

Conducting Research with NGOs: Relevant Counterfactuals from the Perspective of Subjects

David W. Nickerson
Associate Professor
Political Science
University of Notre Dame
dnickers@nd.edu

Susan D. Hyde
Associate Professor
Political Science and International Affairs
Yale University
Susan.hyde@yale.edu

Some field experiments in comparative politics are designed to test hypotheses derived from social science theories, and are created and implemented entirely by researchers. Other field experiments explore the effects of existing interventions undertaken by non-academic organizations, like Intergovernmental Organizations (IOs), International Non-Governmental Organizations (INGOs), or government bodies. These interventions may be intended to reduce poverty, boost political participation, encourage development, promote democracy, or curb the spread of disease. The interventions may be well-designed, have positive effects, and be cost effective. They may also have unintended consequences, adverse effects, or no effects whatsoever. Determining the effects of these interventions is important for public policy, intellectually rewarding, and socially desirable.

In evaluating the ethics of research, the difference between researcher-led interventions and studies of interventions initiated by third parties like IOs and INGOs is a critical factor. The crux of the ethical question depends on the cause of the intervention (or the “treatment”) evaluated in the field experiment, so we distinguish between “the intervention” and “the research,” and focus on counterfactual comparisons between the world with and without the research. Does the researcher cause the intervention to happen? Or would the intervention happen even if the researcher were not working with the organization? In cases where the intervention itself would not occur in the absence of the research, then the ethical questions involve both the intervention and the research, as is typical in human subjects cases. The relevant counter-factual involves the benefits of and costs of both the intervention and the research: both would not occur in the absence of the study. However, in cases where the field experimental treatment would occur even if the study of it did not, then the ethical questions

primarily concern the research component, and should be less concerned with the intervention itself, as we detail below. If an NGO, foreign government, or other third party will pursue the intervention independently; the relevant (ethical) counter-factual comparison is between a world in which the treatment occurs and a world in which the treatment occurs and its causal effects are (at least in principle) better understood.

Questions about the ethics of collaborative field experimental research are relevant to both scholars and Institutional Review Boards (IRBs), which are housed at all US-based universities, and that scholars must seek approval from for all research involving human subjects.¹ Many questions about the ethics of field experiments are addressed by IRBs, which are charged with protecting human subjects from unethical conduct by university researchers. The Federal Policy for the Protection of Human Subjects (FPPHS) provides guidelines and rules for IRB's across universities in the United States. These rules focus on three actors: subjects; researchers; and government departments. Although this focus is understandable given outrage over infamous research conducted by governments (e.g., the Tuskegee Syphilis study) and researchers (e.g., Milgram's *Obedience to Authority* experiments), an increasing amount of contemporary research is produced by scholars working in conjunction with programmatically focused non-governmental organizations (NGOs). This essay offers practical guidelines for evaluating the appropriateness of collaborative experiments between researchers and NGOs, both for IRBs specifically and for the scholarly community more generally.

In particular, this essay argues that when researchers work with organizations already engaged in programmatic activity, IRBs and observers interested in the ethics of field

¹ In general, when we reference IRBs in this chapter we are referring to the researcher's home institution.

experiments should focus less on the nature of the intervention and more on the changes introduced by the researcher's involvement in the project. The intervention made by the NGO (or that would be made by the NGO even in the absence of the researcher) should be taken as the baseline state of existence for subjects in the study and the IRB should ask whether the research-driven component adheres to the principles of the Belmont report such as autonomy and beneficence. Even if a study falls outside of IRB scrutiny, researchers should be particularly concerned with the ethics of research when they cause an intervention to happen in the real world, even if they are working in cooperation with an organization or government entity.

This essay is concerned both with how IRBs view field experimental research when scholars collaborate with NGOs and with the more general ethical relationships surrounding researcher-NGO cooperation. It considers IRB approval of a field experiment as a special case of the more general ethical concerns for cooperative research. It then explains why government studies are typically viewed as exempt from IRB oversight. The heart of the essay then discusses the middle ground occupied by cases where researchers partner with NGOs rather than governments, which concludes with a set of practical questions that researchers and IRB's should ask about research conducted in conjunction with NGOs. These points are then illustrated by a case study of an academic-NGO research study in Cambodia and working with a partisan political campaign in the United States. .

Standard Paradigm

The most common scenario for an IRB to consider is one where the researcher is acting directly on the subjects in the experiment. In these instances, the FPPHS provides very clear guidance. The IRB should give the design careful scrutiny under three broad principles: autonomy; beneficence; and justice. *Autonomy* advances the claim that subjects should have the right to decide whether or not they participate in the experiment. Regardless of the scientific merits of the study, subjects should be explained the costs and benefits of participating in the study and pro-actively opt into the study of their own free will. Just as a researcher should not surreptitiously slip a drug into a person's coffee without her consent, nor should they reveal sensitive information about friends or construct roadblocks to make voting in an election more difficult without the explicit approval of the people involved. The requirement for informed consent is not absolute. The FPPHS makes exceptions for studies of "minimal risk", which is how many social psychology experiments are allowed (e.g., Cialdini Reno and Kallgren 1990; Norton et al. 2012), but the researcher needs to make a clear and compelling statement about why subject autonomy is being violated.

The principle of *beneficence* holds that the researcher has a duty to minimize the amount of harm (or risk of harm) that a subject in the study faces and maximize the amount of benefit. While assiduously adhering to this principle involves a great deal of effort and careful planning, the intuition behind the dictum is straightforward. If a researcher is going to intervene in a subject's life, the researcher has an obligation not to cause harm to the subject unnecessarily. Even if the risk of harm is low, the researcher should take every reasonable precaution to prevent the harm and rectify the situation quickly should harm occur. When someone is under your care or observation, playing roulette with their well-being is not

acceptable. For instance, an intervention distributing new energy efficient stoves to villages (Ludwinski, Moriarty and Wydick 2013) should verify that the stoves will not increase the risk of house fires, create noxious fumes, or be misused in a manner that harms subjects in a way that is foreseeable. The principle of beneficence also precludes the researcher from denying subjects basic goods to which they have a reasonable expectation of enjoying. For instance, while a researcher may not have a moral obligation to construct bridges in remote areas, they cannot prevent the usage of existing bridges by residents for the purpose of research.

Participating in the study should be as beneficial and as harmless as possible.

The principle of justice mandates that the costs of the study are borne by people who will eventually benefit from the results of the study. The goal of science is to benefit humanity, but if the costs of doing science are inflicted on subgroups who have no reasonable expectation of ever gaining the benefit of the results, the principle of justice forbids this asymmetry. The canonical example of this dynamic is with drug testing in developing countries. Recruiting human subjects for drug trials is much less expensive in poor countries, so drug companies may be tempted to run human trials in these locales. Once the treatment is proven safe and effective, the company could then release the drug in developed markets where insurance companies and patients can afford the newly developed and now proven drug. Meanwhile the residents of the poor country where the drug was tested are unlikely to be able to afford the new drug, even though they provided a collective benefit to the rest of the world while shouldering the cost. The principle of justice forbids this unequal dynamic.

The purpose of this essay is not to explore these three principles of evaluation. Instead, this essay identifies the locus where these principles apply and examines the consequences for

academic studies where researcher involvement is not direct. One way to think about the relationship is by considering counterfactual circumstances from the point of view of subjects in the experiment.

The subjects exist in some sort of status quo. They have an expected level of psychological, physical, and emotional well-being. The researcher is not responsible for this status quo and under no obligation to improve the situation of potential subjects. However, the researcher is directly responsible for the effect of their intervention on the subject. When the researcher acts on a subject -- no matter how innocuously -- it shifts the subject from her status quo to a state at least partially created by the intervention. The change is often desirable: subjects are often compensated for their time; learn something from the experience; and may have been exposed to a beneficial treatment. However, it is the role of the IRB to determine that the intervention does not worsen the state of subjects in the experiment and require that the researcher take reasonable steps to minimize risks to subjects. As noted in the first row of Table 1, the effect of the intervention on the subject is the key area for scrutiny.

Table 1: Focus of Ethical Questions about Experiments

Cause of Treatment	Intervention	Research
Researcher	Full review	Full review
Government	Exempt	Exempt
NGO	Minimal scrutiny	Full review
Hybrid Researcher-NGO Model	Minimal scrutiny / Full review if Researcher Causes Treatment	Full review

The situation is markedly different when a researcher studies a US government program. According to the FPPHS §46.101 b5, as long as the project is approved by an agency head the research is considered “exempt” and suggests that IRB’s provide the most minimal levels of scrutiny. There are five primary reasons for this reduced level of scrutiny. The first

and foremost is that the government represents legitimate authority. The autonomy principle is important, but the government's monopoly on the legitimate use of force subordinates the autonomy principle on a daily basis. Given that governments have the authority to arrest and jail citizens (or even use deadly force), seize property, and wage war, the right to research the effect of policies is minor in comparison.

The second is that public accountability provides a presumption of goodwill on the part of the government. Researchers within universities are largely insulated from public scrutiny. Not only is the public unaware of researcher activities but they have no mechanism for removing researchers who engage in unethical behavior. In contrast, government policies and pilot tests are observed by the public and processes exist by which objections can be noted. While the heads of government agencies are rarely directly elected, the people appointing the heads are elected. The desire of elected officials to avoid controversy causes agency heads to self-censor and pursue only those policies and tests that they think will enhance the public well-being. Government abuses can occur (e.g., Tuskegee Airmen; Internment of the Japanese) and some policies have negative consequences (e.g., the construction of high density public housing), but the vast majority of activities are pursued with the public interest explicitly in mind. Unlike agency heads, university researchers have neither public accountability nor a presumption of good will.

Third, the government agencies have to make collective decisions and enact policies over the objections of some individuals. Nearly every government action creates winners and losers, so a set of people is likely to object to the new policy implemented by the agency. Governing by its very nature entails violating the autonomy of certain individuals when

implementing policy. Even the act of doing nothing (i.e., adhering to the status quo) is a decision with winners and losers, so there is no path where government actions can engage in behaviors that benefit everyone. The exemption for the evaluation of government policies recognizes this reality.

Fourth, policy evaluation is something that a government agency may legitimately desire. Although a government policy may be intended to advance the public welfare and accomplish particular policy goals, there is no guarantee that the policies will be successful. Careful and rigorous evaluation of policies can help to inform future policy decisions, so agencies can view research as an integral part of the policy making process rather than a peripheral activity that deserves special and isolated scrutiny.

Finally, researcher evaluation of government agency policies is a public service that should be encouraged. Although policies implemented by government agencies are public relative to the inner-workings of universities, many members of the voting public will be unaware of the decisions and their consequences. Allowing researchers to partner with government agencies provides unusual transparency into operations. Not only can the researcher document the activities of the agencies, but it is also possible for her to measure their effectiveness. Much as the media is considered the “watchdog of the government”, researchers can ensure that the government is behaving appropriately and effectively.

Underlying each of these justifications for the exemption enjoyed by research involving government agencies is the notion that the government has the legitimate authority to act on its own even if its programs are not systematically evaluated, but that understanding the consequences of government policies is important. This assumption is much more problematic

in non-democracies, and does not guarantee that government actions are ethical or harmless. Again, let us approach the possible counterfactuals from the perspective of subjects in the experiment. The government agency is implementing a policy. This policy (with or without research) moves subjects from the status quo to a new position under the intervention being studied. The presence or absence of the researcher has nothing to do with this shift. Moreover, as noted earlier, research and evaluation can be considered a part of the policy changed being imposed by the government agency. That is, any involvement by the researcher is completely subsumed by the actions of the government agency. From the perspective of the subjects in the study, the researcher is irrelevant to their well-being and autonomy. Thus, research of policy interventions undertaken by US government agencies is already deemed exempt under human subjects guidelines, even when the introduction of research of an ongoing program influences which individuals receive the treatment (see Table 1, row 2).²

Hybrid Model with Interventions by NGOs³

In the space between research involving interventions made directly by the researcher and interventions undertaken by a government agency lies research involving interventions made by non-governmental organizations. NGOs differ from governments in a few key respects. Chief among these differences is that NGO's can claim no legitimate authority over and are not directly accountable to subjects in an experiment. Re-examining the justification for

² Of course, there may be government programs that a given researcher deems to be unethical, and individual researchers should not use the study of government programs as a blanket excuse to avoid thinking about any ethical considerations. The fact that an activity is allowed or legal or sanctioned by a government does not mean it is laudatory. Researchers should always follow their moral compass.

³ Note that NGOs are not subject to IRB approval. All discussion of IRBs pertains to their oversight over researchers who are based at universities, in the United States and in other countries that have similar bodies governing academic research with human subjects.

government research enjoying exempt status from IRB oversight, these differences in authority and accountability mean that the goodwill of the NGO cannot be assumed.

Of course, it is possible to argue that goodwill cannot be assumed on the part of many governments, particularly those that are willing to run harmful “experiments” on their citizens, and non-democratic governments may lack the type of legitimacy discussed above. In practice, some well-established NGOs may have significantly greater accountability and goodwill than some “bad” governments. Nevertheless, the goodwill of an NGO cannot be taken for granted, and the starting point for research conducted in collaboration with an NGO should be whether an intervention by an NGO (absent the research) is manifestly undesirable.

What is meant by manifestly undesirable? The IRB, when evaluating the researcher’s role in the study, should ensure that researchers are not participating in programs that clearly violate the rights of the individuals involved, but that is a lower threshold than ensuring all principles of the Belmont Report are adhered to. NGOs are created to accomplish programmatic goals (e.g., reduce poverty; improve health; educate the public) and pursuing those goals requires taking action. There will be instances where the NGO commits to a course of action without knowing the ultimate consequences of the decision. The world is sufficiently complicated that it is impossible to know *a priori* the outcomes of all interventions – hence the need for research. NGOs implement policies that they think will further their programmatic goals, but may be mistaken. The goal of the IRB should not be to judge the efficacy or wisdom of the intervention.

For instance, voter intimidation campaigns where citizens are threatened, beaten, and otherwise coerced into abstaining from elections are clear violations of the basic rights of the

people subjected to such campaigns of terror. Although understanding the effects of such intimidation campaigns is interesting and important, one of the roles of IRBs should be to ensure that researchers are not involved in morally repugnant behavior. Although the researcher may not be directly intimidating voters, her presence and active involvement offers a tacit endorsement of the activity.

In contrast, suppose the researcher is working with an anti-corruption campaign that distributes information on which candidates have been convicted on corruption charges and which candidates are “clean” (e.g., Banerjee et al. 2011; Figueiredo, Hidalgo, and Kasahara 2011). The campaign is intended to allow voters to cast a more informed ballot and increase participation by highlighting clean politicians. Of course, it is possible that the anti-corruption campaign could have negative consequences such as decreasing voter turnout by highlighting corrupt politicians or serving as a catalyst for street fights among partisans. But such consequences are not clear *ex ante* and whether the outcome of the intervention is positive or negative is an empirical question and part of the reason that the experiment is interesting.

The outcome of the anti-corruption experiment is also irrelevant to the review by the IRB. The intervention will happen regardless of whether the researcher participates or not. From the perspective of the human subjects involved in the research, the researcher is not harming the subjects by observing precinct-level electoral returns for the number of votes cast and support for clean candidates. If the researcher were directly implementing the anti-corruption campaign, then she would need to demonstrate that she had taken precautions to minimize the possible risks posed by the intervention. However, since the intervention is taking place anyway, the IRB’s question about the intervention outcome should be “could a

reasonable person believe that this intervention would be helpful, or is it clearly designed to harm subjects?”

This “reasonable person” standard only applies to the intervention itself. The IRB should apply all the usual criteria about how the researcher involvement affects subjects in the proposed study. However, it is important to note that the baseline for comparison is not the status quo but the change in the conditions for subjects under the NGO’s intervention with and without the activity of the researcher.

This change in baseline does not give researchers a free pass to conduct experiments. Researchers can often analyze human subjects data after an NGO has finished a program under the “found data” provision. However, if the researcher is active in the creation of the data, the found data clause does not apply. The autonomy and beneficence principles must be adhered to, but it is the researcher component rather than the intervention component that is the relevant consideration. Returning to the anti-corruption campaign example from above, if the NGO is going to intervene with the anti-corruption materials, then the subject autonomy to be exposed to the campaign or not has already been compromised and the researcher is irrelevant. However, autonomy applies just as strictly to the pure research component. The subjects should be free to decline interviews, survey instruments, blood pressure readings or any other measurement that the researcher seeks to apply. Just because the NGO plans to intervene does not mean that the researcher has *cart blanche* to do as she sees fit.

The application of the beneficence principle when an NGO is staging the intervention is similar. The risks associated with the intervention planned by the NGO may not be relevant to the IRB’s consideration, but risk is still a consideration because the presence of the researcher

may introduce risks for subjects. For instance, surveying residents of neighborhoods where an anti-bribery intervention is taking place could potentially link illegal behaviors to particular individuals and put them at risk of being arrested. The IRB should demand that the researcher develop and implement strategies to assure the anonymity of survey responses and protect the identity of people surveyed. In some areas the mere presence of survey teams in a neighborhood can attract unwanted attention from local gangs and put residents at risk – even if they decline to answer the survey themselves. Researchers need to openly discuss any possible risks associated with the research component and explain how they plan to minimize such risks. However, the risks associated with the intervention conducted by the NGO are not the responsibility of the researcher (assuming that the manifest purpose of intervention is not to impose harm on individuals). From the perspective of the subjects involved in the study, the introduction of the researcher is the only relevant change. Thus, the intervention should receive limited scrutiny, while the major ethical question concerns the addition of the research component (see Table 1, row 3).

The application of the justice principle to research changes very little from this perspective. Research testing new policies that were unlikely to ever be implemented in the experimental setting simply to save money (e.g., measuring the effect of television advertising campaigns in regions with very low television penetration into the market) remains in potential violation of human subjects protections. However, the partnership with an NGO should generally ameliorate concerns about justice. NGOs implement pilot programs in settings where they plan to roll out the full scale versions of programs. Given all the difficulty, local knowledge, and infrastructure required to implement sizable field experiments, it would make little sense

to conduct pilots in irrelevant settings. Thus, questions about justice should continue to be asked about the research, but the partnerships with NGOs actually help to address these concerns.

This shift in the baseline for the purposes of IRB review is predicated on the fact that the NGO is in fact separate from the researcher. However, there are cases where the line between researcher and NGO is blurry. Some researchers are principals in consulting firms and not for profit organizations. A few NGOs like the Poverty Action Lab exist for the sole purpose of conducting research. When the researcher also has decision making authority over the NGO, there is effectively little difference between the researcher and the NGO. From the point of view of subjects in the study, the NGO is acting as an extension of the researcher rather than as an independent intermediary. In these instances, the relevant point of comparison is between the status quo pre-intervention and the intervention with the research component. IRBs should apply the full battery of criteria to the intervention itself, not just the research component.

This paradigm means that IRBs should inquire about the nature of the relationship between the researcher and NGO performing the intervention. There is no bright line standard, which is why we added the bottom row of Table 1 to address these hybrid situations. The researcher's relationship to the NGO and the NGO's decision making process will have to guide the IRB in determining the extent to which the NGO is acting independently. Serving the as the CEO of the NGO probably grants the researcher a large degree of authority over the NGO's activities and completely muddies the line between researcher and NGO. In contrast, acting as an advisor to the NGO will probably not be sufficient to trump all decision making authority. Serving on the board of directors to an NGO most likely falls into a middle ground that the

researcher will have to explain carefully and the IRB should consider. The core concept of interest is who makes the final decision to authorize and implement the program.

Funding of the program is a related issue. If the researcher supplies funds for implementing the intervention (apart from the separate cost of the research), then she is responsible for the intervention and the relevant point of comparison is the status quo versus risks under the intervention. She is responsible because the intervention would be unlikely to occur on its own because funds from the researcher were necessary. Money is not a smoking gun, however. If the researcher supplies funds for just the research component of the intervention, then the intervention sans research may have been possible and the intervention should be treated as exogenous to the researcher. The researcher could also help write grants for the organization; lending expertise to the organization to raise money or improve proposals and operations is not the same thing as making decisions on behalf of the NGO. The key decision is programmatic. If the NGO has capacity to execute the intervention and the program is within the broader mission of the NGO, then there is no particular reason to be suspicious that the researcher is dictating the terms behind the scenes. However, the researcher should clearly address her relationship with the NGO.

To summarize our core argument, Institutional Review Boards (and those interested in ethics of experiments more generally) should approach the beneficence and autonomy criteria from the point of view of the subjects in a study. When an NGO is planning to implement a new program and the researcher is not directly the cause of that program, then the effect of the program becomes the baseline for comparison. The subjects have already lost their choice on whether they are exposed to the program and its consequences. The relevant question is

whether researching the program crafted by the NGO compromises autonomy or beneficence independently. The proposed research needs to adhere strictly to these two principles, but the nature of the intervention should not be the focus of the review board because the intervention by the NGO was going to happen even if the research did not.

One potential objection to this line of reasoning is that researchers would be passing the buck to NGOS, and the dynamic will lead to a slippery slope in which researchers coerce NGOs into providing cover for ethically questionable actions. At the extreme a researcher could create a fake NGO in order to dodge the ethical standards that academic research normally faces. Although this is possible, it is unlikely because it assumes INGOs are very malleable and researchers are devious. Not only are these traits likely to be rare, it is far from clear that IRB oversight will mitigate these two problems. Ethically compromised scholars and unprincipled, profit seeking INGOs are likely to do other ethically objectionable things: falsify results, lie about funding sources, harass their staff, or otherwise violate the spoken and unspoken norms of the academic profession. Overall, the “fake NGO” or “coerced NGO” possibility is a very labor intensive way of avoiding IRB approval to conduct research and unlikely to be the normal state of affairs.

Perhaps more likely is the possibility that researchers would take on work as consultants or advisors to an organization and then use that relationship to create opportunities for field experimental research. This can be a grey area, but should be governed by the same general principle. If the researcher, even when acting outside his or her “day job” as a university professor, causes something to happen in the world that would otherwise not occur *and* relies on that intervention for field experimental research, the full intervention should be subject to

human subjects review. If a researcher is working with an NGO in a capacity that is not at all related to field experimental research, does the researcher still have an obligation to bring up concerns that might arise in the course of university IRB review? In this type of case, the researcher may feel obligated to bring up concerns related to the intervention and they are more than welcome to do so, but it is not necessarily their role to bring up IRB standards in all aspects of their life. The central ethical concerns, as we view them, mainly arise if the researcher is involved in using the NGO's work for their own research.

To summarize our argument and make the proposed considerations more concrete, when IRBs and researchers are confronted with proposed research evaluating the effects of NGO interventions, we suggest the following questions as a helpful starting point:

- 1) Would the intervention occur without the researcher?
 - a. What role does the researcher play in the organization (i.e., does she have decision making authority?)
 - b. Who is paying for the intervention?
- 2) Is the intervention obviously harmful or otherwise unethical that a university should not tacitly endorse?
- 3) Does the research component respect the autonomy of the subjects?
 - a. Do subjects have the ability to opt out?
 - b. Is any information being collected directly from individual subjects?
- 4) Does the research component comply with the demand for beneficence?
 - a. Are subjects put at risk by researching the intervention?
 - b. What steps are being taken to protect the welfare of the subjects?
- 5) Does the research component comply with the justice requirement?
 - a. Will the subjects benefit from the results of this research?

Although these questions will not address every potentially problematic relationship or troublesome dynamic, they provide a framework for discussing the major ethical concerns that will commonly occur. Additionally, we recommend that any researcher engaged in a cooperative relationship with an NGO for field experimental research provide a clear

explanation of the nature of the relationship between the researcher and the organization, including who pays for what, in the first footnote to any related publication. Researchers should also define whether they were paid to conduct the study (and research for which the cooperating organization pays the researcher for their work should potentially be discounted, as in other fields).

Case study #1: Cambodia

To illustrate the proposal in this chapter, this section describes a democracy assistance intervention studied by one of the authors (hereafter the “researcher”) and then answers the questions proposed in the preceding sections. The goal of the section is to illustrate how the principle of relevant counterfactuals can be practically applied.

Democracy assistance is a common component of foreign policy, and democracy assistance activities are funded by dozens of governments, implemented by numerous NGOs and IGOs, and have been carried out in more than one hundred countries in the world. At its core, democracy assistance is a targeted intervention aimed at encouraging democracy in target states. Democracy assistance organizations most often have permission for their work from host governments, but in some cases do not. As highlighted by a 2008 report from the National Research Council, most US funded democracy assistance activities are carried out absent the type of evaluation that would provide “compelling evidence of whether those projects had their intended effects.” (National Research Council 2008, p 5) The same is true for most actors in democracy assistance, and it is relatively standard practice for INGOs to

intervene within sovereign states without systematically evaluating the consequences of these interventions.

Motivated by the 2008 report, several NGOs receiving USAID funding were mandated by USAID to conduct rigorous impact evaluations of some of their democracy assistance programs. The program that was ultimately one of the pilot studies had been run by an International NGO (INGO) in Cambodia since 2004. Although it had been documented in periodic reporting and anecdotal evidence suggested that the program had positive effects, it had not previously been evaluated using rigorous methods for impact evaluation.

The program brings elected Cambodian Members of the National Assembly (MNAs) to rural villages in their district for multiparty town hall meetings. Once elected, it is rare for MNAs to visit rural parts of their districts. It is even more uncommon for rural villagers to be exposed directly to any situations in which opposition political parties share the same stage, and are on more or less equal footing, with the ruling Cambodian People's Party. The specific aims of the program are to encourage the MNAs to be more accountable to their constituents and to educate rural villagers about their roles and responsibilities as voters with the broader aim of encouraging them to hold their elected MNAs accountable.

The intervention in Cambodia consists of a town hall meeting designed to bring national representatives face to face with their constituents in a question and answer session. These town hall meetings are held in rural villages, are open to the public, and attended by 400 to 1,000 community members. The town halls are advertised using a loudspeaker on the back of a motorbike, which travels around all accessible roads and paths in the area surrounding the

village, thereby reaching most villagers. The meetings are conducted with the permission of national, district, and local officials, who typically help publicize the events.

Each meeting begins with remarks by the senior program officer or director of the INGO, who explains the purpose of the meeting to the audience. Each introduction includes a statement that Cambodia is a democracy and that they, the audience, have elected their representatives and have a right to share their concerns with them; that the purposes of their elected representatives are to form policies and exercise oversight on the government, but not to provide gifts or personal favors; that these representatives must represent the interests of their constituents and that they work for and are accountable to their constituents; and a reminder that everyone, including the audience, has the right to speak and be heard and to criticize and disagree with the policies or positions on any issue.

Participating Members of the National Assembly (MNAs) are seated at a table in front of the participants and offer brief opening remarks to the audience. The floor is then open to participants to voice their concerns, raise questions directly to members of the National Assembly, and request that actions be taken to resolve problems. During their allocated response time, MNAs use the opportunity to update citizens on the activities of the legislature and government and provide other information relevant to their constituents. The dialogues encourage two-way communication and are unscripted, allowing for sometimes challenging questions and demands from citizens.

To ensure equity and neutrality and foster constructive dialogue, each town hall meeting follows a format and requires that MNAs, participants, and moderators follow a code

of conduct.⁴ The code specifies such items as the allotment of speaking time, appropriate and inappropriate topics for discussion, and the rules regarding audience participation. Local authorities such as commune councilors, village chiefs, and district and provincial officials are often present at the town halls, and in 2010, at the request of the parties, time was allotted for local authorities to speak at the end of the dialogues as well. Citizens who attend the town hall meetings are given water and bread, but no other remuneration or incentive to attend.

The town hall meeting program is explicitly designed to promote citizen interaction with multiple parties, and to demonstrate to citizens that opposition parties in Cambodia do exist and that debate between them is a normal part of multiparty democracy. Representatives from five parties in the National Assembly have been involved in the meetings, including the Cambodian People's Party (CPP), the Sam Rainsy Party (SRP), FUNCINPEC,⁵ the Human Rights Party (HRP), and Norodom Ranariddh Party (NRP).⁶ For each town hall meeting, each party that has one or more elected Member of the National Assembly in that province is invited to the meeting. Thus, fewer than five parties are represented at each town hall event.⁷ In addition, town

⁴ A code of conduct was developed in 2008 to clarify misunderstandings and avoid confusion about the town hall process and allegations of bias. It also serves to prevent verbal attacks and personal insults during the dialogues. This method of preventing and resolving disputes was agreed to by the five political parties represented in the program. The code is regularly reviewed with the parties to ensure its relevancy and make modifications, if necessary.

⁵ FUNCINPEC is an acronym for the National United Front for an Independent, Neutral, Peaceful, and Cooperative Cambodia. It represents the party name in French, "Front Uni National pour un Cambodge Indépendant, Neutre, Pacifique, et Coopératif."

⁶ From 2004 to March 2008, only three political parties – the Cambodian People's Party (CPP), the Sam Rainsy Party (SRP), and FUNCINPEC – had seats in the Assembly. With two new parties, the Human Rights Party (HRP) and Norodom Ranariddh Party (NRP), gaining seats in the 2008 elections, five parties have now participated.

⁷ Each party chooses which representative will attend the town hall event. Through conversation with the NGO director, it appears that MNAs are trained for attendance to the town hall event.

hall events are not held in provinces where the ruling CPP party is represented in the National Assembly, which excludes half of the provinces in Cambodia.⁸

The forum provides a rare opportunity for policy debate and showcases the distinctions between legislators from different parties and their approaches to both local and national concerns. Citizens interact with their elected MNAs, and see MNAs from multiple political parties interact with one another. The meetings provide an opportunity to enhance MNAs' knowledge of and relations with their constituencies and educate citizens on the roles and responsibilities of MNAs in a democracy. Citizens, through their attendance in the town hall event, have an opportunity to learn about their form of government including, not only the roles of the MNAs, but also about their own roles and rights as citizens. In theory, through their participation or the participation of others, citizens may also be motivated to take part in politics either through political party activity or other political behavior. The research component of the study involved helping to randomize where the community meetings took place. The randomization created communities that were on average identical and differed only with respect to the occurrence or non-occurrence of the meeting. The researcher could then conduct public opinion surveys in treatment and control communities to determine whether the meetings made residents more informed, more likely to participate, feel more efficacious, value democracy as a form of government, and support the regime. These questions are routinely asked in comparative public opinion surveys and no extremely sensitive questions were asked.

⁸ These provinces are excluded from this program because the INGO elected not to work there for a variety of reasons, including that there are no elected representatives from multiple political parties representing those provinces eligible to participate in the events.

The first set of questions to consider about the research is whether the intervention would have occurred in the absence of the researcher. In this case, the Constituency Dialogue program would have taken place regardless of the researcher's involvement in the impact evaluation, and had been ongoing for several years. The (donor-mandated) addition of the impact evaluation to the program added a pre and post-test survey to the existing programmatic activities, and expanded and systematized a focus group activity that took place the day before each Constituency Dialogue. The researcher helped implement and analyze the results from the impact evaluation. As further evidence that the researcher did not "cause" the intervention, the researcher does not otherwise work for the INGO, and is not a decision-maker in the organization. She chose not to be paid as a consultant on this project in exchange for access to the data, though travel expenses for the researcher were paid for by the INGO. The intervention itself was, critically, paid for by the INGO, not the researcher. In this particular case, the research component was also paid for by the INGO, though this, in our view, would not make an important ethical difference.

Second, is there evidence that, *ex ante*, the intervention would be judged by a reasonable person to be obviously harmful? In this case, the answer in advance of the study was no, as most of the anecdotal evidence was positive. Note that negative findings about the program as a result of the study were possible, but it was not clear before the evaluation the program was obviously harmful or unethical.

Third, IRBs should also consider whether the research (not the intervention) compromises the autonomy of subjects. The central added component introduced by the research was the survey. The survey contained a standard introduction and respondents could

opt out. Respondents were also free to skip any questions that they preferred not to answer. In short, while subjects were unable to opt-out of the intervention implemented by the NGO (i.e., the presence or absence of the community meeting in their region), they were free to participate or not in the research component of the study.

The fourth set of questions focus on beneficence in relation to the research component. The researcher-introduced component of the experiment is the survey, so an evaluating IRB should run through the questions it applies to most comparative survey research. Most surveys are deemed “minimal risk” and exempt under FPPHS guidelines except when surveys involve vulnerable populations or sensitive questions. In this instance, none of the questions were sensitive and answering would never put a respondent at risk. Furthermore, personal information was separated from the survey results, so although the data were individual in nature, they are anonymous. Since the survey constitutes minimal risk, the next question regarding beneficence requires weighing the gains from the research. Given that the intervention is likely to continue to occur even in the absence of the study (and that any risks that apply to the research would likely apply to the intervention), and the research component adds significant knowledge about the effects of the intervention, the added value is not trivial, and could lead to future modifications of the program that make it more beneficial to participants.

The fifth and final questions relate to justice, and whether it is likely that the subjects would benefit from the research. Since the participating NGO was conducting this study with an eye toward future implementation of the program in Cambodia, there is little question that

subjects in the experiment would be the ones to directly benefit from the added knowledge about the effects of the program provided by the research.

In short, the researcher involvement in the Constituency Dialogues experiment in Cambodia should be a relatively easy case for outside observers to deem acceptable. Objections that people may have about the overall program (e.g., exposing community residents to partisan violence; potentially affecting electoral outcomes) do not apply to the purely research component. Helping to randomize the intervention and implementing a survey to measure the attitudes of community residents are minimally intrusive and help to make the entire enterprise more valuable by measuring the effects of the program with an aim of improving future iterations of the project.

Case Study #2: Michigan

To illustrate that the same issues can arise in the developed world, we describe an experiment conducted by a political organization in the United States in conjunction with a team of researchers. In 2002, the Michigan Democratic Party was interested in measuring the effectiveness of mobilization efforts aimed at younger voters. The ability to engage people under the age of 35 was a priority for the party leadership and they wanted to ensure that traditional methods of mobilizing voters (i.e., door hangers, volunteer phone calls, and door knocks from volunteers) were effective on this younger target group (for a full description see Nickerson, Friedrichs, and King 2006). The scripts and campaign literature employed in the mobilization effort were relatively anodyne persuasive appeals encouraging voting for the entire Democratic ticket. While subjects in the experiment (i.e., targeted young voters) would

not be offended by the scripts, they also had no opportunity to opt out of either the experiment or the outreach effort by the party.

The 2002 election in Michigan was close up and down the ballot. For instance, the gubernatorial election was decided by 4 percentage points, the Democrats received more of the US Congressional vote but won only 7 seats compared to 9 seats won by the Republicans, and the Attorney General race was decided by a mere 5000 votes. In such a competitive environment, it was a bold decision by the party to work with a team of researchers in order to conduct a fully randomized experiment with a control group that would not be mobilized by the coordinated campaigns.

This decision was made somewhat easier by the fact that one member of the research team raised a considerable sum of money to fund the experiment. While the state party contributed the funds for the printed materials (flyers, door hangers, and scripts), the researcher raised funds from private donors to conduct the experiment. Most of the money went to paying for staff to oversee the experiment, but this had the additional effect of increasing the capacity of the campaign itself. A major job of the staff to manage the experiment was volunteer recruitment, retention, and oversight to assure that the hangers were hung, phones were called, and doors were knocked. This effort probably resulted in a larger volunteer outreach effort aimed at youth than would have been possible otherwise. The researchers also improved efficiency of the mobilization by assisting with the targeting and managing databases. So while the researchers did not have the power to authorize or initiate the mobilization experiment on their own, their efforts did change the nature of the youth engagement by the Michigan Democratic Party in 2002.

After Election Day passed, the researchers collected information on the turnout behavior of the subjects in the experiment by referencing official voter turnout records maintained by the Michigan Secretary of State and publicly available (for a nominal fee). The researchers also conducted a post-election survey (funded by an academic institution) via telephone to measure the vote choice for both the treatment and control groups. The survey was extremely brief and asked questions typical of political surveys like who the respondent voted for and whether they recalled any campaign outreach (see Nickerson 2005 for a full description).

If the researchers had attempted this experiment on their own, it would potentially run afoul of ethical guidelines for research. The biggest concern is that the autonomy principle is violated since subjects have no choice of whether to participate in the experiment. Subjects were not forced to read the door hangers or engage in conversation with volunteers – indeed, many threw out the hangers, hung up the phone, or refused to answer the door in the first place – but there was no option to opt out of the study.⁹ Some people might also object to conducting a mobilization experiment in an environment where it is possible that the outreach could decide the election. Whether or not Democratic victory would be a positive or negative consequence of the election is a matter of opinion and it is not up to researchers to make that decision for voters. Given that the intervention involves minimal risk, such a researcher initiated experiment would likely be approved, but it would deserve careful consideration.

If the Democratic Party shouldered the costs of the entire program, then the intervention would require less scrutiny and the study would be easy to approve. Campaigns

⁹ Exceptions can be made to the mandate to obtain informed consent when informed consent is impossible and the risk to subjects is minimal.

have a right to engage in their electioneering, which means that residents do not have a right to not be targeted by the campaign. The mobilization effort may affect electoral outcomes; in fact, that is the entire point from the perspective of the political party studying its outreach. The question for the researcher and an IRB, then, is whether the intervention is inherently inappropriate. A reasonable person could believe that the mobilization tactics in question would pose no risk to the subjects in the experiment. Campaigns around the world utilize similar tactics without incident. Since the intervention is harmless to individual wellbeing and well within the programmatic goals of the organization, the intervention would not be a relevant factor in considering the appropriateness of the experiment.

The ethical question would be whether the research component adheres to the standards outlined in the Belmont report. The research component consists of three primary components and each is unobjectionable.¹⁰ First, the researchers randomized the state party's list. Just as political parties have the right to target whom they think appropriate, they also have the right not to target individuals. The control group is not being deprived of any expected right or benefit. Other groups may still knock on their door or make phone calls. Second, the researchers referenced public data bases to determine who voted in the treatment and control groups. Although it is true that subjects cannot opt out of this data collection exercise, researchers are merely referencing publicly available data and not forcing individuals to give up information. Finally, the researchers attempted to survey subjects in the experiment. This survey was entirely voluntary and subjects could easily opt out of this part of

¹⁰ Volunteers recorded the disposition of their attempts to contact subjects in the treatment group, but this would occur in the absence of the experiment. The disposition of attempts to contact voters is valuable information for campaigns for volunteer oversight and better understanding the electorate. Care needs to be taken, however, to ensure that the research does not cause this information to become public.

the experiment by declining to take the survey or answering particular questions. Thus, the purely research component of the experiment poses no cause for concerns and would be approved quickly.

The actual experiment, however, fell between the two extremes of researcher initiated invention and the organization lead experiment. Understanding this relationship in some degree of depth is necessary to determine the degree of scrutiny the intervention itself should receive. A key question is whether the organization chose to engage in the experiment independently or whether the researchers made the decision for them or exercised some form of coercion. The power dynamics in this case are relatively straightforward. Two graduate students can propose the experiment to a state party chair, but there is no way that they have the authority to force her/him to act in a certain way. So in this case, the locus of the ultimate decision making authority squarely lies with the organization making the intervention and not the researchers.

The next question is whether the intervention would have taken place without the researchers. While it is true that the idea for the experiment originated with the researchers and not the organization, the proposed experiment manipulated outreach that was taking place regardless of researcher intervention. Given that the question is whether the intervention should receive full scrutiny, the fact that the researchers are not greenlighting the intervention is the key detail.

That said, the researchers did increase the scope of the operations. By hiring dedicated staff and assisting with technical details, the researchers allowed the organization to apply the treatment to more individuals than would otherwise have been possible. And this is an

excellent example of why bright line standards may be impossible apply in settings where NGOs and researchers collaborate and split costs. It is relatively obvious that if the researchers are the proximate cause of the lion's share of the outreach, then the experiment should be treated as if the researchers are responsible for the intervention. One could adopt a "one drop" rule where any researcher contribution taints the experiment and the intervention must receive full scrutiny, but that would preclude the researcher from supplying any labor to see that the research occurs, which strikes the authors as unreasonable. So how much labor and expense sharing can the researcher do before the intervention flips from full to minimal scrutiny? We do not think there is a good one-size-fits-all answer to this question, although the questions we outlined above should help IRBs and individual scholars work through important questions associated with researcher collaboration with real-world entities.

Conclusion

Researchers and Institutional Review Boards should be focused on the welfare of the human participants in research. This chapter argues that the effect of research should be evaluated from the perspective of the human subjects whose welfare and dignity are being protected. When a researcher assists an NGO evaluate the effectiveness of a program, it is the researcher's activities and not the NGOs that should receive the most careful scrutiny. From the perspective of the subjects in the experiment, the NGO program is being implemented regardless of whether the researcher participates or not. The relevant question is whether the research component respects subject autonomy and adheres to the principles of beneficence and justice.

As with nearly all ethical questions, there will be gray areas and reason to debate. For instance, it may not be possible to cleanly separate the research component from the program implemented if the program is altered to accommodate the research. In these cases, the review board should focus on the effect of the change in the program introduced by the research compared to the pure program rather than a counterfactual where no program took place. The goal of this essay is to provide a framework for asking these questions and focusing the discussion on the most ethically relevant component of the research.

Bibliography

- Banerjee, Abhijit V., Selvan Kumar, Rohini Pande, and Felix Su. 2011. "Do Informed Voters Make Better Choices? Experimental Evidence from Urban India." Working paper, November 11, 2011. http://www.hks.harvard.edu/fs/rpande/papers/DoInformedVoters_Nov11.pdf
- Cialdini, Robert B., Raymond R. Reno, and Carl A. Kallgren. 1990. "A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places." *Journal of Personality and Social Psychology* 58(6):1015-1026. doi: 10.1037/0022-3514.58.6.1015
- De Figueiredo Miguel FP, F Daniel Hidalgo, and Yuri Kasahara. 2011. "When do voters punish corrupt politicians? Experimental evidence from Brazil." Unpublished manuscript, February 15, 2011. http://www.law.berkeley.edu/files/SPCorruptionv52_for_Berkeley_Law_and_Economics_Workshop.pdf
- Ludwinski, Daniel, Kent Moriarty, and Bruce Wydick. 2013. "Environmental and health impacts from the introduction of improved wood stoves: evidence from a field experiment in Guatemala." *Environment, Development and Sustainability* 13(4):657-676.
- National Research Council, Committee on Evaluation of USAID Democracy Assistance Programs. 2008. *Improving Democracy Assistance: Building Knowledge Through Evaluations and Research*. Washington, DC: The National Academies Press.
- Nickerson, David W. 2005. "Partisan Mobilization Using Volunteer Phone Banks and Door Hangers," *The Annals of the American Academy of Political and Social Science* 601:10-27.
- Nickerson, David W., Ryan F. Friedrichs, and David C. King. 2006. "Partisan Mobilization Experiments in the Field: Results from a Statewide turnout experiment in Michigan," *Political Research Quarterly* 34(1):271-292.
- Norton, Michael I., Elizabeth W. Dunn, Dana R. Carney, and Dan Ariely. 2012. "The persuasive "power" of stigma?" *Organizational Behavior and Human Decision Processes* 117(2):261-268. [<http://dx.doi.org/10.1016/j.obhdp.2011.08.002>.]