Field Experiments in the Global South: Assessing Risks, Localizing Benefits, and Addressing Positionality

Biz Herman ⓡ Amma Panin ⓡ Nicholas Owlsley ⓡ Graeme Blair ⓡ Alex Dyzenhaus ⓡ Elizabeth Iams Wellman ⓡ Allison Grossman ⓡ Ken Opalo ⓡ Anisha Singh ⓡ Hannah Alarian ⓡ Lindsey Pruett ⓡ and Yvonne Tan

Accepted, *Perspectives on Politics*

Introduction

Field experiments, also known as randomized controlled trials (RCTs), have emerged as a leading methodological tool to strengthen causal inference in the social sciences (Banerjee & Duflo 2009). Designed to make causal claims on research questions of interest to scientists, policymakers, and the public, RCTs often require significant resources, substantive interventions in participants' lives, and partnership with non-academic implementers (Davis & Michelitch 2021; Haas et al. 2021; Teele 2014;). Consequently, RCTs carry significant costs and risks for participants, research staff, and researchers themselves (Kaplan et al. 2020). Although all political science studies require ethical and cost-benefit evaluations, RCTs invoke special considerations of identity, positionality, and power dynamics (Haas et al. 2021, Kim et al. 2021), especially for studies conducted in the Global South. RCTs' resource-intensive nature means they are often conceived and led by scholars from the Global North (Panin 2020, Cordenau-Huci et al. 2021). This has implications for question choice (Thachil & Vaishnav 2018), site selection (Porteous 2020), and relationships with Global South research partners (Fujii 2012; Bleck et al. 2018; Mwambari 2019).

We do not take for granted the Global North concentration of human capital leading RCTs.[1] In this piece, we draw on our interdisciplinary experiences to develop a set of questions for RCT research in the Global South[2], suggesting ways to involve scholars and research staff who hail from the study site at every research stage. We maintain these interactions are not one-off exchanges, but rather opportunities to foster meaningful collaboration. We see such efforts as complementary to institutional efforts to recruit and retain graduate students and junior faculty from the Global South. We organize this piece by four distinct, yet interrelated research stages: idea generation, planning, implementation, and dissemination.

Idea Generation

Ideas behind RCTs emerge from various contexts, often removed from communities being studied. Most proposed experiments are concentrated not only within five Global North countries, but within five academic institutions (Cordenau-Huci et al. 2021). Regardless of an idea's source, prioritizing early and ongoing inclusion of community voices is the first imperative and challenge for the study (Davis 2020).[3] It is challenging to determine who counts as a relevant stakeholder; addressing questions of positionality, belonging, and "insider/outsider" status can be as complex and frustrating as they are essential (Kim et al. 2021).[4] Yet, all researchers planning work in the Global South — whether five thousand miles or five blocks away from home — should ask themselves tough questions before proceeding with any RCT (Cowen 2019): Is the question relevant to the communities where research will take place? Can I

[1] We also acknowledge that the majority of RCTs are conducted within the Global North.
[2] See online appendix for case studies detailing RCT experiences from authors and a question checklist for conducting RCTs in the Global South
[3] See Stakeholder Engagement case study in Appendix
[4] Decisions over who represents the "community" is not straightforward and often a site of contestation (Haas et al. 2021; Harrison & Michelson 2021).

answer this question without intervening in human lives? Are my career goals and research priorities leading me to overlook ethical or moral concerns? Answers to these questions may reveal that the costs of a study do not exceed the benefits.

Community collaboration during idea generation cultivates relevant research, shared interest in preserving long-term partnerships, and indicates respect for stakeholders. Occasionally, prioritizing ideas and research goals of in-country partners can come with trade-offs in later stages of research. Yet, the possibility of generating useful insights for host communities, while still providing theoretically important, internally valid academic contributions should be seen as central in modern social science research (Bleck et al. 2018).

Including community input is especially crucial when questions are generated externally. Such input can come from potential participants, activists, researchers, academics, policymakers, and more, and can orient researchers toward salient issues and challenges (Thachil & Vaishnav 2018, Asiamah et al, 2021). Over time, investments in community relationships allow for frank pushback, responsiveness, and active collaboration with in-country partners, which can improve design, measurement, conceptualization, and buy-in.[5]

Businesses, governments, international agencies, and civil society organizations often choose RCTs for program evaluation, viewing it as the "gold standard." Evaluation may be of a pre-planned program, or a novel design by the research and implementation teams. In either case, developing implementation partnerships requires evaluating the compatibility of objectives,

[5] See Election case study in the Appendix.

benefits, risks, and ethical concerns (Levine 2021). Researchers should assess the demand for evidence related to the proposed intervention, considering how their own inclusion could alter the program's design, data gathered, and overall goals of the project. Researchers should scrutinize implementing partners' track records and cease partnerships if they find abuses.

Planning

Planning an RCT in the Global South requires thoughtful assessment of the intervention and research protocols. If costs to participants and study communities outweigh potential benefits, researchers should not proceed with the project. Many experiments have distributional effects, reallocating benefits from one set of people to another. As a result, the benefits and harms of an experiment between groups are often in conflict.[6] Resolving these tensions requires value judgments, such as prioritizing the interests of citizens over politicians. We suggest researchers consult and engage with a wide array of stakeholders when making these choices, and report on the values they discover, as well as their own in writeups.

Researchers can consider relevant stakeholders as persons affected both directly and indirectly. The identity of those directly affected depends on research design, including sampling, treatment assignment, and measurement. This includes persons responsible for administering treatments and collecting data. It is also important to consider individuals who are indirectly impacted due to their links to those directly affected (Humphreys 2015), including via social, geographic, political, and economic networks. This includes those potentially affected by release of study

[6] For example, experiments that publicize politician scorecards help politicians with sterling records and the constituents of poorly-behaved politicians who can then vote them out, but they harm politicians with poor records (Grossman and Michelitch 2018).

results. Talking to potential research participants in pre-experiment surveys and focus groups, as well as local experts, may be needed to identify the set of affected groups and the benefits and risks they face.

Ex-post cost-benefit analyses of interventions are increasingly common (Brown & Tanner 2019) based on the program's costs and experiment's estimate of benefits. These tools could be adapted to provide informed ex-ante approximations of costs and benefits across different experimental designs, guiding decisions of whether to undertake the study. Costs should include not only those related to the intervention, but to the experiment itself, including to participants and persons indirectly affected. Benefits should include welfare changes caused directly by the program and experiment, as well as subsequent outcomes, including policy change if programs are demonstrated effective and adopted more widely. Values ascribed to the status quo can be elicited from community stakeholders before research begins to identify how to best maximize benefits, minimize harms, and build in risk-reduction strategies.

In such cost-benefit analyses, various value dimensions and to whom they accrue should be considered. Scientific value can be assessed by potential learning relative to past studies; where publication bias is not an issue, this can be done using meta-analysis. Simulation-based power analyses can quantify expected learning (Blair et al. 2021).

Additionally central to planning an RCT is considering what and how to randomize—as well as what cannot and should not be randomized. Many persons in highly-studied communities understand aspects of experimental design and may view intervention as unfair or adopt negative

attitudes if they perceive themselves as excluded from a study and its benefits (Karim 2020). Alternatively, randomization can be seen as a fairer allocation mechanism than other methods. Although such circumstances are relatively rare, when interventions are known to have positive impacts validated across a range of contexts, randomization by researchers constitutes withholding known benefits and is thus unethical.

Design of research instruments should also incorporate stakeholder feedback to minimize harm, meet scientific goals, and maximize learning (Thachil, 2018). Supplementing surveys with qualitative data, including interviews, program facilitator notes, participant reflections, and psychophysiological data facilitates validation and depth.

Hiring and collaborating with local researchers and project managers furthers ground-up planning, including monitoring of research integrity and risks. Staff training should be expansive and responsive, with an eye toward developing sustainable skills for community partners and researchers. Training modules can be determined in partnership with enumerators and program staff to determine what is valuable in local labor markets. Researchers should create an open environment where all team members can share ideas and push back on designs that threaten scientific, practical, and ethical integrity.

Implementation

Things go wrong in the field, presenting both risks and opportunities. Researchers may have to halt or pivot studies at any point in implementation, so active monitoring is essential. This decision should consider that implementing partners may already be deeply involved,

participants' lives altered, and lessons generated to disseminate. Researchers can improve the quality and speed of identifying problems or unexpected opportunities by centering community voices throughout. Consulting with participants, with the understanding that power imbalances may affect willingness to disclose harm, can help researchers identify whether red-lines have been crossed.

Addressing ethical implications and positionality is a dynamic process throughout the project life cycle, especially as contextual or circumstantial shifts change implementation plans. Violence makes it unsafe to collect data; new leadership of a partner organization cancels a treatment arm; a global pandemic multiplies stay-at-home orders for involved stakeholders. Such internal and external factors can often threaten completion of an experiment. In such cases, what should researchers do?

First, researchers should not simply leave. Foreign researchers often have greater travel flexibility, including diplomatic support for evacuations and financial resources. However, they also have responsibilities to implementation partners, research staff, and RCT participants. Even if they halt a study, researchers should continue to provide promised support, responsibly hand over ongoing duties,[7] and share insights generated thus far. Second, researchers should consider how to leverage resources for research staff. Recommendations and connections to facilitate future jobs, creating mechanisms to safely and securely transfer outstanding funds, and establishing a network for all partners can help preserve benefits. Further, researchers should determine how they can engage and support ongoing concerns even after departure. Finally,

[7] See WhatsApp case study in Appendix.

sharing what can be learned from a project that "fails" at the implementation stage is imperative. This can be done by sharing protocols, survey instruments, and/or a short report on the failure itself to provide future recommendations. Silent failures beget repeats.[8]

Throughout, research teams should reflect on positionality and center participant well-being, including mental health. This can be done by conducting additional analyses to understand context-specific factors among vulnerable communities beyond IRB standards (Cronin-Furman & Lake 2018) and ensuring relevant care is accessible to participants throughout with resources such as referral cards and connection to reliable healthcare providers. Local partners can highlight potential problems of accessing care, including stigma or logistic barriers (Khedari et al. 2021). Care for physical and mental health of research staff should also be prioritized during planning and budgeting, via longer data collection schedules, regular debriefing, and providing mental health counselors.

Scientific Communication

A significant justification for RCTs is they generate credible causal evidence that can inform policy. Therefore, the ability to clearly communicate research findings to policymakers and affected communities is central to the ethical and empirical frameworks guiding RCT design and implementation. Effective scientific communication should target key stakeholders—including participants, communities directly and indirectly affected, policymakers, implementation partners, and in-country academics—with tailored outputs, including memps, pamphlets, and

[8] See Learning from Failure case study in Appendix.

short videos, publicized via local partners and media (Thachil & Vaishnav, 2018). Findings should be presented throughout the research cycle using context-specific language and with comprehensible research transparency. Research outputs relevant to study communities may differ from academic outputs: reporting on secondary analyses may be as useful as the experimental results themselves (Herman, 2021). Funding and capacity for communication efforts should be incorporated into budgeting and hiring considerations from the outset. The creation of compensated advisory committees with community stakeholders can provide feedback on inputs and outputs at regular intervals and help with dissemination throughout the research cycle.

Conclusion

Assessing risks, localizing benefits, and addressing positionality requires having an explicit understanding of what not to do. Research questions that do not reflect priorities of community stakeholders, designs that fail to align with local contexts, or an over-reliance on implementation partners and prioritization of scientific validity (Bedecarrats et al. 2019) can all serve as potential barriers to effective and ethical RCTs. Time horizons between researchers eager to register outputs may conflict with policymakers who want longitudinal insights. Journal reviewer priorities may misalign with policy applicability. Reports that use disciplinary jargon may dissuade non-academic audiences from engaging. Centering community contexts, stakeholders, and demands at each research stage is key to ensuring that RCTs in the Global South are ethically sound and generating insights that serve the populations they investigate.

For the above changes to take hold within academia, it is incumbent on journal reviewers, tenure committees, and other institutions of academic power to institutionalize incentives that center collaborative processes. Reporting guidelines that document who was consulted at each research stage should be instituted for papers, including discussion of ex-ante cost-benefit analyses and steps taken to identify and mitigate risks. Institutional change is vital, as individual researchers' efforts cannot alone produce changes to entrenched norms.

References:

Asiamah, G., Awal, M. & MacLean, L. (2021) Collaboration for Designing, Conceptualizing, and (Possibly) Decolonizing Research in African Politics. PS: Political Science & Politics.

Banerjee, A. & Duflo, E. (2009) The Experimental Approach to Development Economics. Annual Review of Economics 1(1) 151-178.

Bedecarrats, F. and Guerin, I. and Roubaud, F. (2019) All that Glitters is not Gold: The Political Economy of Randomized Evaluations in Development. Development and Change, 30(3) 735-762

Bleck, J., Dendere, C. & Sangaré, B. (2018) Making North–South Research Collaborations Work. PS: Political Science & Politics, 51(3).

Blair, G., Coppock, A. and Humphreys., M. *Research Design: Declaration, Diagnosis, Redesign*. Book manuscript.

Brown, E. D., and J. C. Tanner. (2019). "Integrating Value for Money and Impact Evaluations." World Bank Policy Research Working Paper 9041.

Corduneanu-Huci, C., Dorsch, M. T., & Maarek, P. (2021). What, where, who and why? An Empirical Investigation of Positionality in Political Science Field Experiments. PS: Political Science and Politics, Forthcoming.

Cowen, N. (2019). For Whom Does "What Works" Work? The Political Economy of Evidence-Based Education. Educational Research and Evaluation 25(2) 81-98

Cronin-Furman, K., & Lake, M. (2018). Ethics Abroad: Fieldwork in Fragile and Violent Contexts. PS: Political Science & Politics, 51(3).

Davis, J. Manipulating Africa? (2020) Perspectives on the Experimental Method in the Study of African Politics, African Affairs, 119 (476) 452-467

Davis, J. & Michelitch, K. (2021). Field Experiments: Thinking through Identity and Positionality. PS: Political Science and Politics, Forthcoming.

Fujii, L. (2012). Research Ethics 101: Dilemmas and Responsibilities. PS: Political Science & Politics, 45(4), 717-723.

Grossman, G., & Michelitch, K. (2018). "Information Dissemination, Competitive Pressure, and Politician Performance between Elections: A Field Experiment in Uganda." American Political Science Review, 112(2): 280-301.

Haas, N., Haenschen, K., Kumar, T., Panagopoulos, C., Peyton, K., Ravanilla, N., & Sierra-Arévalo. (2021). Organizational Identity and Positionality in Randomized Control Trials: Considerations and Advice for Collaborative Research Teams. PS: Political Science & Politics, Forthcoming.

Harrison, B. & Michelson, M. (2021). LGBTQ Scholarship: Researcher Identity, RCTs, & Ingroup Positionality. PS: Political Science & Politics, Forthcoming.

Herman, B. (2021). Responses to Trauma: the Consequences of Forced Migration on Social Stability Among Syrian Refugees in Jordan, Working Paper.

Kaplan, L., Kuhnt, J. & Steinert, J. (2020) Do No Harm? Field Research in the Global South: Ethical Challenges Faced by Research Staff. World Development. 127.

Karim, S. (2020). Relational State Building in Areas of Limited Statehood: Experimental Evidence on the Attitudes of the Police. American Political Science Review 114(2), 536-551.

Khedari, V., Niconchuk, M., Beranbaum, S., Herman, B., Alkareem, Z., Jaber, B., & D'Andrea W. (2021) The Field Guide for Barefoot Psychology: Impact of a Biology-centered Psychoeducation Curriculum on Refugee Mental Health and Stigma. Working Paper.

Kim, E., Badrinathan, S., Choi, D., & Zhou, Y. (2021). Navigating 'Insider' and 'Outsider' Status as REsearchers Conducting Field Experiments. PS: Political Science & Politics, Forthcoming.

Levine, A. S. (2021). How to Form Organizational Partnerships to Run Experiments. Advances in Experimental Political Science, 199-216.

Mwambari, D. (2019). Local positionality in the production of knowledge in Northern Uganda. International Journal of Qualitative Methods, 18.

Panin, A. (2020). Economics Experiments in Africa: How Many and by Whom? Codesria Bulletin 4: 23-28

Porteous, O. (2020). Research Deserts and Oases: Evidence From 27,000 Economic Journal Articles on Africa, Working Paper

Sheely, R. (2018). Institutional Diversity and Public Goods Maintenance: Experimental Evidence from Kenya, Working Paper

Teele, D. (2014). Reflections on the ethics of field experiments. Field experiments and their critics: Essays on the uses and abuses of experimentation in the social sciences, 115-140.

Thachil, T. & Vaishnav, M. (2018) The Strategic and Moral Imperatives of Local Engagement: Reflections on India. PS: Political Science & Politics, 51(3).

Thachil, T. (2018). Improving surveys through ethnography: Insights from India's urban periphery. Studies in Comparative International Development, 53(3), 281-299.

Online Appendix:

A. Case Studies

Election Case Study: Benefits and Obstacles

Due to long-standing research networks and extensive connections in a West African country by a team member, an in-country partner was able to approach researchers with ideas for a collaborative program prior to an upcoming election. Because the idea was generated directly from the relevant experience of in-country partners within a state bureaucracy, the research team was able to conduct a field experiment with communities who are often difficult to access. In this way, academic researchers gained unique insights and were able to use their skills and training to help ensure the neutrality and quality of the field experiment, while in-country partners were also able to build capacity, develop skill sets, and address issues that were most important to them at that time. Further, after the experiment was completed, the in-country partner was able to leverage the support of the internationally-based research team members to expand elements of the project that were deemed particularly successful and helpful.

However, the trade-offs within the project are also noteworthy. From the outset the project suffered from a lack of funds. Even after multiple conversations, international donors were skeptical of a project that did not necessarily fall in line with their specific interests, goals, and networks. Without these funds, the field experiment itself was difficult to scale up to a large enough sample size. Nevertheless, this case is illustrative of how researchers with access to

funding, grant writing experience, and useful methodological training can lend their skill sets to support mutually beneficial collaborative research projects.

Stakeholder Engagement in Policy Research

As part of a large team of researchers evaluating different aspects of policy reform in an eastern African country, one author has had first hand experience on the benefits of creating partnerships and communicating frequently with stakeholders with a view to producing actionable knowledge. This study involved working with government bureaucrats at the national and local levels, recipients of public services, representatives of bilateral and multilateral donors, and an international project implementing organization.

Core to the stakeholder engagement strategy was the creation of a Steering Committee, which met at least once a year for the duration of the research period. At first the benefit of a Steering Committee was not readily apparent to the entire research team. However, over time, the research team was appreciative that the funder prioritized serious stakeholder engagement from the outset. It sharpened the research questions asked, forced close attention to be paid to the political and institutional context, and enabled anticipation of potential challenges in research implementation. Most importantly, it increased the relevance of the overall research.

The Steering Committee consisted of representatives from all key stakeholders groups - including different government departments, academics from the main national university (some of whom were part of the research team), grassroots civil society groups, and a leading national think tank. These stakeholders were selected on the basis of the objective of the study - to understand the challenges in a specific sector, with a view of informing government reform efforts. At meetings, policy and research questions relevant in stakeholders' daily work were solicited, and ideas and findings from the ongoing research were shared by the research team.

An interesting byproduct of this arrangement was the fact that senior government officials, bureaucrats, and the beneficiaries of public services would routinely discuss the research design and findings from different perspectives. These discussions enriched both the research team's study design, including its experimental component and the quantitative and qualitative work done using administrative data. Finally, the research team's rapport with government officials and other stakeholders ensured research continuity even as various government officials were transferred in the middle of the study period.

Leaving the WhatsApp Group: A Failure to Communicate

After two months of fieldwork in a West African country, one author got on a plane back to the Global North and replaced her SIM card with her usual Global North service provider. This

meant that she no longer received notifications from the WhatsApp group that project staff had been using to communicate on the ground. She assumed they would be able to contact her by email. It later turned out that project staff had assumed she would stay part of the WhatsApp group and had been sending questions and updates to the group. Because of this miscommunication, issues that arose after the main fieldwork had been implemented went unaddressed for days (e.g. payments to implementing partners that hadn't been properly processed, questions from research participants).

Learning from Failure

Examples of reporting on the failure of an RCT include the World Bank's Development Impact blog (https://blogs.worldbank.org/impactevaluations/failure), which maintains a page of "Learning From Failures." Researchers submit short reports about RCTs that went wrong and lessons learned.

B. Checklist

In this checklist, the authors gather the issues and suggestions discussed into a series of questions that researchers can interrogate while designing, planning, implementing and communicating about RCTs in the Global South. This checklist can be seen as supplemental to the more detailed discussions in the paper.

1. Have I sought input on my research questions and created spaces where critical feedback throughout the research process can be received from a variety of stakeholders—potential participants, activists, researchers, academics, policymakers—whose positionalities with respect to insider/outsider status or local/foreign identity are different from mine?

2. Have I assessed the demand for evidence related to the proposed study, the value of the study as designed, and weighed it against the impacts of the proposed study, particularly to communities directly affected by the study?

3. Have I considered alternatives to and evaluated potential costs of randomization—such as allocations of resources that might be considered unfair or the reduced autonomy of participants? Have I established red-lines around issues that should not be randomized and have I ensured that my study design does not cross these lines?

4. Have I ensured that training of research staff is expansive and responsive, with an eye toward developing sustainable skills for local partners and communities?

5. Have I planned for adequate support of research staff throughout the project cycle, including adequate compensation, manageable work schedules, and access to mental health care?

6. Have I planned ways to disseminate insights and findings from the project throughout the entire research cycle, not only at the end?

7. Does my project have systems in place to continuously receive feedback that can help me evaluate whether it is safe and prudent to continue data collection?

8. Have I set up systems and clearly communicated how project staff can easily contact me after I leave the field?

9. Have I used my power and connections to generate a measure of safety and continuity for research staff? For example, have I written reference letters or provided clear documentation of their participation in the projects and outputs on the research?

10. In case a project "fails," have I shared findings generated—such as protocols and survey instruments, or a short report on the failure itself that provides recommendations for addressing research limitations and challenges?

11. Have I prepared my results to be disseminated in a variety of non-academic channels, including by creating narratives, pamphlets, or short videos and publicizing via local organizations and media?