

Can Randomized Controlled Trials Be Remedied?

Rabab El-Mahdi & Samer Atallah

As the Middle East and North Africa (MENA) is becoming a lab for randomized controlled trials in the social sciences, this essay reflects on the use of this methodology in the region. Reflecting on earlier critiques of the method's deployment in the Global South, we argue that the perils of its use outweigh the benefits. Unlike many of the existing critiques, we maintain that this methodology's intrinsic flaws are further exacerbated in the context of MENA, and that its shortcomings cannot be mitigated by safeguards. So instead of trying to further refine it, researchers should seek alternative methods.

As corporations move away from testing products on animals, many social science researchers are taking pride in testing the impact of proposed socioeconomic interventions on humans through randomized controlled trials (RCTs). Introduced by the Abdul Latif Jameel Poverty Action Lab (J-PAL) as the best way to find efficient poverty reduction interventions, RCTs have been widely adopted in development economics and are advancing in other social sciences as the gold-standard methodology for evidence-based findings.¹ This advent of RCTs is now reaching the Middle East and North Africa (MENA), where J-PAL established a regional office in 2020. Reflecting on the ethical concerns and power dynamics underlying the use of this methodology, we question its usefulness for advancing knowledge in and about MENA, whether its benefits outweigh its shortcomings, and if the latter could be mitigated.

We argue that despite their adherence to principles of ethical conduct and institutional review board (IRB) protocols from *The Belmont Report*, the increasing use of RCTs will cause more harm than good within the MENA research enterprise, further warping knowledge about the region. Ontologically, RCTs adopt a “medical gaze” reenacting the doctor-patient power dyad, which further exacerbates the already skewed parameters of how knowledge is being extracted, filtered, framed, and used in MENA.² Built into that dyad is the distinction between the researcher as a holder/fixer of knowledge and the researched as the unknowing/sick to be “treated.” Devoid of the necessary understanding of power complexities, the use of RCTs reflect a technicalized worldview stemming from the assump-

tion that human behavior and social relations can be “fixed” by the right technical “treatment” and that the “doctor/expert” would know and test that treatment. In that sense, the researchers test “treatments” that they choose, using the people they include at random, then produce what they deem “objective” generalizable knowledge.

In MENA, this medical gaze is further compounded by the “Orientalist gaze,” where the knowledge of and about the region has long been and continues to be a product of its objectification by the colonial oppressors and their “civilizing” mission.³ With RCTs’ benefactors and owners located in powerful institutions of the Global North, having no prior knowledge of the region, and in most cases even conducting their trials remotely, this work risks enabling another version of the extractive knowledge production projects associated with colonialism. The produced knowledge often interprets the region through unknowing eyes, and so, in the existing context of the policing state(s), where citizens of the region are restricted in terms of what and how information is used to create dominant narratives about their cultures and societies, the use of RCTs further compartmentalizes and decontextualizes knowledge about them and their lives.

In the past two decades, the use of RCTs in the social sciences has sparked a heated debate about the advantages and disadvantages of this methodology, which MENA social scientists will now need to think intently about. Despite ethical and methodological concerns such as withholding of “treatment” options from control groups and limitations on the validity and scalability of their findings, the use of RCTs has been on the rise in the Global North and is catching up in MENA.⁴ Advocates emphasize a number of strengths, including requiring little or no prior information and minimal underlying assumptions while enjoying independence from “expert” knowledge, which can sometimes be manipulative or politically biased, as well as increased immunity to research or publication bias since trial registration and pre-analysis plans are often mandatory.⁵ In order to benefit from these advantages, others have been pushing for the development of safeguards that might minimize the potential negative implications of RCTs.⁶ Within this debate, we argue that the inherent flaws of this methodology cannot be remedied by safeguards and that we should be very wary of its use in the social sciences in MENA.

Powerful benefactors – a rich university in the Global North, a donor agency, a government ministry, an international nongovernmental organization (INGO), or a private corporation – have traditionally used experiments in general and RCTs in particular on the poor, consolidating the inherent skewed power dynamic inherent in its doctor-patient dyad. As economist Sanjay Reddy notes, “nearly every RCT involves treating poor people, usually also in poor countries.”⁷ A random sample of 130 interventions in low- and middle-income countries shows that “50% of all authors, and 59.2% of first authors are from countries in North America and West-

ern Europe.”⁸ In political science, examining the repository of MENA-related articles in the top journals of the discipline, the overwhelming majority of the twenty-two articles using experiments as a primary method were by authors based in the Global North.⁹ In line with this pattern, the use of RCTs in MENA has been patronized by the MIT-based J-PAL, which by 2024 had conducted around forty RCTs, constituting the overwhelming bulk of these kinds of experiments in the region.¹⁰ While most of the topics under investigation necessitate a deep understanding of the local dynamics (for example, sectarian relations in Iraq, or interpretations of social norms in Saudi Arabia), all forty or so studies were led by one of twelve hundred J-PAL researcher-affiliates based outside the region.¹¹

In 2020, despite not having any partners in the region, J-PAL established its MENA regional office (J-PAL MENA) at the American University in Cairo (AUC). Its expansive mandate includes conducting randomized evaluations, building partnerships for policymaking and scaling up programs, and offering training on RCTs. According to the J-PAL website,

Our research team evaluates the impact of social programs and policies in MENA, covering a wide range of sectors including social protection, employment, education, and gender. Through online and in-person courses, we train implementers, policymakers, donors, and advocates on how to generate and use rigorous evidence. Our policy team works to institutionalize learning from evidence and disseminate research results to governments and other partners.¹²

Yet, except for the director of the MENA office, all the lead researchers who would offer any of these services are researchers who live and work elsewhere and have almost no training or expertise in the region. Despite these limitations, in 2022, the J-PAL office at AUC established the Egypt Impact Lab (EIL) with the Ministry of Planning and Economic Development. In their own words, the EIL “works to build a culture of evidence-informed decision-making across government by building partners’ capacity to use evidence in program design and delivery and using administrative data to facilitate evidence generation.”¹³ The statement reflects the very common yet false belief that evidence is exclusive to the use of RCTs.

Even as RCTs take hold as the gold standard for conclusive findings, they distort knowledge production toward a particular set of researchers who can afford this methodology, which by definition is not MENA-based. Given the regional context, in which researchers suffer from lack of publicly available data, inadequate financial resources, and severely constrained academic freedom, the advent of costly RCTs will further bias the choice of research topics toward areas of interest to wealthy donors, governments, and researchers in big institutions. This could already be observed in J-PAL MENA, where the implemented RCTs have to be either initiated or picked up by one of their research affiliates – all of which are based outside the region – in case they’re supported by one of their local partners,

the majority of which are government ministries, UN agencies, and sometimes corporations.¹⁴ Hence, the knowledge generated and the evidence deployed may be, in fact, little more than the mirror image of the policy agenda of regimes or the funding priorities of donor institutions. At the same time, the favoring of RCTs in submissions to prestigious and widely cited journals disadvantages researchers who do not adopt it, whether as a matter of choice or due to the lack of access to research funding. This creates a vicious circle in which knowledge creation and dissemination can be done only by researchers with access to funding and who adopt RCTs, reflecting their own set of priorities and narratives about the region.

This institutional and epistemological power imbalance between those who initiate RCTs, the local research institutions, and the researched makes this methodology especially problematic in MENA. The region has long been defined through the “imperial gaze”: the privileged observer’s authority, domination, and control determine what is observed, and hence what is visible.¹⁵ This imperial gaze has dictated not only how the oppressed are seen but also how they perceive themselves, with the Western white subject being the normative standard. The extent to which this gaze plays a role in choosing the research topics and deciding what will be worth emphasizing cannot be underestimated.

Unlike claims about RCTs freeing the research from expert knowledge and *a priori* bias, the trials test a specific predetermined treatment of the researcher’s choice reflecting preconceived notions and, in the case of MENA, a predominant image about the region. For example, J-PAL’s only study on Saudi Arabia is led by three white researchers who are not MENA specialists, and who confidently establish associations about social norms and female labor participation based on their slim sample of five hundred Saudi men.¹⁶ Focusing on female labor participation, in a country classically portrayed as not very women-friendly, the researchers opt to show how “correcting beliefs lead to more women working outside the home.”¹⁷ The choice of the topic and the trial’s findings center very specific cultural understandings about outside work being more advantageous and ignore a whole range of political and economic factors that affect that decision. Rather, it chooses to start from and spotlight social norms, a topic long fetishized in Western studies of MENA. In that regard, the use of RCTs allows for this kind of reductionism, and gives it a scientific allure.

Similarly, another study evaluated the impact of mixed Christian-Muslim soccer teams on social cohesion and interactions between these groups in ISIS-affected Kurdish areas of Iraq.¹⁸ The researcher framed the issue as tensions and mistrust resulting from differences in religion, dismissing other equally plausible explanations. For instance, since the majority of the Christians in these Kurdish areas are displaced from Arab regions of Iraq, ethnic rather than religious tensions might have been at play. Or perhaps the lack of trust was simply one between

longtime inhabitants and newcomers, with no religious or ethnic dimension to it. By its design, however, the experiment eliminated these very plausible explanations before they were tested.

Even when the research is thought to be in partnership with a local collaborator, the *a priori* knowledge and “expertise” of outside researchers always take precedence over that of the local. In an example reported by J-PAL cofounders Esther Duflo and Abhijit Banerjee, a local NGO sought out their expertise, and they changed the intervention originally proposed by the NGO:

when Seva Mandir, an NGO in Rajasthan, India, with whom we have had a longstanding relationship, was interested in improving the quality of their informal schools, their initial idea was to implement a teacher incentive program based on test scores. However, they were persuaded by the results from Glewwe et al. (2003) that showed that teacher incentives could result in teaching to the test or other short-run manipulations of test scores. They then decided to implement an incentive program based on teacher presence. To measure attendance in very sparsely populated areas where schools are difficult to access, [we] (Duflo and Hanna 2007) proposed the use of cameras with date and time stamps. Although Seva Mandir was initially surprised by the suggestion, they agreed to try it out.¹⁹

The authors do not reflect on the kind of power imbalances that might have pushed the local NGO to accept the suggested change. Whether it is the authority of expert knowledge, hopes for funding, the need for whatever the intervention offered, or even the effect of government involvement in RCTs, the power imbalance is too big to be ignored.²⁰ Yet, given the overly technical nature of RCTs and their anti-reflexivity, the researchers do not usually pause to reflect on such nuances.

This example reveals another defining flaw in RCTs: the lack of trust in the research participants (subjects) compared with the unbounded confidence that the researchers (experts) have in themselves. The intervention was changed based on the presumption that teachers would deliver “teaching to the test or other short-run manipulations.” Accordingly, the intervention became essentially the surveillance of research subjects who could not be trusted to deliver quality education otherwise. Due to that lack of trust in the research participants, an invasive and demeaning treatment – “the use of cameras with date and time stamps” – was proposed. Banerjee and Duflo do not seem to have stopped to reflect on the implications of the fact that, as they report, Seva Mandir was “initially surprised by the suggestion.”²¹ Nor do they do so in hindsight; when analyzing this case, the NGO’s reaction is fleetingly mentioned and quickly dismissed.

The inherent belief in RCTs that research participants are not fully competent is also reflected in policies that authorize withholding proposed treatment details to prevent what is known as the Hawthorne effect, a phenomenon whereby subjects

modify their behavior when they know they are being studied. David B. Cameron, Abhijit Mishra, and Alan N. Brown's examination of interventions in low- and middle-income countries found that only 35 percent of the published papers mention whether people are aware of being part of an experiment.²² While researchers are expected to be truthful and forthcoming in theory, they are often trained and incentivized to be dishonest in practice for the experiment to yield the best results.

In isolating measurable variables to establish causality, RCTs often blur "the distinction between evidence and policy" to validate a particular view of the world.²³ While Duflo urges the use of RCTs as a device by which to tinker in crafting tailored solutions to policy problems, the working of RCTs can easily gloss over precisely the details that make such customization possible. Like any experiment, an RCT also stipulates that the average difference between the observed outcome in a treatment group and that of the control group is a result of the intervention commonly referred to as the treatment effect.²⁴ Based on the assumptions of randomization among the two groups and the linearity of expectation, the estimated difference is then assumed to be driven by the treatment, which, as we have seen, is not necessarily true.²⁵ Moreover, the treatment effect is an average that does not capture the heterogeneous effect within the treatment group. The underlying assumption of randomness leads to estimating the mean treatment effect. It basically suggests what intervention works, on average, without in fact proposing or detailing what Duflo claims to be after: the mechanism of *how* it works. Given that RCTs are inherently less constrained than nonrandomized methods in their design, establishing causality between the intervention and the observed outcome is even more problematic. At the same time, the rigid structure of RCTs offers a reductionist understanding, neglecting the nuanced, contextual factors underlying any intervention in MENA.

In that sense, the RCT is not simply an objective scientific methodology as its proponents claim; it is a political and ethical choice advancing a technicalized apolitical worldview under the false pretense of scientific merit.²⁶ Urging economists to take a detail-oriented approach similar to those of more practical professions like plumbing, Duflo, the pioneer and champion of RCTs, explains,

the economist-plumber stands on the shoulders of scientists and engineers, but does not have the safety net of a bounded set of assumptions. She is more concerned about "how" to do things than about "what" to do. In the pursuit of good implementation of public policy, she is willing to tinker. Field experimentation is her tool of choice.²⁷

In drawing the distinction between the "how" and the "what" in the economist's mandate, Duflo ignores the underlying "why," reiterating a long-held neoliberal doctrine that assumes there is an ultimate economic technocratic fix waiting to be discovered. From this perspective, only "field experimentation" counts, not ethnographies, life histories, or interviews because only experiments involve a "tested

treatment.” Her analogy is built around those with technical expertise: engineers, scientists, and plumbers. Absent from Duflo’s image is a storyteller, a philosopher, a painter, or a politician, because the power of interpretation, imagination, and politics are forcefully removed from this technicalized vision of the world.

The misalignment of incentives among the intervention’s researchers, funders, and owners also raises a number of ethical concerns. Researchers may be driven by academic or career motivations; funders may prioritize cost-effectiveness or scalability; and RCT owners, often local governments or INGOs, may be seeking political gain or fulfilling donor requirements. These divergent goals can lead to interventions being implemented in ways that prioritize many other interests over the needs and rights of local populations. A J-PAL study in Morocco, where the researchers partnered with the local affiliate of an international, private utility company to sell in-home water connections to low-income households in Tangier, illustrates many of these ethical questions.

The J-PAL researchers framed the study as an evaluation of a social program that “offered low-income households a chance to get an in-home water connection”; the “treatment” constituted what amounts to a sales campaign for this corporation.²⁸ Blurring the line between marketing and impartial research, the treatment offered by the research team included:

- Preapproval for the loan from the authorities;
- Project officers assisting households in preparing all paperwork and identification documents required for the application;
- Branch officers visiting households to collect the down payment; and
- A door-to-door awareness campaign to encourage households to buy a connection.

The benefactor was a private French corporation and the so-called study was facilitating “a connection to the water and sanitation network at full price,” which is “a substantial expense for these households.”²⁹ The “treatment” could easily implicate these local households in unsustainable financial burdens and debt, and in absence of strong consumer-rights groups and community organizations, the influence of the corporation and its hired research team was not counterbalanced by any other entity. Yet the RCT was guided by an IRB and none of the involved found that it conflicted with the principles laid out in *The Belmont Report*, which lists respect for persons, beneficence, and justice as the three pillars of ethical RCTs.³⁰

Using the three principles in *The Belmont Report*, which stem from the governing of medical trials, the research enterprise has developed guidelines for researchers using RCTs. For example, “respect for persons” has been operationalized into “informed consent,” which is a necessity in the majority of these trials.³¹ But in most

MENA countries, as in the Moroccan case, there are significant disparities in power between the local populations and researchers that cast doubt on their ability to willingly offer consent. The RCT researchers are often backed by a government agency or an international institution and come with some kind of material offer, be it the water connection in Morocco or joining a football team in Iraq, making it difficult for these populations to decline. Oftentimes, vulnerable communities are incentivized into participating in trials such as the one in Morocco without fully understanding the implications. For example, in Egypt, a sizable number of all RCTs testing social provision programs are conducted by EIL. Actual and hopeful beneficiaries of these programs do not have the leverage to decline the related RCTs, for fear of being excluded from those who receive some provisions. In such contexts, the ethical principle of informed consent is compromised, raising serious concerns about its validity.

Similarly, beneficence and justice – the two other ethical principles that are supposed to guide RCT ethics – ultimately depend on the researchers' individual judgment and conscience, which again leaves the research participants beholden to the will of the benefactor, who is also an extractor. Beneficence directs researchers to “do no harm” and “maximize benefits, minimize risks.” Justice deals with participant selection and requires that those who take the risks should receive the benefits.³² Yet while harm in medicine can usually be physically identified and quantified, in social relations, it is not so clear. For example, many would argue that subjecting refugees to narratives about the ongoing civil strife that uprooted them might be traumatic. But in an article published in the *American Journal of Political Science*, one of the leading journals in the field, the authors conducted an experiment to test the impact of competing narratives about the Syrian civil war on Syrian refugees in Lebanon.³³ For the authors, the IRB at their institution, and the editors and reviewers of the journal, this experiment seems not to have breached the “do no harm” principle, despite the obvious likelihood that exposure to messages about the war might cause emotional distress or psychological harm. That risk of triggering the trauma was not counterbalanced by any benefit for the participants.

Given the wide and sometimes divergent interpretations of these ethical considerations, it is not clear how the risks and ethical concerns associated with RCTs could be remedied by developing guidelines or more safeguards. This is further complicated considering that the ethical questions manifesting in RCTs are very different and specific to the trial and the context in which they are conducted, and cannot be subsumed under a set of fixed safeguards. The futility of attempts to develop guidelines is apparent in current shortcomings of the IRBs, which are the most important safeguard of RCTs.

Rather than protecting the research participants or subjects, IRBs “provide legal protection for institutions.”³⁴ With the power balance already in favor of the bene-

factor, the only technical safeguard available is designed to work to their advantage. IRBs “rely on self-regulation and trust,” which in turn “rely almost entirely on information (on risks, etc.) submitted to it by the researchers themselves.”³⁵ This self-regulation and trust do not wane, despite the multiple times when “researchers appear to understand that they are on sticky ground, but instead of reconsidering the intervention, they are at pains to find a way around it.”³⁶ So, instead of attempting to go beyond IRBs and adopt more ethical guidelines, social scientists should consider replacing RCTs with more reflexive and less rigid methodologies.

There is a need for context-sensitive, ethically sound social research in the Middle East and North Africa that respects the complexities and specificities of the different peoples and societies of the region. While RCTs are widely regarded as the gold standard in evaluating interventions, their use does not meet that need. Assuming that an intervention from an Indian village or a Brazilian town can be “translated” into a policy initiative in Morocco or Lebanon is basically to assume that the poor are the “same” everywhere by virtue of their poverty. Such homogenization not only dismisses the context in which the tinkering that justifies RCTs takes place but also flattens the knowledge and narratives about the people subjected to the study. People are reductively defined by their poverty without regard to the many other dimensions of their existence or what got them there. These flaws are not remediable by using stopgap guardrails, or by resorting to exhortations to transparency in research design. Understanding the kinds of policy intervention that might improve the lives of the people of MENA will require much more subtle, nuanced research – work that actually reflects respect, beneficence, and justice.

ABOUT THE AUTHORS

Rabab El-Mahdi is Associate Professor of Political Science at the American University in Cairo and Founding Director of Alternative Policy Solutions, a research center focused on producing progressive alternatives to neoliberal policies in Egypt and the Global South. She is the author of *Political Manipulation? Civil Society and the State in Egypt and Bolivia* (2012) and editor of *Arab Spring in Egypt: Revolution and Beyond* (with Bahgat Korany, 2012) and *Egypt: The Moment of Change* (with Philip Marfleet, 2009).

Samer Atallah is Associate Professor of Economics at the Onsi Sawiris School of Business at the American University in Cairo. He has published in such journals as *European Journal of Political Economy* and *International Journal of Development and Conflict*.

ENDNOTES

- ¹ Founded in 2003 at the Massachusetts Institute of Technology by Nobel Prize laureates Abhijit Banerjee and Esther Duflo, J-PAL is a global research center working to reduce poverty by ensuring that policy is informed by scientific evidence anchored by a network of more than twelve hundred researchers around the world.
- ² Michel Foucault describes the “medical gaze” as the lens doctors use to fit a patient’s story into a biomedical paradigm, filtering out what they deem as irrelevant information, highlighting the dominance of medical knowledge in society and the power dynamics between doctors and patients. Michel Foucault, *The Birth of the Clinic: An Archaeology of Medical Perception* (Pantheon Books, 1973).
- ³ The Orientalist gaze refers to how the “Orient” has been historically constructed in opposition to the “Occident” by its Western colonizers. Edward W. Said, *Orientalism* (Pantheon Books, 1978).
- ⁴ Stéphane J. Baele, “The Ethics of New Development Economics: Is the Experimental Approach to Development Economics Morally Wrong?” *Journal of Philosophical Economics* 7 (1) (2013): 10653; Rose McDermott and Peter K. Hatemi, “Ethics in Field Experimentation: A Call to Establish New Standards to Protect the Public from Unwanted Manipulation and Real Harms,” *Proceedings of the National Academy of Sciences* 117 (48) (2020): 30014–30021; and Reetika Khera, “Some Questions of Ethics in Randomized Controlled Trials,” *Review of Development Economics* 28 (4) (2023): 2072–2087.
- ⁵ Abhijit V. Banerjee and Esther Duflo, “The Experimental Approach to Development Economics,” *Annual Review of Economics* 1 (1) (2009): 151–178; Angus Deaton and Nancy Cartwright, “Understanding and Misunderstanding Randomized Controlled Trials,” *Social Science & Medicine* 210 (2018): 2–21; and Esther Duflo and Abhijit V. Banerjee, *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty* (Public Affairs, 2011).
- ⁶ For example, Scott Desposato, “Subjects and Scholars’ Views on the Ethics of Political Science Field Experiments,” *Perspectives on Politics* 16 (3) (2018): 739–750; and Trisha Phillips, “Ethics of Field Experiments,” *Annual Review of Political Science* 24 (1) (2021): 277–300.
- ⁷ Sanjay G. Reddy, “Economics’ Biggest Success Story Is a Cautionary Tale,” *Foreign Policy*, October 22, 2019, <https://foreignpolicy.com/2019/10/22/economics-development-rcts-esther-duflo-abhijit-banerjee-michael-kremer-nobel>.
- ⁸ See David B. Cameron, Abhijit Mishra, and Alan N. Brown, “The Growth of Impact Evaluation for International Development: How Much Have We Learned?” *Journal of Development Effectiveness* 8 (1) (2016): 18.
- ⁹ Melani Cammett and Isabel Kendall, “Political Science Scholarship on the Middle East: A View from the Journals,” *PS: Political Science & Politics* 54 (3) (2021): 448–455.
- ¹⁰ Interview with J-PAL MENA regional director and economist Ahmed El-Sayed, September 15, 2024, Cairo, Egypt.
- ¹¹ Ibid.
- ¹² “J-PAL Middle East and North Africa,” Abdul Latif Jameel Poverty Action Lab J-PAL, <https://www.povertyactionlab.org/middle-east-and-north-africa> (accessed September 17, 2024).

- ¹³ “Egypt Impact Lab : An Embedded Lab in the Ministry of Planning, Economic Development, and International Cooperation,” Abdul Latif Jameel Poverty Action Lab J-PAL, <https://www.povertyactionlab.org/egypt-impact-lab> (accessed September 4, 2024).
- ¹⁴ Interview with J-PAL MENA regional director and economist Ahmed El-Sayed, September 15, 2024, Cairo.
- ¹⁵ Edward W. Said, *Culture and Imperialism* (Vintage Books, 1994).
- ¹⁶ Leonardo Bursztyn, Alessandra L. González, and David Yanagizawa-Drott, *Misperceived Social Norms : Women Working Outside the Home in Saudi Arabia* (Abdul Latif Jameel Poverty Action Lab J-PAL, 2020).
- ¹⁷ *Ibid.*, 1.
- ¹⁸ “Inter-Religious Soccer Leagues to Promote Social Cohesion in Post-ISIS Iraq,” Abdul Latif Jameel Poverty Action Lab J-PAL, <https://www.povertyactionlab.org/evaluation/inter-religious-soccer-leagues-promote-social-cohesion-post-isis-iraq> (accessed September 5, 2024).
- ¹⁹ Banerjee and Duflo, “The Experimental Approach to Development Economics,” 155. See also Seán Mfundza Muller, “The Unacknowledged Normative Content of Randomised Control Trials in Economics and Its Dangers,” in *The Positive and the Normative in Economic Thought*, ed. Sina Badieli and Agnès Grivaux (Routledge, 2022), 167.
- ²⁰ Khera, “Some Questions of Ethics in Randomized Controlled Trials.” See also Banerjee and Duflo, “The Experimental Approach to Development Economics,” quoted in Muller, “The Unacknowledged Normative Content of Randomised Control Trials in Economics and Its Dangers,” 174.
- ²¹ Banerjee and Duflo, “The Experimental Approach to Development Economics,” 155.
- ²² See David B. Cameron, Abhijit Mishra, and Alan N. Brown, “The Growth of Impact Evaluation for International Development : How Much Have We Learned?” *Journal of Development Effectiveness* 8 (1) (2016) : 18.
- ²³ McDermott and Hatemi, “Ethics in Field Experimentation.” See also Drèze, “Evidence, Policy and Politics,” 9.
- ²⁴ Douglas MacKay, “The Ethics of Public Policy RCTs : The Principle of Policy Equipose,” *Bioethics* 32 (1) (2018) : 59–67.
- ²⁵ Michel Abramowicz and Anna Szafarz, “Ethics of RCTs : Should Economists Care about Equipose?” in *Randomized Control Trials in the Field of Development : A Critical Perspective*, ed. Florent Bédécarrats, Isabelle Guérin, and François Roubaud (Oxford University Press, 2020), 280, <https://doi.org/10.1093/oso/9780198865360.003.0003>. See also Angus Deaton, “Instruments, Randomization, and Learning about Development,” *Journal of Economic Literature* 48 (2) (2010) : 424–455 ; Lant Pritchett and Jessica Justin Sandefur, *Context Matters for Size : Why External Validity Claims and Development Practice Don’t Mix* (Center for Global Development, 2013) ; Hunt Allcott, “Site Selection Bias in Program Evaluation,” *The Quarterly Journal of Economics* 130 (3) (2015) : 1117–1165 ; Tessaim Bold, Mwangi Kimenyi, Germano Mwabu, and Justin Sandefur et al., “Experimental Evidence on Scaling Up Education Reforms in Kenya,” *Journal of Public Economics* 168 (2018) : 1–20 ; and Esther Vivalt, “How Much Can We Generalize from Impact Evaluations?” *Journal of the European Economic Association* 18 (6) (2020) : 3045–3089.

- ²⁶ Lisa Anderson, “Risk & Responsibility: Social Science Research as a Modern ‘Anti-Politics Machine,’” *Dædalus* 154 (2) (Spring 2025): 243–262, <https://www.amacad.org/daedalus/risk-responsibility-social-science-research-modern-anti-politics-machine>.
- ²⁷ Esther Duflo, “The Economist as Plumber,” *American Economic Review Papers & Proceedings* 107 (5) (2017): 3.
- ²⁸ The researchers use the words “offer” and “a chance to buy,” never the word “selling.” “Household Water Connections in Tangier, Morocco,” Abdul Latif Jameel Poverty Action Lab J-PAL, <https://www.povertyactionlab.org/evaluation/household-water-connections-tangier-morocco> (accessed September 5, 2024).
- ²⁹ Ibid.
- ³⁰ The first principle includes allowing people to decide for themselves what to participate in. The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research* (U.S. Department of Health, Education, and Welfare, 1979), <https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html>; and “Protection of Human Subjects, Code of Federal Regulations, Title 45, Part 46,” U.S. Department of Health and Human Services, <https://www.hhs.gov/ohrp/regulations-and-policy/regulations/45-cfr-46/index.html> (accessed September 5, 2024). See also Cameron, Mishra, and Brown, “The Growth of Impact Evaluation for International Development.”
- ³¹ The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, *The Belmont Report*. See also Cameron, Mishra, and Brown, “The Growth of Impact Evaluation for International Development,” 18.
- ³² Scott Desposato, “Public Impacts from Elite Audit Experiments: Aggregate and Response Delay Harms,” *Political Studies Review* 20 (2) (2022): 217–227. See also Cammett and Kendall, “Political Science Scholarship on the Middle East”; and Phillips, “Ethics of Field Experiments.”
- ³³ Daniel Corstange and Elizabeth A. York, “Sectarian Framing in the Syrian Civil War,” *American Journal of Political Science* 62 (2) (2018): 441–455.
- ³⁴ Khera, “Some Questions of Ethics in Randomized Controlled Trials,” 11.
- ³⁵ Ibid., 8.
- ³⁶ Ibid.