

Ogden, Timothy N. (2017), *Experimental Conversations: Perspectives on Randomized Trials in Development Economics*, Cambridge, MA: MIT Press.

In *The Logic of Scientific Discovery*, Karl Popper (1959) famously noted how

[A] theory of induction is superfluous. It has no function in a logic of science. [...] The best we can say of a hypothesis is that up to now it has been able to show its worth, and that it has been more successful than other hypotheses although, in principle, it can never be justified, verified, or even shown to be probable. (p.315)

For those of us who have studied the philosophy of science, the use of randomized controlled trials (RCTs)—or any other empirical method—to show what “works” in development has always been somewhat of a head-scratcher. After all, who is to say that a statistically significant finding is not simply the result of a type I error?

Timothy Ogden’s *Experimental Conversations*, a collection of interviews on the use of RCTs in economics—particularly in development economics—features no interview with anyone at an agricultural and applied economics department (or with anyone at a land-grant university, for that matter), and it cites only one article published in this journal.

So why should agricultural and applied economists read it?

First and foremost, because amid the many interviews with people whose positions and opinions on RCTs are well-known and need no further publicity, there are important insights, both for and against the modern-day use of RCTs. And perhaps more importantly, because Ogden’s questions and his interviewees’ answers are often surprisingly honest in admitting that RCTs are simply one tool among others useful to begin answering questions of the form “Does X cause Y?”

Many key players of the “credibility revolution” (Angrist and Pischke, 2010) in development economics are interviewed, but instead of having to look through a glass darkly and having to figure out what those key players are thinking through their academic output or popular-press accounts, the words of Michael Kremer, Abhijit Banerjee, Esther Duflo, Dean Karlan, Angus Deaton, Lant Pritchett, and Rachel Glennerster are presented here verbatim, or slightly edited from several conversations over the years. This question-and-answer style is beneficial in that what is typically seen as a heated debate is presented here in a more nuanced way. That said, nothing surprising is said by any of the usual suspects and, for the most part, they hold to their previously held positions and opinions.

It is thus no surprise that the most interesting of Ogden’s interviews are with unusual suspects. The interview with Judy Gueron of MDRC, a think-tank, is especially enlightening when Gueron, an avowed *randomista*, admits that RCTs should not determine policy, because policy ultimately depends on values and politics. Scientific evidence, however rigorous, is only one of the many inputs to inform policy decisions. Similarly, the interview with the Ford Foundation’s Frank DeGiovanni—an outsider to academic debates—discusses how RCTs are most useful when trying to effect a paradigm shift, but how they should also be accompanied by ethnographic evidence.

Of those who are actual RCT practitioners, Ogden often asks why they first started doing RCTs. The answers are pragmatic. Jonathan Morduch explains that he was motivated to move beyond the never-ending debates over the validity of instrumental variables and onto more policy-oriented discussions.

Similarly, David McKenzie explains that his audience was often not satisfied with the non-experimental methods he presented, so he began to implement RCTs.

It would be wrong to mistake this pragmatism with the view that RCTs are somehow the “gold standard” of empirical strategies. However much this sentiment might have been promoted by some of the usual suspects over the last two decades, it is largely absent from Ogden’s interviews. Dean Yang not only suggests that specific statistical methods shouldn’t be blindly privileged over others, but also that RCTs aren’t for everyone. Successfully running an RCT, particularly in a developing country, requires specific skills—managerial ability, in particular—that are neither taught in economics or agricultural and applied economics graduate programs nor necessary for doing good research. From personal experience, we know how painful it is to invest in an RCT whose results simply do not yield any significant impact.

Finally, the interview with Tyler Cowen—a name not typically associated with RCTs—is especially useful in how it hints at the political economy of research surrounding RCTs. That is, anyone wanting to use a novel methodology (running RCTs over the past decade, or collecting panel data 25 years ago) faces high entry barriers, and the presence of those entry barriers leads to a centralization of influence.

Yet Ogden’s book is not without some blind spots and important omissions. Though the book portrays Michael Kremer as being the father of RCTs in development economics, the contribution of Immink and Viteri (1981a,b), who published a two-part article in the *Journal of Development Economics* 35 years ago in which they looked at the efficiency wage hypothesis in Guatemala using an RCT, is overlooked. More generally, R.A. Fisher’s (1935) agricultural experiments, which were conducted in the 1930s and were influential in the development of experimental social science, are only mentioned in passing.

Still, Ogden’s treatment of the subject remains refreshingly even-handed, and he himself admits that he went from being an RCT zealot to someone whose views are a bit more tempered. Ogden reflects that performing these interviews has made him more skeptical of *all* research, including the evidence generated by RCTs. In this sense the book leads to a worthwhile conclusion about the philosophy of science, and one that aligns with Popper’s argument: a theory can never be proven by empirical science; it can only be falsified. This critical reasoning is foundational to the vocabulary of hypothesis testing, but is often overlooked when we claim that RCTs or other methods allow us to know what “works” in development—or any other field of human inquiry.

Marc F. Bellemare and Jeffrey R. Bloem

University of Minnesota

References

Angrist, Joshua, and Jörn-Steffen Pischke (2010), “The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics,” *Journal of Economic Perspectives* 24(2): 3-30.

Fisher, Ronald (1935), *The Design of Experiments*, 9th ed., New York: Macmillan.

Immink, Maarten D.C., and Fernando E. Viteri (1981a), “Energy Intake and Productivity of Guatemalan Sugarcane Cutters: An Empirical Test of the Efficiency Wage Hypothesis, Part I,” *Journal of Development Economics* 9(2): 251-271.

Immink, Maarten D.C., and Fernando E. Viteri (1981b), "Energy Intake and Productivity of Guatemalan Sugarcane Cutters: An Empirical Test of the Efficiency Wage Hypothesis, Part II," *Journal of Development Economics* 9(2): 273-287.

Popper, Karl R. (1959), *The Logic of Scientific Discovery*, New York: Basic Books.