PROFESSION SYMPOSIUM

The Risk of Over-Reliance on the Institutional Review Board: An Approved Project Is Not Always an Ethical Project

Melissa R. Michelson, *Menlo College*

The Institutional Review Board (IRB) process is designed to protect the rights and welfare of human subjects, particularly vulnerable populations including undocumented immigrants and children. However, the process as currently designed and implemented can lead to approval of studies that have unethical and harmful outcomes or that violate the law, which with 20/20 hindsight never should have been conducted. In part, this is due to reliance on the IRB process as the sole check on whether a research project should proceed rather than also pausing to consider whether the research is ethical or might cause broader harm. For example, will it undermine the integrity of an election by changing the outcome, despite not causing harm to individual voters? This article explores this topic of the risk of over-reliance on IRB approval using a recent controversial get-out-the-vote (GOTV) experiment as a focal example, and suggests how to supplement the IRB process to better determine which experiments should be conducted and which should find their way to the circular file.

THE PURPOSE OF THE IRB

Government regulations regarding the acceptable conduct of research involving human subjects were developed in the 1960s in light of research that abused the rights of vulnerable populations (e.g., the Tuskegee experiments). In 1978, the *Belmont Report* (National Commission) set forth three ethical principles as guides: respect for persons, beneficence, and justice. Respect for the personal dignity and autonomy of individuals guides the requirements for informed consent. Beneficence requires the minimization of risk and engaging in risk–benefit analyses, including not only potential risks and benefits to individuals but also to society as a whole. Justice entails fair selection processes for individuals as well as for social, racial, sexual, and ethnic groups to protect vulnerable populations including undocumented immigrants, children, prisoners, and patients. Some language in the *Belmont Report* and other guidelines address broader social considerations; for example, the Nuremberg Code specifies that an experiment "should be such as to yield fruitful results for the good of society."

Robert J-P. Hauck testified in the 1990s before a National Science Foundation panel regarding the ways in which the IRB system "ill-fit and ill-served political science research" (Hauck 2008, 475). Various reforms followed, but Hauck argued in his opening essay to a 2008 symposium in *PS: Political Science & Politics* that IRBs continued to pose notable problems for the profession. Essays in that symposium included recommendations for improvement including (1) creating subunits staffed by faculty more familiar with the specific area of research and thus better able to assess risk; and (2) making the process more transparent by shifting the focus of IRBs to educating researchers (Levine and Skedsvold 2008).

These recommendations notwithstanding, little has changed since 2008. Meanwhile, in the past 15 years, there has been an explosion of interest in conducting field experiments. These projects typically are vetted by academic IRBs and pose no discernable harm to the individuals involved, either physically or psychologically. However, these experiments may cause unintended psychological distress to citizens, and political science field experiments may affect politics and society more broadly (e.g., by increasing the turnout of specific types of voters, thereby potentially changing the outcome of close elections). Despite guidelines mandating avoidance of harm to society overall, the IRB process as interpreted at many institutions is solidly for the protection of individuals rather than broader populations. This contributes to the inability of IRBs to sufficiently serve as a check on the ethics of political science experiments.

Increasing numbers of scholars have explored the ethical questions raised by this work, as they did at a 2013 two-day meeting in San Diego.[1] As the use of experiments continues to expand, we must continue the conversation.

POLITICAL SCIENCE IRB IN PRACTICE

I am not aware of anyone who enjoys or looks forward to the IRB process. It is perceived as a necessary but annoying hurdle between inspiration for an interesting project and putting it into practice. Citing "countless conversations with colleagues," Yanow and Schwartz-Shea claimed that many political scientists "actively decide to sidestep [IRBs] by not submitting their proposals for review" (Yanow and

doi:10.1017/S104909651600024X
© American Political Science Association, 2016

Schwartz-Shea 2008, 483). I have had many of these same conversations with colleagues who lament the degree to which IRB members do not understand their work, who see risks where none exist, or who delay or reject their research. These scholars do not want to cause harm; indeed, they believe their work is important and valuable, that it will lead to a healthier democracy or a better world. They see avoiding (or misleading) their IRB as a necessary evil—as a means justified by the end.

The IRB process can be time-consuming and frustrating. IRBs may seem to make unreasonable demands for revisions or to take too long to approve time-sensitive ideas. I personally have conducted more than 300 field experiments, some of which had to be changed or canceled because of IRB concerns. Recently, however, I have been wondering whether I am truly convinced that my research deserves approval or if I simply continue to pressure my IRB for approval because it has been resistant. In other words, does the focus on obtaining IRB approval blind me to possible flaws in my proposals?

Many scholars see IRB approval as a nuisance rather than as a useful check on the ethics of a proposed experiment. They look for ways to avoid the process or to cut corners to obtain the green light for moving forward; some, as noted by Yanow and Schwartz-Shea, simply find a rationale for not submitting their proposal for review at all. This fairly common attitude toward the IRB is disturbing because it suggests that research that perhaps should not be conducted is not being properly vetted. The degree to which this is true is difficult to estimate—scholars rarely boast publicly of their decision to avoid proper IRB procedures—but it is an understandable result of a process viewed by many political scientists as inadequate and unreasonable. This attitude logically leads researchers to view the IRB as the only necessary check on their research plans—or a check to be avoided—rather than a minimal bar to be surmounted.

This attitude logically leads researchers to view the IRB as the only necessary check on their research plans—or a check to be avoided—rather than a minimal bar to be surmounted.

THE MONTANA MAILERS

A powerful example of the result of scholars avoiding proper IRB review is that of the GOTV postcards sent to voters in Montana before the 2014 elections. The political scientists who designed the experiment were interested in knowing whether providing more information about candidates in certain types of races might increase voter turnout. That is, they theorized that registered voters who received the mailers would be more likely to vote and more likely to complete their ballot.

The mailers provided a visual illustration of an ideological spectrum (liberal to conservative), including the placement of President Barack Obama and former Republican presidential candidate Mitt Romney on that continuum. Also included were candidates in the targeted down-ballot election contest, based on "candidates' ideological positions based primarily on their fundraising activities" and using publicly disclosed campaign-finance records (more information about the ideological ratings is available at http://data.stanford.edu/dime).

Postcards were sent to voters in three states (California, New Hampshire, and Montana); however, by far, the majority of criticism aimed at the experiment related to the Montana mailers. Those mailers included ideological labeling of candidates for two nonpartisan elections to the Montana Supreme Court. One of those judicial races, between incumbent Justice Mike Wheat and challenger Lawrence VanDyke, was hotly contested. Third-party groups spent almost three quarters of a million dollars on the race—mostly from conservative groups including the Koch brothers' Americans for Prosperity—in favor of VanDyke.[2] Many observers, including this author, questioned the decision to insert ideological labels and partisanship into a nonpartisan judicial race (Michelson 2014). Although existing work belies the popular understanding of these races as truly nonpartisan (Bonneau and Cann 2013; Hall 2014), the blatant insertion of partisanship into the election was broadly criticized by residents of the Glacier State, who questioned the researchers' motives.

Gronke (2014) noted the possibility that the large number of mailers sent in such a small-scale election—that is, about 15% of the electorate, possibly half of the voting population—indicated that the experiment could have affected the outcome of the election. In the *Citizens United* era, many independent, dark-money groups engage in campaigning for and against candidates and ballot measures. Does that mean that political scientists also should engage in research that might change election outcomes? In the language of the IRB, perhaps the appropriate question is whether the benefit of what will be learned outweighs the risks. However, IRB procedures focus on possible harm to individuals, not election outcomes; it is unlikely that the researchers considered this potential risk.

A final concern is related to illegal aspects of the experiment. Montana Secretary of State Linda McCulloch charged the researchers with violating multiple state election laws because of the use of the state seal on the mailers and the wording, "2014 Montana General Election Voter Information Guide." A formal complaint was filed with the Commissioner on Political Practices, Jonathan Motl, on October 24, 2014, which detailed charges that the mailer had violated four separate sections of Montana law.[3]

The presidents of Stanford and Dartmouth quickly responded to the complaint and sent letters of apology to all postcard recipients, but the damage was already done: the

mailers were later found to have been sent in violation of election laws.

To summarize, the project raised concerns that outsiders were attempting to influence the election, that it might have altered the outcome of the election had it gone unnoticed until after Election Day, and that it broke laws. The project harmed the reputations of the involved scholars and their institutions, and it cost those institutions a considerable amount of trouble and expense. Motl's final report of May 11, 2015, left open the possibility of future fines or litigation.

This is where IRB procedures fall short as an ethical check, despite the lack of proper IRB review of this research. Albertson and Gadarian (2014) noted, "IRBs are not guardians of ethics; they are protectors of institutional legal liability." The moral obligation is on the researcher. We as individuals have a responsibility to be moral, ethical, and cautious, and to think through the consequences of our research. Is the knowledge that will be gained worth tinkering in a real-world election? Is there possible harm to the emotional stability of participants? Is there potential damage that will result from interfering in the political process? I doubt that the Stanford and Dartmouth researchers considered these questions.

However, IRB procedures focus on possible harm to individuals, not election outcomes; it is unlikely that this potential risk was considered by the researchers.

The fact that the Montana mailers were never submitted for IRB review, either at Stanford or at Dartmouth, may be interpreted as a strong argument for better control of researchers to ensure IRB involvement. Yet, as noted in Motl's final report, the Dartmouth IRB found the parallel experiment in New Hampshire to be exempt from IRB review, and it is plausible that it would have similarly exempted the Montana mailers.[4] Without involvement from local elected officials or nonpartisan community organizations familiar with Montana's political context and local election laws, it is unlikely that the experiment would have been blocked.

There were clear potential consequences that the researchers should have considered. Sending postcards to a large portion of the electorate in such a close race should have raised red flags about the possibility of altering the outcome of the election. Sending postcards with partisan information in a nonpartisan judicial race in a state that had recently endured dark-money scandals should have raised red flags about how the recipients of the mailers would react to outside money coming into their state. Sending postcards linked to elite out-of-state universities should have raised red flags about how the voters of Montana would react to being the guinea pigs in a research experiment.

None of these concerns was likely to be revealed through the IRB process. IRB members are not likely to be experts in electoral law, and it is unreasonable to assume that they would have noted the possible illegal aspects of the experimental proposal. They were also unlikely to know much about Montana politics or that the size of the experiment raised the possibility of changing the outcome of the election.

The Stanford and Dartmouth professors may have believed that what they were doing was harmless research and that no harm would result. I do not believe they acted maliciously. The project posed no harm to individual subjects and seemed likely to generate information that would contribute to a legitimate social good: broader voter turnout. The ability to send mailers to a large proportion of the electorate was seen as a benefit of locating the experiment in Montana because it would likely lead to smaller standard errors and statistically significant effects.

Criticism of the Montana mailers has been rapid and widespread, but it constitutes only one example. Other experiments—with proper IRB approval—also have raised concerns. A 2007 GOTV effort informing citizens that their name would appear in the newspaper after the election indicating whether they had voted generated sufficient concern from local election officials that the names were never published (Panagopoulos 2010). An experiment that colleagues and I conducted in 2010 to test the effect of providing postage-paid envelopes to a random sample of vote-by-mail voters (Michelson et al. 2012) generated calls of concern to the local registrar (with whom we were partnering). The local newspaper ran a front-page story that criticized the registrar for secretly experimenting on voters (Riddle 2010).

RECOMMENDATIONS

Our profession is on the frontier of a new direction in experimental research (if we have not already crossed it). The rapid growth of field experiments has outpaced considerations of how such research should be held to moral and ethical standards. Much like medical research had to be reined in after the excesses of the 1960s, experimental research also may need to be better controlled. Seeking IRB approval should not be equated with a decision that a proposed experiment is ethical. Instead, researchers should start by asking *themselves* whether they are behaving ethically. As discussed previously, this can be difficult; our enthusiasm for our own ideas can blind us to possible risk and harm.

Another route is for researchers to supplement their own internal check on the ethics of a GOTV experiment by involving partners. This may mean working with candidates, local election officials, or a local community organization. Involving actors who are not political scientists but rather real-world practitioners provides a different perspective and, usually, relevant expertise about legal constraints and possible unintended effects. Working with a partner in Montana, for example, might have helped the Stanford–Dartmouth team to better understand the local political context regarding dark money or led them to design postcards that did not break

local laws. This type of pracademic work has other advantages as well (Gillespie and Michelson 2011). Of course, partnering with candidates and elected officials also can introduce disadvantages and ethical concerns, particularly if those partners have something to gain from the research. When possible, scholars should seek to partner with independent local actors, such as nonprofit and nonpartisan organizations.

I personally have conducted hundreds of GOTV experiments, the majority of which were in cooperation with local officials, candidates, and community organizations. In 2013, I worked with my students at Menlo College to increase their participation in the November 2014 election, focusing on the local race for Atherton City Council. When voter registration closed, 288 people were registered at our campus address, about 6% of the town's 4,900-person electorate. The election was a close race among four candidates for three seats, and those votes potentially held the power of victory. In fact, the race was narrowly decided, with only 29 votes separating the third- and fourth-place candidates; the second-place candidate won a seat by only five more votes than the third-place candidate. Yet, I do not believe that we were unethically interfering. We conducted the work with IRB approval and with the full knowledge and cooperation of all four candidates, as well as with the local registrar of voters.

Other GOTV work that I conducted in 2014 was in cooperation with candidates. They chose the messages and conducted the outreach; my role was to evaluate the effect of their efforts. This work is more clearly pracademic, allowing for more robust external validity. It also provides an ethical check because the actual contact with voters was conducted by those with more experience about what that contact can and should include. The messages to targeted voters were delivered by the candidate's volunteers or via short recorded videos of the candidates.

In both examples, a key factor was including all of the relevant real-world candidates, either members from all of the local parties (in the latter example) or all of the candidates in the council election. Informing and involving all relevant actors allows for input from those directly affected by the research, mitigates the risk of introducing bias or unethically interfering in the election results, and increases the likelihood of concerns being raised about potentially unethical or illegal activities.

In fact, partnering may allow political scientists to conduct experiments that otherwise would be illegal. In 2009, my colleagues and I conducted a cold text-messaging GOTV experiment with registered voters in San Mateo County, partnering with the local registrar (Malhotra et al. 2011). If we had sent these unsolicited text messages on our own, they would have been illegal SPAM; partnering with the registrar meant that the messages were legal and could be sent without obtaining consent from voters.

Working with partners or local officials does not guarantee that experiments will be ethical; however, these supplements to the IRB process may ensure that enthusiasm for our ideas does not unwittingly lead us to cross ethical or legal boundaries. If possible, it is best to partner with a local organization or local candidate(s), especially when working remotely or in an area where the research team is less familiar with local history and political culture. Local officials should be notified about the project; for GOTV work, this might include the local registrar of voters or the secretary of state. Doing so may generate helpful feedback about legal requirements as well as unexpected assistance or opportunities. In this way, we can ensure that our research conforms to the three core principles of the *Belmont Report*: respect for persons, beneficence, and justice. ■

Seeking IRB approval should not be equated with a decision that a proposed experiment is ethical. Instead, researchers should start by asking themselves *whether they are behaving ethically.*

NOTES

1. Available at http://polisci2.ucsd.edu/polisciethics/schedule.html.
2. Available at www.nytimes.com/2014/11/03/us/montana-judicial-race-joins-big-money-fray.html.
3. Available at http://politicalpractices.mt.gov/content/2recentdecisions/McCullochvStanfordandDartmouthComplaint
4. *McCulloch* v. *Stanford and Dartmouth*, No. COPP 2014-CFP-046. Available at www.montana.gov.

REFERENCES

Albertson, Bethany, and Shana Gadarian. 2014. "Was the Facebook Emotion Experiment Unethical?" *The Washington Post*, July 1. Available at www.washingtonpost.com/blogs/monkey-cage/wp/2014/07/01/was-the-facebook-emotion-experiment-unethical.

The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research.* Washington, D.C.: Department of Health, Education, and Welfare. DHEW Publication OS 78-0012. 1978. Available at http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.html.

Bonneau, Chris W., and Damon M. Cann. 2013. "Party Identification and Vote Choice in Partisan and Nonpartisan Elections." *Political Behavior*, Online first. Available at http://dx.doi.org/10.1007/s11109-013-9260-2.

Gillespie, Andra, and Melissa R. Michelson. 2011. "Participant Observation and the Political Scientist: Possibilities, Priorities, and Practicalities." *PS: Political Science & Politics* 44 (2): 261–5.

Gronke, Paul. 2014. Available at http://blogs.reed.edu/earlyvoting/commentary/more-on-glaciergate-if-non-partisan-judicial-elections-are-not-a-democratic-good-can-we-conduct-research-intended-to-actively-undermine-them. Accessed October 27, 2014.

Hall, Melinda Gann. 2014. *Attacking Judges: How Campaign Advertising Influences State Supreme Court Elections.* Stanford, CA: Stanford University Press.

Hauck, Robert J. P. 2008. "Protecting Human Research Participants, IRBs, and Political Science Redux: Editor's Introduction." *PS: Political Science & Politics* 41 (3): 475–476.

Levine, Felice J., and Paula R. Skedsvold. 2008."Where the Rubber Meets the Road: Aligning IRBs and Research Practice." *PS: Political Science & Politics* 41 (3): 501–5.

Malhotra, Neil, Melissa R. Michelson, Todd Rogers, and Ali A. Valenzuela. 2011. "Cold Text Messages as Mobilization Tools: Implications for Theories of Political Participation." *American Politics Research* 39 (4): 664–81.

Michelson, Melissa R. 2014. "Messing with Montana: Get-Out-the-Vote Experiment Raises Ethics Questions." *The New West*, October 25, 2014. Available at https://thewpsa.wordpress.com/2014/10/25/messing-with-montana-get-out-the-vote-experiment-raises-ethics-questions.

Michelson, Melissa R., Neil Malhotra, Andrew Healy, Donald P. Green, Allison Carnegie, and Ali Adam Valenzuela. 2012. "The Effect of Prepaid Postage on Election Turnout: A Cautionary Tale for Election Administrators." *Election Law Journal* 11 (3): 279–90.

Panagopoulos, Costas. 2010. "Affect, Social Pressure and Prosocial Motivation: Field Experimental Evidence of the Mobilizing Effects of Pride, Shame, and Publicizing Voting Behavior." *Political Behavior* 32 (3): 369–86.

Riddle, Ryan Thomas. 2010. "Election Experiment Slammed." *Palo Alto Daily Post*, December 16, p. 1.

Yanow, Dvora, and Peregrine Schwartz-Shea. 2008. "Reforming Institutional Review Board Policy: Issues in Implementation and Field Research." *PS: Political Science & Politics* 41 (3): 483–94.