

HOME / VICTORY LAB : THE NEW SCIENCE OF WINNING CAMPAIGNS.

## The Death of the Hunch

Campaigns used to guess which ads were most effective. Now they can prove it. How Obama's embrace of empiricism could swing the 2012 race.

By [Sasha Issenberg](#) | Posted Tuesday, May 22, 2012, at 12:37 PM ET



Still from "The Road We've Traveled"  
<http://www.barackobama.com>

Two weeks ago, top Obama campaign advisers Jim Messina and David Axelrod announced a \$25 million national television buy, a figure rightfully [acknowledged](#) with a sense of wonder, given that there were still six months to go before Election Day. But anyone waiting for coast-to-coast shock-and awe must be disappointed. The ads have rolled out at a desultory trickle: a [nine-state buy](#) for a 60-second overview of Obama's first-term successes; a [Spanish-language health-care ad](#) running in Florida and another in English about higher-education costs appearing there and in Nevada; and a long ad about Bain Capital that reportedly cost [less than \\$100,000](#) to place in markets across five states. In other words, the Obama team has broken nearly every piece of received wisdom that media consultants like to offer about the intensity and duration necessary for television ads to be successful in the modern era.

But scattered, unsustained messaging has become the unlikely hallmark of the well-funded Chicago campaign. The strategy was put into play even before Romney emerged as the Republican nominee. There was the late-November [advertising run on satellite systems](#) that the campaign called "tiny," and then silence until a brief January [broadcast-buy](#) across six states focusing on energy, ethics, and the Koch brothers. An isolated flight of brochures about health-care legislation hit mailboxes in March, timed to Supreme Court arguments on the subject. In voluminous (if not easily audited by outsiders) online ads and targeted email blasts, the campaign has addressed seemingly every topic or theme imaginable: taxes paid by oil companies, the "war on women," and a variety of local issues of interest in battleground states.

If these forays seem random, it's because at least some of them almost certainly are. To those familiar with the campaign's operations, such irregular efforts at paid communication are indicators of an experimental revolution underway at Obama's Chicago headquarters. They reflect a commitment to using randomized trials, the result of a flowering partnership between Obama's team and the [Analyst Institute](#), a secret society of Democratic researchers committed to the practice, according to several people with knowledge of the arrangement. (Through a spokeswoman, Analyst Institute officials declined to comment on the group's work with Obama and referred all questions to the campaign's press office, which did not respond to an inquiry on the subject.)

Advertisement

The Obama campaign's "experiment-informed programs"—known as EIP in the lefty tactical circles where they've become the vogue in recent years—are designed to track the impact of campaign messages as voters process them in the real world, instead of relying solely on artificial environments like focus groups and surveys. The method combines the two most exciting developments in electioneering practice over the last decade: the use of

randomized, controlled experiments able to isolate cause and effect in political activity and the [microtargeting statistical models](#) that can calculate the probability a voter will hold a particular view based on hundreds of variables.

Obama's campaign has already begun rolling out messages to small test audiences. Analysts then rely on an extensive, ongoing microtargeting operation to discern which slivers of the electorate are most responsive, and to which messages. This cycle of trial and error offers empirically minded

electioneers an upgrade over the current régime of approaching voters based on hunches.

"In the first experiment you probably have no idea," says Avi Feller, a Harvard graduate student and former Obama White House aide who has written about political experiments. "But by the 20<sup>th</sup> randomized trial you can start to say 'we've seen this group be more responsive.' You can start to do better than just wild guesses."

\*\*\*

The Analyst Institute was formed in 2007 to organize an expanding research portfolio produced by liberal consultants and institutions that were adopting techniques from medicine and the social sciences to better run campaigns. Many of the



Illustration by Robert Donnelly.

group's early experiments focused on voter turnout, often tracking the impact of motivational techniques that were informed by behavioral psychology. Experimenters would randomly assign voters to different get-out-the-vote treatments and measure after an election whether one group turned out at a higher rate than the other. This was relatively straightforward and inexpensive—whether someone votes can be tracked on publicly available electoral rolls—and required only a campaign or institution willing to hold out a control sample for tests. But even this was too demanding a burden for many political players: While institutions like the AFL-CIO and Rock the Vote signed up, candidates were typically unwilling to make such a commitment for research that wouldn't yield insights until after the election.

While turnout experiments were good for isolating whether an individual phone call, door knock, or piece of mail could mobilize citizens, they couldn't track how voters chose between

candidates. For that, campaigns continued to rely on many of the same techniques they had used for measuring public opinion for a half-century. They would look at the issues that self-described “undecided” voters said mattered most to them, or what those people saw as the strengths and weaknesses of each of the candidates. Researchers would typically gather small samples of those voters for focus groups, to get a feel for the language and images that they would respond to.

From those qualitative and quantitative sources, media strategists would develop specific messages. To test them, pollsters would present voters with an argument or piece of information (the deficit has increased under Obama, Romney put a dog on the roof of his car) and ask if it made them “more or less likely” to support the candidate. In some cases, pollsters would ask respondents how they planned to vote, then read them descriptions of the candidates and ask the vote-choice question again to see who moved.

The messages that were most persuasive in polls typically became the stuff of television spots, candidate speeches, online ads, direct-mail pieces, and robocalls. In the heat of a race, campaigns of any significant size would run tracking polls, which allow strategists to spot daily movement they could attribute to campaign activity. But the polls lacked the ability to account for cause and effect. Did the candidate’s numbers move because of her new TV ad about the economy or her new mail piece about abortion—or despite them both?

The Analyst Institute convinced many of the left’s leading institutions that randomized-control trials could be adapted to answer such questions empirically. In March 2008, after John McCain had become his party’s nominee and as Democrats still struggled to pick theirs, the AFL-CIO wanted to determine how to most effectively define the Republican in the eyes of its membership. Working with the Analyst Institute, the AFL’s political department developed three different direct-mail attacks on McCain. One highlighted the senator’s economic-policy agenda and one (called “McBush”) portrayed him as a clone of the unpopular incumbent. A third was presented a testimonial from an old white union electrician and navy veteran who conceded a McCain strength at the outset. “War hero? Absolutely,” the veteran says. “Voice for working families? No way.”

The AFL assigned Ohio union members to one of the three programs, and after mailing them conducted polling interviews with around 1,000 people in each group. Recipients of the “policy” and “McBush” mailers seemed unmoved by the messages they contained: around 38 percent of each universe supported McCain, almost indistinguishable from his support within a control group that had received no contact at all from the AFL. But the “testimonial” left its mark on the Republican candidate: only 32 percent of its recipients said they supported McCain, a drop of 5.6 percent against the control. The AFL made the testimonial a central part of its mail program nationwide.

But the AFL was measuring only the average impact of each message across the entire swath that received it. What if certain types of people were more likely to respond to specific messages than others? Elsewhere, political statisticians had succeeded in developing new methods of disaggregating the electorate so that campaigns could target individual voters

instead of entire precincts and media markets or broad demographic categories. Many settled on the statistical models known collectively as microtargeting: algorithms weighing as many as 1,000 different personal variables to generate probabilities predicting whether individual citizens would vote, whom they would support, and their views on specific issues.

But many attending the Analyst Institute's monthly lunch sessions were bothered by the fact that those models were still built on the spine of traditional polling, which relied on voters to describe how open-minded they would be to new arguments. They wondered whether it would be possible to fuse the real-world empiricism of experiments with the granular profiles made possible by microtargeting. What if campaigns tried out their messages on voters, then used their databases to identify the distinctive characteristics of the people whose minds changed?

\* \* \*

In the summer of 2010, the Democratic women's group EMILY's List was eager to help state treasurer Robin Carnahan in her run for an open Missouri senate seat, but wasn't sure what types of arguments they should make on her behalf. The group was interested in communicating with rural independent women in the state, but wasn't sure what it ought to say. Would it be more effective to present voters with hard-edged attacks on Carnahan's opponent, Congressman Roy Blunt, or a more balanced account of the candidates' contrasting positions?

EMILY's List had been one of the groups involved in the Analyst Institute's launch, and the two collaborated again on an experiment-informed program to refine its pro-Carnahan tactics. The design was straightforward: EMILY's List would have its consultant prepare two different direct-mail flights of four pieces each. One would be comparative ("Here's where Congressman Roy Blunt and Robin Carnahan stand on working families") and the other purely negative focusing on the Republican's known vulnerabilities ("Blunt has proven he's not on our side" and "Blunt doesn't know the difference between lobbying and legislating"). A sample of rural independent women voters would be randomly selected, and assigned to receive one of the two flights of mail.

On Aug. 31, after all four pieces of mail had been delivered, the Analyst Institute commissioned polling interviews with 5,912 voters in the state. Among those who had received the negative mail, 38.3 percent supported Carnahan—one point ahead of those who received the comparative message and three points ahead of a control group that received no mail at all from EMILY's List.

Experimenters then set out to identify the attributes that distinguished voters who had been moved by the negative message toward supporting Carnahan. Using census data, they learned that almost all the movement had come from voters in neighborhoods in the third socioeconomic quartile. Women living in precincts with an average annual household income between \$37,500 and \$45,000 had increased their support for Carnahan by over 10 percent, while the other three quartiles moved barely at all. Other predictive characteristics were

revealed as well: Those in areas with the densest populations moved most, as did those in the precincts with the highest concentration of single parents.

By Sept. 5, EMILY's List not only had the confidence to know that its negative mail would have more impact than the comparative material, but could begin trawling through a Missouri voter file to pluck the targets most likely to be persuaded by it: women in upper-middle class towns crowded with single parents.

\* \* \*

Four years ago, the Obama campaign used experimental methods to test nearly all of its online communications, randomizing the design of Web pages, the timing of text message blasts, and the language of email solicitations to measure their relative effectiveness. (Dan Siroker, who worked online analytics for Obama in 2008 and now counts the re-election campaign as a client of his company Optimizely, describes the process known as A/B testing [here](#).)

But that ethic never fully translated offline, where effects are much harder to measure than tallying clicks. During the summer of 2008, Obama advisers had casual interactions with Analyst Institute officials and ultimately integrated many of the group's best practices for get-out-the-vote tactics. The campaign briefly considered including an experimental component into its otherwise robust data efforts, but the compressed period between the primaries and the general election offered little time to upend a national communications strategy for the sake of testing.

This campaign is a different story. The experimental ethic was embraced by campaign leadership at the outset of the re-election effort. The formal arrangement with the Analyst Institute, which appears (according to federal filings) to cover a \$22,000 monthly retainer, marks the group's most significant engagement ever with a candidate's campaign. An institute analyst is now based at the Chicago headquarters.

The Obama campaign's long reach and big budget should significantly expand the frontiers of experimental politics, which have been limited by a tax code that prevents academic and nonprofit researchers from disseminating partisan messages. A presidential campaign faces no such restriction, and political operatives familiar with testing methods believe it should be possible to randomize Obama's messages not only by household (as in the EMILY's List test) but by larger political units—like media markets or cable systems—to track the effects of mass media. ([Rick Perry's 2006 gubernatorial re-election](#) randomized its broadcast buys over a three-week period, but the goal of the project was to test the impact of advertising at different levels and not the effectiveness of specific messages.)

Plenty of instinct and art remain in the Obama campaign's approach to message development. The early stages of the process resemble the traditional model, with media strategists relying on massive amounts of conventional polling from outside firms to track the electorate's mood and campaign dynamics, and on focus groups to add impressionistic

texture and a venue to audition specific images and language. The ads and direct-mail brochures that emerge from this process can then be assigned to small groups of voters under experimental conditions, pitted against one another in various combinations and across different audiences.

That full testing cycle can take around two weeks. In the case of mail, that includes the time it takes to design, print, and mail a piece—and a window for polling before and after to see what impact it had on opinions. Then analysts can model the attributes of those who were moved by the mail. Is an ad about the auto bailout more likely to persuade upscale or downscale voters? Did younger voters respond differently than older ones to information about particular provisions of the health-care bill? Are attacks on Romney's Bain record more salient with those leaning toward Obama or those leaning toward Romney?

Before making strategic adjustments based on the experimental findings, however, analysts have to consider whether the differences they find among voters really reflect the workings of the campaign's messages and not just statistical noise. "The key issue when dealing with subgroup analysis is it gets very easy to keep looking until you find something—what statisticians call 'data-dredging,' " says Feller. "I could go through each variable: Do women respond differently than men? Do 85-year-old people respond differently than 75-year-olds? Do cat owners respond differently than dog owners?"

There is still, then, room at Obama's Chicago headquarters for old-fashioned political intuition. What looks like a spring of experimentation will soon give way to a summer of analysis and strategic adjustments. Statisticians will find patterns, and political hands—relying in part on findings from other, more traditional methods—will discern whether those patterns can be exploited, and perhaps test them again. By fall, the hypotheses will outnumber the hunches.