

## Consulting the People – Thoughtfully

James Fishkin

There are many ways to consult the public. In this overview, I will consider a simple classification of the most common approaches and argue that certain versions best fulfill some key democratic values.

The simple classification below focuses on two issues: what and who? The first has to do with what form of public opinion is being assessed, the second has to do with whose opinion it is that is being assessed. When we ask about forms of public opinion, we are asking about the thoughtfulness of public input. When we ask who is included, we are asking about how the aspiration for inclusion is implemented.

When considering forms of public opinion, let us say that opinion is “refined” if it is the product of deliberation exposing it to a wide range of alternative views supported by sincere arguments and reasonably accurate information. Refined opinion is informed – informed about competing views and facts sincerely viewed as relevant by proponents of different positions. People are aware of the arguments and have reflected on them or thought about them. By contrast, we will say that opinion is “raw” if it is not the product of such deliberation.

The other distinction concerns *whose* opinion is being consulted. While the classifications I will focus on do not exhaust all the possibilities, they cover the principal practical alternatives. The people consulted can be self-selected; they can be selected by some method of sampling that attempts to be representative without probability sampling; they can be chosen by random sampling; or they can constitute virtually all voters (or members of the group being consulted). When these two dimensions are combined, then the eight possibilities in the chart emerge.

### I. Eight Forms of Public Consultation

Method of Selection:			
1. Self-selection	2. Non-random	3. Random Sample	4. "Everyone"

#### *Sample: Public Opinion*

<b>A. Raw</b>	1A SLOPS	2A Some Polls	3A Most Polls	4A Referendum Democracy
<b>B. Refined</b>	1B Discussion Groups	2B Citizens Juries, etc	3B Deliberative Polls	4B "Deliberation Day"

The first category, 1A is already being implemented, especially on the internet. Norman Bradburn of the University of Chicago has coined an acronym SLOP for "self-selected listener opinion poll". Before the internet, radio call-in shows would commonly ask for responses by telephone to some topic. The respondents to SLOPS are not selected by scientific random sampling as in public opinion polls. The respondents instead, simply select themselves. They are predominantly those who feel more intensely or feel especially motivated. Sometimes, they are organized. The SLOP, it is thought, gets "grass roots" opinion. However, in the parlance of American lobbyists, sometimes the response is something more organized and synthetic – the impression of grass roots that is really "astroturf".

A good example of the dangers of SLOPS came with the world consultation that Time magazine organized about the "person of the century". Time asked for votes in several categories, including greatest thinker, greatest statesman, greatest entertainer, greatest captain of industry. Strangely, one person got by far the most votes in every category, and it turned out to be the same person. Who was this person who towered above all rivals in every category? Ataturk. The people of Turkey organized to vote, by post card, on the internet, by fax and produced millions more votes, as a matter of national pride than the rest of the world could muster for any candidate, just through individual, unorganized voting.<sup>1</sup>

Media organizations routinely conduct SLOPS on the internet on a wide range of political or social matters. A SLOP involves visitors to

1 "Is this the Man of the Century?" Guardian, October 30, 1997, page 1.

a web site, gives people a sense of empowerment (they are registering their opinions) but it produces data that is misleading, that offers only a distorted picture of public opinion. To take just one example, SLOPS, at the time of impeachment in the US routinely showed large majorities in favor, while scientific polls showed a completely different picture. Those feeling most intensely bothered to register their views, sometimes more than once.

The difficulty with category 1A is that it offers a picture of public opinion that is neither representative nor deliberative. It offers a picture of uninformed opinion that is also distorted and partial in whom it includes. If it is a mirror of public opinion, it is more like a carnival fun house mirror than one that reproduces what it reflects.

An alternative to the SLOPS of category 1A is the possibility of serious deliberation, refined public opinion, produced among self-selected groups. Discussion groups fill out Category 1B. If the discussion groups offer the opportunity to weigh the main alternative arguments that fellow citizens would want raised on an issue, then they can achieve a measure of deliberation on an issue even if the participants are not a good mirror of the entire population. The Kettering Foundation supports a large network of “National Issues Forums” (NIF) in the US and in several other countries, in which thousands of self-selected participants deliberate conscientiously and sincerely with briefing materials that offer a balanced and accurate basis for discussion.<sup>2</sup> These participants meet in churches, schools, neighborhood venues and spend hours in serious consideration of the alternatives. However, their conclusions, while filtered or deliberative are not representative of the views of the entire public.

Indeed, it is arguable that self selected samples face a more difficult task in deliberating. To the extent that they are similar, demographically and attitudinally, contrary points of view and contrary interests may not get voiced in the discussions. Random sampling has an advantage over self selection, not just in representativeness, but also, potentially, in facilitating deliberation by providing access to a diversity of points of view among people who might never normally have a serious discussion. For example, in the Deliberative Poll on crime in Britain, a random sample deliberating about criminal justice policy included criminals in the sample. They enriched the discussion by being able to talk about their experience. Mid-

---

<sup>2</sup> For a good overview of these activities and the vision behind them, see Mathews 1994.

dle class respondents would not, in the ordinary course of life, talk to car thieves about criminal justice policy. While the sample had access to many experts (in the question and answer process in plenary sessions) its small group discussions were also enriched by the sheer diversity generated by random sampling. Self selection, by contrast, would likely yield only the homogeneity of the like minded and those from similar social locations.

Category 2A combines raw public opinion with methods of selection attempting to achieve some degree of representativeness – but that do not employ probability sampling. Some public opinion polls fall into this category. Those employing quota sampling justify their method as an attempt to approximate probability sampling. Some spectacular failures, such as the 1948 Dewey / Truman debacle and the 1992 British General Election have been blamed at least in part on the use of quota sampling (for the latter see Jowell et al. 1995).

Category 2B employs non-random methods of selection with attempts to arrive at more deliberative public opinion. There are a variety of methods of public consultation that fit this category. So-called “citizens juries” use quota samples to select small numbers of participants (typically 12 or 18) to deliberate for several days or even weeks on public issues. Consensus Conferences begin with self-selection (soliciting respondents through newspaper ads) and then use quotas to attempt to approximate representativeness. These methods often suffer from the same problem noted above. They begin with self-selection and then employ such small numbers that any claims to representativeness cannot be credibly established.<sup>3</sup>

Category 3A, combining probability samples with raw opinion is exemplified, of course, by the public opinion poll, in its most developed form. It avoids the distorted representativeness of SLOPS as well as the more modest distortions of non-random sampling. Just as Gallup vanquished the *Literary Digest* by using quota sampling for the effective launch of the public opinion poll in the 1936 US Presidential election, this category, 3A, trumps the SLOPS of 1A as well as the quota sampling of 2A.<sup>4</sup>

---

3 Another problem is that these research designs do not permit evaluation of how those agreeing to participate compare to those who do not.

4 Gallup abandoned quota sampling after the 1948 election. The advantage of probability sampling was demonstrated by the success of the Survey Research Center at Michigan in that election.

Public opinion polling reflecting raw public opinion offers a thin “top of the head” expression of the public voice. On complex policy or political questions, the views represented by polls are crippled by what Anthony Downs called “rational ignorance” (Downs 1957). If I have only one vote in millions, why should I spend a lot of time and effort becoming informed (as we would like ideal citizens to do) when my individual vote or opinion will not make any appreciable difference? In addition, the views reported by polls on complex political or policy matters are often crippled by a second factor – the tendency to report opinions that are not only based on little thought or reflection, but that may not exist at all. Phantom opinions or “non-attitudes” are reported by polls because respondents almost never wish to admit that they do not know, even when offered elaborate opportunities for saying so. Building on the classic work of Phil Converse of the University of Michigan, George Bishop and his colleagues at the University of Cincinnati dramatized this issue with their study of attitudes towards the so-called “Public Affairs Act of 1975”. Large percentages of the public offered an opinion even though the act was fictional. The Washington Post more recently celebrated the twentieth anniversary of the non-existent “Public Affairs Act of 1975” by asking respondents about its “repeal”. The sample was split, with half being told that President Clinton wanted to repeal the act and half being told that the “Republican Congress” wanted its repeal. While such responses were based on a minimal amount of information (or misinformation provided to the participants, since the act did not exist in the first place) the information base was really just a response to a cue about who was for the proposal and who was against it.<sup>5</sup>

Deliberative Polling<sup>6</sup>, which fits in our category 3B, was developed explicitly to combine random sampling with deliberation. Deliberative Polls are meant to include everyone under conditions where the public can think. Deliberative Polling attempts to employ social science to uncover what deliberative public opinion would be on an issue by conducting a quasi experiment, and then it inserts those deliberative conclusions into the actual public dialogue, or, in some cases, the actual policy process.

---

5 For a good overview of this work by George Bishop and the replication by the Washington Post under the direction of Richard Morin, see Bishop 1995.

6 Deliberative Polling® is a trade mark of James S. Fishkin. Any fees from the trade mark are used to support research at the Center for Deliberative Democracy at Stanford. See <http://cdd.stanford.edu>.

Deliberative Polls have been conducted around the world. Nationally televised projects include the US, Britain, Australia, Denmark, Bulgaria, Hungary and local projects include China, Canada and various parts of the US. For details see <http://cdd.stanford.edu>.

Deliberative Polling begins with a concern about the defects likely to be found in ordinary public opinion – the incentives for rational ignorance applying to the mass public and the tendency for sample surveys to turn up non-attitudes or phantom opinions (as well as very much “top of the head” opinions that approach being non-attitudes) on many public questions. At best, ordinary polls offer a snapshot of public opinion as it is, even when the public has little information, attention or interest in the issue. Deliberative Polling, by contrast, is meant to offer a representation of what the public would think about an issue under good conditions. Every aspect of the process is designed to facilitate informed and balanced discussion. After taking an initial survey, participants are invited for a weekend of face to face deliberation; they are given carefully balanced and vetted briefing materials to provide an initial basis for dialogue. They are randomly assigned to small groups for discussions with trained moderators, and encouraged to ask questions arising from the small group discussions to competing experts and politicians in larger plenary sessions. The moderators attempt to establish an atmosphere where participants listen to each other and no one is permitted to dominate the discussion. At the end of the weekend, participants take the same confidential questionnaire as on first contact and the resulting judgments in the final questionnaire are usually broadcast along with edited proceedings of the discussions throughout the weekend.<sup>7</sup> The weekend microcosm tends to be highly representative, both attitudinally and demographically, as compared to the entire baseline survey and to census data about the population. In every case thus far, there have also been a number of large and statistically significant changes of opinion over the weekend. Considered judgments are often different from the top of the head attitudes solicited by conventional polls. Looking at the full panoply of Deliberative Polls we believe that about two thirds of the opinion items change significantly following deliberation.

But what do the results represent? Our respondents are able to overcome the incentives for rational ignorance normally applying to the

---

<sup>7</sup> For a good see Fishkin 1997. For more detailed analysis see Lushkin, Fishkin and Jowell 2002.

mass public. Instead of one vote in millions, they have, in effect, one vote in a few hundred in the weekend sample, and one voice in fifteen or so in the small group discussions. The weekend is organized so as to make credible the claim that their voice matters. They overcome apathy, disconnection, inattention and initial lack of information. Participants from all social locations change in the deliberation. From knowing that someone is educated or not, economically advantaged or not, one cannot predict change in the deliberations. We do know, however, from knowledge items, that becoming informed on the issues predicts change on the policy attitudes. In that sense, deliberative public opinion is both informed and representative. As a result, it is also, almost inevitably, counter-factual. The public will rarely, if ever, be motivated to become as informed and engaged as our weekend microcosms.

The idea is that if a counterfactual situation is morally relevant, why not do a serious social science experiment – rather than merely engage in informal inference or arm chair empiricism – to determine what the appropriate counter-factual might actually look like? And if that counterfactual situation is both discoverable and normatively relevant, why not then let the rest of the world know about it? Just as Rawls's original position can be thought of as having a kind of recommending force, the counterfactual representation of public opinion identified by the Deliberative Poll also recommends to the rest of the population some conclusions that they ought to take seriously. They ought to take the conclusions seriously because the process represents everyone under conditions where they could think.

The idea may seem unusual in that it melds normative theory with an empirical agenda – to use social science to create quasi experiments that will uncover deliberative public opinion. But most social science experiments are aimed at creating a counterfactual – the effect of the treatment condition. In this effort to fuse normative and empirical research agendas, the trick is to identify a treatment condition that embodies the appropriate normative relevance.

Two general questions can be raised about all research designs – questions of internal and external validity (see Campbell and Stanley 1963). Sample surveys are relatively high on external validity: we can be fairly confident about generalizing the results to larger populations. By contrast, most social science experiments done in laboratory settings are high in internal validity: we can be fairly confident that the apparent effects are, indeed, the result of the experimental treatments. However, experiments done with college students, for example, lack a basis for

external validity if the aim is to find out something about the general population.

If a social science experiment were to have relatively high internal validity, where we could be confident that the effects resulted from the normatively desirable treatment, and if it were also to have relatively high external validity where we could be confident about its generalizability to the entire citizen population, then the combination of those two properties would permit us to generalize the consequences of the normatively desirable property to the entire citizenry. We could be confident in the picture of a counterfactual public reaching its conclusions under normatively desirable conditions. In other words, if an experiment with deliberation were high on internal validity, then we could be confident that the conclusions were the result of deliberation (and related factors such as information). And if such an experiment were high on external validity then we could be confident about generalizing it to the relevant public of, say, all eligible voters. Only with both kinds of validity would the quasi experiment called Deliberative Polling have any claim to represent the considered judgments of the people.

Even in the best case for realizing category 3B there is a limitation to what is accomplished. Deliberative Polling, whether on-line or face to face, involves only a scientific random sample of the population. The thoughtful and informed views created in the experiment are not widely shared because the bulk of the public is still, in all likelihood, disengaged and inattentive because it is subject to the incentives for rational ignorance that routinely apply to citizens in the large scale nation state. Deliberative Polling overcomes those incentives for a microcosm, but leaves the rest of the population largely untouched (we say largely since the rest of the population may well witness the process through the media).

The last two categories, 4A and 4B, parallel the previous ones, except that when ideally realized, they would offer the full embodiment of the kind of result represented by scientific sampling in 3A and 3B. If everyone somehow participated in mass consultations such as voting or referendum democracy, then 4A would represent the same views as those offered by public opinion polls in 3A. Of course, one problem with referendum democracy and other forms of mass consultation that attempt to involve the bulk of the mass public, is that turnout is often so defective that only a portion of the public participates. Sometimes the participation in referendums or national elections is so low, in fact, that the distinction between mass plebescitary democracy and self-selected samples in SLOPS becomes difficult to draw. Of course, there are pos-

sible institutional remedies for low turnout. Australia has a long tradition of compulsory voting, fining non-voters, that has worked quite well to provide one of the highest turnouts in the world in national elections. However, it is well established that compulsory voting has done little or nothing to improve the level of knowledge or engagement among voters, just the level of participation.

The last possibility, 4B, is the most ambitious. Just as conventional polling (3A) models actual top of the head opinion in the mass public, which is represented by plebescitary democracy (4A) in our scheme, in the same way, Deliberative Polling 3B, models mass deliberative public opinion 4B. The latter, however, is usually counterfactual. The mass public, in other words, is usually not deliberating; it usually does not have considered judgments on most policy issues. How could this counterfactual possibility be realized?

Bruce Ackerman and I have a proposal. We call it “Deliberation Day” (see Ackerman and Fishkin 2004). The problem for the Deliberative Poll was to motivate a microcosm of the entire population to overcome the incentives for rational ignorance and to engage in enough substantive face to face discussion to arrive at informed judgments – informed about the issues and the main competing arguments about them that other citizens would offer. But it is one thing to imagine doing this for a microcosm; quite another to imagine doing it for the entire population. Gallup’s vision was that the combination of the media and polling could turn the entire country into “one great room.” The media would send out competing views and the polls would report the public’s judgments and it would be as if the entire country were in one town meeting.<sup>8</sup> This vision foundered, however, on the lack of a social context that would encourage small group deliberation. If everyone is in one great room in the large scale nation state, the room is so big that no one is listening. A different, more decentralized strategy is required.

We propose a national holiday in which all voters would be invited to participate in local, randomly assigned discussion groups as a preparation to the voting process a week later. Candidates for the major parties would make presentations transmitted by national media and local small group discussions would identify key questions that would be directed to local party representatives in relatively small scale town meetings held

---

<sup>8</sup> For a summary of this original vision and an argument that it is better achieved by the Deliberative Poll, see my 1997.

simultaneously all over the country. Incentives would be paid for each citizen to participate. The cost, while massive, would make democracy far more meaningful as it would provide for an input from the public that involved most people and that also lead to a large mass of citizens informed on the issues and the competing arguments. If the incentives for participation in this national holiday activity, "Deliberation Day", worked and people actually became well informed, it would make real the counterfactual deliberative opinion represented by the quasi-experiment of the Deliberative Poll. Candidate behavior and advertising would have to adjust to the fact that voters would have become informed on the issues. The anticipation of such a deliberative public could do a great deal to transform the rest of the public dialogue.

While full scale realization of this idea is only a far off possibility, it is meant to dramatize a different way of thinking about democratic reform. The major cost of the reform is the new holiday. We propose to take an existing holiday, Presidents Day and devote it to picking our next president. We have actually piloted the idea in this 2004 Presidential Election. In 17 cities, locally televised Deliberative Polls were conducted, mostly on the same day, with statistical microcosms that represented what the local publics would think if they were all deliberating. In many cases these local deliberations produced significant knowledge gains and changes of opinion. The local/national project also dramatized the value of putting a human face on opinion change to enrich political communication as well as the prospects for creating civic engagement through discussion in local communities around the country.

There are two categories in our scheme that achieve both values – 3B and 4B--Deliberative Polling and Deliberation Day. Deliberative Polling achieves inclusiveness through a form of political equality – everyone has an equal chance of being selected through random sampling. The latter achieves inclusiveness through mass participation. Ideally, everyone does actually participate. In both cases, an important new increment of thoughtfulness is added by the deliberative process itself – briefing materials, small group discussions, questions and answers from competing experts, opportunities to reflect together on new information and competing arguments in a safe public space. Both strategies – Deliberative Polling and Deliberation Day – combine inclusiveness and greater thoughtfulness. Both are meant to be antidotes to shrinking sound bite democracy and disaffected mass participation. Both are realizations of the same pattern of deliberative practice – small group discussions alternated with plenary sessions with competing experts. The difference is whether this

kind of experience is undertaken by scientific samples or by something approaching the entire mass public. The former achieves inclusiveness via scientific sampling; the latter achieves it via mass participation. The former is practical but only a representation of the deliberative public voice. The latter is a full scale realization but far more expensive. In the absence of the political will to achieve the latter, we hope that practical applications of the former will pilot the expressions of “we the people” – but under conditions where the people can think seriously about the views they are expressing.

### References

- Ackerman B., J. S. Fishkin, (2004). *Deliberation Day*, New Haven / London, Yale University Press
- Bishop, G. (1995). “Leaders, the Public and Democracy”. *Society* 35 (July), No. 5, 2
- Campbell, D., J. Stanley (1963). *Experimental and Quasi-Experimental Designs for Research*, Chicago, Rand-McNally
- Downs, A. (1957). *An Economic Theory of Democracy*, New York, Harper and Row
- Fishkin, J. S. (1997). *The Voice of the People: Public Opinion and Democracy*. New Haven/London, Yale University Press, 76-80, 161-76
- Luskin, R. C., J. Fishkin, and R. Jowell (2002). “Considered Opinions: Deliberative Polling in Britain”. *British Journal of Political Science* 32.
- Mathews, D. (1994). *Politics for People*, Chicago, University of Illinois Press
- Jowell, R., et al., (1993). “The 1992 British General Election: The Failure of the Polls”, *Public Opinion Quarterly* 57, 238-63.

