

Rip, Ike, and a statistician named Twiddle, come to Grandview portraying themselves as insurance agents. Of course, the truth comes out. It came out in the NIC because Jim Fishkin boasted of it; it came out in Magic Town because Rip got heavy with the town's leading journalist, Mary Peterson [played by Wyman]. Through her reporting, Smith is revealed as a polling expert, and Grandview as a miracle town that has been the perfect barometer of national opinion for a half century.

That's heady stuff for the locals. They lose their heads. And, if Grandview is a perfect reflection of what all America at its best thinks, an investigator has only to take up residence and imbibe. A dim-witted but enterprising mayor sets up over 20 polling booths around the city. Anticipating Fishkin, each is equipped with a reference library, because "we want to make sure that folks are thinking right." Of course, Grandview opinion is valuable proprietary data: "Everybody will be instructed not to give opinions to outsiders, only to our own official polltakers."

The town's fall comes quickly. The first Grandview poll (in the self-conscious new order) yields results obviously out of touch with national thinking. Grandview responses, compared to sensible Gallup readings of the day, became the object of national derision. The town's ultimate disgrace is administered by the country's leading radio commentator (Lowell Thomas), who delivers the essential homily: "So this little town that's always been right, turned out to be ridiculously wrong. They were so completely out of tune with the country, people are beginning to wonder where Grandview is. Certainly it can't be in the United States."

That is, once made self-conscious, this "mathematically perfect community" quickly becomes representative of nothing—a doomed social science experiment.



The Emperor Has No Clothes

by Warren J. Mitofsky

Now that the National Issues Convention is history, we should take note of what we have learned—if anything. The key question, James Fishkin told us, was whether the country would come to different conclusions about issues if those issues were seriously deliberated by the American public. Presumably, the country would be better off if its citizens were more thoughtful about important matters of general concern. That is a wonderful hypothesis and I take his point quite seriously.

Fishkin attempted to test his hypothesis by bringing a representative sample of citizens to Austin, Texas, in mid-January to deliberate "the issues" and to kick off the presidential campaign. All the presidential candidates were supposed to assist in those deliberations. Fishkin's point was to create "a more thoughtful and representative way of launching the [presidential] primary season." Only 12% of Democrats and 8% of Republicans participated in primaries in 1992, he told us, which gave a lot of influence to a handful of voters. He led us to believe his televised classroom would produce a better showing in 1996.

Let's look at the record and see how well Fishkin succeeded. What did his civic-minded backers achieve for their \$4,000,000? And what about NORC—how did it do for its \$500,000, (which may be a record for the most expensive public opinion survey per respondent ever conducted)?

Americans were supposed to watch as their televised fellow citizens in the NIC classroom discussed issues among themselves and with the presidential candidates. This new-found thoughtfulness about the issues would take them and us beyond sound-bite journalism, all for the good of the country. To this end, the Public Broadcasting Service presented 5 hours and 30 minutes of air time over three days, an hour less than originally planned due to President Clinton's absence. The leading presidential candidates did not cooperate. In addition to the President, Bob Dole and Pat Buchanan declined to appear. Steve Forbes, Lamar Alexander and Phil Gramm appeared via satellite. Only Richard Lugar and Vice President Gore came in person.

Victory is Declared

In a post-NIC press release, the sponsors used the approach George Aiken (R-Vt.) proposed for ending US involve-

“*Average citizens are not likely to change their opinions very much based upon a discussion with strangers of topics they have not thought much about and really do not care about strongly. Opinion change is an evolutionary process. It takes place over time and is aided by relatives, friends and associates, the kind of people we talk to on a regular basis.*”

Polling America—The NIC Survey

ment in Vietnam—they declared victory and went home. “Deliberative Poll Reveals Significant Change in Delegates’ Views on Key Issues,” they announced. Their press release offered the before and after answers to 31 of the 81 items in their questionnaire. What I found curious was the paucity of change. Remember, the point of all this was “whether the country would come to *different* conclusions about issues” [emphasis added]. Given all the so-called

“*We expect something less than perfection from survey research, but a 36% response rate cannot be thought of as representative of the population from which it was drawn. It borders on self-selection.*”

deliberation and attention by the sponsors, the pollsters and the media that the participants were exposed to I would have expected significant change on many items. Instead, according to the NIC press release, we find that of the 27 items the sponsors claimed “changed,” only 13 differed by more than 10 points from the pre-NIC measurement. The largest change was 19 points and that was on the question: “Public officials care a lot about what people like me think.” The “agree” side went from 41% to 60%. (I am amazed that *any* of the 459 pampered guests who participated in the NIC disagreed with this statement after their weekend in Austin.) The change for the other 14 items which were noted as different was 10 points or less.¹

For some observers, change of opinion was not the issue. As Phillip Meyer said on AAPORNET, opinion “change isn’t the important dependent variable here. If you did a dozen of these deliberative polls simultaneously, you might get [opinions] changing in as many different directions, depending on the mix and the dynamics in each group. What should change uniformly is the quality

of opinions, e.g., as, defined by Yankelovich in... *Coming to Public Judgment*. There should be more internal consistency, less mushiness, more ideological constraint, more ability to connect one issue to another.” The questions in the deliberative poll did not measure mushiness or any other notion of the importance of the topic to the respondent. If the issues were of low intensity for the respondents it is not surprising that there was little net change regardless of any gross switching respondents may have done.

Another thing missing is an analysis of how these issues and the minimal change in opinion noted fit into the 1996 presidential campaign. The showbiz part of the NIC, the part that attracted the air time and the sponsors with big bucks and noble intentions, was to associate issues with the 1996 presidential campaign. If they expect to replace sound bites and all those superficial media polls it will be necessary for Fishkin and his allies to show not only a change in issue positions from the deliberative experience but some influence on the presidential election. To date they haven’t shown anything. Furthermore, it is not likely they will. As for increasing participation in the nominating process, turnout through Super Tuesday had been only 7%, down from the 12% turnout in 1992 that Fishkin abhorred.

The NIC isn’t a “Field Experiment”

Any notion that the NIC delegates represented a cross-section of American voters in 1996 is a mistaken notion at best, and downright misleading at worst. In an article that appeared in *The Public Perspective* prior to the NIC, I said a representative sample was a requirement in order to generalize from these results to all Americans. This requirement is a convention Philip Converse, chair of the technical advisory committee for this project, finds “bemusing.” He said the sample does not have to be representative if it is a field experiment. “Naturalistic observations and mere statistical controls,” are satisfactory for him. Fishkin’s project is a field experiment,

according to Converse. Or is it? A field experiment, as I understand it, has two elements the Fishkin effort does not have: (1) a random selection of participants and (2) a control group.

The Response Rate Approached Self-Selection

If we are going to generalize from a sample about the population from which it was selected statistical theory requires error-free measurement of *all* elements in the sample. It also assumes random selection of the elements. At an absolute minimum, we must rule out self-selection for both surveys and experiments. The imperfect world of the social sciences tolerates something less than the perfection called for by statistical theory. In order to draw conclusions about the Fishkin experiment the least we should expect is a sample comparable in quality to other work done by NORC. When a leading research organization puts its reputation before the public we expect results we can count on. Indeed, we got that for stage one of this project. NORC achieved a response rate of 72% for the initial survey. Not quite the 80% Norman Bradburn said he was aiming for, but just slightly less than the 75% NORC gets with its General Social Survey.

But this project was not a one-stage project. Its object was to measure the *change* in opinion as a result of participating in the deliberations in Texas. Causing change was the main reason for having a deliberative meeting. Determining change in opinion requires a second measurement. That second measurement had a response rate of only 36%, almost 20 points less than the goal NORC had set for itself. We expect something less than perfection from survey research, but a 36% response rate cannot be thought of as representative of the population from which it was drawn. It borders on self-selection. Nor can we justify it by claiming the demographics of those interviewed are similar to the larger sample. Matching demographics is not germane unless the organizers want to assert that there is a high correlation between the demographics and

the *change* in opinion. We need assurance that the change in opinions of the 36% who responded represents a comparable change in the whole sample, and making that assertion is an impossibility.

Effect of Reinterviewing the Participants

A field experiment, just as a laboratory experiment, requires a control group. If we introduce an experimental condition (the whole Austin experience, in this case), we need to know if any change we observe is due to the experimental condition or to some other condition. The most obvious other condition for participants in this experiment is the act of being interviewed twice. Asking people questions is bound to raise their consciousness about the topics covered in the interview. We need to know that any change in a respondent's opinion is not just the result of talking for an hour to a stranger who wrote down all his/her answers about foreign affairs, the economy and family values. The act of being interviewed more than once is what makes panel participants give different answers as a group than one-time participants in a survey. For example, an effect is noticeable for people interviewed during the first month and succeeding months in the Current Population Survey's estimates of the labor force; it also is noticeable in political polls when voters are interviewed both pre- and post-election and the "after" group is compared to people only interviewed post-election. The act of interviewing changes the future behavior of respondents.

Fishkin's participants were interviewed twice. Without a control group we do not know if any change in response is due to the experience of being recruited, flown to Austin, treated like a celebrity by being asked their opinions on national television and having participated in the deliberations, or just due to being interviewed twice. We will never know, because there was no control group and it was not a field experiment as claimed by Converse. It has been suggested that the people inter-

viewed at the first stage by NORC, who chose not to come to Austin, be reinterviewed as a control group. The obvious flaw in this suggestion is that these non-participants also are a self-selected group. They are not a random part of the original sample any more than the Austin participants. Treating this group as a control will waste more money and serve no point.

“ *As for [the deliberative experiment] increasing participation in the nominating process, turnout through Super Tuesday had been only 7%, down from the 12% turnout in 1992 that Fishkin abhorred.* ”

Reform: We Need More Than Good Intentions

We all hope issues can be a more central part of campaigns. That is why there was \$4 million worth of support for this project. It is difficult to get the public and the media to focus on issues, much less to help voters deliberate about them, if the candidates will not publicly discuss issues as part of their campaign. I suggest that candidates do not discuss issues because it is not in their interest to do so. Every issue position has proponents and opponents. The more specific a candidate is the more potential opponents will be accumulated over the course of a campaign. That's why candidates talk in vague generalities about issues, at best. Most seem to prefer to duck them all together. And what a candidate says about issues during a campaign often differs from what he or she does once elected. Nonetheless, it is the job of journalists and educators to advance the public dialogue about issues in a way that will better inform the public about how a candidate will govern, if elected.

Fishkin's attempt to stimulate the discussion of issues in a presidential

campaign has served to heighten interest in the topic for many people. He deserves credit for that. Unfortunately, the ones we most wish had heard it—the candidates—have not responded in the nominating campaign to date. But NIC was bound to fail. It is not possible to recruit a representative sample. NORC did a magnificent job trying, but it failed. The candidates had nothing to gain by showing up in person and exposing themselves by being specific about issues. Average citizens are not likely to change their opinions very much based upon a discussion with strangers of topics they have not thought much about and really do not care about strongly.

Opinion change is an evolutionary process. It takes place over time and is aided by relatives, friends and associates, the kind of people we talk to on a regular basis. It also is aided by the media, which facilitate discussion even when the material they present is not as thought-provoking as we might like. The NIC took place during a relatively short time period and excluded most influence from family, friends, associates and the media. This contributed to the artificiality of the process.

It is unfortunate that Fishkin and his sponsors made his experiment into a media event full of promises that could not be kept, which were swallowed by gullible supporters from academia, journalism and business. The sponsors staged an event—that much is clear. But nothing in their press release justifies the conclusions drawn about the success of the deliberative poll.

Endnote:

¹ I combined categories such as “strongly agree” and “agree somewhat,” and “very useful” and “somewhat useful,” for the purpose of this exercise.



*Warren J. Mitofsky is president,
Mitofsky International*