

Experimental Designs for Political Communication Research:
From Shopping Malls to the Internet

Shanto Iyengar
Stanford University
Political Communication Lab
<http://pcl.stanford.edu>

Workshop in Mass Media Economics
Department of Political Science
London School of Economics
June 25-26, 2002

Introduction

Twenty-five ago, the use of experimental methods was a rarity in the discipline of political science. In the early 1980s, a surge of interest in the interdisciplinary area of political psychology led experimental methods to trickle into several sub-fields of political science, including political communication. Despite this boost, longstanding concern over the limited realism and generalizability of experimental findings has maintained the position of survey research as the dominant paradigm in the field of political communication.

In the presentation that follows, I describe the inherent advantages of experimental design in disentangling cause from effect, using recent examples from political communication research. Next, I argue that the use of field techniques and the technological advances associated with the growth of the Internet go a long way towards neutralizing the traditional weaknesses of experimentation. First, experiments administered online can prove just as realistic and generalizable as conventional experiments. Second, issues of sampling bias -- previously endemic to experiments -- can be overcome through the greater “reach” of online experiments and, in addition, by the application of standard probability sampling techniques to the recruitment of online experimental participants. Finally, experimental design can be adapted to permit voluntary exposure to the experimental stimulus, thereby reflecting the inherent selectivity of real-world media audiences. Thus, as described below, experiments now represent a “dominant” methodology for political communication researchers.

The Issue of Causal Inference

In the field of political communication, the major advantage of the experiment over the survey -- and the focus of the discussion that follows -- is the ability to isolate and test the effects of specific components of political messages. Consider the case of political campaigns. At the aggregate level, campaigns encompass a concatenation of messages, channels, and sources, all of which may influence the audience, often in inconsistent directions. The researcher's task is to identify specific causal factors and delineate the range of their relevant attributes. Even at the relatively narrow level of campaign advertisements, for instance, there are virtually an infinite number of potential causal factors, both verbal and visual. What was it about the infamous "Revolving Door" advertisement that is thought to have moved so many American voters in 1988? Was it, as widely alleged, Mr. Horton's race? Or was it the violent and brutal nature of his behavior, the fact that he was a convict, the race of his victim, or what? Experiments make it possible to isolate the explanation, whether it be text-based, or in the form of audio-visual cues. Surveys, on the other hand, can only provide indirect evidence on self-reported exposure to the causal variable in question.

The weaknesses of survey design for isolating the effects of mass communication have been well documented. In a widely cited paper, Hovland (1959) identified several artifacts of survey research including unreliable measures of media exposure. The use of self-reports to measure exposure to political messages is especially problematic. People have notoriously weak memories for their political experiences (see, for instance, Bradburn, Rips and Shevell, 1987; Pierce and Lovrich, 1982). In the Ansolabehere and Iyengar experiments on campaign advertising, over fifty percent of the participants who

were exposed to a political advertisement were unable, *some thirty minutes later*, to recall having seen the advertisement (Ansolabehere and Iyengar, 1999). Of course, the considerable error in self-reports necessarily attenuates survey-based estimates of the effects of political campaigns (see Bartels, 1993, 1996).

In addition to the problem of measurement error, self-reported media exposure is typically endogenous to political attitudes, including candidate preference. That is, those who claim to tune in to politics differ systematically (in ways that matter to their vote choice) from those who do not. The endogeneity issue has multiple ramifications for political communication research. First, consider those instances where self-reported exposure is correlated with political predispositions, but actual exposure is not. This is generally the case with televised political advertising. Most voters encounter political ads not deliberately, but in the course of watching their preferred tv programs in which the ad breaks contain a heavy dose of political messages. Thus, actual exposure is idiosyncratic (based on the viewer's preference for particular programs), but self-reported exposure is based on political predispositions.

The divergence between actual and self-reported exposure has predictable consequences for "effects" research. In experiments that manipulated the tone of campaign advertising, the researchers found that exposure to negative messages discouraged turnout (Ansolabehere and Iyengar, 1995). However, on the basis of self-reports, survey researchers concluded that exposure to negative campaign advertising stimulated turnout (Wattenberg and Briars, 1999). But was it recalled exposure to negative advertising that prompted turnout, or the greater interest in campaigns among likely voters responsible for their higher level of recall? When recall of advertising in the

same survey was treated as endogenous to vote intention and the effects re-estimated using appropriate two-stage methods, the sign of the coefficient for recall was reversed: those who recalled negative advertisements were less likely to intend to vote (see Ansolabehere et al., 1999). Unfortunately, most survey-based analyses fail to disentangle the reciprocal effects of self-reported exposure to the campaign and partisan attitudes/behaviors. As this example suggests, in cases where actual exposure is less selective than reported exposure, self-reports will prove especially biased.

In other scenarios, the tables may be turned and the experimental researcher may be at a disadvantage. Actual exposure to political messages in the real world is typically not analogous to random assignment. Unlike advertisements, news coverage of political events can be avoided by choice, meaning that exposure is limited to the politically engaged strata. Thus, as Hovland (1959) pointed out, manipulative control weakens the ability to generalize to the real world where exposure to politics is typically not involuntary. In these cases, as I describe in the section on experimental realism below, it is important that the researcher use designs that combine manipulation with self-selection of exposure.

In summary, the advantage of the experimental approach is the ability to isolate causal variables, which become the basis for experimental manipulations. In the next section, I describe manipulations designed to isolate the effects of the tone of advertising campaigns and the race of the criminal suspect in television news reports.

Negativity in Campaign Advertising

Experimental tests of negativity require variation in the tone of a campaign advertisement while holding all other attributes constant. (Although the contrast with

survey-based measures of exposure to negative advertising could not be more stark, the inherent imprecision of surveys has not discouraged survey analyses of campaign negativity. For recent examples, see Geer and Finkel, 1998; Freedman and Goldstein, 1999; Kahn and Kenney, 2000.) In the Ansolabehere/Iyengar experiments, the researchers manipulated negativity by varying the text (the soundtrack) of an advertisement while preserving the visual backdrop. The negative version of the message typically placed the sponsoring candidate on the unpopular side of a salient policy issue. Thus, during the 1990 gubernatorial campaign between Pete Wilson and Dianne Feinstein, our ads described the candidates either as opponents or proponents of offshore oil drilling and thus as either friends or enemies of the environment. This manipulation was implemented by substituting the word “yes” for the word “no.” In the positive conditions, the script began as follows: “When federal bureaucrats asked for permission to drill for oil off the coast of California, Pete Wilson/Dianne Feinstein said no . . . “ In the negative conditions, we substituted “said yes” for “said no.” An additional substitution was written into the end of the ad when the announcer claimed that the sponsoring candidate would work to either “preserve” or “destroy” California’s natural beauty. Given the consensual nature of the issue, negativity could be attributed to candidates who claimed their opponent was soft on pollution.¹

¹ Of course, this approach assumes a one-sided distribution of policy preferences and the tone manipulation would be reversed for participants in the experiment who actually favored off shore drilling.

Positive Environment Ad



Negative Environment Ad



Racial Cues in Local News Coverage of Crime

As any regular viewer of television will attest to, crime is a frequent occurrence in broadcast news. In response to market pressures, television stations have adopted a formulaic approach to covering crime, an approach designed to attract and maintain the highest degree of audience interest. This “crime script” suggests that crime is invariably violent and those who perpetrate crime are disproportionately nonwhite. Because the crime script is encountered so frequently (several times each day in many cities) in the course of watching local news, it has attained the status of common knowledge. Just as we know full well what happens when one walks into a restaurant, we also know -- or at least think we know -- what happens when crime occurs.

In a series of recent experiments, Frank Gilliam and I have examined the effects of both elements of the crime script on audience attitudes. For illustrative purposes, I focus here on the racial element. In essence, we manipulated the racial appearance of the alleged suspect depicted in a typical news report while maintaining all other visual characteristics. The original stimulus was a local news report, which included a close-up “mug shot” of the suspect. The picture was digitized, “painted” to alter the perpetrator’s

skin color, and then re-edited into the news report. As shown below, beginning with two different perpetrators (a white male and a black male), we were able to produce altered versions of each individual in which their race was reversed, but all other features remained identical. Participants who watched the news report in which the suspect was thought to be non-white expressed greater support for “punitive” policies (e.g. imposition of “three strikes and you’re out” remedies, treatment of juveniles as adults, etc.). Given the precision of the design, these differences in the responses of the subjects exposed to the white or black perpetrators can only be attributed to the perpetrator’s race.

Suspect Comparison



In short, as these examples indicate, the experiment provides unequivocal causal evidence because the researcher is able to isolate the causal factor in question, manipulate its presence or absence, and hold all other potential causes constant.

The Issue of Generalizability

The problem of limited generalizability, long considered the Achilles Heel of experimental design, manifests itself at three levels: mundane realism, sampling bias, and endogeneity of exposure.

Mundane Realism

Because of the need for tightly controlled stimuli, the setting in which the typical laboratory experiment occurs is often quite dissimilar from the setting in which subjects ordinarily experience the “target” phenomenon. Concern over the artificial properties of laboratory experiments has given rise to an increased use of so-called field experiments in which the procedures and settings more closely reflect ordinary life.

A common strategy in field experiments is the reliance on interventions with which subjects are familiar. The Ansolabehere/Iyengar campaign experiments were relatively realistic in the sense that they occurred during ongoing campaigns characterized by heavy levels of televised advertising (see Ansolabehere and Iyengar, 1996). The presence of a political advertisement in the local news (the vehicle used to convey the manipulation) was hardly unusual or unexpected since candidates advertise most heavily during news programs. The advertisements featured real candidates -- Democrats and Republicans, liberals and conservatives, males and females, incumbents and challengers -- as the sponsors. The material that made up the experimental stimuli were selected either from actual advertisements used by the candidates during the campaign, or were produced to emulate typical campaign advertisements. In the case of the latter, the researchers spliced together footage from actual advertisements or news reports making the treatment ads representative of the campaign-advertising genre. (Of

course, the need for control made it necessary for the test ads to differ from actual political ads in several important attributes including the absence of music and the appearance of the sponsoring candidate.)

Realism also depends upon the physical setting in which the experiment is administered. Asking subjects to report to a location on a university campus may suit the researcher, but may make the experience of watching television equivalent to the experience of visiting the doctor. A more realistic strategy is to provide subjects with a milieu that closely matches the setting of their living room or den. To that end, the Ansolabehere/Iyengar experimental “laboratory” was designed to resemble, as closely as possible, the natural “habitat” of the television viewer. Comfortable couches and chairs were arranged in front of a television set, with houseplants and wall hangings placed around the room. Respondents had access to refreshments and reading matter (newspapers and magazines) during the viewing sessions. In most cases, a family member or friend took part in the experiment at the same time, so that subjects did not find themselves sitting next to a stranger while viewing the political advertisements.²

A further step toward realism concerns the power of the manipulation (also referred to as experimental realism). Naturally, the researcher would like the manipulation to be noticed by the subject. At the same time, it is important that the manipulation not overpower the subject (as in the Milgram obedience studies where the task of administering electric shock to a fellow subject proved overpowering to many). In the case of the campaign advertising experiments, we resolved the experimental realism-

² It is possible, of course, for the experimental setting to be too realistic. During the early days of our campaign experiments, we provided subjects with access to a remote control device, only to discover that a subject used it to fast forward the tape during the commercial breaks.

mundane realism tradeoff by embedding the manipulation in a commercial break of a local newscast. The political ad appeared with other non-political ads (we excluded other political ads from the newscast) and because subjects were led to believe that the study was about “selective perception of news,” we gave them no rationale to pay particular attention to ads. Overall, the manipulation was relatively small, amounting to thirty seconds of a fifteen-minute videotape.

In general, the mundane realism - experimental control trade off is significant. The fact that subjects watch treatments in the company of others means that their level of familiarity with fellow subjects is subject to unknown variation. And producing experimental ads that more closely emulate actual ads (e.g. ads with musical background included and featuring the sponsoring candidate) introduces a series of confounded variables associated with the appearance and voice of the sponsor.

Sampling Bias

The most widely cited limitation of experiments concerns the composition of the subject pool (Sears, 1986). Typically, laboratory experiments are administered upon "captive" populations -- college students who must serve as guinea pigs in order to gain course credit. College sophomores may be a convenient subject population, but are they comparable to "real people?"

In conventional experimental research, it is possible to broaden the participant pool, but at considerable cost/effort. Locating experimental facilities at public locations and enticing a quasi-representative sample to participate proves both cost- and labor-intensive. Typical costs include rental fees for an experimental facility in a public area (such as a shopping mall) where it is possible to attract a wide range of participants;

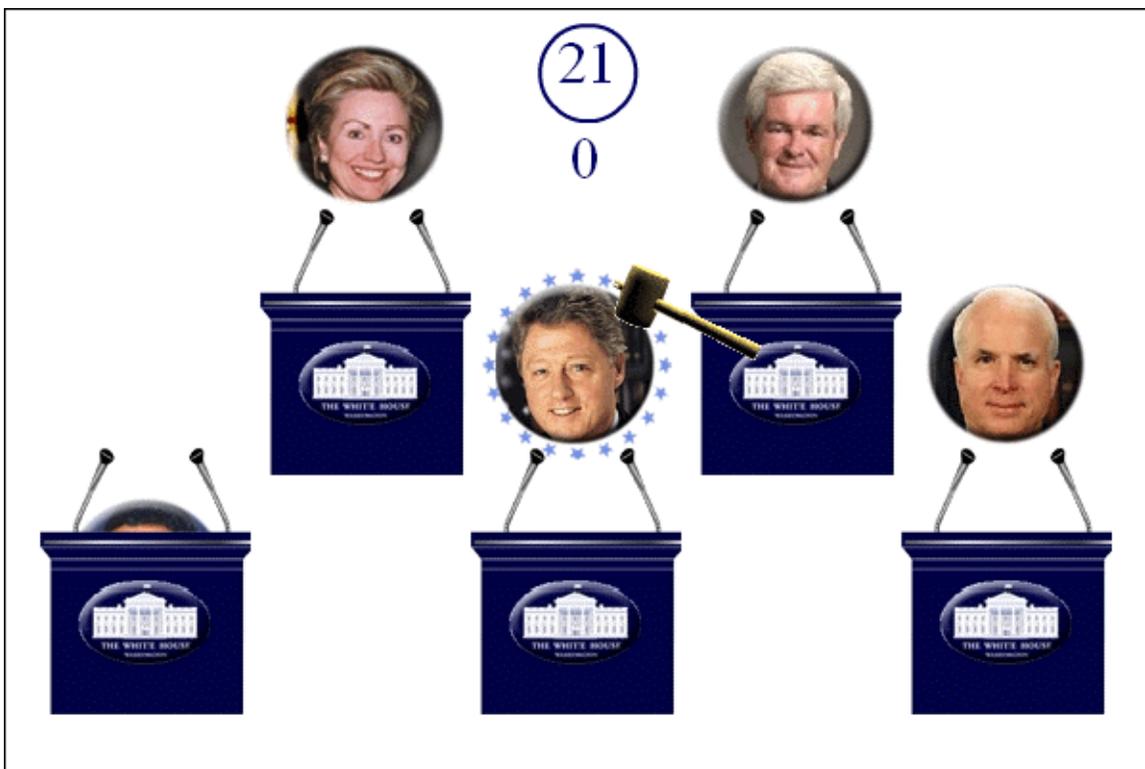
recruitment and compensation of subjects; and recruitment, training and compensation of research staff to administer the experiments. In the local news experiments conducted in Los Angeles in 1998 and 1999, the total costs per subject amounted to approximately \$45. Fortunately, technology has both enlarged the pool of potential participants and reduced the per capita cost of administering subjects.

Today, traditional experimental methods can be rigorously and far more efficiently replicated using on-line strategies. The advantages of using the Internet as the experimental “site” include the ability to reach diverse populations without geographic limitations. The rapid development of multimedia-friendly browsers makes it possible to bring video presentations to the computer screen. Indeed, all the major presidential candidates now “air” their televised ads at their web sites, and it is common practice for people to follow news reports from online sources such as CNN. The technology is so accessible that subjects can easily “self-administer” experimental manipulations. Compared with conventional shopping mall studies, therefore, the costs are minimal. Moreover, with the ever-increasing use of the Internet not only are the samples more diverse, the setting in which participants encounter the manipulation (surfing the Web on their own) is also more realistic.

The Political Communication Laboratory at Stanford University has been administering experiments over the WWW for nearly three years. These experiments feature text, audio, and video manipulations. One of the Lab’s more popular online experiments is “whack a politician,” (<http://pcl.stanford.edu/exp/whack/polm>) modeled on the well-known whack-a-mole arcade game. Ostensibly, the game provides subjects the opportunity to “bash” well-known political figures. Before playing the game,

subjects complete a consent form and brief pretest questionnaire. After playing the game, they self-administer the posttest. Since the game imposes severe time and attention constraints (subjects see five different moving faces, each hittable for a period of between two and three seconds), the whacking task provides an unobtrusive measure of group identity. That is, subjects are expected to target “out-group” figures for more extensive whacking.

Whack-a-Pol



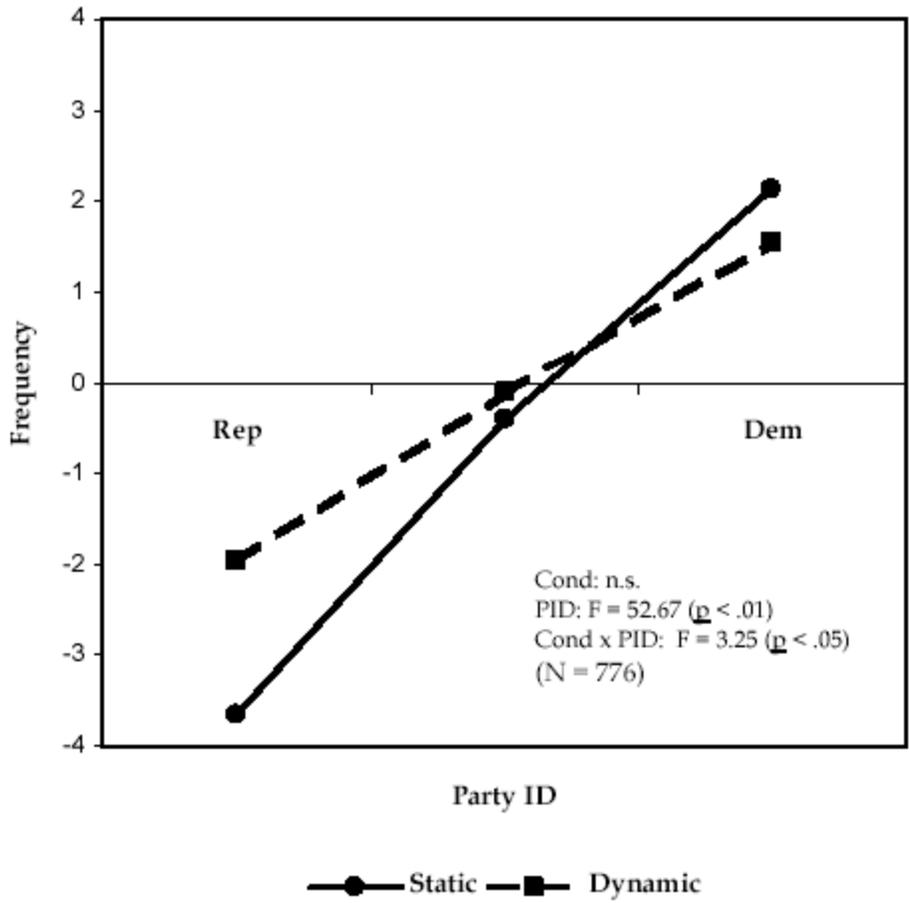
Our initial study was designed to test this prediction using political party affiliation as the basis for identity. The targets were five well-known American politicians (President Bush, Bill Clinton, Hillary Clinton, Newt Gingrich, Jesse Jackson, and John McCain). Two conditions were created so as to make party-based whacking more or less difficult. In the “static” condition, the five targets consistently appeared in

the same location; in the “dynamic” condition, their location on the screen was randomized. Naturally, we expected stronger effects of subjects’ party affiliation on their whacking behavior in the former condition.

780 subjects participated in the first study. The results, as summarized in Figure 1, demonstrate very strong effects of party affiliation. Republican subjects selected Democratic targets and vice versa. Second, the significant interaction between party affiliation and static versus dynamic condition revealed the weakened proclivity of partisans to pick on out-group targets in the dynamic condition. Note, however, that partisan whacking survived (at a significant level) in the dynamic condition; despite the additional level of difficulty, subjects behaved as partisans. Thus, these results demonstrate the power of party affiliation as a political cue.³

³ In a second study, we have replicated these findings using nationality instead of party as the basis for group identity. American players tended to hit John Kennedy and Martin Luther King Jr. less frequently than residents of other countries. This particular game can be played at <http://pcl.stanford.edu/exp/whack/polhm>.

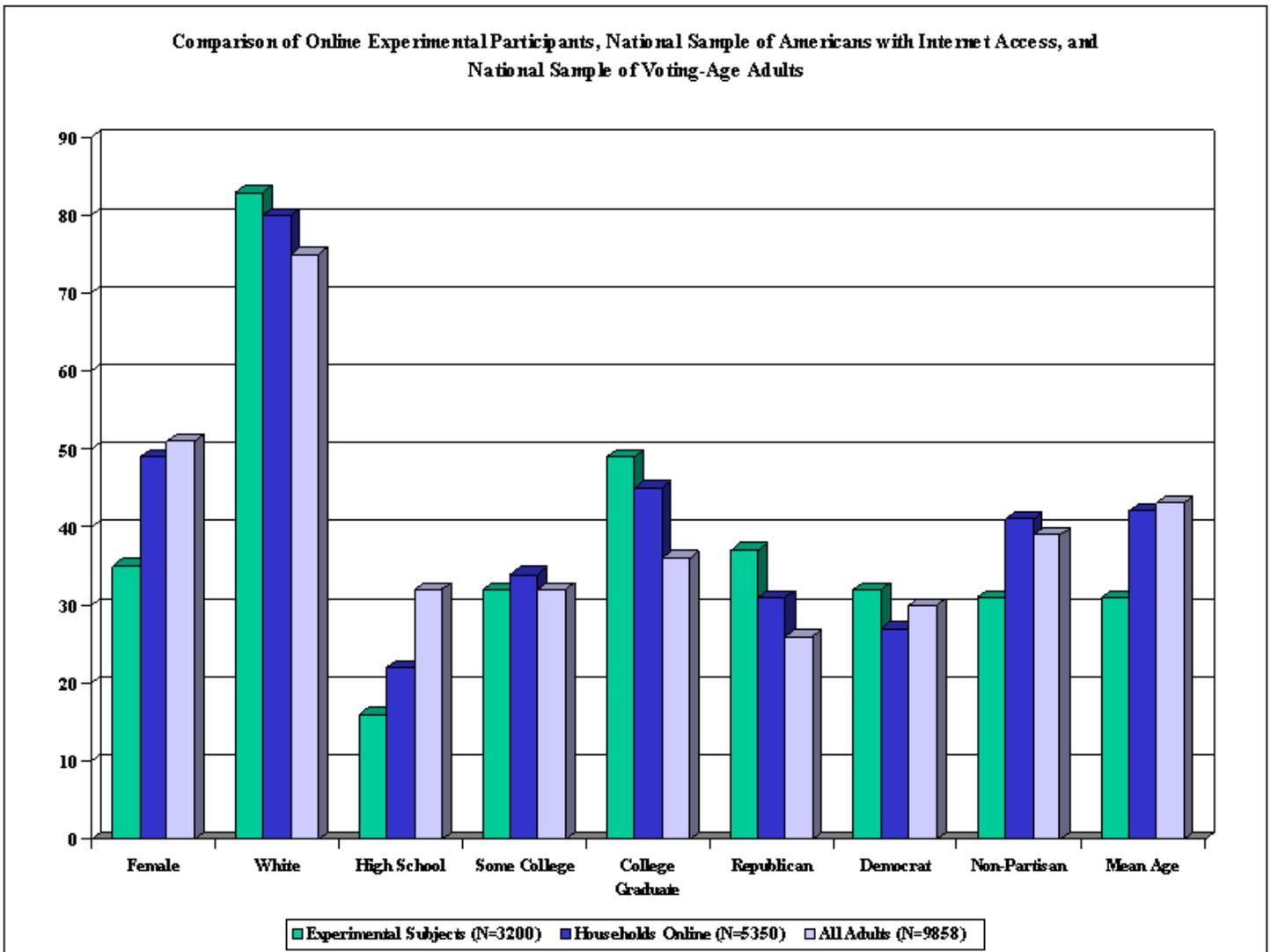
Figure 1
Effects of Party Identification in “Whack-a-Pol”



The subjects who participated in the whack-a-pol study found the PCL site on their own initiative and then chose to play the whacking game. How do such “drop-in” subjects compare with typical Internet users? We have compiled the demographic profiles of our online study participants and the results are shown in Figure 2. For context, we compare our participants with a representative sample of adult Americans with home access to the Internet as well as a representative sample of all voting-age adults.⁵ (The survey was administered in March 2000.) The first comparison enables us

to gauge the degree of online self-selection; that is, the degree of divergence between drop-in participants and typical Internet users. The second comparison indicates the degree of discrepancy between our self-selected online samples and a representative sample of voting-age adults.

Figure 2⁴



⁴ I am grateful to my colleague Doug Rivers, then CEO of Knowledge Networks, for making these data available.

The comparisons in Figure 2 support two broad patterns. The more surprising of the two is that the participants in the online experiments reasonably approximate the online user population. The digital divide, however, is still sufficiently wide to exclude major categories of the population from participating in online studies.

The comparison between our experimental subjects and the representative sample of Americans with home Internet access was closest with respect to race/ethnicity and education. The predominance of whites and the college-educated was about the same in the subject pool and among national survey respondents with Internet access.

Of the background factors considered, political communication researchers are most interested in the variable of party identification. Among experimental subjects, Republicans made up the largest group (37 percent), followed by Democrats and Independents. Thus, the participant pool was somewhat less Democratic (and more Republican) than the broader online population. Notwithstanding these minor differences, it appears that party identification does not enter into the decision to participate in online experiments with political content.

The clearest evidence of experimental selection bias emerged with age and gender. Study participants were much younger (on average, by nearly ten years), and much more likely to be male than the online population. The sharp divergence in age may be attributed to the fact that our studies are launched from on an academic server that is more likely to be encountered by college students, and also to the general “surfing” proclivities of younger users. The gender gap is more puzzling and may reflect differences in political interest. Our studies are explicitly political in focus, which may act as a disincentive to potential women subjects.

In summary, if the population of interest consists of Americans with online access, the experimental participants comprise a reasonably representative sample at least with respect to race, education, and party affiliation. The experiments deviate from the online population on the attributes of gender and age, drawing disproportionately male and younger participants.

Turning to the digital divide, the access threshold remains a strong liability for online research. In relation to the broader adult population, our experimental participants are younger, more educated, more likely to be white males, and less apt to identify as a Democrat. With the exception of age and gender, these differences are just as stark when the comparison is between all adults and the representative sample of adults with online access (for evidence of the scale of differences between Internet users and nonusers, see Moss and Mitra, 1999; Papadakis, 2000).

Demographic differences associated with Internet access are well known. The question of parallel attitudinal differences, however, has received less attention. Our data suggest that online access is correlated with political party identification. Study participants (and online users in general) are more likely to identify as Republican than Democrat. The plurality of Republicans online may reflect access differences associated with economic standing and ethnicity. Whatever the reason, online researchers interested in the effects of campaigns will need to make special efforts to over sample Democrats.

Although these data make it clear that people who participate in online media experiments are no microcosm of the adult population, the fundamental advantage of online over conventional field experiments cannot be overlooked. Conventional experiments recruit subjects from particular locales; online experiments draw subjects

from across the country. The Ansolabehere/Iyengar campaign experiments, for example, recruited subjects from a particular area of southern California (greater Los Angeles). The online experiments, in contrast, attracted subjects from thirty different American states and several countries.

To this point, our data are limited to experiments with drop in participants. It is possible to broaden the online participant pool by using banner ads that offer potential subjects a financial incentive to participate. One anticipates that subjects who respond to a monetary incentive will prove more diverse than those who volunteer. Should the researcher be fortunate enough to have grant support, it is possible to go one step further by contracting with the market research firm of Knowledge Networks which has pioneered the use of web-based surveys and experiments with fully representative samples.

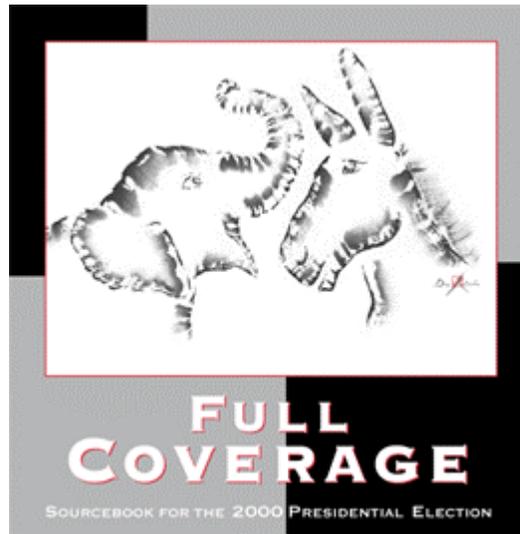
In essence, KN has solved the thorny problem of selection bias inherent to conventional online surveys (which reach only that proportion of the population that is online) by recruiting a nationwide panel through standard telephone methods. This representative panel (over 150,000 Americans between the ages of 16 and 85) is provided free access to WebTV. In exchange, they agree to participate (on a rotation basis) in market and research studies being conducted by KN. The surveys are administered over a WebTV platform. Thus, by providing free Internet access to their panelists, KN is able to deliver samples that meet the highest standards of probabilistic sampling. Because their panelists have an incentive to participate, KN also provides significantly higher response rates than telephone-based surveys.

From Exogenous to Endogenous Media Manipulations

The final drawback of experimental design for political communication researchers is the misfit between exposure based on random assignment in the experiment, and exposure based on self-selection in the real world. To cite an extreme case, I might design a manipulation in which one group of participants watches Public Television's "News Hour with Jim Lehrer." No more than a handful of the participants assigned to this condition would be members of the regular audience for the program; to the degree the effects of exposure are conditioned by socio-economic status or other such attributes, generalizing these experimental results would appear hazardous.

Self-selected exposure can be incorporated into experimental designs by providing participants with choice or control over their exposure to the treatment. As a recent case in point, Simon Jackman and I designed an experiment in which two different compact disks were mailed to a representative sample of American voters just before the 2000 presidential election (for details, see Iyengar, Jackman, Hahn, and Markus, 2001). Participants assigned to the "unmediated" CD condition received a CD containing the campaign speeches and televised advertisements (along with the respective party platforms) of the two major presidential candidates, but which was devoid of any media reports. In contrast, the "mediated" CD, identical in appearance and user interface, presented a collection of typical print, radio, and television news reports on the presidential campaign. The two treatments were calibrated to be roughly equivalent in terms of length, ratio of multimedia to text content, and amount of coverage aimed at the two main candidates.

Mediated CD



TK3 - sourcebook2000.tk3
File Edit Tools Window Notebook Help

The Issues Overview and Ideology Full Coverage

CNN Inside Politics
Sep 19, 2000, 5 pm EST
CNN's Brooks Jackson Compares Bush's and Gore's TV Ads and Press Releases (Video)



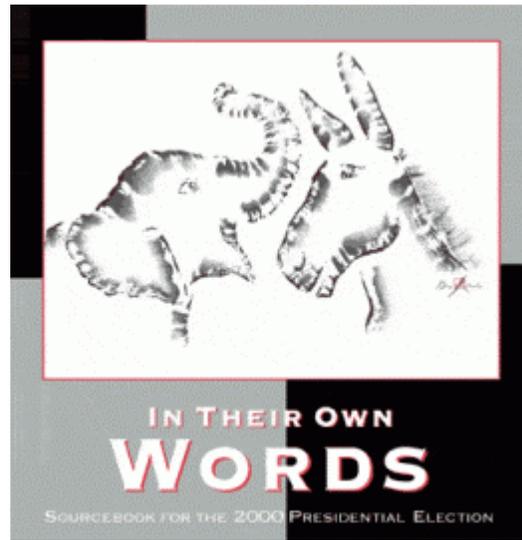
Transcript:
SHAW: Well, beyond prescription drugs, the Bush and Gore camp seem to have plenty to quibble about these days. Our Brooks Jackson has been checking out some areas of dispute on the trail and on the airwaves.
(BEGIN VIDEOTAPE)
JACKSON (voice-over): Not so long ago, they promised

to take the high road.
BUSH: We are going to herald what we stand for in a positive and constructive way.
GORE: I'm not going to say a single negative personal thing about my opponents. You will not hear from me in this entire campaign.
JACKSON: Well, that's what they said then. But look at what their campaigns are doing now. Republicans run misleading descriptions of Gore's Medicare proposal.
(BEGIN VIDEO CLIP, RNC AD)
NARRATOR: But his prescription drug plan forces seniors into one HMO selected by the federal government.
(END VIDEO CLIP)
JACKSON: That's wrong, of course. Gore's government-run plan is voluntary. And seniors who sign up could still buy their medication from pharmacies.
Democrats countered with this about Bush's prescription plan:
(BEGIN VIDEO CLIP, DNC AD)
NARRATOR: And Bush forces senior he does include to go to HMOs and insurance companies for coverage.
(END VIDEO CLIP)
JACKSON: That's wrong, too. Bush's approach is also

find
hiliter
stickies
notebook
marks
contents
retrace
history
help
quit

115

Unmediated CD



TK3 - sourcebook2000.tk3
File Edit Tools Window Notebook Help

Health Care & Retirement Medicare and Prescription Drugs - Gore In Their Own Words

call on Congress to give the FDA unequivocal power over this issue, including the power to impose tough financial penalties on companies that market to children.

We must match the tobacco companies' big advertising campaign with national counter-advertising about the dangers of smoking and the risks of cancer. And we must double our investment in efforts to prevent smoking – so we can prevent more cancer, and so we can find new ways to break the grip of nicotine addiction.

The issue isn't easy—and there are entrenched interests on the other side. But it's an issue where we can never give up, and never give in. I promise you: I never will.

The steps I am announcing today are realistic, and the goals I am setting are achievable. We can and must harness the wonders of today's scientific discoveries, to keep people healthy and alive. We can and must bring the best of treatment and prevention to millions of Americans.

...

Al Gore
June 13, 2000
TV Ad: "Prescription Drugs"

detach video

[Announcer]:
Every week, Bob Darthez has to afford his groceries and prescription drugs. He's worked a lifetime, but now he's at the mercy of the big drug companies. They're using money and lobbyists to stop progress in Washington.

Al Gore is taking them on. Fighting for a Medicare prescription drug benefit for seniors like Bob Darthez.

419

find
highlighter
stickies
notebook
marks
contents
retrace
history
help
quit

The CDs were mailed to a sample of 950 adults with Internet access, selected at random by Knowledge Networks.⁵ The CDs, which were delivered to the study participants two weeks in advance of the election, were programmed to enable “usage tracking.” That is, each time the CD was accessed, the host PC activated a data file in which subsequent CD usage was recorded. Of course, participants were fully informed of this feature and were given instructions (at the time of the survey) concerning the transmission of this data to Knowledge Networks. In both conditions, the CDs were described as an educational product of Stanford University, which sought feedback on user experience and reactions. Participants were asked to use the CD as they saw fit.

The rationale underlying the CD intervention was two-fold. First, we intended to provide users with a greater range of choices concerning their exposure to campaign information. We anticipated that participants’ ability to control their access to personally relevant information would stimulate interest in the campaign and the likelihood of voting. Second, because the unmediated CD allowed voters to bypasses the strategic and often contentious nature of news coverage, we further expected CD use to boost voter engagement in the campaign. Moreover, the mere act of using an election-related CD might trigger a process of self-inference (Bem, 1972) by which people come to see themselves as relevant political actors. Use of the CD may cue participants to consider themselves politically motivated, which, in turn, carries over to political participation (for a discussion of such “foot in the door” effects, see Schneider et al., 1979, pp. 102-03).

This “empowerment” hypothesis, of course, applies to both treatments, although we

⁵ We screened participants on the basis of home Internet access for two reasons. First, access was required for participants to transmit the CD usage data back to Knowledge Networks. Second, we wished to emulate the effects of using the unmediated CD on voters who actually ordered a copy of the CD during the campaign from an online “store.”

expect the effect to be more pronounced in the case of the unmediated CD (mainly because of the content differences noted below).⁶

Our second prediction was more subtle, and based on fundamental differences in the content of the two treatments. By definition, participants in the mediated condition encountered more references to the strategy and horse race frames, which typically dominate news media coverage. In contrast, the unmediated CD was primarily issue-oriented. The two conditions differed further in terms of their tone. In keeping with the demands of interpretive journalism, news reports in the media CD were apt to present a somewhat critical, adversarial perspective on the candidates. The speeches, ads, and platform planks, on the other hand, were characterized by generally self-promotional or positive presentations from both candidates. Finally, and most obviously, participants in the candidate condition encountered the candidates' own voices; those in the media condition encountered the voices of pundits, experts and reporters. Each of these content differences, we anticipated, would alter participants' evaluations of the electoral process. Cumulatively, we expected that exposure to issue-oriented, generally positive, and first-person (from the candidate's perspective) accounts of the campaign would increase public support for campaigns and the candidates who contest them. Conversely, we expected that exposure to news media coverage of the campaign would have precisely the opposite effect. Thus, while exposure to either CD was expected to boost interest or

⁶ The interview schedule included several indicators of political involvement including self-reported turnout and a set of attitude items measuring participants' sense of political efficacy (defined as the perception that one is capable of exercising influence over the process).

involvement, evaluations of the process were expected to diverge depending upon exposure to unmediated or mediated information.⁷

As in any experiment, we assigned participants to the media and candidate CD conditions at random. Also typical of experiments where participation is voluntary, however, a considerable number of participants who received the CD did not complete the survey or return the CD tracking data. Many of the latter, we presume, chose not to use the CD. Non-compliance, of course, is a result of self-selection. Respondents who used the CDs mailed to them and those who did not are likely to differ on a variety of characteristics known to influence the outcome variables of interest. For example, perhaps people less interested in politics were less likely to view the CDs than respondents more engaged by the election. We also found clear evidence that CD use was guided by personal experience; people were more likely to look at sections of the CD that addressed issues that affected them (see Iyengar et al., 2001a). Elderly participants, for instance, paid more attention to the candidates' discussions of social security. These differences suggest that CD users and non-users differ systematically, and a naïve analysis of the effects of the CD intervention, one that ignores the fact that compliance with the treatment and the dependent variables of interest are jointly determined by the same underlying predispositions, risks overestimating the effects of CD use.

An important feature of the CD study design was the ability to treat exposure to direct campaign communication as endogenous to voter predispositions. All participants who were sent the CD were interviewed immediately following the election. Comparing the attitudes of this group with the control group allows us to estimate the effects of

⁷ As indicators of public enthusiasm for campaigns, we included an index of political cynicism. Participants rated candidates for elective office (and public officials in general) according to their sincerity,

“intent to treat” (see Angrist et al., 1996; Gerber and Green, 2000). That is, we can observe the effects of receiving the CD on participant attitudes, an effect that is based almost purely (give or take some self-selection in the non-response to the survey) on random assignment. We can also assess the effects of the treatment itself by examining the responses of participants who actually used the CD (defined as the 435 participants who were able to return their CD usage data). This group, clearly, is more strongly contaminated by selection bias.

Table 1 compares the effects of intent to treat with actual treatment for reported turnout, the sense of political efficacy, and the index of political cynicism. Note that the effects of intent to treat, although statistically significant in two of the three cases, were much weaker than the effects of actual CD use. The turnout gains associated with intention to treat, for instance, were only six and eight points (compared with twelve and thirteen points for actual treatment) for the media and candidate CDs, respectively. Similarly, the gains in political efficacy, although still robust at the level of intention to treat, were weakened when non-compliers were added to compliers. Finally, in the case of cynicism, we could not replicate the significant effect of exposure to the unmediated CD at the level of intention to treat. When compliers and non-compliers were pooled, exposure to the unmediated CD had no significant effect on level of cynicism.

honesty, and concern for the public good.

Table One

Treatment Versus Intention to Treat Effects: 2000 CD Study

	Control Group	Med. CD/ Intent to Treat	Unmed. CD/ Intent to Treat	Med CD/ Treatment	Unmed. CD/Treatment
Turnout	.77 (642)	.83** (496)	.85** (523)	.89*** (207)	.90*** (226)
Sense of Efficacy	1.74 (639)	1.99*** (423)	2.01*** (451)	2.02** (194)	2.11*** (223)
Trust in Politicians	1.92 (602)	1.83 (417)	2.01 (438)	1.87 (194)	2.17* (217)

***p < .001; **p < .01; *p < .05

The presence of instability in the magnitude of the experimental effects between actual treatment and intent to treat may be taken as evidence of selection bias. Thus, a more precise estimation of the effects of either CD requires that we adjust for the built-in, motivational advantages of participants who complied with the intervention (see Iyengar et al., 2001 for an analysis of the “discounted” effects of CD use). Fortunately, the econometrics and statistics literature offer several solutions to the problem of selection bias (for recent reviews, see Little and Rubin, 2000; Angrist et al., 1996; Heckman and Robb, 1985). Taking self-selection into account, we concluded that actual CD use significantly boosted voter involvement in the case of the mediated CD, whereas mere receipt of the CD accounted for most of the effects of the unmediated CD (see Iyengar et al., 2001b).

Conclusion

The standard comparison of experiments and surveys favors the former on the grounds of precise causal inference and the latter on the grounds of greater generalizability. As I have suggested, however, traditional experimental methods can be

effectively, and just as rigorously replicated using online strategies which have the advantage of reaching a participant pool that is more far-flung and diverse than the pool relied on by conventional experimentalists. Online techniques also permit a more precise “targeting” of recruitment procedures so as to enhance participant diversity. Banner ads publicizing the study and the financial incentives for study participants can be placed in portals or sites that are known to attract underrepresented groups. Women subjects or African Americans, for instance, could be attracted by ads placed in sites catering to these groups.

Online experiments also provide strong metrics for monitoring the effects of the manipulation. Most web browsers record information on user behavior, a feature that can provide precise, yet unobtrusive measures of participant attitudes. In the case of the CD study, we were able to test the hypothesis of selective attention by comparing page visits across different groups of participants. Participants whose personal circumstances placed them at risk of incurring significant medical costs were found to spend more time considering the materials on health care. Participants with young children were drawn similarly to the issue of education. Such behavioral measures of attention, of course, are preferable to conventional survey self-reports.

The most compelling argument in favor of online experiments, however, is the inexorable diffusion of information technology. The latest government figures for the U.S. show that two-thirds of the population uses computers and 54% use the Internet (U.S. Dept. of Commerce, 2002). More revealingly, 90 and 75 percent of Americans between the ages of 5-17 use computers and the Internet respectively. As online communication penetrates classrooms across America and as the market share of online

communication sources grows, the external validity gap between experimental and survey methods can only close.

References

- Angrist, J. D., G. W. Imbens, and D. R. Rubin. 1996. "Identification of causal effects using instrumental variables," Journal of the American Statistical Association, 91: 444-55.
- Ansolabehere, S., and S. Iyengar. Going Negative. New York: Free Press.
- Ansolabehere, S., Iyengar, S., and A. Simon. 1999. "Replicating experiments using aggregate and survey data." American Political Science Review, 93: 901-10.
- Bartels, L. M. 1993. "Messages received: the political impact of media exposure." American Political Science Review, 87: 267-85.
- Bradburn, N., L. J. Rips, and S. K. Shevell. 1987. "Answering autobiographical questions: the impact of memory and inference in surveys." Science, 236: 157-61.
- Finkel, S. E., and J. Geer. 1998. "A spot check: Casting doubt on the demobilizing effect of attack advertising." American Journal of Political Science, 42: 573-95.
- Freedman, P., and K. M. Goldstein. 1999. "Measuring media exposure and the effects of negative campaign ads." American Journal of Political Science, 43: 1189-208.
- Gerber, A. S., and D. P. Green. 2000. "The effects of canvassing, telephone calls, and direct mail on voter turnout: A field experiment," American Political Science Review, 94: 653-63.
- Gilliam, F., and S. Iyengar. 2000. "Prime suspects: the influence of local television news on the viewing public," American Journal of Political Science, 44: 560-73.
- Heckman, J. and R. Robb. 1985. Alternative methods of solving the problem of selection bias in evaluating the impact of treatments, in H. Wainer ed., Drawing Inferences from Self-Selected Samples. Berlin: Springer-Verlag.

- Hovland, C. L. 1959. "Reconciling conflicting results derived from experimental and survey studies of attitude change," American Psychologist, 14: 8-17.
- Iyengar, S., and A. Simon. 2000. "New perspectives and evidence on political communication and campaign effects," Annual Review of Psychology, 51: 149-69.
- Iyengar, S., S. Jackman, K. Hahn, and M. Prior. 2001. Does direct campaigning empower voters? Paper presented at the Annual Meeting of the American Political Science Association.
- Kahn, K. F., and P. J. Kenney. 1999. "Do negative campaigns mobilize or suppress turnout? Clarifying the relationship between negativity and participation." American Political Science Review, 93: 877-90.
- Moss, M. L., and S. Mitra. "Net equity." <http://urban.nyu.edu/research/net-equity/>.
- Papadakis, M. C. 2000. "Complex picture of computer use in the home emerges." National Science Foundation Issue Brief, March 31.
- Pierce, J. C., and N. P. Lovrich. 1982. "Survey measurement of political participation: selective effects of recall in petition signing," Social Science Quarterly, 63: 164-71.
- Sears, D. O. 1986. College sophomores in the laboratory: Influences of a narrow data base on the social psychology view of human nature, Journal of Personality and Social Psychology, 51: 515-30.
- Wattenberg, M. P., and C. L. Briens. 1999. "Negative Campaign Advertising: Demobilizer or Mobilizer?" American Political Science Review, 93: 891-900.