



Ethics in field experimentation: A call to establish new standards to protect the public from unwanted manipulation and real harms

Rose McDermott^a and Peter K. Hatemi^{b,c,1}

Edited by Margaret Levi, Stanford University, Stanford, CA, and approved October 1, 2020 (received for review June 10, 2020)

In 1966, Henry Beecher published his foundational paper “Ethics and Clinical Research,” bringing to light unethical experiments that were routinely being conducted by leading universities and government agencies. A common theme was the lack of voluntary consent. Research regulations surrounding laboratory experiments flourished after his work. More than half a century later, we seek to follow in his footsteps and identify a new domain of risk to the public: certain types of field experiments. The nature of experimental research has changed greatly since the Belmont Report. Due in part to technological advances including social media, experimenters now target and affect whole societies, releasing interventions into a living public, often without sufficient review or controls. A large number of social science field experiments do not reflect compliance with current ethical and legal requirements that govern research with human participants. Real-world interventions are being conducted without consent or notice to the public they affect. Follow-ups and debriefing are routinely not being undertaken with the populations that experimenters injure. Importantly, even when ethical research guidelines are followed, researchers are following principles developed for experiments in controlled settings, with little assessment or protection for the wider societies within which individuals are embedded. We strive to improve the ethics of future work by advocating the creation of new norms, illustrating classes of field experiments where scholars do not appear to have recognized the ways such research circumvents ethical standards by putting people, including those outside the manipulated group, into harm’s way.

ethics | field experiments | research

There has been a rapid and dangerous decline in adherence to the core foundations of ethical research on human participants when it comes to field experiments in the social, behavioral, and psychological sciences (1–7). For example, just looking at one discipline, a review of all articles published in the preeminent political science journals from 2013 to 2017 found that almost none of the field experiments in that period reflected compliance with the current ethical requirements that

govern research with human participants*; it is common knowledge that many field experiments are conducted without the consent, knowledge, or debriefing of participants (7, 8). Critically, even when researchers adhere to ethical guidelines, they are following principles developed for a different era, designed to protect and limit risk to individual participants in controlled laboratory settings. The basic principles of “respect for persons,” “justice,” and “beneficence,”[†] while clearly applicable to

^aPolitical Science, Brown University, Providence, RI 02912; ^bPolitical Science, Pennsylvania State University, University Park, PA 16802; and ^cMicrobiology and Biochemistry, Pennsylvania State University, University Park, PA 16802

Author contributions: R.M. and P.K.H. designed research, performed research, and wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Published under the PNAS license.

¹To whom correspondence may be addressed. Email: Phatemi@gmail.com.

First published November 23, 2020.

*R. McDermott, C. Crabtree, P. K. Hatemi, *Ethics and Methods*, May 4, 2018, London, England.

[†]The Nuremberg Code of 1947 (9) and the 1964 Declaration of Helsinki (10) adopted by the World Medical Association form a set of principles widely regarded as the cornerstone of ethical research on human subjects. They declared that participants must give informed consent, there must be a substantial scientific basis for the study, and experiments should yield findings that cannot be obtained any other way. High-profile cases of questionable research, including the Tuskegee Syphilis Study, Milgram’s Obedience Study, and the Stanford Prison Experiment, led to the National Commission for the Protection of Human Subjects and the Belmont Report (11), which codified a set of three basic principles to protect human participants: respect for persons, justice, and beneficence. In this article, we focus on the functional ability of these principles, not their underlying foundations. Nevertheless, we believe it important to recognize that these principles are grounded in basic principles of ethics and human rights, with a foundation in moral philosophy going back to Socrates. For instance, respect for persons rests on the principle of autonomy where people should not be treated as a means to a researcher’s end, but as individuals in their own right, and derives justifications from categorical imperative and moral law theories. Similarly, justice is grounded in contractualist perspectives, especially Rawlsian types, but is also informed by feminist care ethics types of approaches that focus attention on the experiences of individuals. Beneficence reflects an integration of virtue ethics and consequentialist perspectives and is designed to make sure that group benefits do not neglect individual welfare.

the design of field experiments, were not written to address the influence that large-scale manipulations can have on entire societies. The rise of large-scale real-world interventions raises new ethical dilemmas because experimenters now routinely target outcomes that affect whole societies, and often do so without the public's consent, knowledge, debriefing, or any means to identify or reverse long-term real-life negative effects. Manipulation effects routinely influence both the target population as well as the wider public who are equally likely to be harmed by an intervention, without their awareness. Indeed, as far as we could find, no work in any discipline has even attempted a review of the long-term effects of real-life manipulations from social science field experiments. Efforts to change the outcome of real elections, purposely stoke intergroup resentment and sectarian conflict, retraumatize people in conflict zones, and increase corruption represent only a few examples of recent real-world social science field experiments (12–17).[‡]

Such experiments have become mainstream, and pose a fundamental challenge to the ethical principles enshrined in the Declaration of Helsinki, the Belmont Report, and other established ethical guidelines (19). Remarkably little attention has been given to the societal harms that result from real-world manipulations. As such, there is an acute need for both enforcement of current ethical norms as well as updated ethical standards to protect individuals and populations from societal harms resulting from field interventions. Such an advance offers the additional benefit of working to protect the credibility and value of the scientific enterprise itself. It is also important to recognize the distinction between adhering to universal research standards and institutional ethics approval. While institutional review board (IRB) guidelines may increase ethical awareness, they exist to provide legal protection for institutions. As we illustrate below in a Cornell and Facebook experiment, IRBs may be necessary, but they are not sufficient to ensure the ethical treatment of subjects or the protection of broader communities because they do not primarily exist for the purpose of protecting participants, nor do they necessarily accomplish this goal (20). Additionally, there is great variation in IRB standards, both within and across institutions and countries. Therefore, we do not advocate for increased responsibility on the part of IRBs. Rather, we focus on universal ethical standards, with a goal of updating those standards to shape appropriate ethical principles for field experiments going forward.

Here we discuss only some of the kinds of risks and contraventions of established ethical guidelines resulting from large-scale real-world experiments. Our examples are not provided to render any judgment on intent. Rather, just the opposite: We assume that all of these cases did not intend to bypass ethical concerns. Science is an undertaking of learning and trial and error, but, often, as an enterprise, it forgets the lessons of the past. Mistakes, including those retroactively declared as such, are how we learn, and discussing new dilemmas openly and honestly is how science improves. In that spirit, we strongly admonish those who seek to blame, shame, or play "gotcha." We make these observations not to throw stones from afar, but rather in an attempt to aid from within, raise these concerns, and encourage a new consensus around the protection of populations during field

experiments. Indeed, we too have come to learn as we have made our own mistakes; those are, in part, what led us to raise these concerns more broadly. With deep humility and respect for all those seeking to conduct good research, we recognize the need for correction and hope to change minds for the future, not to place blame for past decisions or judge anyone's intentions. We strive to improve the ethics of future work by illustrating classes of field experiments where the broader academy does not appear to have fully recognized the ways such research circumvents ethical standards by putting people, including those outside the manipulated group, into harm's way. In this way, we identify new risks that require the creation of new norms explicated below.

New Risks from Large-Scale Social Science Research

In our discussion of field experiments that appear to violate principles of respect for persons, justice, and beneficence, as well as our introduction of novel concerns, we do not provide a systematic review of problematic studies, since no such analysis exists. Rather, we selected classes of experiments that: 1) appeared in high-impact top-tier field journals and interdisciplinary journals such as *PNAS*, *Science*, and *Nature*; 2) have been highly cited; 3) are common; 4) are carried out in conjunction with large state entities, governments, or corporations; 5) affected large populations; or 6) caused real public harms. These are not the only studies that demonstrate the concerns we raise, but instead represent classes of studies that set trends for work in the future or follow a problematic trend now. Indeed, the types or experiments we selected do not constitute outliers, nor are they extreme or rare. However, it is important to keep in mind that the studies we discuss here represent only a handful of examples from hundreds of such studies.

One might be concerned that the classes of experiments we discuss, or the cases where violations of ethical guidelines are apparent, are the result of cherry-picking. The classic example of cherry-picking would be if we were claiming the barrel of cherries were all bad, and then we picked out only the handful of bad cherries to make the case, but this is not what we are doing here. Rather, we are picking out the bad cherries to save the barrel, and we think this is a critical difference. Nevertheless, there are a lot of bad cherries that are easy to find. Discussing them openly allows us to identify the dangers that such systematic ethical disease presents. To be clear, we are describing the kind of problems that have arisen because many experiments in certain classes are increasingly being conducted without adherence to basic research ethics. As a result, new problems have arisen from technological advances not covered by current ethics. Our goal is to facilitate potential solutions going forward.

Social Pressure Manipulations. One of the most common contraventions of respect for persons and beneficence, including lack of informed consent and debriefing and disregard of the do-no-harm axiom, involve social pressure experiments. In seeking to identify what increases or depresses voter turnout, for example, scores of studies have undertaken large-scale interventions in real elections (21–25). Some explicitly state the intent is to change outcomes, generate feelings of group conflict, or pursue activist and partisan goals. These studies use a variety of tactics, including mailers, phone calls, and door-to-door visits, including from fake candidates. Several studies have targeted very large minority populations in such ventures, large enough to change electoral outcomes, by sending racially charged group-conflict messages and other anxiety-inducing stimuli. For example, one study targeted

[‡]Purposefully unethical behavior, including acts of libel, hate speech, fraud, or electoral violations [see Bonica, Rodden, and Dropp's (18) attempt to influence elections in Montana, for example], while important to reduce, is not the focus of this discussion.

approximately half of the registered Black voters in a southern state in the United States with a history of racial inequality and sent a quarter of this population a racially charged group-conflict message (15). This produced a reduction in minority turnout in a real election. Such manipulations are reminiscent of the types of action that led to the installation of voting rights laws requiring Supreme Court supervision of protected classes during the Civil Rights era.

Shifting the actual outcome of an election has real effects on local and national society. This alone should merit discussion on what experimenters can ethically do. However, more critical to our concerns is the public's welfare. Reducing a minority population's turnout and representation will have negative consequences for that community for years to come. Tens if not hundreds of thousands of people were manipulated. This outcome violates the principle of beneficence where "persons are treated in an ethical manner not only by respecting their decisions and protecting them from harm, but also by making efforts to secure their well-being." Two explicit rules serve to underscore beneficent actions: 1) "do no harm"; and 2) "maximize possible benefits and minimize possible harms." Given historical racial inequality in access to voting, generating feelings of interethnic hostility plausibly adds to the discord, disharmony, and racial tension in a population that has not yet recovered from past (and current) transgressions. Furthermore, even those experiments that claim to increase voter turnout disproportionately advantage those with the means to vote, and thus enhance negative societal effects by reducing the relative turnout for the least advantaged in society, specifically minorities, resulting in further electoral inequality (26).

None of the scores of studies we found in this class reported obtaining informed consent prior to the manipulation or debriefed the unknowing participants, letting them know that they had been manipulated. By not doing so, the experimenters did not follow the basic principle of respect for persons that requires researchers to: 1) inform participants of the potential risks related to their participation and 2) acquire permission before conducting research on anyone. Informed consent allows participants to avoid potential harms by opting out and is a "moral prerequisite" for any study to take place (9); it constitutes "the fundamental principle of human-subjects protection" (27) where "... a researcher is only (ethically and/or legally) justified in using a research subject if the research subject has consented to being so used" (28). This applies to all experiments, laboratory or field. Debriefing also constitutes an integral component of respect for persons that serves two critical functions: to remediate negative consequences and to revert participants to their prior state, including allowing people to return to how they felt about themselves and others before the study began. This process gives researchers a chance to correct unintended harms that may have accrued through participation and to inform participants about the purpose of the study, thereby removing any confusion the experiment might have caused (29).

Participants did not have a chance to opt out through informed consent or to return to their prior psychological state subsequent to the intervention through debriefing (i.e., some attempt to demobilize group hostility or remove harmful consequences). The principles of respect for persons and informed consent rest on expectations of individual autonomy. Such self-determination is fundamentally violated when field experiments manipulate people and elections without consent or debriefing. As a result, participants' access to an important public good (i.e., voting), critical for democratic governance, was influenced without their knowledge or approval. Indeed, if made aware of it, more than half the

subjects would likely oppose the research, much less consent to taking part in it (8, 30). Nor do such groups receive any benefit from the research. These two factors lie in perfect opposition to respect for persons and beneficence.

Other common manipulations explicitly "threatened" as many as tens of thousands of people with "exposure" to their peers and community if they did not vote (31), with the specific goal to shame the public or induce anxiety and negative emotions if they did not engage in the behaviors the researchers desired (32, 33). Yet we could find no social pressure study in this group that addressed the effects of surveillance manipulation on public health, particularly regarding the effects that social pressure can have on vulnerable individuals. Social threats, such as posting one's name or telling one's neighbors about their personal life, are likely to produce higher rates of mental health trauma, especially for people with anxiety disorders. According to the National Institutes of Mental Health, 18% of the US population suffers from some type of anxiety condition (34). People with anxiety and others often experience severe stress as a result of believing that some negative trait has been publicly exposed. This means that, for every 1,000 people pressured, the health of 180 of them is likely to have been negatively affected as a direct result of this manipulation. We do not claim that all such individuals will inevitably experience enough anxiety to put them at a health risk as a result of this manipulation (35). Yet, it remains important to consider the more than minimal risk to vulnerable populations prior to such manipulations; this does not appear to have happened, despite evidence from medical and public health studies that indicate threatening anxious people can precipitate symptoms. While some healthy people can become more resilient following major crises (36), the opposite tends to be true for highly anxious people. Such individuals may feel they cannot say no to social pressure manipulations because of fear of social stigma, and, as such, these people are not only denied the option of not participating, but they can also be pressured to act in a manner that the experimenter wants while still being more likely to suffer negative outcomes (37). Placing high-anxiety individuals under social pressure is equivalent to placing undue influence on at-risk populations, such as prisoners, children, or vulnerable others; even if unknowingly, it takes advantage of them. These experiments run counter to the principles of beneficence and justice that require fair and equal distribution of the risks and benefits of participation, including in the recruitment and selection of participants. Of critical importance, justice forbids exposing one group of people to risks solely for the benefit of another group.

As the number of unknowing participants increases, so too does the magnitude of unintended spillover. People are embedded in social networks and share their experiences, leading to a greater number of individuals affected, posing long-term consequences for the larger society, without any means for prevention or correction. In all of the field experiments of this class we could find, none conducted or reported assessments of informed consent or postexperimental checks on health or neighbor relations of the intended (and forced) participants or the wider affected communities (e.g., increased rates of suicide, hospitalization or other medical treatments, burden on friends and family). The guiding documents of ethical research tell us the public should not be manipulated without consent, debriefing (respect for persons), and a full understanding of potential health risks consequent to intervention (beneficence and justice), yet this guidance is not reflected in many published studies in the highest-impact journals.

Social Media. There is no more salient example of the influence field experiments can exert on the wider society than studies using social media platforms (6). Tens of millions of people in single events have been manipulated in academic experiments (12, 38). We have already learned in a short time the negative effects of such manipulations, including the ability of domestic and foreign powers to weaponize social media and manipulate democratic elections. Basic truths are now questioned, and trust in public institutions is at an all-time low. The ability to engage in micro-targeting, and the rapid way in which negative and hostile information, real or fake, is shared on social media, only serves to increase the potential danger of manipulating large groups of people without the ability to manage or understand the widespread effects that occur outside the investigator's control. Facebook and their academic colleagues' now-infamous experiments that manipulated the mood of hundreds of thousands of people by randomly pushing positive or negative posts to their feed investigated how human emotional states are transferred to others by contagion. The studies, however, did not consider all of the untold negative events that occurred from this manipulation. How many people were put over the edge, thrown into a bad mood, engaged in domestic violence, caused emotional distress to others, or lost their jobs due to the manipulation? Such follow-up was never undertaken.

A recent emotional-contagion study (39) conducted on hundreds of thousands of people by researchers at Cornell University simply did not obtain any ethics approval (3). Cornell's IRB decided that the study did not need approval because the data had been collected by Facebook. According to the defenders of these studies, users consent to this kind of manipulation when they agree to a company's terms of service. This is factually untrue. Terms of service for social media platforms do not meet the standards of informed consent for ethical research; rather, they are designed for purposes of civil liability. Others argue that such manipulations reflect nothing more than what people encounter every day (40, 41). First, this is a common misinterpretation and misapplication of Common Rule 45CFR 46. This rule defines minimal risk as "the probability and magnitude of harm or discomfort anticipated in the research are not greater in and of themselves than those ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests." The class of experiments clearly rise above minimal risk, however, and thus require review. Indeed, a consensus has formed that this does not mean it is relative to the population under study (42). Rather, for example, experimenters cannot put people who are normally at risk for death or corruption in a situation where those might occur. Second, and critically important, even if minimal risk is determined, it does not obviate the ethical requirements for consent. Third, it is questionable whether researchers have a right to influence someone's mood for their own self-interest. Finally, claims that "things like this happen every day" must be taken to their logical conclusion. Rape happens every day; racism happens every day; sexism happens every day; homophobia happens every day. Frequency does not provide ethical justification. If we follow the "every day" argument, this means that researchers have a right to conduct studies that launch racist profanity at others, that inspire sexist behavior, that create homophobic fear, undermine public trust, and delegitimize science. Terrible things do happen every day and people endure them, but to argue that the public must endure such things in the name of social science, without their consent, especially when such studies have yet to prove any tangible benefit

to the manipulated public, is not an ethically defensible position. Clearly, it is beyond the purview of academics to try to regulate social media platforms; however, this does not abnegate scholars of the responsibility to establish and police ethical principles for our own work. Indeed, no serious scholar should ever look to Facebook or Twitter for the ethical standard by which to guide their field research. Cambridge Analytica provides all of the evidence we need to demonstrate the folly of such an undertaking.

Resource Allocation. Scholars increasingly partner with international organizations, governments, and others to examine the effects of various processes on outcomes such as electoral accountability or support for the government. These efforts are almost always proclaimed to be designed for the public good, but they often produce negative side effects. In one study, half of the "subjects"—people who were behind on rent and in danger of eviction—were denied monetary assistance for more than 1 y so researchers could determine who ended up homeless (43). Perhaps the most poignant example is a class of studies where scholars worked with activists and nongovernmental organizations (NGOs) to empower women in developing countries through microfinance or direct cash infusion. Women benefitted in numerous ways; however, domestic violence against the women also increased substantially, as they were seen to violate prevailing norms of patriarchal rule (44, 45). They also upset the patriarchal hierarchy that existed among the women, fracturing support systems in the future when the money ran out. These consequences go unaddressed, yet ethically should be addressed before any intervention takes place through thorough, context-specific assessments drawn from observational and qualitative research. Insufficient thought and attention to negative downstream consequences appears common in the design of the intervention in these types of studies where field experimenters do not engage the population to anticipate what effects their "good deeds" might have.

There are at least four sets of interrelated problems that emerge from these designs. First, when such experiments offer rewards that far exceed average monthly incomes, the design is coercive, since individuals do not have true freedom to refuse such a large influx of cash. This violates the principle of respect for persons. Second, giving life-altering benefits to some people and not to others, no matter how random the assignment, can often result in resentment and anger in the larger community toward those who do receive the benefit, including the stimulation of tribal warfare in developing countries. As any learned scientist knows, relative gains matter, and such effects can and do exacerbate inter- and intragroup conflict. Imposed inequality can and does have negative consequences, particularly when investigators are unaware of the history of tribal rivalries and familial hierarchies that their interventions exacerbate. Third, as some receive benefits and others do not, imbalances and inequity often wreak mayhem on social networks, families, and communities. These latter two violate the principles of justice and beneficence. Finally, when researchers partner with governments, NGOs, and other organizations, they are compromised. No matter how well-intentioned, the design and execution of research is influenced by the goals and resources provided those organizations, who are not bound by the same standards of professional research ethics. Scholars cannot rely on the ethical requirements of such organizations any more than they can on the regulations of social media platforms.

Conflict Generation. Some studies have stoked sectarian fighting, others have encouraged protest at the risk of jail and other real harms to the public (17), and still others increase ideological, ethnic, and racial polarization (14). Some studies have presented subjects with videos or other media of actual violence being perpetrated against members of their in-group in order to investigate the effect on group identity and behavior toward members of the perpetrating out-group (16). It was no surprise that exposure to violent repression pushed subjects toward stronger in-group identification and out-group hostility. None of these studies reported any consideration of the downstream consequences for the larger society affected by these studies or the health of those exposed to such videos in these vulnerable populations. None of those studies reported a follow-up or plan to monitor whether or not the study itself generated prolonged hatred or engagement in subsequent violent retaliation generated by what they observed. Such effects are likely and could last for years, especially in the absence of debriefing. Rarely, if ever, do such studies report clinical professionals on staff to address these risks. Such unnecessary and disturbing exposure challenges the principles of respect for persons, justice, and beneficence.

Experiments that manipulate and change larger societies without consent, controls, proper testing, debriefing, and dialogue with the population are unethical regardless of whether the motivation, intent, or result is “good” or “bad.” Studies that seek to justify the means by the ends for the good of society ignore that their good is often very different from the subjects’ definition of good, and one researcher’s good can constitute another person’s notion of evil. This is particularly true for moral, political, religious, cultural, and social beliefs, where ideas on what is right, just, fair, or positive can be highly contentious and dependent on individual and local norms and culture.

Corruption. Another increasingly common domain of field experiments involves corruption. The argument for such experiments is obvious; it is very difficult to study dishonest behavior openly (7). However, these studies pose significant risks to individuals who are peripheral to the subjects. This class of experiments is often conducted in underdeveloped countries. For example, one group of experimenters, in attempting to understand and reduce government officials’ demands for bribes, raised salaries in Ghana, believing higher incomes would reduce corruption. The intervention actually increased police demands for bribes and the amounts given by truck drivers to the police (46). The public, while not the target, was and will be negatively affected for the foreseeable future. Other studies involved creating false businesses and agencies. In a region where government corruption is high, the effects of these interactions reduced public trust in societies where trust in institutions is already low, but necessary in order to maintain stable governments and societies. In these and other studies, the principles of respect for persons and beneficence are violated as individual subjects’ and society’s welfare become superseded by investigator interest. Equally important, the effects on society from contagion cannot be controlled. Similar concerns apply to many other types of studies where international organizations and scholars seek to impose their own personal value-laden outcomes, all the while ignoring the negative societal effects on the affected population.

Life-Course Manipulation. If there is a culmination of all of the preceding classes of field experiments surrounding what happens when experimenters alter the lives of subjects without their

knowledge, consent, or debriefing, or without adherence to principles of respect for persons, justice, or beneficence, it is the class of studies that purposely seek to change life path. Some appear innocuous at first, such as researchers and OKCupid using dating sites to create intentional mismatches to see what happens in mating behavior (5). But imagine finding out years later that your relationship was based on a “lie.” How might that disrupt a life? Other examples, reminiscent of Watson’s “Little Albert,” are far more sinister. The case of the “three identical strangers” (Yale University) separated at least five sets of twins and triplets at birth, purposely placing children into families with different socioeconomic status and other characteristics to see what would happen to their lives. For years, researchers conducted home visits while lying to the families, stating that they were part of routine monitoring after adoption (4). The families were never told that their child had siblings. The study specifically targeted the vulnerable biological parents who could not take care of their children, children who needed to be adopted, and parents who deeply desired a child. Yale sealed all details until 2065, when the likelihood of all injured parties being dead or unable to recover damages is high, while the probability of finding surviving biological family members is low. None of the “subjects” provided informed consent. This is neither an extreme example nor an outlier. In fact, until very recently, researchers at Yale were still following up on the siblings. The fact that such deception is ongoing demonstrates that this is the type of study we invite if we rest our arguments on researchers as activists, putting faith in their own beliefs about what is good for others, as opposed to allowing people to choose for themselves if they want to be part of an experiment or not. A natural question is “could this happen again?” Sadly, we believe so. Even if Yale was to change their approach, many institutions, including those in wealthy advanced democracies, either do not require ethics approval for social science research or simply do not have an IRB at the institutional level. More than IRB adherence is needed to protect the public from such experiments in the future.

Only a Small Sample of the Ongoing Harms. All of the above examples are real studies that have been conducted. These examples illustrate only a tiny portion of the classes of already-realized harms that have and will continue to result from large-scale real-world public manipulations without proper adherence to appropriate ethical standards. The lack of adherence to the basic principles of respect for persons, beneficence, and justice are particularly problematic when vulnerable populations, who are at the highest risk for serious psychological, economic, physical, sexual, and social harms, are involved.

New Requirements and a New Standard: Respect for Societies

We join a growing list of scholars across disciplines who argue we must respect the voice of the public who are asking not to be experimented on without consent (30, 47, 48). To promote that end, we strongly encourage greater interdisciplinary teaching of research ethics at the graduate level across the social sciences, including principles applied to both laboratory and field settings. In addition, we suggest that it is time for academic professional associations, journals, and institutions to update our policies to adhere to existing ethical norms and formulate new requirements to address potential harms raised by large-scale field experiments that impact entire populations, especially in light of technological advances. If we are to believe authors’ ostensible claims about the

significant nature of their results, then we have all the evidence we need to know that these studies affect individuals and societies in negative ways, without their knowledge, consent, or the possibility for remediation of effects. It is not logically possible, nor ethically defensible, to have it both ways. Manipulations have effects, and the consequences of such widespread effects must be properly considered by the researcher at the design stage and evaluated again at the publication stage. If ethical procedures are not followed, or unnecessary harms occur, then publication should not follow.

Field experiments are a powerful tool, and some argue that they should not be subject to the same minimal ethical standards as other research (41, 49). This is akin to arguing that pistols should have a 10-round limit on magazine capacity but assault weapons can have unlimited rounds. Both the realized and potential risks to larger populations from field experiments are far greater than those in social science lab experiments. To argue that, since they pose higher risk, they should be subject to lower ethical thresholds is not reasonable or scientific. There is an overwhelming consensus in the scholarly world that the benefits of research must outweigh the costs.

We further advocate extending ethical guidelines to include an additional standard providing for societies. This proposed fourth basic principle, respect for societies, requires addressing the potential effects manipulations can have on both local and large-scale societal outcomes. This consideration represents more than an aggregation of individual rights. Ethics designed for the protection of individuals are not designed to protect groups or to address the effects of manipulating entire communities and social structures. When manipulations are conducted in a living society, effects are unpredictable and influence more than the target population through contagion. When individuals have not been given the opportunity to consent, or are not in the group under direct manipulation, researchers are still ethically obligated to respect their rights and welfare.

We are not arguing that field experiments should be abandoned; just the opposite. Indeed, we recognize their value, and therefore wish to highlight the need for more responsible and stringent adherence to ethical guidelines designed for their particular effects and challenges. We encourage a more robust debate about how best to accomplish this goal. There is great value in understanding how small-scale processes can affect large-scale outcomes through real-world investigation, and no other method can outperform field experiments for external validity. However, they must be justified first, and at the very least adhere to the same minimal principles required of all other forms of experimental research. From an ethical standpoint (see the Nuremberg Code), an experiment should not be conducted if there are more appropriate ways to explore existing phenomena than to create real-life situations that harm actual populations. Simply put, there is no need to run an experiment on millions of people when a sample size of a 1,000 will provide all of the power needed for a meaningful effect size.

In cases where the larger society might be affected through large-scale intervention and experimentation, additional protections for the wider publics should be included. Manipulating real public outcomes should not occur without broader public discussion, debate, approval, and sanction. Indeed, in all other types of studies or efforts that affect the greater public or living societies, there are strict procedures. For example, the Food and Drug Administration has well-established guidelines for releasing interventions into the public that include at least four phases over

the course of years, where small-scale controlled studies are scaled up gradually until the intervention is deemed safe to release into society. Such standards provide a valuable template for experimenters who seek to manipulate a large-scale living society. Just as no ethical researcher would release a new medication without major testing, even for the social good of eliminating horrible diseases or viruses, similarly impactful social manipulations should undergo equivalent scrutiny.

Yet, when it comes to social science field experiments, we have somehow entered into a Wild West where anything goes when it takes place in the public sphere in large populations, while small controlled laboratory experiments must follow established guidelines. Thus, we suggest that, for any field experiment on a real-life society, the relevant publics should be consented and debriefed. Otherwise, scholars will be engaging in public manipulation without public protection of a kind that, if it were conducted by a foreign government, would be considered a violation of international law, if not an act of war. To those who advocate that such consent is not possible, we argue, if one can manipulate millions, then one can consent millions. Even if individual consent is difficult, technology allows for a number of ways to inform the public that a large-scale experiment is about to be released. Local, state, and federal governments do this when making public service announcements, including notifications regarding road closures, risk of fire, and Amber alerts. Radio, Internet, billboards, phone notifications, and television do work. Giving the public the ability to be aware of, and potentially find a way to opt out of, a manipulation and the means to report negative side effects is enshrined in the Nuremberg Code, the Declaration of Helsinki, and the Belmont Report. Simply offering public notice is a low-threshold means for acquiring at least passive consent. In the face of such warning, large-scale public outcry might warn researchers that their design may pose serious risks to the welfare of the wider population and should be abandoned.

Some argue such processes are too onerous, too costly, or too time-consuming (1, 41). Manipulating people's real lives and changing outcomes in a real society, according to the basic standards of human protections, should be onerous and meticulous and take time. It should require substantial public debate and scrutiny. That is the purpose of individual and public protections. If we abandon such principles, how far should academic investigation be able to go? Should scholars be allowed to start a riot to see how violence spreads? This appears to be close to the case in Hong Kong (17). Or should they be allowed to place transgender people at risk to see how the public engages with them, as was done recently in the United States (10)?

What matters is the standards we adopt, not simply the effects of any given study. Otherwise, we are placing our own welfare over that of the subjects and populations we purport to be investigating and often claim to be helping. If we advocate for unlimited and unlicensed real-world manipulations, we open a door that is not controllable, where there is little ability or avenue for people to recover and return to their original state and no ability to stop unintended spillover effects, which an investigator may not be able to anticipate or recognize in advance.

It is critically important to recognize the inherent conflict of interest in creating a real-world outcome, analyzing the results, benefiting from the findings, and then serving as judge and jury on the social value of such studies. Intentions to manipulate the public for the sake of changing a societal outcome may be the enterprise of private corporations, campaigns, foreign governments, and other entities, but manipulating a person's real life or

an entire living society without consent or notification, or proper preliminary testing, ethically cannot, and should not, be the goal of legitimate scholarly research. Scholarly research is intended to understand, not change, public outcomes. Activism is a personal choice, not a scholarly one. One can be both, but a declaration that the scholar seeks to change opinion in a given study must be declared and be part of the approval process for the study, as well as any publication, particularly if funding comes from an interested source. This important distinction is what lends credibility to academic research and heightens legitimacy, and any study that could negatively affect the public's health or incite violence and long-term discord should be regarded with increased scrutiny.

How can we institute these changes? The most proven avenue is through enhanced training and education, journals, external funding, and professional associations, which set the guidelines for each field and offer a potent mechanism to institutionalize norms and provide oversight of ethical issues. Once high-impact journals and funding opportunities require adherence to particular ethical standards, research incentives shift quickly, as has been the case with data transparency and replication. Such policies might include mandated ethics statements as part of the submission process, which is common in many psychology and health-related outlets. Changes in training may also help institutionalize the protections we advocate. Most researchers at US institutions typically only complete an online training module; this is no substitute for the kind of extensive, discipline-wide, consensual education that can take place through mentorship, apprenticeship, and coursework.

Conclusion

We hope to inspire greater discussion, debate, and the eventual emergence of a consensus around appropriate policy to address ethical concerns for wider public welfare in field experiments. What constitutes a well-designed study and what constitutes an ethical one can be contiguous or mutually contradictory, and serious thought must be given to the relevant trade-offs. Participant and societal protection, however, should never be sacrificed

solely to advance individual research interests or professional success. When we began this work and circulated our first paper on this topic in 2013, the response was some curiosity and confusion, but often hostility. Most field experimenters appeared unaware of the ethical issues, and we were even told that field experiments were exempt from consent (we have yet to find this blanket exemption). A few years later, especially in the wake of the upheavals over the Montana, Facebook, and other experiments, a wave of recognition identifies that serious problems often result from widespread social interventions. The public has made clear they consider this to be a problem. High-profile news articles, public debate, and admonishment of experimenters, including by legislators, alongside a social-media firestorm, has provided ample evidence that the public does not want to be experimented upon without their consent. We see this article as one step forward in an attempt to address these concerns and explore ways to improve the ethical consensus surrounding field experiments. We also suggest that all types of research will benefit from more self-conscious ethical review. Ultimately, the welfare of participants and the public depends on knowledgeable, caring, and responsible investigators who place participant well-being and the public welfare ahead of all other aspects of the research enterprise.

Data Availability. There are no data underlying this work.

Acknowledgments

We thank the American Academy of Arts and Sciences, especially President David Oxtoby, for their sponsorship of a meeting on ethics in field experiments held in Cambridge, MA, in November 2019; we also thank all participants. We thank the committee on revising Human Subjects Guidelines for the American Political Science Association; we are grateful to chair Scott Desposato for additional comments. We thank Margaret Levi and the Stanford Center for Advanced Study in the Behavioral Sciences for sponsoring a meeting on ethics in March 2018. We thank Ann Arvin and other participants at that meeting for helpful comments. We also thank participants at a meeting on ethics and methods at London School of Economics organized by Denisa Kostovicova and Ellie Knott; Dara Kay Cohen offered helpful additional comments after this meeting.

- 1 D. L. Teele, *Field Experiments and Their Critics: Essays on the Uses and Abuses of Experimentation in the Social Sciences* (Yale University Press, 2014).
- 2 D. Baldassarri, M. Abascal, Field experiments across the social sciences. *Annu. Rev. Sociol.* **43**, 41–73 (2017).
- 3 R. Meyer, Everything we know about Facebook's secret mood manipulation experiment. *The Atlantic*, 28 June 2014. <https://www.theatlantic.com/technology/archive/2014/06/everything-we-know-about-facebooks-secret-mood-manipulation-experiment/373648/>. Accessed 9 November 2020.
- 4 B. Lerner, Three identical strangers: The high cost of experimentation without ethics. *Washington Post*, 27 January 2019. <https://www.washingtonpost.com/outlook/2019/01/27/three-identical-strangers-high-cost-experimentation-without-ethics/>. Accessed 9 November 2020.
- 5 M. Wood, OKCupid plays with love in user experiments. *NY Times*, 28 July 2014. <https://www.nytimes.com/2014/07/29/technology/okcupid-publishes-findings-of-user-experiments.html>. Accessed 9 November 2020.
- 6 J. Jouhki, E. Lauk, M. Penttinen, N. Sormanen, T. Uskali, Facebook's emotional contagion experiment as a challenge to research ethics. *Media Commun.* **4**, 75–85 (2016).
- 7 G. H. McClendon, Ethics of using public officials as field experiment subjects. *Newslett. APSA Exp. Sect.* **3**, 13–20 (2012).
- 8 K. Peyton, Ethics and politics in field experiments. *Newslett. APSA Exp. Sect.* **3**, 13–20 (2012).
- 9 Nuremberg Military Tribunals, "Permissible medical experiments" in *Trials of War Criminals Before the Nuremberg Military Tribunals CCLN, October 1946-1949* (US Government Printing Office, Washington, DC, 1949), pp. 181–184.
- 10 World Medical Association, "Recommendations guiding medical doctors in biomedical research involving human subjects: Revised 52nd WMA, 2000" (WMA General Assembly, Edinburgh, Scotland, 1964).
- 11 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" (Department of Health, Education and Welfare, Bethesda, MD, 1978).
- 12 C. A. Bail et al., Exposure to opposing views on social media can increase political polarization. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 9216–9221 (2018).
- 13 D. Broockman, J. Kalla, Durably reducing transphobia: A field experiment on door-to-door canvassing. *Science* **352**, 220–224 (2016).
- 14 J. Barceló, Mobilizing collective memory: A large-scale field experiment on the effects of priming collective threat on voting behavior, SPSA 2019 Preliminary Program, version 3.0 (2018).
- 15 D. W. Nickerson, I. K. White, *The Effect of Priming Racial In-Group Norms of Participation and Racial Group Conflict on Black Voter Turnout: A Field Experiment* (Ohio State University, 2013).
- 16 G. Nair, N. Sambanis, Violence exposure and ethnic identification: Evidence from Kashmir. *Int. Organ.* **73**, 329–363 (2019).
- 17 L. Bursztyjn, D. Cantoni, D.-Y. Yang, N. Yuchtman, Y. J. Zhang, Persistent political engagement: Social interactions and the dynamics of protest movements. https://home.uchicago.edu/bursztyjn/Persistent_Political_Engagement_July2019.pdf. Accessed 9 November 2020.

- 18 D. Willis, Professors' research project stirs political outrage in Montana. *NY Times*, 28 October 2014. <https://www.nytimes.com/2014/10/29/upshot/professors-research-project-stirs-political-outrage-in-montana.html>. Accessed 9 November 2020.
- 19 J. Kleinsman, S. Buckley, Facebook study: A little bit unethical but worth it? *J. Bioeth. Inq.* **12**, 179–182 (2015).
- 20 M. R. Michelson, The risk of over-reliance on the institutional review board: An approved subject is not always an ethical project. *PS Polit. Sci. Polit.* **55**, 299–303 (2016).
- 21 A. Chong, O. A. L. De La, D. Karlan, L. Wantchekon, Does corruption information inspire the fight or quash the hope? A field experiment in Mexico on voter turnout, choice, and party identification. *J. Polit.* **77**, 55–71 (2014).
- 22 A. S. Gerber, D. P. Green, Do phone calls increase voter turnout?: A field experiment. *Public Opin. Q.* **65**, 75–85 (2001).
- 23 A. S. Gerber, D. P. Green, The effects of canvassing, telephone calls, and direct mail on voter turnout: A field experiment. *Am. Polit. Sci. Rev.* **94**, 653–663 (2000).
- 24 D. P. Green, A. S. Gerber, D. W. Nickerson, Getting out the vote in local elections: Results from six door-to-door canvassing experiments. *J. Polit.* **65**, 1083–1096 (2003).
- 25 C. Panagopoulos, Affect, social pressure and prosocial motivation: Field experimental evidence of the mobilizing effects of pride, shame and publicizing voting behavior. *Polit. Behav.* **32**, 369–386 (2010).
- 26 R. D. Enos, A. Fowler, L. Vavreck, Increasing inequality: The effect of GOTV mobilization on the composition of the electorate. *J. Polit.* **76**, 273–288 (2013).
- 27 L. D. Walker, National science foundation, institutional review boards, and political and social science. *PS Polit. Sci. Polit.* **49**, 309–312 (2016).
- 28 S. Holm, S. Madsen, "Informed consent in medical research—a procedure stretched beyond breaking point?" in *The Limits of Consent: A Socio-Legal Approach to Human Subject Research in Medicine*, K. Dams-O'Connor, J. M. Ketchum, J. P. Cuthbert, Eds. (Oxford University Press, Oxford, UK, 2009), pp. 11–24.
- 29 American Psychological Association, *Ethical Principles in the Conduct of Research with Human Participants* (American Psychological Association, Washington, D.C., 1973).
- 30 S. Desposato, Subjects and scholars' views on the ethics of political science field experiments. *Perspect. Polit.* **16**, 739–750 (2018).
- 31 A. S. Gerber, D. P. Green, C. W. Larimer, Social pressure and voter turnout: Evidence from a large-scale field experiment. *Am. Polit. Sci. Rev.* **102**, 33–48 (2008).
- 32 A. S. Gerber, G. A. Huber, A. H. Fang, A. Gooch, The generalizability of social pressure effects on turnout across high-salience electoral contexts: Field experimental evidence from 1.96 million citizens in 17 states. *Am. Polit. Res.* **45**, 533–559 (2017).
- 33 A. S. Gerber, D. P. Green, C. W. Larimer, An experiment testing the relative effectiveness of encouraging voter participation by inducing feelings of pride or shame. *Polit. Behav.* **32**, 409–422 (2010).
- 34 NIMH, *Anxiety Disorder Among Adults* (National Institutes of Mental Health, 2016).
- 35 R. W. Booth, B. Mackintosh, D. Sharma, Working memory regulates trait anxiety-related threat processing biases. *Emotion* **17**, 616–627 (2017).
- 36 M. D. Seery, E. A. Holman, R. C. Silver, Whatever does not kill us: Cumulative lifetime adversity, vulnerability, and resilience. *J. Pers. Soc. Psychol.* **99**, 1025–1041 (2010).
- 37 M. L. Hatzenbuehler, J. C. Phelan, B. G. Link, Stigma as a fundamental cause of population health inequalities. *Am. J. Public Health* **103**, 813–821 (2013).
- 38 R. M. Bond et al., A 61-million-person experiment in social influence and political mobilization. *Nature* **489**, 295–298 (2012).
- 39 A. D. Kramer, J. E. Guillory, J. T. Hancock, Experimental evidence of massive-scale emotional contagion through social networks. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 8788–8790 (2014).
- 40 D. Broockman, *Panel Discussant on Ethics, Institutional Review Boards and Conflicts of Interests in Field Experiments* (The Center for Advanced Study in the Behavioral Sciences, 2018).
- 41 R. B. Morton, K. C. Williams, *Experimental Political Science and the Study of Causality: From Nature to the Lab* (Cambridge University Press, Cambridge, UK, 2010).
- 42 National Research Council, Population Co, *Proposed Revisions to the Common Rule for the Protection of Human Subjects in the Behavioral and Social Sciences* (National Academies Press, 2014).
- 43 C. Buckley, To test housing program, some are denied aid. *NY Times*, 8 December 2010. <https://www.nytimes.com/2010/12/09/nyregion/09placebo.html>. Accessed 9 November 2020.
- 44 S. R. Schuler, S. M. Hashemi, S. H. Badal, Men's violence against women in rural Bangladesh: Undermined or exacerbated by microcredit programmes? *Dev. Pract.* **8**, 148–157 (1998).
- 45 N. S. Murshid, F. M. Critelli, Empowerment and intimate partner violence in Pakistan: Results from a nationally representative survey. *J. Interpers. Violence* **35**, 854–875 (2020).
- 46 J. D. Foltz, K. A. Opoku-Agyemang, Do higher salaries lower petty corruption? A policy experiment on West Africa's highways. https://cega.berkeley.edu/assets/miscellaneous_files/118_-_Opoku-Agyemang_Ghana_Police_Corruption_paper_revised_v3.pdf. Accessed 2 November 2010.
- 47 G. D. Ruxton, T. Mulder, Unethical work must be filtered out or flagged. *Nature* **572**, 171–172 (2019).
- 48 H. Algahtani, M. Bajunaid, B. Shirah, Unethical human research in the field of neuroscience: A historical review. *Neurol. Sci.* **39**, 829–834 (2018).
- 49 D. P. Green, Panel discussant on ethics, institutional review boards and conflicts of interests in field experiments (The Center for Advanced Study in the Behavioral Sciences, Stanford, CA, 2018).