy domain(s). Send them balanced, accessible briefing materials to help inform them and get them thinking more seriously about the same subject(s). Transport them to a single site, where they can spend several days grappling with the issues, discussing them with one another in randomly assigned, moderated small groups and putting questions generated by the small group discussions to

<sup>15</sup> See Paul M. Sniderman, Richard A. Brody and Philip E. Tetlock, *Reasoning and Choice: Explorations in Political Psychology* (Cambridge: Cambridge University Press, 1991); Arthur Lupia, 'Busy Voters, Agenda Control, and the Power of Information', *American Political Science Review*, 86 (1992), 390–403; Arthur Lupia, 'Shortcuts versus Encyclopedias: Information and Voting Behavior in California Insurance Reform Elections', *American Political Science Review*, 88 (1994), 63–76; Arthur Lupia and Mathew D. McCubbins, *The Democratic Dilemma: Can Citizens Learn What They Really Need to Know?* (New York: Cambridge University Press, 1998); Samuel L. Popkin, *The Reasoning Voter: Communication and Persuasion in Presidential Campaigns* (Chicago: University of Chicago Press, 1991).

<sup>16</sup> Benjamin I. Page and Robert Y. Shapiro, *The Rational Public: Fifty Years of Trends in Americans' Policy Preferences* (Chicago: University of Chicago Press, 1992). The logic is faulty, of course, since individual-level errors are unlikely to cancel out. See Larry M. Bartels, 'Uninformed Voters: Information Effects in Presidential Elections', *American Journal of Political Science*, 40 (1996), 194–230, and Luskin, 'Political Psychology, Political Behavior, and Politics'.

<sup>17</sup> See Larry M. Bartels, 'Post-Cold War Defense Spending Preferences', *Public Opinion Quarterly*, 58 (1994), 479–508; Bartels, 'Uninformed Voters'; Robert C. Luskin and Suzanne Globetti, 'Candidate versus Policy Considerations in the Voting Decision: The Role of Political Sophistication' (unpublished, Department of Government at the University of Texas at Austin, 1997); Delli Carpini and Keeter, *What Americans Know About Politics*; Scott L. Althaus, 'Information Effects in Collective Preferences', *American Political Science Review*, 92 (1998), 545–58. The work of John R. Zaller, *The Nature and Origins of Mass Opinion* (New York: Cambridge University Press, 1992), though not framed as challenging this extenuationist literature, is also worth mentioning as forceful counterevidence.

<sup>18</sup> See Fishkin, *Democracy and Deliberation* and *The Voice of the People*.

carefully balanced panels of policy experts and political leaders. At the end, question the participants again, using the same instruments as at the beginning.<sup>19</sup> Excerpt portions of the proceedings to make a television programme.<sup>20</sup>

All this makes for a complex, multi-faceted event. The Deliberative Poll is an innovative means of gauging and communicating mass preferences, not as they are but as they would be with fuller information and reflection. It is a television programme, with its own educative value (apart from any impact of the poll results). And it is a social-scientific quasi-experiment, capable of shedding light on questions central to both political behaviour and democratic theory.

From a scholarly perspective, the most important aspect of the Deliberative Poll is as quasi-experiment. The one grand treatment consists of everything that happens from the moment of recruitment, immediately following the pre-deliberation questionnaire, to the post-deliberation questionnaire at the end. Notably, it includes:

- The anticipation of the event, of participating in discussions, of getting to put questions to experts and politicians, of possibly being televised in the process. Well before the deliberations proper, in the weeks between recruitment and arrival on-site, a good many participants have already begun paying more attention to relevant news stories, talking more about relevant topics with family and friends, and thinking about them more on their own.
- The exposure to information in a carefully balanced briefing booklet laying out the main proposals being discussed by political leaders and the arguments being made for and against them. Note by contrast that materials consumed under natural conditions are generally skewed by selective exposure.
- The small group discussions. The participants spend the largest portion of the working time on-site in groups of ten to fifteen or twenty, led by trained moderators, to discuss the probable consequences of the various policy proposals and how those consequences should be weighed. Random assignment, following on random sampling, means that the discussions feature a far wider variety of perspectives than most participants are likely to encounter in real life, where they talk mostly with others very like themselves. The moderators also help ensure that competing arguments are heard, again to a degree rare or unknown in real life.
- The opportunity to hear and question balanced panels of policy experts and

<sup>19</sup> The idea thus bears some resemblance to Dahl's notion of a 'minipolous', whose thousand or so randomly chosen members would deliberate electronically for up to a year. See Robert A. Dahl, *Democracy and Its Critics* (New Haven, Conn.: Yale University Press, 1989).

<sup>20</sup> See Fishkin, *The Voice of the People*. We should not go so far as to say that an event otherwise fitting this description would not be a Deliberative Poll without television, but every Deliberative Poll so far has been televised in some fashion. Apart from its potential for influencing great-world opinion, the televising of the proceedings may help raise the participants' sense of effectuality, increasing their felt incentives to learn about the issues. See Geoffrey Brennan and Loren Lomansky, *Democracy and Decision: The Pure Theory of Electoral Preference* (Cambridge: Cambridge University Press, 1993).

politicians. Yet again the balance is important. It is much harder than in real life to 'tune out' Tories, Labour supporters or others with whom one expects to disagree.

—The conversational spillover into corridors, dining rooms and bars, where the discussion of the issues often continues.<sup>21</sup>

Technically, this is only a *pre-* or *quasi-experiment*, to follow Campbell and Stanley's distinctions, both because it lacks the full measure of experimental control characteristic of laboratory experiments and because it lacks a true, i.e. randomly assigned, control group.<sup>22</sup> The first limitation is unavoidable, given the field setting and extended treatment period. The inability to exclude extraneous influences is largely shared by all field experiments, however, and the inevitable lack of time to digest and work through the implications of large quantities of new information in the laboratory means that anything resembling the Deliberative Poll can only be a field experiment. The second limitation may be minor, depending on the precise design. Several Deliberative Polls have now been accompanied by telephone interviews, at more or less the same time as the post-deliberation questionnaire was being completed by the participants, with separate random samples.<sup>23</sup> Attrition from interview to participant sample aside, this set-up is essentially equivalent to random assignment, bringing the telephone survey sample very close to a true control group.<sup>24</sup>

The immediate object of the experimental treatment is to make the participants more like ideal citizens, at least with respect to the topics under discussion. Thanks to the anticipation and experience of participation, they should become more interested in, think more carefully about, and learn more about the policy issues under discussion. Conceptually, these are three distinguishable gains – in *involvement*, *thought* and *information* – all of which may in turn affect policy preferences. All are closely related, however, and we

<sup>21</sup> The deliberation in Deliberative Polling thus engages what Brennan and Lomasky, in *Democracy and Decision*, call the 'expressive function of politics'. The process of publicly providing reasons for one's views affects the nature and complexity of those views. Unlike Brennan and Lomasky, however, we should not wish to abandon the secret ballot (see below).

<sup>22</sup> See Donald T. Campbell and Julian C. Stanley, *Experimental and Quasi-Experimental Designs for Research* (Chicago: Rand-McNally, 1963).

<sup>23</sup> In the National Issues Convention of January 1996 the telephone survey included the whole of the Deliberative Poll questionnaire; in the British General Election Deliberative Poll of April 1997 and the Australian national Deliberative Poll of October 1999 it was confined to vote intention and one or two other questions. The National Issues Convention and the Australian Deliberative Poll also reinterviewed as many as possible of either all or a random subsample of the non-participants (those initial interviewees who did not attend the deliberative weekend). The Australian Deliberative Poll was also closely followed by the Australian Election Study, using many of the same questions, with its own independent random sample, queried by mail rather than telephone.

<sup>24</sup> One other qualification concerns non-response in the recruitment of both treatment and control samples. But this is a reservation attaching to all samples, and a far graver one for the typical field experiment, where nobody frets over non-response because there is no probability sampling, nor even any well-defined universe to be sampling from.

confine our attention here to information, arguably the most important and certainly the most readily measurable.<sup>25</sup>

The scientific value of the Deliberative Poll is that it provides a way of addressing the effects of information (and thought and involvement) on policy preferences. Would people's policy preferences be appreciably different if they knew appreciably more about politics? Would only individual-level preferences be affected, with changes in opposite directions cancelling each other out in the aggregate, or would the distribution of opinions be appreciably different as well? These are critical questions, not only for our understanding of public opinion but for our understanding and evaluation of representative democracy.

Some approach to these questions may be made with survey data, and some work of this genre has already begun to suggest that information does affect both individual-level and aggregate preferences.<sup>26</sup> But these analyses must make do with cross-sectional comparisons of more with less informed respondents, controlling for some limited set of possible confounds statistically. With Deliberative Polling, in contrast, we can actually increase given respondents' levels of information (and thought and involvement) and observe the results.

#### THE FIRST DELIBERATIVE POLL

As of this writing, there have been eighteen Deliberative Polls, nine of them national. Five of the national ones have been conducted in Britain, one each in the United States and Denmark, and two in Australia. In Britain, the deliberative weekends have been held in Manchester and televised in Channel 4's 'Power and the People' series. The topics have been crime policy (April 1994), Britain's role in the European Union (June 1995), the future of Britain's monarchy (July 1996), the May 1997 General Election (April 1997), and the future of the National Health Service (July 1998). The American national event was the National Issues Convention, whose deliberative weekend took place in Austin, Texas, in January 1996. The topics there were the state of the American economy, America's role in the world and the American family. The Danish Deliberative Poll, deliberating in Odense in August 2000, concerned Denmark's national referendum on whether to adopt the euro. The first Australian Deliberative Poll, deliberating in Canberra, in October 1999, concerned the national referendum two weeks later on Australia's becoming a republic with

<sup>25</sup> On the relationship between information and thought, see Luskin, 'Measuring Political Sophistication'; on that between information-cum-thought or *sophistication*, on the one hand, and interest, on the other, Robert C. Luskin, 'Explaining Political Sophistication', *Political Behavior*, 12 (1990), 331–61 and Robert C. Luskin and Joseph C. Ten Barge, 'Education, Intelligence, and Political Sophistication' (paper presented at the annual meeting of the Midwest Political Science Association, Chicago, 1995).

<sup>26</sup> Bartels, 'Uninformed Voters'; Delli Carpini and Keeter, *What Americans Know About Politics*; Luskin and Globetti, 'Candidate versus Policy Considerations in the Voting Decision'; Althaus, 'Information Effects in Collective Preferences'.

a parliamentarily appointed president replacing the Queen as head of state. The second, deliberating in February 2000, also in Canberra, concerned 'aboriginal reconciliation'. Both received coverage on the ABC, the first on Australia's version of the American programme *Sixty Minutes* as well.

The remaining nine Deliberative Polls, all in the United States, have been regional. Eight have been 'Town Meetings', chiefly in Texas, sponsored by regulated electric utilities and focused on ways of providing for future energy needs.<sup>27</sup> The ninth was held in November 1999, drawing from the ninth through twelfth grade public school student population in Portland, Oregon. The topic was school and educational policies.

All the national Deliberative Polls but the Australian ones have been sponsored heavily or entirely by media organizations, all the regional Deliberative Polls but the one in Portland by electric utility companies. The Australian Polls have been sponsored primarily by Issues Deliberation Australia, a private foundation, the Portland Poll by the Nike Foundation. All eighteen have drawn probability samples, been partially televised in one fashion or other, and otherwise adhered closely to the outline above.

This article examines the first Deliberative Poll, held in Manchester in April 1994. It was designed by the present authors, and the survey work was done by the National Centre for Social Research (formerly SCPR). The subject was what to do about crime, widely and accurately perceived as a serious and worsening problem.

The sampling followed a three-stage probability sampling design from a universe consisting of all registered electors in all parliamentary constituencies south of the Caledonian Canal (i.e. omitting the north of Scotland). The first step was to draw forty parliamentary constituencies, with probability proportional to the number of registered electors; the second was to draw one polling district within each constituency, also with probability proportional to the number of registered electors; the third was to draw electors within each polling district, in this case with equal probability.<sup>28</sup> The first stage was stratified by standard region, population density, and the percentage of owner-occupiers. We call attention in passing to the contrast between this design and that of most British media polls, which at that point used quota samples.<sup>29</sup>

<sup>27</sup> The relevant populations have been the customers in the utilities' service territories, which in some cases have spilled over into other states. The first deliberative weekend in this series was hosted by Central Power and Light in Corpus Christi between 31 May and 2 June 1996, and the last by Texas Utilities in Dallas 16–17 October 1998. The others, in chronological order, were held by West Texas Utilities in Abilene; Southwestern Electric Power Company in Shreveport, Louisiana; El Paso Electric in El Paso; Houston Light and Power in Houston; Entergy in Beaumont; and Southwestern Public Service Company in Amarillo. Since the recent deregulation of the Texas electric power industry, substituting the verdicts of the market for the Deliberative Poll and other means of gathering public input, this series now seems to have come to an end.

<sup>28</sup> Using the February 1993 electoral registers, since those for 1994 were not yet available.

<sup>29</sup> The reliance of British pollsters on quota rather than probability samples may have played a major part in the generally gross underprediction of the Tory vote in 1992. See Roger Jowell *et al.*, 'The 1992 British Election: The Failure of the Polls', *Public Opinion Quarterly*, 57 (1993), 238–63.



The pre-deliberation survey was a face-to-face interview sandwiched around a self-administered questionnaire. An initial interview segment of about fifteen minutes was followed by a self-administered questionnaire of another fifteen minutes or so and then by a further interview segment of a few minutes more. In all, we obtained 869 respondents, for a response rate of 74 per cent. This *interview sample* as we shall call it, looks only trivially different, with respect to sociodemographic characteristics like gender, age and class, from the Census and the General Household Survey (GHS). For example, 49 per cent of the interview sample and 50 per cent in the Census were under age 45; 49 per cent of the interview sample and 52 per cent in the Census were women, and 49 per cent of the interview sample and 48 per cent in the GHS held non-manual occupations.

At the end of the second interview segment, the interviewer invited the respondent to participate in the Deliberative Poll a few weeks later. The inducements included the prospects of an expense-paid weekend at a good hotel in Manchester, a tour of the Granada Television Studios there, being on television, getting to learn and talk with others about an important topic, and getting to put questions directly to political leaders. There was also a monetary inducement of £50. With the aid of numerous phone calls to pursue the undecided, reassure the doubtful and keep in touch with the already committed, a total of 301 respondents appeared on the day. This admittedly self-selected subsample was also highly representative, as we shall see.

Meals, breaks, local transport time and limited free time aside, the working portions of the weekend broke down as follows. On Friday evening, the participants spent 45 minutes in plenary session being welcomed, watching a brief documentary describing the issues they would be discussing, and being reminded of what lay before them. On Saturday, they spent three-and-a-half hours in small group discussions, then three hours in large group exchanges with panels of experts fielding questions, then another hour back in the small groups.<sup>30</sup> On Sunday, they spent two hours in a plenary session questioning MPs representing each of the three major parties, then one last hour in the small groups. At the very end, they filled out the same self-administered questionnaire they had completed during the initial interview.<sup>31</sup>

The briefing materials were drafted and the moderators provided by *The*

<sup>30</sup> Due to logistical constraints, the sample was split in three for the large group (as distinct from plenary) sessions. Each expert panel gave three performances, rotating among the three rooms holding the large groups. Subsequent Deliberative Polls have avoided this arrangement. All the large group sessions have been plenary, and the plenary and small group sessions, limited to roughly 60 to 90 minutes apiece, have alternated through the day.

<sup>31</sup> This immediately post-deliberation measurement was followed by a third administration of the same questionnaire ten months later. We make no use of these third-wave observations here. Although they show some persistence, the question of persistence is secondary. The most important question, from the standpoint of the Deliberative Poll's rationale, is what changes result from placing the participants in this counterfactually stimulating environment, not what changes persist when they return to their everyday lives.

*Independent* newspaper. The Kettering Foundation, drawing on its experience with its National Issues Forums, a global network of small group discussions to facilitate citizen discussion of public policy issues, helped train the moderators and craft the briefing materials, which were also informed by our observation of several focus groups conducted by SCPR. The proceedings were videotaped by Granada Television, then distilled down to a two-hour programme broadcast two weeks later on the independent network Channel 4.

#### THE DATA

Between them, the interviewer- and self-administered questionnaires afford the usual range of sociodemographic variables, including measures of region and urbanness of residence, sex, race, marital status, family composition, occupation and education; media usage; party affiliation; experiences with crime; factual information about crime, crime policies and the criminal justice system; factual information about British politics generally; policy preferences related to crime; values that might underlie them; and empirical premises bearing on crime policies. Questions on which we foresaw little change – the sociodemographic and value items, especially – were assigned to the interviewer-administered questionnaire, whose contents were asked only once, before the event. All the items in the self-administered questionnaire, notably including all the policy items, were asked both before and after.

By far the most numerous questionnaire items, forty-nine in all, gauge *policy preferences*, the main dependent variables. All but six have five-point scales, running mostly from either strongly agree to strongly disagree, strongly in favour to strongly against, or very effective to not at all effective. Several items run from definitely preferring one of a pair of opposing alternatives to definitely preferring the other. The highest scores, of five, generally denote the most affirmative responses (agree strongly, very effective, etc.). The remaining six items stem from a question asking respondents to identify three of six alternatives as the most effective, next most effective, and third most effective ways of reducing crime. We scored the corresponding three items as 4, 3 and 2, respectively, and the remaining three items, corresponding to the three unchosen alternatives, as 1.

Substantively, all but one of the policy items concern either: (a) *root causes*: economic or social conditions, including, in the economic case, unemployment and poverty and, in the social case, the amount of time parents spend with their children, the level of violence and crime on television, and the firmness of discipline in the schools; (b) *policing*: putting more police on the beat, giving them more leeway and letting them carry guns; (c) *punishment*, including such issues as the tradeoff between punishment and rehabilitation, whether the proportion of convicts sent to prison should be higher or lower, how harshly they should be treated there, how long their terms should be, and whether the death penalty should be reinstituted for some crimes; (d) *procedural rights*, including both the broad tradeoff between the risks of convicting the innocent and of

letting the guilty go free and more specific questions about ways of either preserving or extending suspects' rights or, *per contra*, improving the efficiency of the courts and raising the odds of conviction by curtailing the rights of those suspected or accused of committing crimes; the treatment of *juvenile offenders*, including more and less severe alternatives for dealing with first-time juvenile burglars; and *self-protection*: including the ideas of people's 'making their property more secure', instituting 'more schemes like neighbourhood watch', and undertaking 'street patrols'.<sup>32</sup>

Also important, to both the Deliberative Poll and this article, are the items measuring information. Eight involve purely *factual information*, like the size of Britain's prison population, another three *empirical premises* of a more debatable sort, like the proposition that 'sending people to prison makes them *more*, not less, likely to become hardened criminals' (italics in question text). In the tables below, these latter are grouped with the policy preferences, with which they also share psychological ground, but literally, at least, they are closer to factual information – more debatable, but less so than policy preferences, and still matters of 'is' rather than 'ought'.

Half of the factual information items concern the British legal system, the other half British politics generally. All are true–false items scored 1 for correct answers and 0 for incorrect answers or don't-know's. The legal items (and their right answers) are: 'British courts are allowed to sentence a murderer to death' (false); 'Britain has a larger prison population than any other country in Western Europe' (true); 'Britain has more people serving life sentences than the rest of the European Community put together' (true); and 'it is possible to be tried by a jury in a local magistrate's court' (false). The corresponding political items, taken partly from the British Election Study knowledge scale,<sup>33</sup> are: 'the number of members of parliament is about 100' (false); 'the longest time allowed between general elections is four years' (false); 'Britain's electoral system is based on proportional representation' (false); and 'no-one is allowed to be on the electoral register in two different places' (false).<sup>34</sup>

#### REPRESENTATIVENESS

It is important to Deliberative Polling to start with a representative sample, the point being to gauge the opinions of a hypothetical public differing from the one we actually see only in respect of deliberation and variables affected by it. Thus

<sup>32</sup> The only policy item not fitting under these headings asks how far the respondent agrees or disagrees with Britain's legalizing 'certain so-called "soft drugs" like cannabis'. Further description of the policy and other items will emerge as we proceed, and full questionnaire texts are available on request from the Center for Deliberative Polling, Department of Government, University of Texas, Austin, TX 78712, USA.

<sup>33</sup> Jean Martin, Anthony Heath, Karl Ashworth and Roger Jowell, 'Development of a Short Quiz to Measure Political Knowledge', *CREST Working Paper No. 21* (London: Centre for Research into Elections and Social Trends, 1993).

<sup>34</sup> This last item was poorly put. See below.

we worried that the self-selected subsample who came to Manchester would be significantly biased in some material respect. Perhaps the very young or very old or people with small children would be especially unlikely to come. Perhaps women would be more reluctant than men. Quite possibly, we thought, respondents who had some personal experience of crime or cared particularly strongly about the topic would be especially likely to come. Any such bias would be worrisome to the extent that it was both sizeable and correlated with opinions about crime – or, still more so, with the propensity to change one's opinion or to change one's opinion in some given direction.

By any reasonable standard, these worries proved groundless. On the great majority of items the 301 participants differed only insignificantly from the other 569 interviewees.<sup>35</sup> In all, only fourteen of 102 items exhibit a statistically significant difference at the 0.05 level. Virtually all these differences, moreover, even the statistically significant ones, are relatively small.

The only real exceptions have to do with legal and political knowledge. On three of four legal knowledge items and two of four political knowledge items, participants were 7–11 per cent more likely to know the right answer. While we should naturally prefer that these differences be zero, we hasten to point out that: (a) they are still fairly modest; (b) they obtain in distinctly lesser degree for the one remaining legal knowledge item and not at all for the two remaining political knowledge items, so that participants and non-participants look somewhat less different on our summary measures of legal and political knowledge; (c) ordinary polls generally possess the same sort of bias (in generally unknown degree); and (d) the likely impact of starting with a more knowledgeable than average sample of participants-to-be is to attenuate participation's apparent effect by reducing the potential for information gains.

#### CHANGES IN POLICY PREFERENCES

It is still more centrally important that Deliberative Polling affect the participants' policy preferences. On at least some policy items (more than would be expected by chance), the post-participation distribution, of relatively informed preferences, should look noticeably different from the pre-participation distribution, of relatively uninformed ones. This need not be true of every Deliberative Poll, but it should be of most, for otherwise Deliberative Polling, while very likely still contributing to the wisdom and engagement of its participants, would be revealing nothing different from what ordinary polling does.

<sup>35</sup> The comparisons can be found in Appendix A on the website version of this article, which presents difference-of-means (*t*) tests for binary and other interval items and ordinal items of more than three categories (most have five or more) and difference-of-distributions ( $\chi^2$ ) tests for ordered trichotomous and nominal items.

### *Net Change*

By this criterion, this first Deliberative Poll fared very well. Do increased information and thought alter the multivariate distribution of opinion? Yes, they do. On a sizeable majority of policy items, opinions underwent statistically significant net change. Table 1 presents the means, before and after participation, of the participants' positions on every policy item (including the three empirical premises), along with the *p*-value from a paired comparison test of the significance of the difference. Of fifty-two items, thirty-five or 67.3 per cent showed statistically significant net change at the 0.05 level. No policy type is exempt.

Substantively, moreover, many of these changes are impressive.<sup>36</sup> In principle, if everyone started at one extreme of the scale and ended up at the other, the net change could be as large as  $\pm 4$  on a five-point scale. For reasons closely allied to those sketched in our discussion of gross change below, however, nothing close to this will ever actually happen. Many participants, starting not at 1 or 5 but at 2, 3 or 4, can change only 3 points (from 2 or 4) or 2 points (from 3). Many, starting at whatever point, will in fact change less, either 1 point or not at all, either because they are already close to where they 'should be' (many are, even if many are not) or by erring consistently (by either misguided striving or chance). And many of the changes in one direction, finally, will be negated by changes in the other direction. Taking all this into account, a mean change of say  $\pm 1.0$  on a five-point scale would be extraordinarily, verging on unattainably, large. None of the changes in Table 1 is quite that seismic, but ten of the five-point scales show changes of  $\pm 0.35$  or more, while three of the four-point scales show changes of  $\pm 0.26$  or more.<sup>37</sup>

In particular, there seem to have been four or five major shifts. The first was an increased sense of the limitations of prison as a tool for fighting crime. Twelve of sixteen items under the 'punishment' heading showed statistically significant net change, and all twelve of the significant changes were towards less punitive attitudes. Participants left wanting to send fewer criminals to prison, less inclined to stiffen criminal sentences, and more willing to support alternatives to prison for 'criminals who are not a big threat to society'. They became more hesitant about imposing a life sentence automatically for all murders. Forced to choose between punishing criminals and reforming them, they left more inclined to try reform. Regarding 'juvenile offenders', similarly, the significant changes were that two proposals for dealing leniently with first-timers – that they be just warned or sent to a special institution – received increased support, while one

<sup>36</sup> They also run distinctly larger than the differences between participants and non-participants in Appendix A in the website version of this article, despite there being ample reason (outlined in this paragraph) to expect them to be ceilinged quite low and no reason, apart from good sampling and the absence of much self-selection bias, to expect the latter to be similarly low.

<sup>37</sup> These latter must be interpreted cautiously, however, given that, missing data aside, these six items must average 2.0 for each respondent (and thus for the sample) by construction. A net decrease in one must therefore be balanced by a net increase in at least one of the others.

TABLE 1 *Net Changes of Position in Policy Preference*

Variable	Mean $P_1$	Mean $P_2$
<i>Punishment</i>		
Punishment vs. reform	3.29	2.97 *
Prisons reform, not just punish	3.99	4.06
Tougher sentences	4.15	3.92 **
Stiffer sentences generally	4.18	3.79 **
More offenders to prison	3.68	3.19 **
Fewer people to prison	2.79	3.14 **
Stiffer sentence and more prisons†	2.17	1.95 **
Only dangerous criminals to prison	3.07	3.44 **
More offenders out of prison	3.41	3.58 *
More offenders: community service	3.76	3.99 **
More offenders: military service	3.64	3.29 **
More offenders get training	3.64	3.94 **
Prison life tougher	3.87	3.88
All murderers life sentence	4.33	3.99 **
Life sentence means life	4.48	4.45
Death most appropriate for some crimes	3.80	3.79
<i>Root Causes (Economic)</i>		
Reduce unemployment	1.83	1.93
Reduce unemployment and poverty†	3.10	2.35 **
Causes vs. apprehension	1.97	1.86
<i>Root Causes (Social)</i>		
Spend more time with children	4.37	4.61 **
Parents more time with children†	2.36	2.25
Teach children right from wrong	4.57	4.72 **
Reduce TV violence and crime	3.79	3.97 **
Firmer school discipline	4.28	4.44 **
Schools responsible for moral guidance†	1.74	2.19 **
<i>Police</i>		
More police on the beat	4.50	4.33 **
On-duty police carry guns	2.67	2.36 **
Police more powers to catch criminals†	2.28	2.14
<i>Procedural Rights</i>		
Convict innocent. vs. fail to punish guilty	2.54	2.16 **
Police allowed to bend rules	1.93	1.73 **
Fewer jury trials	3.04	3.12
Court rules less on side accused	3.25	3.47 **
End presumption of innocence	2.66	2.64
Silence mentionable in court	3.39	2.96 **
Right to silence in police questioning	2.81	3.19 **
Confessions alone not convict	3.67	3.93 **
Ind. investigations of complaints against police	4.13	4.35 **
<i>Juvenile Offenders</i>		
1st-x child burglar: just warn	3.24	3.60 **
1st-x child burglar: community service	3.99	4.07

TABLE 1 *Continued*

Variable	Mean $P_1$	Mean $P_2$
1st-x child burglar: juvenile institution	2.80	2.47 **
1st-x child burglar: ordinary prison	2.04	1.74 **
Special tough institution for repeat juvenile offenders	3.76	3.69
What age treat as adults	3.43	3.44
<i>Self-Protection</i>		
People vs. police to protect	3.67	3.53
People take responsibility for prop.†	1.64	1.90 **
People make property more secure	4.26	4.11 *
More Neighborhood Watch schemes	4.04	3.85 **
Street patrols vs. police	3.06	3.00
<i>Miscellaneous</i>		
Legalize soft drugs	2.71	2.63
<i>Empirical Premises</i>		
Illegal drugs major crime cause	4.00	4.37 **
Prison life too soft/tough	2.05	2.10
Prison hardens criminals	3.35	3.48 *

Note: Mean  $P_1$  and Mean  $P_2$  denote sample mean positions before and after deliberation. The ranges are given in Table A1 in Appendix A of the website version of this article.

† By construction, these six items must average 2.0.

\* Significant at the 0.05 level (two-tailed).

\*\* Significant at the 0.01 level (two-tailed).

proposal for dealing strictly with them – that they be sent to ordinary prisons – received decreased support.

At the same time, the deliberative experience scarcely made the participants soft on crime. Large and unchanged majorities supported restoring the death penalty, ensuring that ‘life sentences ... *mean* life’ (italics original), and making ‘prison life ... tougher and more unpleasant’. And on the items most directly related to punishment or prison, apart from those about ‘criminals who are not a big threat to society’, the net changes were mostly from highly punitive to somewhat less punitive. The two items concerning the proportion of criminals who should be sent to prison ended up near the middle of the five-point scale, but the two tougher-sentencing items and the question about whether murderers should always receive life sentences ended up – after significant mellowing – close to 4 (with 5 being the sternest response). In all, it does not seem too far from the mark to describe the participants as having begun the Deliberative Polling process wanting to ‘hang ’em high’ and having ended it wanting, on the whole, to hang ’em somewhat lower.

The second major shift was towards increased sensitivity to defendants’ procedural rights. Seven of the nine items under this rubric showed significant change. The participants did leave somewhat more likely to endorse the general notion that ‘the rules in court should be less on the side of the accused’, but all the other significant net changes were in the direction of supporting procedural

rights. There was decreased willingness to wink at police rule-bending and increased support for having 'complaints against the police ... investigated by an independent body'. There was increased support for 'the right to remain silent under police questioning', decreased willingness to see the exercise of that right count against the defendant in court, and increased support for discounting confessions obtained under police questioning in the absence of other evidence. Forced to choose between the risks of convicting the innocent and letting the guilty go free, an increased percentage opted for the latter.

The third major shift was away from faith in the value of *policing*, meaning surveillance by the police, individual citizens or citizens' organizations, and, relatedly, of *obviation*, meaning the reduction of such temptations as visibly unattended and easily removed property. Thus our participants largely edged away both from reliance on self-protection, as reflected in diminished support for 'neighbourhood watch schemes' and for 'people making their property more secure', and from proposals for increasing the numerical strength or powers of the police.<sup>38</sup> Their enthusiasm for putting 'more police on the beat', while remaining high, diminished somewhat, as did their enthusiasm, never high, for the proposals that 'on-duty police officers ... normally carry guns' and that the police be given 'greater powers to catch criminals'.

The fourth major shift was towards ameliorating social root causes, towards an emphasis on what might be thought of as 'family values'. All but one of the six items under this rubric exhibited statistically significant net change of this sort. Participants warmed to the ideas of 'parents spending more time with their children', teaching children right from wrong, having 'less violence and crime on television', having 'firmer discipline in the schools', and 'schools being more responsible for moral guidance' as effective ways of reducing crime.

Finally, the fifth major shift, of which we are somewhat less confident, given that there are only two relevant items, was away from dealing with the economic root causes of crime, specifically unemployment and poverty. Of the two relevant items, one showed significantly decreased support and the other nearly so. This may in some sense have been complementary to the increased emphasis on dealing with *social* root causes, stemming from a change in perception of where the root causes of crime actually lie.

Why these particular changes? We can offer a few impressions, based on firsthand observation and statistical bits and pieces. With respect to punishment, many of the participants seemed to absorb the information that imprisonment is very expensive and the arguments that it may be ineffective, even counterproductive. One statistical glint of this may be seen (in Table 1) in the increased endorsement of the empirical premise that 'sending people to prison makes them *more*, not less, likely to become hardened criminals' (*italics*

<sup>38</sup> The first half of this observation is somewhat gainsaid by the *increased* support for 'people taking more responsibility for keeping their property safe', but we suspect that the operative word in this formulation is 'responsibility', rendering it more of a piece with the proposals regarding social root causes.



original). Another may be seen (in Table 3) in the huge increases (both over 30 percentage points) in the percentages knowing that 'Britain has a larger prison population than any other country in Western Europe' and that 'Britain has more people serving life sentences than the rest of the European Community put together'. Juxtaposed with rising crime, these and similar facts and arguments seemed to give those who initially favoured sending more criminals to prison and treating them more harshly while there some pause.

We suspect that most of our participants had never before given much thought to procedural issues or had much idea of what protections existing procedures afford, much less of their rationales. We know that they heard and participated in considerable discussion of these matters in Manchester (and had the opportunity of reading about them in the briefing materials beforehand). The 10 percentage point increase in the percentage knowing that 'it is [not] possible to be tried by a jury in a local magistrate's court' (see Table 3) suggests some relevant learning. But this shift may also have been partly a matter of thought (specifically, framing), not just information. Certainly the expert arguments favouring procedural rights asked participants to imagine themselves in the role of the (unjustly) accused, something many of them may not frequently have done.

For the rest, we have little or no statistical evidence, for want of relevant items, but our impressions from firsthand observation generally accord with the thrusts of the policy attitude changes we have been describing. Participants heard how difficult it is to stop a determined thief, which may have curtailed support for policing and obviation. Many of the discussions featured accounts of the well-enough-off who commit crimes and of the genuinely poor who do not, along with assertions of the importance of values, and of the family's role in instilling them. And it may also be worth mentioning that there was considerably increased agreement with the empirical premise that 'some people's need for *illegal drugs* is a major cause of crime in Britain' (see Table 1, italics original), which may resonate more with family values than with economic distress as a root cause of crime.

### *Gross Change*

But net change is not the only way of reckoning a Deliberative Poll's success. As quasi-experiment, Deliberative Polling is also successful to the extent that thought and information change individual preferences, even if all the changes cancel each other out, leaving the aggregate distribution of preferences unchanged. From this perspective, net changes we have just been examining may be understatements. There was virtually no net change, for example, in support for either the death penalty or the presumption of innocence, but only 58 per cent gave precisely the same answer both times to the first question, while only 45 per cent did so both times to the second.

In Table 2, therefore, we present several measures of gross change: the proportions of the sample who *change position* on our five-point scales, who

*change side* (counting movement in or out of neutrality), and who *change side completely* (disregarding movement in or out of neutrality). The first proportion is necessarily the highest, the third necessarily the lowest.<sup>39</sup>

So how much gross change is there? The percentage changing position never dips below 33 per cent, regularly exceeds 50 per cent, and occasionally tops 60 per cent. The percentage changing side runs in the 30–40 per cent range. Only one item, concerning the effectiveness of teaching children right from wrong, comes in under 10 per cent. Even the percentage changing side completely is frequently in the teens, sometimes in the twenties. On the question asking whether ‘only hardened criminals, or those who are a danger to society, should be sent to prison’, it nears 30.

Least these last percentages seem slender, it is worth bearing in mind that they reach 100 per cent only when everyone who agrees beforehand disagrees afterwards and vice versa (and no one is ever in the middle), the sort of thing that ‘happens only in Gilbertian libretti’.<sup>40</sup> In practice, they are ceilinged far lower, since nobody already on his or her ‘full-information’ side is likely to change.<sup>41</sup> Realistically, then, how large can the percentage changing side completely be? Any answer must be rough and dependent on a number of assumptions, but it is hard to find plausible parameters for which the percentage changing sides completely exceeds 30 per cent.<sup>42</sup> In this light, a great many of the percentages changing side completely, even those in the mid-to-high single digits, are impressive. Those in the twenties are astonishing.

As with net change, there are some noticeable variations. Proposals for dealing with social root causes see relatively little gross change, as might

<sup>39</sup> Table 2 includes only five-point scales, the notion of ‘side’ being meaningless for the six four-point scales included in Table 1.

<sup>40</sup> Luskin and Globetti, ‘Candidate versus Policy Considerations in the Voting Decision’.

<sup>41</sup> Saving those who are there merely by chance, having made lucky non-attitudinal guesses at  $t_1$  (pre-participation), and gained sufficiently little from their participation that they are still guessing at  $t_2$  (post-participation). See fn. 42.

<sup>42</sup> Suppose illustratively that the  $t_1$  distribution on the five-point scale is uniform, with 20 per cent in each category. Immediately we are down to 80 per cent who can change side completely, since change from the neutral category is excluded by definition. Assume further, pessimistically for the Deliberative Poll and conservatively for this argument, that 30 per cent of this 80 per cent choose sides randomly, even at  $t_2$ . Among the remaining 70 per cent, assume that everyone who is not already on his or her ‘full-information’ side at  $t_1$  moves there at  $t_2$ , while no one else changes side. If the correlation (Kendall’s  $\tau_{au-b}$  = Pearson’s  $r$ ) between ‘full-information’ and  $t_1$ , side is 0.4, while five times as many people change in one direction as the other, roughly 32 per cent of the 70 per cent will change side, as of course will 50 per cent of the 30 per cent. The percentage of the sample changing side completely, therefore, will be approximately  $80\% * [(30\% * 50\%) + (70\% * 32\%)] = 28.9\%$ . If we assume a larger percentage at the midpoint at  $t_1$ , fewer than 100 per cent of those on the ‘wrong’ side at  $t_1$  moving to the ‘right’ side at  $t_2$ , a higher correlation between ‘full-information’ and  $t_1$  side, a lower percentage still guessing at  $t_2$ , or a less imbalanced traffic between sides, this ceiling is still lower. For parallel calculations, see Robert C. Luskin, ‘From Denial to Extenuation (and Finally Beyond)’, in Kuklinski, ed., *Thinking about Political Psychology*, and Luskin and Globetti, ‘Candidate versus Policy Considerations in the Voting Decision’.

TABLE 2 *Gross Changes in Policy Preference*

Variable	Percentage changing position	Percentage changing side	Percentage changing side completely
<i>Root Causes (Economic)</i>			
Reduce unemployment	50.0 **	26.6 **	9.8 **
Causes vs. apprehension	46.7 **	26.0 **	21.9 **
<i>Root Causes (Social)</i>			
Spend more time with children	41.7 **	14.5 **	2.4 **
Teach children right from wrong	33.0 **	6.7 **	2.0 *
Reduce TV violence and crime	48.7 **	26.7 **	7.1 **
Firmer school discipline	41.7 **	14.6 **	2.4 **
<i>Police</i>			
More police on the beat	40.7 **	17.2 **	3.7 **
On-duty police carry guns	51.3 **	32.8 **	5.7 **
<i>Punishment</i>			
Punishment vs. reform	57.0 **	33.7 **	23.4 **
Prisons reform, not just punish	53.3 **	29.9 **	8.9 **
Tougher sentences	52.0 **	26.1 **	6.6 **
Stiffer sentences generally	53.3 **	34.0 **	8.5 **
More offenders to prison	64.7 **	46.6 **	12.3 **
Fewer people to prison	60.3 **	45.7 **	13.7 **
Only dangerous criminals to prison	61.7 **	44.9 **	28.9 **
More offenders out of prison	59.0 **	42.2 **	21.3 **
More offenders: community service	54.0 **	27.2 **	12.9 **
More offenders: military service	56.3 **	35.5 **	16.9 **
More offenders get training	57.0 **	35.9 **	13.6 **
Prison life tougher	51.0 **	29.8 **	10.2 **
All murderers life sentence	46.7 **	25.3 **	12.0 **
Life sentence means life	38.0 **	14.3 **	5.5 **
Death most appropriate for some crimes	42.3 **	20.8 **	8.2 **
<i>Procedural Rights</i>			
Convict innocent. vs. fail to punish guilty	52.3 **	32.6 **	22.3 **
Police allowed to bend rules	43.3 **	17.3 **	6.4 **
Fewer jury trials	58.3 **	44.5 **	16.6 **
Court rules less on side accused	61.0 **	48.8 **	13.4 **
End presumption of innocence	54.7 **	36.1 **	17.7 **
Silence mentionable in court	59.7 **	46.8 **	26.4 **
Right to silence in police questioning	60.0 **	43.2 **	26.2 **
Confessions alone not convict	54.0 **	35.4 **	14.3 **
Ind. investigations of complaints against police	43.0 **	14.2 **	4.4 **
<i>Juvenile Offenders</i>			
1st-x child burglar: just warn	63.0 **	44.2 **	23.4 **
1st-x child burglar: comm. service	55.7 **	22.9 **	9.5 **
1st-x child burglar: juvenile institution	63.3 **	43.6 **	16.1 **
1st-x child burglar: ordinary prison	52.7 **	20.7 **	5.9 **
Tough institution for repeat juvenile offenders	57.7 **	37.4 **	14.5 **

TABLE 2 *Continued*

Variable	Percentage changing position	Percentage changing side	Percentage changing side completely
<i>Self-Protection</i>			
People vs. police to protect	56.7 **	37.2 **	27.1 **
People make property more secure	47.3 **	18.6 **	4.5 **
More Neighborhood Watch schemes	55.0 **	31.2 **	7.1 **
Street patrols vs. police	55.3 **	41.6 **	9.5 **
<i>Miscellaneous</i>			
Legalize soft drugs	53.0 **	33.2 **	11.7 **
<i>Empirical Premises</i>			
Illegal drugs major crime cause	46.0 **	23.8 **	14.1 **
Prison life too soft/tough	42.7 **	22.7 **	2.4 **
Prison hardens criminals	57.0 **	44.9 **	10.5 **

\* Significant at the 0.05 level (one-tailed).

\*\* Significant at the 0.01 level (one-tailed).

be expected of items tapping relatively well-formed attitudes, towards family, morality, children and schools (though again 'relatively little' is still quite a lot). In contrast, questions about punishment and procedural rights show particularly widespread gross change, as to a lesser extent do those about the treatment of juvenile offenders. Perhaps excepting punishment, these three are domains in which pre-deliberation attitudes were probably not very well crystallized, and in all three, notably including punishment, the participants encountered a great many facts and empirical premises capable of altering opinions.

#### INFORMATION GAINS

It is still not enough, however, to show that many participants' policy preferences changed. The rationale of Deliberative Polling requires that such changes be not merely random bouncing around, nor the outgrowth of some purely social dynamic, but rather the product of learning and reflection, of coming to see consequences for valued goals more clearly and weigh them more carefully – in short, that the post-deliberation opinions be, for many participants, at least, more *considered*.

We may begin by establishing that the participants did learn. Table 3 gives the proportions getting each of the legal and political knowledge items correct, as well as the means of the summary measures averaging these proportions within each domain.

On the surface, the participants may seem not to have learned much about politics beyond the issues at hand. Three of the four more general political knowledge items show slight, statistically insignificant *decreases*, and the mean percentage of correct answers across the four items also decreased, also insignificantly, from 42.3 to 40.1 per cent. It is worth noting, however, that other

TABLE 3 *Legal and Political Information Gains*

Variable	Mean 1	Mean 2	Diff.
<i>Legal Information</i>			
Death sentence in Britain?	85.0	79.7	- 5.3
Britain largest prison population	49.7	80.0	30.3 **
Britain more life sentences	20.3	58.0	37.7 **
Jury trial in local magistrate's court?	58.3	68.0	9.7 **
<i>Political Information</i>			
Number of MPs about 100?	69.0	67.0	- 2.0
Max 4 yrs between general elections?	27.3	24.7	- 2.7
Britain has proportional representation?	54.7	54.7	0.0
Can't be on electoral register in 2 places	18.3	14.0	- 4.3
<i>Summary Indices</i>			
Legal information	53.3	71.4	18.1 **
Political information†	50.3	48.8	- 1.6
Overall information†	52.0	61.7	9.7 **

*Note:* Each item is scored {0,1}, and the means of the individual items are thus the percentages of participants answering them correctly. Diff. denotes Mean 2 - Mean 1. The summary variables 'Legal information', 'Political information' and 'Overall information' are the means of the respondents' scores on the items in the legal domain, the political domain and the two domains combined. Each mean is the percentage of respondents who answered the item correctly.

† These summary indices omit the electoral register item.

\* Significant at the 0.05-level (one-tailed).

\*\* Significant at the 0.01-level (one-tailed).

Deliberative Polls have shown larger increases in general political information and that even in the present case a third interview some ten months later, whose results we keep for a later paper, does show significant increases. The participants do seem to have been stimulated to learn discernibly more about politics generally, if mainly in the longer run.

Even in the short run these numbers may be too low. The electoral register item, showing both the smallest percentage correct and the largest decrease, amounted to a trick question: you *cannot* vote but *can* be on the electoral register in two different places in the same election. It would not be unreasonable to reverse this item, counting the technically right answer (false) as wrong and the technically wrong answer (true) as right. The percentage getting this item right would then increase from 82 per cent at time 1 to 86 per cent at time 2, and the mean percentage getting the four political information items right would increase, still insignificantly, from 54.8 per cent at time 1 to 58.5 per cent at time 2. A safer solution, however, and the one we adopt, is simply to discard the item. That leaves a decrease, but a still smaller and still less significant one, in the mean percentage getting the remaining three general information items right, as reported in Table 3.

The participants learned much more – indeed a very great deal – about the British legal system. Three of the four relevant items show significant

increases, the lone exception being the question about the possibility of jury trials in local magistrates' courts. The percentages knowing that Britain has the largest prison population in Western Europe and that it has more people serving life sentences than the rest of the European Community combined increased by nearly 40 per cent! The mean percentage of the four items answered correctly increased by nearly 20 per cent.

The final line of Table 3 shows the mean percentage of all seven general and crime-specific factual information items answered correctly (with the electoral register item discarded). The mean percentage answering correctly increases by 9.7 percentage points, from 52.0 per cent at time 1 to 61.7 per cent at time 2.<sup>43</sup> The increase is statistically significant beyond the 0.01 level. This is the information index we use in the further analysis below.

#### WHO CHANGED?

The remaining challenge is to show that the preference changes and information gains were linked, that it was indeed those who learned most who changed most (and in the same direction as the participant sample as a whole), so that it was they who were driving the net changes. To keep the analysis to manageable proportions, we reduce individual policy items to summary indices. The items fall naturally into just five multi-item clusters, gauging the extent to which the respondent favoured imposing stricter and more certain *punishment*, giving greater attention to ameliorating the *social root causes* of crime, relaxing *procedural protections* for those suspected or accused of crimes, giving the *police* more power and resources, and placing greater emphasis on *self-protection*. In each case, we have used all or virtually all the relevant five-point scales, jettisoning only the handful that on the basis of correlations and factor analyses did not cohere very well with the rest. We reflected items as necessary, so that higher scores always pointed towards approval of the policy positions just stated, then projected the average on to the [0, 1] interval.<sup>44</sup>

Table 4 shows the Time 1 and Time 2 means and the net changes in these five policy indices. Consistent with the item-level results already presented, these figures show diminishing support for imposing stricter and more certain punishment, relaxing procedural protections, giving the police more power and resources, and placing greater emphasis on self-protection and increasing support for addressing social root causes. The Time 1–Time 2 differences may look surprisingly small but that is because these indices, in contrast to the individual items in Table 1, have been transformed to the 0–1 scale. All the Time 1–Time 2 differences in Table 4 are statistically significant at the 0.0001 level.

A first question is the extent to which opinion change was driven by small

<sup>43</sup> This conclusion does not greatly depend on the choice of how to treat the electoral register item. If the item is left as originally scored the mean increase is still 7.9 percentage points, still significant beyond the 0.01 level. If it is reversed, the mean increase is 10.9 percentage points.

<sup>44</sup> Appendix B in the website version of this article lists the ingredients.

TABLE 4 *Descriptive Statistics for Policy Indices*

	Time 1 ( $t_1$ )		Time 2 ( $t_2$ )		Mean $t_2 - t_1$
	Mean	Std. Dev.	Mean	Std. Dev.	
Punishment	0.597	0.157	0.537	0.180	- 0.059**
Social Root Causes	0.787	0.176	0.835	0.136	- 0.048**
Police	0.648	0.198	0.586	0.205	- 0.062**
Procedural Rights	0.445	0.169	0.406	0.173	- 0.039**
Self-Protection	0.696	0.162	0.662	0.171	- 0.034**

\* Significant at the 0.001 level (two-tailed test).

\*\* Significant at the 0.0001 level (two-tailed test).

group mechanisms. These could be a matter of the balance of the arguments in the discussions, tilting this way in one group, that way in another. Or, less charitably, they could be a matter of simple, unreasoning conformity. In either event, they should cause the group mean to become more extreme with respect to either the scale midpoint (0.5) or the whole participant sample's initial centre of gravity, as group members sway in the direction of *their group's* initial centre of gravity. A sizeable literature, nicely reviewed by Sunstein,<sup>45</sup> suggests that many real-world small groups, notably including juries, do tend, under most conditions, to polarize or become more extreme in at least the second of these two senses, and presumably the first as well. Thus Table 5 presents the proportions of the twenty groups whose mean opinions became more extreme in both senses.

TABLE 5 *Small Group Mechanisms: A First Look*

	Percentage of the within-group means changing towards the same side of the midpoint as the group is on at $t_1$	Percentage of the within-group means changing towards the same side of the grand mean as the group is on at $t_1$	Percentage for which within-group variance decreases
Punishment	10.0	60.0	25.0
Social Root Causes	100.0	50.0	85.0
Police	0.0	50.0	40.0
Procedural Rights	90.0	50.0	55.0
Self-Protection	30.0	45.0	60.0
Overall	46.0	51.0	53.0

<sup>45</sup> Cass R. Sunstein, 'Deliberative Trouble? Why Groups Go to Extremes', *Yale Law Journal*, 110 (2000), 71-120. Sunstein misconstrues the evidence from Deliberative Polling, however. Compare his discussion with ours here.

As can readily be seen, these results are quite at odds with what would be expected if small group mechanisms were solely or even largely responsible for opinion change. The small groups do not become consistently more extreme, by either standard. On some policy indices, a sizeable majority of them do indeed move away from the centre, but on others an equally sizeable majority towards it. The variation from issue to issue is greater (indeed quite extreme) when extremity is reckoned from the scale midpoint, but the most important lesson is the same, no matter what the baseline. Of the 100 ( $= 20 \times 5$ ) group-issue combinations only about half show movement away from the centre – 46 per cent when the centre is defined as the midpoint, 51 per cent when it is defined as the overall Time 1 mean. The other half show movement towards the centre.

Conformity mechanisms should also cause the within-group variance to shrink, as group members come to think more alike. Thus Table 5 also presents the proportions of the twenty small groups whose within-group variance diminished on each of the five policy indices. But the story here is similar. Depending on the policy index, the within-group variances may mainly shrink or mainly expand. Interestingly, the issue showing the most uniform shrinkage (punishment) and the issue showing the most uniform increases (social root causes) are both the sort of 'easy', emotion-laden issues on which Time 1 attitudes might be expected to be firmest.<sup>46</sup> Overall, the within-group variances diminish on only just over half (53 per cent) of the hundred group-issue combinations. On the remaining near-half, they increase.

Another question is the extent to which opinion change was confined to members of certain sociodemographic groups. Perhaps only the well educated or otherwise advantaged could make effective use of the information in the briefing materials and discussions, or perhaps only the poorly educated or otherwise disadvantaged could be swayed. The question may be answered by regressing the individual-level changes in the policy indices described in Table 4 on the usual sociodemographic suspects, specifically here: gender, age, urbanness (broken into rural and urban dummies, with city as the omitted category), number of children, marital status (broken into married and formerly married dummies, with never married as the omitted category), employment status (again a dummy, separating the employed from the unemployed), education (on a seven-point scale), race (a dummy separating non-whites from whites), and class (a dummy separating the middle-class from the working-class).

The results, in Table 6, show that the attitude changes we have seen are not well explained by sociodemographic characteristics. Only one of the five equations, the one for social root causes, is significant as a whole. The others miss by miles. Even that one significant equation has an adjusted  $R^2$  of only 0.06.<sup>47</sup> The remaining four  $R^2$ s average 0.004. The social root causes equation

<sup>46</sup> See Edward G. Carmines and James A. Stimson, 'The Two Faces of Issue Voting', *American Political Science Review*, 74 (1980), 78–91.

<sup>47</sup> On the importance of the adjusted  $R^2$ , even when the equation shows 'significant', see Robert C. Luskin, 'Abusus Non Tollit Usum: Standardized Coefficients, Correlations, and  $R^2$ s', *American Journal of Political Science*, 35 (1991), 1032–46.



TABLE 6 *Changes in Policy Attitudes as Functions of Sociodemographic Characteristics*

	Punishment (-)	Social Root Causes (+)	Police (-)	Procedural Rights (-)	Self- Protection (-)
Constant	-0.072*	0.189*	-0.094*	-0.055	-0.006
Gender	0.025	0.009	-0.012	0.007	-0.012
Age	0.000	-0.003*	0.001	0.000	-0.001
Rural†	0.011	-0.005	0.024	0.037	0.015
Urban†	0.025	0.023	-0.001	0.051*	0.040
Number of children	-0.010	-0.019*	0.001	0.002	-0.020
Married‡	0.000	0.059*	-0.009	0.008	0.048
Separated‡	-0.004	0.038	0.047	-0.011	-0.032
Employed	-0.021	-0.034	0.014	-0.018	-0.021
Education	-0.001	0.001	-0.009	-0.003	-0.003
Non-white	0.063	-0.063	0.027	0.043	0.050
Middle-class	-0.022	-0.015	0.007	0.027	0.009
$R^2$	0.052	0.099	0.041	0.033	0.042
adj. $R^2$	0.014	0.063	0.003	-0.005	0.004
$F$	1.383	2.777	1.081	0.871	1.114
Probability	0.180	0.002	0.377	0.569	0.350
$n$	292	291	292	291	292

Note: The parenthetical signs in the column headings indicate the direction of net change for the sample as a whole.

† *Type of area*, originally a trichotomous coding of population density, is broken into the two dummies, *rural* and *urban*, with *city* as the 'omitted category'.

‡ *Marital status*, originally trichotomous, is broken into the two dummies. The *married* dummy contains respondents who are currently married (or living as married). The *separated* dummy contains respondents who are separated, widowed, or divorced. The 'omitted category' is never married.

\* Significant at the 0.05 level (two-tailed).

\*\* Significant at the 0.01-level (two-tailed).

has three significant coefficient estimates, for age, being married and number of children. The sample as a whole moves towards greater support for addressing social root causes (as indicated by the positive parenthetical sign at the top of that column), but younger people and people with more children change less and married people more than others in that direction (as indicated by the signs on the relevant coefficient estimates). Across the table as a whole, however, only 7 per cent of the estimated coefficients (four of fifty-five, not counting the intercepts) are statistically significant, scarcely above the 5 per cent to be expected by chance. It may be particularly worth nothing that the estimated coefficient for education, that most usual of suspects, is nowhere significant.

These first two bits of who-changed analysis help assuage a couple of concerns. The hope of Deliberative Polling is to enable everyone, not merely certain types of people, to come to views based on much fuller thought and information, not merely the preponderance of the opinions of others around them. It would be troubling if the changes of opinion were strictly a matter of

social conformity or confined to certain social strata. But neither, fortunately, seems to be the case. Small group mechanisms (not confined to social conformity) may still have some effect, but there must be more going on.

So what else drove opinion change? What we should really like to show is that it was at least partly driven by learning, in the broadest of senses – by the assimilation of new perspectives and interpretations, as well as the acquisition of more or less factual information. Although we have no means of measuring the former, we can easily measure the latter. To do so, we combine seven of our eight legal and political information items, omitting the electoral register item for reasons cited above, into a single index, consisting of the proportion answered correctly. Let the before and after values of this index be  $I_1$  and  $I_2$ .<sup>48</sup>

The most obvious measure of information gain is of course observed information gain,  $I_2 - I_1$ . A combination of ceiling effects, *item sampling biases* (towards easy items), and well-established propositions about learning, however, imply that  $I_2 - I_1$  must contain a very large dose of perverse, non-random error. Consider: what proportion of all possible information items (including those answerable only by relevant specialists) would those getting all seven information items right actually know? Ten per cent? One per cent? Yet, by our measure, they are at 100 per cent. What sort of gain can they, or for that matter those getting five or six of our seven items right at Time 1, exhibit? Slim to none. Yet these are the participants who should be learning most, not least.<sup>49</sup> As much literature in psychology and communications research suggests, the information-rich get information-richer.<sup>50</sup> Altogether, this implies that those actually gaining the most information appear to be among those gaining the least. Indeed, we shall do better by simply using the Time 2 index,  $I_2$ . The participants with high information scores at Time 2 have all presumably gained a lot of information, either observably, because they started much lower, or, unobservably, because they started high. This intuition can be formalized in

<sup>48</sup> We use political as well as legal information items because the measure works better with the former included. It makes no appreciable difference to include the electoral register item, either as originally scored or reversed. We also experimented with 'corrections for guessing', under which respondents are effectively scored 1 for getting an item right, 0 for not answering, and  $-1$  for getting it wrong. This made only very small differences, in no consistent direction.

<sup>49</sup> In absolute as distinct from proportional terms.

<sup>50</sup> See James D. Dooling and Roy Lachman, 'Effects of Comprehension on Retention of Prose', *Journal of Experimental Psychology*, 88 (1971), 216–22; John D. Bransford and Marcia K. Johnson, 'Contextual Prerequisites for Understanding Some Investigations of Comprehension and Recall', *Journal of Verbal Learning and Verbal Behavior*, 11 (1972), 717–21; Harry L. Chiesi, George J. Spilich and James F. Voss, 'Acquisition of Domain-Related Information in Relation to High and Low Domain Knowledge', *Journal of Verbal Learning and Verbal Behavior*, 18 (1979), 257–73; Susan T. Fiske, Richard R. Lau and Richard A. Smith, 'On the Varieties and Utilities of Political Expertise', *Social Cognition*, 8 (1990), 31–48; Russell W. Neuman, 'Patterns of Recall among Television News Viewers', *Public Opinion Quarterly*, 40 (1976), 115–23; Vincent Price and John R. Zaller, 'Who Gets the News: Alternative Measures of News Reception and Their Implications for Research', *Public Opinion Quarterly*, 57 (1993), 133–64; K. Viswanath and John R. Finnegan Jr, 'The Knowledge Gap: Twenty-Five Years Later', in Bryant R. Burlson, ed., *Communication Yearbook 19* (Thousand Oaks, Calif.: Sage Publications, 1996).

algebra showing, under assumptions just traced, that  $I_2 - I_1$  can be negatively correlated with true information gain, and must be less highly correlated with it than is  $I_2$  alone.<sup>51</sup>

Thus our final, more processual models express the changes in the five policy indices, call them  $P_2 - P_1$ , as functions of Time 2 information  $I_2$  and the difference, call it  $P_1 - G_1$ , between the participant's own Time 1 position and the mean Time 1 position of his or her small group. The first regressor represents learning, the second small group mechanisms.<sup>52</sup> Our expectations are that those who learn the most (as proxied by  $I_2$ ) tend to move the most – in the same direction as the sample as a whole, if they are at least partly driving the overall change – and that participants should tend to narrow the Time 1 gap between their own and their small group's position.<sup>53</sup> The coefficient for  $I_2$  should share the same sign as the mean opinion change ( $P_2 - P_1$ ), while the coefficient for  $P_1 - G_1$  should be negative.

The estimates in Table 7A support both hypotheses strongly. Most importantly for our concerns, Time 2 information has the expected effect for all but one of the five policy attitudes. Its estimated coefficient is both substantively and statistically insignificant in the social root causes equation but substantively and statistically significant – and correctly signed – everywhere else. The significant estimates range in magnitude from roughly 0.08 to roughly 0.10, meaning that participants who get all seven knowledge items right at Time 2 change opinion by 8–10 per cent more of the opinion scale than participants who get none of them right.

On the surface, the results for the Time 1 gap between individual and group are still more impressive. The relevant coefficient estimates are all resoundingly significant, all correctly signed. They range from roughly  $-0.22$  to roughly  $-0.53$ . At the average, roughly  $-0.38$ , a participant who was at one extreme of the policy index (0 or 1) while his or her small group was at the other extreme (1 or 0) would change opinion by 38 per cent more of the opinion scale than one whose time exactly coincided with his or her small group's.

But this calculation is too generous. The maximum-distance half of this contrast never actually occurs, since few respondents and no groups are at either 0 or 1 on any policy index; the maximum (absolute) distances actually

<sup>51</sup> See Robert C. Luskin, 'True versus Measured Information Gain' (unpublished, Department of Government at the University of Texas at Austin, 2001). A similar argument can be made, indeed more simply, on strictly rich-get-richer grounds, for using  $I_1$  alone, and substituting  $I_1$  for  $I_2$  produces essentially the same results as in Table 7. There are no differences in statistical or substantive significance. We prefer  $I_2$  for theoretical/rhetorical reasons. Both  $I_1$  and  $I_2$  proxy true information gain, but each also proxies true information at a given time, and it is less satisfying to show that those who *start* with more information change more than to show that those who *emerge* with more information do so.

<sup>52</sup> The small group means  $G_1$  exclude the participant in question, thus varying slightly from participant to participant within each group.

<sup>53</sup> More precisely, the greater the (necessarily non-negative)  $I_2$ , the larger the opinion change, and the more positive (negative) the Time 1 difference between the individual and his or her group,  $P_1 - G_1$ , the more negative (positive), the opinion change.

TABLE 7 *Attitude Change as a Function of  $t_2$  Information,  $t_1$  Small Group Mean and  $t_1$  Attitude*

Explanatory variable	Punishment (-)	Social Root Causes (+)	Police (-)	Procedural Rights (-)	Self-Protection (-)
<i>Part A: With the Difference between <math>t_1</math> Attitude and <math>t_1</math> Small Group Mean as One Variable</i>					
Intercept	-0.013 (0.019)	0.052** (0.016)	-0.001 (0.024)	-0.057** (0.022)	0.033 (0.024)
$t_2$ Information†	-0.075** (0.029)	-0.007 (0.025)	-0.100** (0.036)	-0.077** (0.032)	-0.066* (0.036)
Distance from $t_1$ group mean‡	-0.214** (0.045)	-0.485** (0.034)	-0.318** (0.044)	-0.322** (0.046)	-0.531** (0.055)
$R^2$	0.076	0.413	0.155	0.150	0.241
adj. $R^2$	0.070	0.409	0.149	0.144	0.236
$F$	12.189	103.933	27.139	25.961	47.102
Probability	0.000	0.000	0.000	0.000	0.000
$n$	299	298	299	298	299
<i>Part B: With <math>t_1</math> Attitude and <math>t_1</math> Small Group Mean as Separate Variables</i>					
Intercept	-0.051 (0.093)	0.325* (0.137)	0.268* (0.136)	0.005 (0.107)	0.229 (0.139)
$t_2$ Information†	-0.075** (0.029)	-0.009 (0.024)	-0.109** (0.036)	-0.088** (0.032)	-0.072** (0.036)
$t_1$ Attitude	-0.209** (0.047)	-0.508** (0.035)	-0.347** (0.046)	-0.387** (0.048)	-0.562** (0.056)
$t_1$ group mean‡	0.273* (0.148)	0.162 (0.164)	-0.060 (0.192)	0.409* (0.221)	0.249 (0.187)
$R^2$	0.077	0.421	0.166	0.207	0.258
adj. $R^2$	0.067	0.415	0.158	0.199	0.250
$F$	8.162	71.349	19.630	25.527	34.248
Probability	0.000	0.000	0.000	0.000	0.000
$n$	299	298	299	298	298

*Note:* Cell entries are coefficient estimates with estimated standard errors in parentheses. The parenthetical signs in the column headings indicate the direction of net change for the sample as a whole and thus the expected sign of the information coefficient.

†  $t_2$  information is the mean of the seven information items, excluding the electoral register question.

‡ The group mean variables are calculated on the *other* group members, excluding the respondent.

\* Significant at the 0.05 level (one-tailed for everything but the estimated intercept, two-tailed for the latter).

\*\* Significant at the 0.01 level (one-tailed for everything but the estimated intercept, two-tailed for the latter).

observed are 0.54 for punishment, 0.77 for social root causes, 0.70 for procedural protections, 0.63 for police and 0.42 for self-protection. The greatest plausible impacts of  $P_1 - G_1$  (these preceding numbers times the corresponding coefficient estimates from Table 7A) are therefore -0.12, -0.37, -0.23, -0.20 and -0.21, respectively. Averaging across the five policy indices, a participant who was at the maximum observed distance from the rest of his or her small group would change opinion by only 23 per cent more of the opinion scale than one whose position exactly coincided with his or her small group's.

A more important qualification lies in the built-in negative correlation between this explanatory variable,  $P_1 - G_1$ , and the dependent variable,  $P_2 - P_1$ . How much are the appropriately negative coefficient estimates due to the presence of  $P_1$  in the dependent variable and  $P_1$  in this explanatory variable? Table 7B addresses this question by breaking  $P_1 - G_1$  into  $P_1$  and  $G_1$  and entering them as separate regressors. The expectation now is that the  $P_1$  coefficient will be negative, for the reason just sketched, while the  $G_1$  coefficient will be positive, with participants tending to increase support for given policies as a function of the extent to which those policies are favoured by their small group.<sup>54</sup>

These results show that the lion's share of the original small group variable's effect stems from its individual rather than its group component. The estimated coefficients for  $P_1$  remain large, negative and overwhelmingly significant. Those for  $G_1$ , however, are greatly eroded. Four of the five carry the appropriate positive sign, but only two (in the punishment and procedural rights equations) achieve statistical significance at the conventional 0.05 level, even by the appropriate one-tailed test. The numbers now range from roughly  $-0.06$  to roughly  $0.41$ , averaging roughly  $0.21$ .

Again, moreover, these numbers are too kind. The range of  $G_1$  is not 0 to 1, but something far narrower: only  $0.46$  to  $0.70$  for punishment,  $0.67$  to  $0.87$  for social root causes,  $0.35$  to  $0.52$  for procedural rights,  $0.52$  to  $0.74$  for police, and  $0.59$  to  $0.79$  for self-protection. The greatest differences  $G_1$  can make to opinion change are therefore only  $0.07$ ,  $0.03$ ,  $0.07$ ,  $-0.01$  and  $0.05$ , respectively. Averaging across the five policy indices, a participant whose small group moved from the lowest to the highest Time 1 group mean would be tugged 4 per cent of the scale in the same direction as a result.

Where, then, do these last analyses leave us? Contrary to our initial bivariate analyses, opinion change does seem to depend, very modestly, on small group mechanisms. More importantly, however, it also depends, rather more, on learning.

## DISCUSSION

Around the world, the combination of television and opinion polling has introduced a plebescitary element into democratic practice, turning up the volume and arguably increasing the distortion in V. O. Key's famous 'echo chamber'. Impressions of sound bites are reflected back by an almost daily stream of opinion polls whose results are reported, in turn, in the mass media, to be reflected back once again. Given a relatively inattentive, relatively ignorant public, the resulting echoes fall well short of democratic aspirations.

The Deliberative Poll harnesses the power of these same technologies to a new and constructive purpose. Television helps draw the participation of both

<sup>54</sup> The model estimated in the table's upper tier can be viewed as a special case of this one, with the coefficient for  $G_1$  constrained to be the negative of that for  $P_1$ .

politicians and citizens, while random sampling ensures that the participants represent the country or region in all its diversity. The deliberations and other aspects of the event provide the opportunity and some incentive for citizens to spend time working towards their own, more considered opinions. The result, in the aggregate, is a picture of a better informed and more thoughtful public opinion.

We say ‘better’ advisedly, claiming only for our results a window on the comparative, not the superlative. The post-event distribution of opinion is almost certainly not what might be expected of an *ideally* informed and thoughtful citizenry (any biases, confounds or other threats to internal validity quite aside). On the grand scale of things, after all, this is a relatively mild intervention. A few weeks’ elevated reading, chatting and thinking, intensified by a couple of days’ focused discussion, can make some difference but cannot fully remedy a lifetime’s inattention. Our only claim is that, on average, our participants emerge looking *more* like ideal citizens than they did beforehand. From this perspective, the changes in mean attitudes may be more important than the means themselves, indicating the directions in which the ideal would diverge from the actual distribution of opinion (although of course we cannot exclude the possibility of non-monotonicities at still higher levels of information and thought).<sup>55</sup>

This account presents much and omits much. Some of the omissions are strictly for want of time and space, but most are of matters that cannot currently be addressed. One theoretically important question, for example, is how much of the event’s effects stem from information gains versus increased thought and thus new or altered perspectives, or simply greater attention. Even without acquiring appreciably more information about crime policy, some participants may come to see the alternatives through different eyes – to focus on different consequences, for different values – or simply to think harder about them and thus to see more, and more clearly, through the same eyes. Another question is how much of the information gains and changes in policy preferences came from the briefing materials, versus talking, reading and thinking about the issues in the gestation period between recruitment and deliberation, versus the small group discussions, versus the large group sessions with policy experts, versus the large group sessions with politicians, etc. A third question concerns effects on other civic attitudes. A more ideal political life may well involve greater tolerance, civility and mutual respect, and participants may well correspondingly emerge from the Deliberative Poll with an enhanced appreciation of opposing arguments and greater willingness to attribute principled if misguided motives to those of unlike mind.

The first British Deliberative Poll leaves these questions hanging, but they are

<sup>55</sup> See Robert C. Luskin, ‘The Heavenly Public: What Would the Ideal Democratic Citizenry Be Like?’ in George Rabinowitz and Michael B. MacKuen, eds, *Electoral Democracy* (Ann Arbor: University of Michigan Press, forthcoming) and Luskin, ‘From Denial to Extenuation (and Finally Beyond)’ and ‘Political Psychology, Political Behavior, and Politics’.

all addressable in principle. Already some more recent Deliberative Polls have addressed perceptions of policy opponents and their motivations. We also hope to devise and include measures of thought and perspective and to elaborate the design in ways that permit the segregation of what we must for now regard as one grand, undifferentiated experimental treatment. It is important to point out, however, that none of this, including the parcelling-out of effects among the elements of the experimental treatment, is essential to establishing some level of success. For that we need only show changes in attitudes associated with increases in knowledge and thought, however produced.<sup>56</sup>

This is indeed part of the pattern sketched here. Specifically, we have shown: first, that, despite the additional layer of self-selection between initial interview and deliberation, our participant sample, *ab ovo*, proved highly representative. On only fourteen of 103 sociodemographic, autobiographic and attitudinal variables were there statistically significant differences between participants and other interviewees, and even these few differences were generally small. Secondly, there was substantial attitude change. On the typical policy item, consisting of a five-point scale, something over half our participants changed position, slightly over 40 per cent changed side (including movement to and from the middle category), and some 10 to 15 per cent changed side completely (excluding movement to and from the middle category). And while many of these changes naturally cancelled out, almost two-thirds of our policy items showed statistically significant net change. Many of these changes, moreover, were substantively impressive. Thirdly, our participants became substantially more knowledgeable about the issues under discussion. Fourthly, the changes were almost completely unrelated to social location. It was not only or disproportionately the well or poorly educated, or people with any particular set of social characteristics, who changed. Fifthly, there was some modest tendency for people to change in the direction of their small groups (perhaps as a function of simple conformism, but perhaps also as a function of legitimate persuasion). And, sixthly, it *was* those who wound up knowing most, and presumably had learned most, who changed most.

A lucky shot? Almost certainly not. While confining this report to the first Deliberative Poll, we should add that the same broad pattern has held in every Deliberative Poll so far. The 'participant sample' is always remarkably similar to the remainder of the original interview sample, both sociodemographically and attitudinally. There is always considerable gross change on almost all the attitude items and considerable net change on most. And the participants always acquire significantly more information.<sup>57</sup>

<sup>56</sup> See Vincent Price and Peter Neijens, 'Deliberative Polls: Improved Measure of "Informed" Public Opinion?' *International Journal of Public Opinion Research*, 10 (1998), 145–76.

<sup>57</sup> See Robert C. Luskin and James S. Fishkin, 'Deliberative Polling, Public Opinion, and Democracy: The Case of the National Issues Convention' (paper presented at the annual meeting of the American Association of Public Opinion Research, Saint Louis, 1998); James S. Fishkin and Robert C. Luskin, 'Bringing Deliberation to the Democratic Dialogue: The NIC and Beyond', in Maxwell McCombs, ed., *A Poll with a Human Face: The National Issues Convention Experiment*

But could the attitude changes we see be due to something outside the deliberative experience, in the nature of what Campbell and Stanley call *history* – influences that affect everyone over the interval of the experiment, participant or no?<sup>58</sup> Perhaps our participants were simply responding to the same influences and registering the same changes as everyone else. This is extremely implausible on the face of it, given the absence of anything obvious in the media. It is rendered still less plausible by the learning coefficients in Table 7, which could be explained by history only if those who learned most would also have changed most anyway, an unlikely scenario for contested issues like these.<sup>59</sup>

There is also relevant side evidence from other Deliberative Polls. While this first Deliberative Poll lacked anything in the nature of a control group, the National Issues Convention in the United States had two: an independent RDD sample interviewed more or less concurrently with the end of the Deliberative Poll and a random half of the non-participant interviewees, reinterviewed by telephone alongside the RDD sample.<sup>60</sup> To budget for history, we might consider not just the  $t_2 - t_1$  differences among our participants but their differences in turn from the  $t_2 - t_1$  differences among a control sample. But whichever control sample we use, the message from the American case is reassuring: the significant  $t_2 - t_1$  differences among the participants mostly survive this comparison.<sup>61</sup>

A considerably more plausible worry is that it was some adventitious and non-replicable aspect of the deliberative experience – a handful of particularly persuasive or off-putting speakers or fellow participants, for example – rather than increased information or reflection that produced the observed changes. Again, however, the learning coefficients in Table 7 suggest otherwise. And, again, there is relevant side evidence from other Deliberative Polls. Since 1996, there have been eight regional Deliberative Polls on electric utility issues in the United States, all addressing the same issues, using almost identical questionnaires, but differing in countless minor and some not-so-minor details. The

(*F*note continued)

in *Political Communication* (Mahwah, NJ: Lawrence Erlbaum, 1999); Robert C. Luskin, James S. Fishkin, Roger Jowell and Alison Park, 'Learning and Voting in Britain: Insights from the Deliberative Poll' (paper presented at the annual meeting of the American Political Science Association, Atlanta, 1999); Robert C. Luskin, James S. Fishkin and Dennis L. Plane, 'Deliberative Polling and Policy Outcomes: Electric Utility Issues in Texas' (paper presented at the annual meeting of the Association for Public Policy Analysis and Management, Washington DC, 1999); Robert C. Luskin, James S. Fishkin, Ian McAllister, John Higley and Pamela Ryan, 'Information Effects in Referendum Voting: Evidence from the Australian Deliberative Poll' (paper presented at the annual meeting of the American Political Science Association, Washington, DC, 2000).

<sup>58</sup> See Campbell and Stanley, *Experimental and Quasi-Experimental Designs for Research*.

<sup>59</sup> See Zaller, *The Nature and Origins of Mass Opinion*.

<sup>60</sup> The British national Deliberative Poll about the general election of 1997 and the Australian national Deliberative Poll about the 1999 referendum on Australia's becoming a republic have also had quasi-control groups.

<sup>61</sup> Luskin and Fishkin, 'Deliberative Polling, Public Opinion, and Democracy: The Case of the National Issues Convention'; Fishkin and Luskin, 'Bringing Deliberation to the Democratic Dialogue: The NIC and Beyond'.



samples were different, as were the small groups and their dynamics. The panels of policy experts and policy makers who fielded the participants' questions overlapped only moderately, and both the participants' questions and the style, organization and precise content of the panelists' answers varied considerably. Yet the results, in broad outline, were everywhere the same.<sup>62</sup>

In all, as we believe we have shown, Deliberative Polling can help illuminate key questions of both mass politics and democratic theory. How 'authentic' are the policy and candidate preferences in ordinary polls and election returns – how much in keeping with the respondents' own fundamental values and interests and accordingly how close to their own 'full-information' preferences? What differences might fuller information make to the distributions of opinions or votes? These are questions Deliberative Polling data can help address, and we hope that the results in these pages have advanced us some small step in that direction.

<sup>62</sup> See Luskin, Fishkin and Plane, 'Deliberative Polling and Policy Outcomes'.

APPENDIX A:1 DIFFERENCES BETWEEN PARTICIPANTS AND  
NON-PARTICIPANTS

Variable	Min	Max	Mean $P_1^\dagger$	Mean $N^\ddagger$
Sex	0	1	0.50	0.48
Age	18	87	45.2	47.5
Number of children	0	5	0.73	0.65
Number adults 18 +	1	2	1.84	1.86
Education, 7 categories	0	6	2.58	2.32
Middle-class or working-class	0	1	0.53	0.48
Socio-economic group	0	6	4.51	4.48
Chief income earner	0	1	0.51	0.52
Strength of party affiliation	1	3	1.65	1.67
Normally read daily AM?	0	1	0.62	0.65
Worry being crime victim?	1	4	2.49	2.50
How safe walking alone after dark?	1	4	2.66	2.60
How much burglary in this area?	1	4	2.55	2.71*
How much street attacks & thefts?	1	4	1.75	1.76
How worried being attacked?	1	4	2.11	2.09
Experience of crime, past 2 yrs.	0	1	0.40	0.38
Car, van, motor stolen?	0	1	0.13	0.09
Car, van, motor broken into?	0	1	0.26	0.23
House broken into, things stolen?	0	1	0.09	0.11
Hit or violently attacked?	0	1	0.06	0.03*
Reported crime/accident?	0	1	0.47	0.47
How confident police would catch burglar?	1	4	1.84	1.84
How confident stolen goods would be returned?	1	4	1.49	1.49
How confident in gov't crime fighting policies?	1	4	1.63	1.77*
Smaller income gap	1	5	3.89	3.73*
Life never to be made fair	1	5	2.85	3.05*
Always obey law	1	5	4.03	4.10
Make more of life	1	5	4.28	4.21
Right to violence	1	5	2.22	2.30
Less sympathy for wrongdoers	1	5	3.63	3.65
Fend for oneself	1	5	3.61	3.68
<i>Root Causes (Economic)</i>				
Reduce unemployment	1	5	1.83	1.93
Reduce unemployment and poverty‡	1	4	3.10	2.91*
Causes vs. apprehension	1	5	1.97	2.10
<i>Root Causes (Social)</i>				
Spend more time with children	1	5	4.37	4.40
Parents more time with children‡	1	4	2.36	2.35
Teach children right from wrong	1	5	4.57	4.58
Reduce TV violence and crime	1	5	3.79	3.72
Firmer school discipline	1	5	4.28	4.40
Schools responsible for moral guidance‡	1	4	1.74	1.74
<i>Police</i>				
More police on the beat	1	5	4.50	4.57
Off-duty police carry guns	1	5	2.67	2.61
Police more powers to catch criminals‡	1	4	2.28	2.45
<i>Punishment</i>				
Punishment vs. reform	1	5	3.29	3.32
Prisons reform, not just punish	1	5	3.99	3.90
Tougher sentences	1	5	4.15	4.20

APPENDIX A:1 *Continued*

Variable	Min	Max	Mean $P_1^\dagger$	Mean $N^\dagger$
Stiffer sentences generally	1	5	4.18	4.18
Stiffer sentence and more prisons $^\ddagger$	1	4	2.17	2.20
More offenders to prison	1	5	3.68	3.76
Fewer people to prison	1	5	2.79	2.72
Only dangerous criminals to prison	1	5	3.07	3.10
More offenders out of prison	1	5	3.41	3.35
More offenders: community service	1	5	3.76	3.71
More offenders: military service	1	5	3.64	3.72
More offenders get training	1	5	3.64	3.58
Prison life tougher	1	5	3.87	4.00
All murderers life sentence	1	5	4.33	4.42
Life sentence means life	1	5	4.48	4.51
Death most appropriate for some crimes	1	5	3.80	3.75
<i>Procedural Rights</i>				
Convict innocent vs. fail to punish guilty	1	5	2.54	2.58
Police allowed to bend rules	1	5	1.93	1.91
Fewer jury trials	1	5	3.04	3.11
Court rules less on side accused	1	5	3.25	3.39
End presumption of innocence	1	5	2.66	2.74
Silence mentionable in court	1	5	3.39	3.38
Right to silence in police questioning	1	5	2.81	2.83
Confessions alone not convict	1	5	3.67	3.58
Ind. investigations of complaints against police	1	5	4.13	4.13
<i>Juvenile Offenders</i>				
1st-x child burglar: just warn	1	5	3.24	3.25
1st-x child burglar: community service	1	5	3.99	3.92
1st-x child burglar: juvenile institution	1	5	2.80	2.80
1st-x child burglar: ordinary prison	1	5	2.04	2.04
Tough institution for repeat juvenile offenders	1	5	3.76	3.79
What age treat as adults	0	5	3.43	3.42
<i>Self-Protection</i>				
People vs. police to protect	1	5	3.67	3.47
People take responsibility for prop. $^\ddagger$	1	4	1.64	1.71
People make property more secure	1	5	4.26	4.15
More Neighborhood Watch schemes	1	5	4.04	3.99
Street patrols vs. police	1	5	3.06	3.06
<i>Miscellaneous</i>				
Legalize soft drugs	1	5	2.71	2.45**
Illegal drugs major crime cause	1	5	4.00	4.20*
Prison life too soft/tough	1	5	2.05	2.01
Prison hardens criminals	1	5	3.35	3.28
<i>Legal</i>				
Death sentence in Britain?	0	1	0.85	0.81
Britain largest prison population	0	1	0.50	0.43
Britain more life sentences	0	1	0.20	0.13**
Jury trial in local magistrate's court?	0	1	0.58	0.47**

APPENDIX A:1 *Continued*

Variable	Min	Max	Mean $P_1$ †	Mean $N$ ‡
<i>Political</i>				
Number of MPs about 100?	0	1	0.69	0.62*
Max 4 yrs. Between general elections?	0	1	0.27	0.27
Britain has proport. representation?	0	1	0.55	0.47*
Can't be on elect. register in 2 places	0	1	0.18	0.19

† Mean  $P_1$  and Mean  $N$  denote sample mean values for participants at  $t_1$  and non-participants. All variables are scored so that higher numbers indicate greater agreement or quantity.

‡ By construction, these six items must average 2.0.

\* Difference ( $P_1 - N$ ) significant at the 0.05 level (two-tailed).

\*\* Difference ( $P_1 - N$ ) significant at the 0.01 level (two-tailed).

APPENDIX A:2 DIFFERENCE BETWEEN PARTICIPANTS AND  
NON-PARTICIPANTS

Nominal variables	Number of categories	Chi-square statistic
Type of area (rural, urban, city)	3	3.36
Region (London, North, Midlands, etc.)	5	13.79**
Age of youngest child (in intervals)	5	2.11
Marital status (married, divorced, never married)	3	6.90*
Daily paper (no paper, Tory tabloids, other)	3	0.55
Party affiliation	4	2.71
Employment status (working, unemployed, etc.)	5	6.53
Employment status of CIE§	3	0.71
Occupation	10	15.4
Education (no quals, some quals, GCE and up)	3	2.60
Race/ethnicity	8	7.27

† Mean  $P_1$  and Mean  $N$  denote sample mean values for participants at  $t_1$  and nonparticipants. All variables are scored so that higher numbers indicate greater agreement or quantity.

‡ By construction, these six items must average 2.0.

§ CIE denotes 'chief income earner'.

\* Difference ( $P_1 - N$ ) significant at the 0.05 level (two-tailed).

\*\* Difference ( $P_1 - N$ ) significant at the 0.01 level (two-tailed).

APPENDIX B: VARIABLES INCLUDED  
IN POLICY INDICES

*Social Root Causes*

Spend more time with children  
Reduce TV violence and crime  
Firmer school discipline

*Police*

More police on the beat  
Off-duty police carry guns

APPENDIX B: *Continued*

---

*Punishment*

Punishment vs. reform  
Prisons reform, not just punish  
Tougher sentences  
Stiffer sentences generally  
More offenders to prison  
Fewer people to prison  
Only dangerous criminals to prison  
More offenders out of prison  
More offenders: community serv.  
More offenders: military serv.  
More offenders get training  
Prison life tougher  
All murderers life sentence  
Life sentence means life  
Death most approp. for some crimes

*Procedural Rights*

Convict innoc. vs. not punish guilty  
Police allowed to bend rules  
Fewer jury trials  
Court rules less on side accused  
End presumption of innocence  
Silence mentionable in court  
Right to silence in police questioning  
Confessions alone not convict

*Self-Protection*

People make property more secure  
More Neighborhood Watch schemes  
Street patrols vs. police

---