

MARIA KONNIKOVA

HOW A GAY-MARRIAGE STUDY WENT WRONG



By Maria Konnikova

May 22, 2015



A study that had excited researchers and journalists alike now looks suspect. What happened?

Photograph by Andy Cross/The Denver Post via Getty

Last December, *Science* published a provocative paper about political persuasion. Persuasion is famously difficult: study after study—not to mention much of world history—has shown that, when it comes to controversial subjects, people rarely change their minds, especially if those subjects are important to them. You may think that you’ve made a convincing argument about gun control, but your crabby uncle isn’t likely to switch sides in the debate. Beliefs are sticky, and hardly any approach, no matter how logical it may be, can change that.

The *Science* study, “When contact changes minds: An experiment on transmission of support for gay equality,” seemed to offer a method that could work. The authors—Donald Green, a tenured professor of political science at Columbia University, and Michael LaCour, a graduate student in the poli-sci department at U.C.L.A.—enlisted a group of canvassers from the Los Angeles L.G.B.T. Center to venture into the L.A. neighborhoods where voters had supported Proposition 8, which banned same-sex marriage. The canvassers followed standardized scripts meant to convince those voters to change their minds through non-confrontational, one-on-one contact. Over the following nine months, the voters were surveyed at various intervals to see what those conversations had achieved. The survey highlighted a surprising distinction. When canvassers didn’t talk about their own sexual orientations, voters’ shifts in opinion were unlikely to last. But if canvassers were openly gay—if they talked about their sexual orientations with voters—the voters’ shifts in opinion were still in evidence in the survey nine months later. The messenger, it turned out, was just as important as the message. The study formed the basis for a segment of “This American Life,” and was featured on Science Friday and in the New York Times, the Wall Street Journal, and the Washington Post. LaCour was offered a job at Princeton.

Five months after the study’s original publication, on Tuesday, May 19th, a PDF was posted to the Web site of David Broockman, a graduate student at U.C. Berkeley at the time. (This summer, he will start working as a professor at Stanford’s Graduate School of Business.) In the document, “Irregularities in

LaCour (2014),” Broockman, along with a fellow graduate student, Joshua Kalla, and a professor at Yale, Peter Aronow, argued that the survey data in the study showed multiple statistical irregularities and was likely “not collected as described.” Almost simultaneously, Green sent a letter to *Science* asking to retract the paper: given the preponderance of evidence, he, too, believed that LaCour had fabricated the study’s results. His letter soon appeared on the site Retraction Watch, and, on Wednesday, *Science* posted an “Editorial expression of concern.” Many media outlets appended editors’ notes to their stories. A study that had excited researchers and journalists alike now looks suspect. What went wrong?

Green and LaCour first met in the summer of 2012, at the University of Michigan. Green was teaching a week-long summer workshop on experimental methods; LaCour was one of the students who signed up. “He was enterprising, he was energetic, [he was] technically able,” Green recalled, when we spoke yesterday. When LaCour proposed an idea for a field study about changing voters’ views on gay marriage through a contact-based approach, Green, who studies prejudice, was receptive.

Good, careful, and controlled field experiments are a relative rarity in political science; it’s difficult to devise and execute rigorous studies outside of the laboratory. But Green happened to know someone with the capability to execute the field intervention: his friend Dave Fleischer, who runs the Leadership Lab at the Los Angeles L.G.B.T. Center. Fleischer’s canvassers were out in the field—collectively, they had already conducted twelve thousand one-on-one conversations with voters—and their efforts seemed to be a good platform for the study. When Fleischer met LaCour, he was impressed. “He was a protégé of Don Green’s. I have enormous respect for Don. His integrity is unimpeachable. If he says someone is a good guy, it means a lot to me. And I respected Mike’s intelligence, his command of statistics, his interest—which was genuine—in what we were doing,” Fleischer said. He was also excited to receive an independent assessment of the program’s impact. “We were eager to measure how effective our work was in terms of magnitude and duration. We

were doing a conscientious self-assessment all along, but that's not a substitute for an honest independent assessment of your work.”

A year later, in the summer of 2013, LaCour told Green that the study they'd discussed back in Ann Arbor was complete: the results suggested that talking with openly gay canvassers could produce a durable attitude shift in favor of gay marriage. “I'm used to studying prejudice, teaching prejudice, thinking about prejudice, and the literature is just suffused with pessimism about any prospect of attitude change,” Green told me. “And here we have a study that shows it has profound effects to have contact with gays.” Green was skeptical, and told LaCour that he needed to replicate the findings by sending out a second wave of Fleischer's canvassers and surveying a second set of voters. LaCour reported back; it appeared, Green said, that “the magic happened again.” The data looked statistically solid, and the analyses seemed to back up LaCour's claims. The survey response rates looked abnormally high, but LaCour claimed to have raised hundreds of thousands of dollars in grant money to offer bonuses to people who responded to the survey; it was reasonable to think that the money was enough to account for their continued participation. Green volunteered to help LaCour draft the study. The next year, it was published.

David Broockman and Joshua Kalla, the Berkeley grad students, were impressed by LaCour and Green's findings. They decided to devote their own resources to pushing the research further. Broockman and Kalla prepared the online surveys, taking the initial steps towards a pilot study on May 6th. Nine days later, they noticed that their response rates were much lower than LaCour and Green's. Hoping to match the earlier study's rates, they looked for more information. They enlisted Peter Aronow, a professor at Yale, and the three began to examine the nuances of the data set. When they began to encounter a number of statistical quirks, Green contacted LaCour's dissertation adviser, Lynn Vavreck. On Monday, Vavreck met with LaCour to demand the raw survey data. After some delay, LaCour told her that he had accidentally deleted it. Later, when the head of U.C.L.A.'s political-science department called

Qualtrics, the online survey platform used for the study, they said that they could find no evidence of a deletion: whatever data was collected in the account LaCour claimed to have used was, presumably, still there. (LaCour was also unable to provide the contact information for any of the respondents.) At Green's behest, Vavreck had also looked further into the study's financing. It turned out to be nonexistent. "He didn't have any grants coming to him. He had a small one that he didn't accept," Green said. "There was no data, and no plausible way of getting the data."

On Tuesday, Green called LaCour. He presented him with the picture as he saw it, and told him it was time to come clean. "You haven't shown us any evidence. It's time to admit the data are fabricated," he said. According to Green, LaCour's response was prompt: "He said, 'No, the data are real.' He was willing to admit he couldn't find the primary data, but that was all." Green gave him until the evening to submit a retraction. When LaCour failed to do so, he went ahead and submitted his own. When I reached out to LaCour for comment, I received the same reply that he posted on his Web site: "I'm gathering evidence and relevant information so I can provide a single comprehensive response. I will do so at my earliest opportunity." Since then, he has updated the site with a new note, which says that he has sent a statement to Marcia McNutt, the editor-in-chief of *Science*, "providing information as to why I stand by the findings in LaCour & Green (2014). I've requested that if Science editor McNutt publishes Professor's Green's retraction request, she publish my statement with it."

If, in the end, the data do turn out to be fraudulent, does that say anything about social science as a whole? On some level, the case would be a statistical fluke. Despite what news headlines would have you believe, outright fraud is incredibly rare; almost no one commits it, and almost no one experiences it firsthand. As a result, innocence is presumed, and the mindset is one of trust. LaCour was an energetic graduate student with the resources to follow through on an ambitious idea; to Green, once the details had been hammered out, there didn't seem to be a need for further monitoring.

When fraud does occur, Green said, it's usually "on the intervention side, not the measurement side." The intervention—in this case, finding people to go from house to house and ask the right questions—is the hard part. People may lie about actually going into the field, or about the number of subjects they've talked to, or about their thoroughness during those interviews. It's especially upsetting to Green that the intervention happened exactly as it should have. "Dave [Fleischer] and his team really did a canvassing intervention where they talked to some people about same-sex marriage," he said. "It was a careful, placebo-controlled design." It's the surveys, apparently, that never took place. "There was an experiment, but the outcomes were never measured. That's the real tragedy here. This is a truly outstanding experimental design and an outstanding group of canvassers," Green said. "I keep getting e-mails from my colleagues saying, 'If we'd known the surveys weren't real, we would've jumped in and done them.' That's the easy part."

In retrospect, Green wishes he had asked for the raw data earlier. And yet, in collaborations in which data is collected at only one institution, it's not uncommon for the collecting site to anonymize and code it before sharing it. The anonymized data Green did see looked plausible and convincing. "He analyzed it, I analyzed it—I have the most ornate set of graphs and charts and every possible detail analyzed five different ways," Green said. Ultimately, though, the system takes for granted that no one would be so brazen as to create the actual raw data themselves.

There's another issue at play: the nature of belief. As I've written before, we are far quicker to believe things that mesh with our view of how life should be. Green is a firm supporter of gay marriage, and that may have made him especially pleased about the study. (Did it have a similar effect on liberally minded reviewers at *Science*? We know that studies confirming liberal thinking sometimes get a pass where ones challenging those ideas might get killed in review; the same effect may have made journalists more excited about covering the results.) Green says that the main factor in his enthusiasm for the study was its elegant design. "There's a literature on attitudes toward gay people and how

they change in the wake of contact, but it's not very good," he said. By contrast, this was "a beautiful study" because it took place outside of the lab, and because "it separated the canvassing from the measurement." Green was especially enthusiastic about the mode of interaction used by the canvassers, which other studies have also shown to be effective. "It's a very high-quality interaction. It's not confrontational," he said. "It's a respectful two-way conversation and the person expressing the view is doing it in a way that accentuates shared goals." Perhaps because of his enthusiasm, Green took what he saw as the best precaution—asking LaCour to replicate his results—but, he said, "didn't realize that strategy had one gaping flaw: the same guy generated the data both times."

In short, confirmation bias—which is especially powerful when we think about social issues—may have made the study's shakiness easier to overlook. But, perhaps ironically, it was enthusiasm about the study that led to its exposure. The events of the past few days were the result not of skepticism but of belief. Red flags were raised because David Broockman and Joshua Kalla liked the study and wanted to build on it. "One way to spin this story is [to talk about] scientific fraud and the field's susceptibility to it," Green said. "But if there's a silver lining here it's the robust self-correcting aspects of science. David and Josh couldn't redo the study in the same way, so they started to ask deeper and deeper questions about the first study. They are the heroes of this story." Broockman and Kalla are now working with Fleischer to give the story a happier ending: when we spoke, Fleischer said that the three of them were in the process of running a new version of the study.



Maria Konnikova is the author of "The Confidence Game" and "Mastermind: How to Think Like Sherlock Holmes."

More: [Gay Marriage](#) [Science](#) [Surveys](#)

THE DAILY

The best of *The New Yorker*, every day, in your in-box, plus occasional alerts when we publish major stories.

E-mail address

Your e-mail address

Sign up

By signing up, you agree to our [User Agreement](#) and [Privacy Policy & Cookie Statement](#).

Read More



DEPT. OF JUSTICE

TO HAVE AND TO HOLD

Reproduction, marriage, and the Constitution.

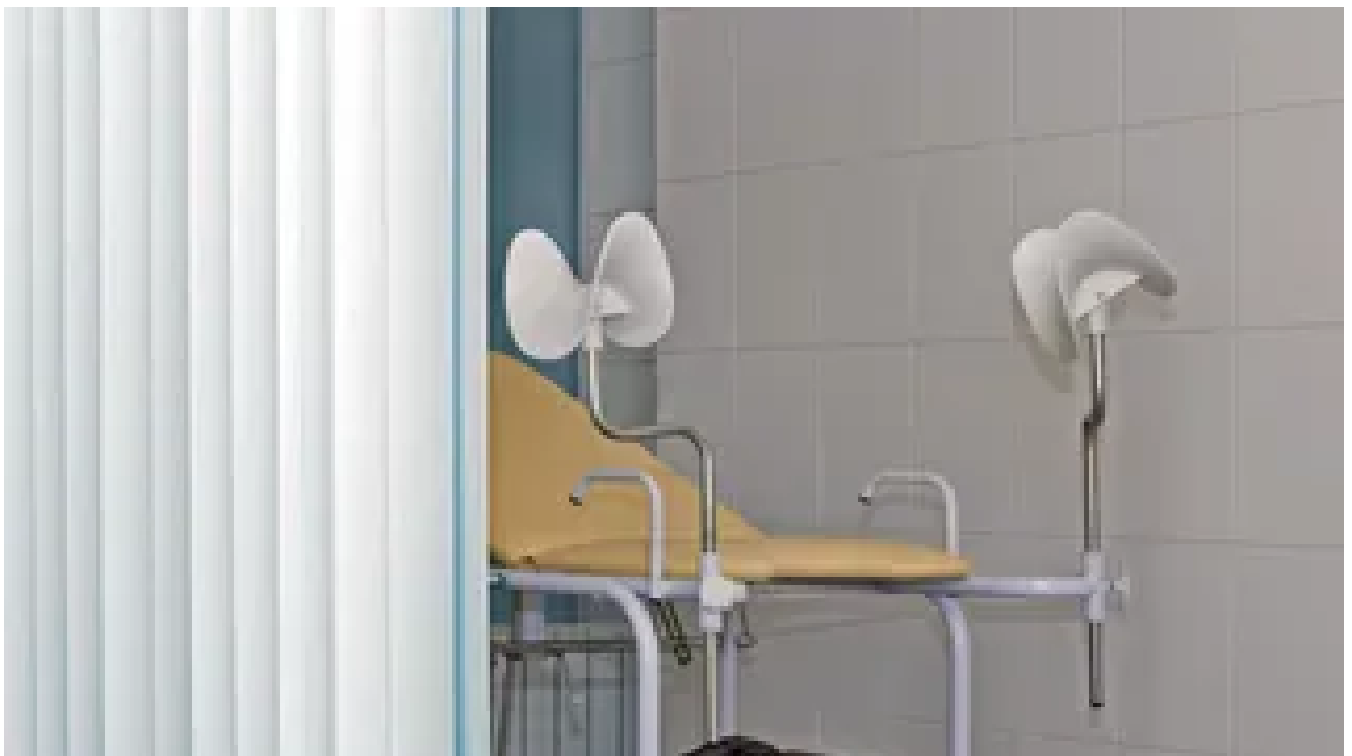
By Jill Lepore



MARIA KONNIKOVA

I DON'T WANT TO BE RIGHT

By Maria Konnikova



UNDER-REVIEW

THE STUDY THAT DEBUNKS MOST ANTI-ABORTION ARGUMENTS

For five years, a team of researchers asked women about their experience after having—or not having—an abortion. What do their answers tell us?

By Margaret Talbot

By Margaret Talbot



PROFILES

JUSTICE ALITO'S CRUSADE AGAINST A SECULAR AMERICA ISN'T OVER

He's had win after win—including overturning *Roe v. Wade*—yet seems more and more aggrieved. What drives his anger?

By Margaret Talbot

Manage Preferences