



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

The rise of the randomistas

On the experimental turn in international aid

Citation for published version:

Donovan, KP 2018, 'The rise of the randomistas: On the experimental turn in international aid', *Economy and Society*, vol. 47, no. 1, pp. 27-58. <https://doi.org/10.1080/03085147.2018.1432153>

Digital Object Identifier (DOI):

[10.1080/03085147.2018.1432153](https://doi.org/10.1080/03085147.2018.1432153)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Economy and Society

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



The rise of the randomistas: on the experimental turn in international aid

Kevin P. Donovan

Kevin P. Donovan, Program in Anthropology & History, University of Michigan, Room 1029, 435 South State St., Ann Arbor, MI, 48109-1003, United States. E-mail: kevinpd@umich.edu

Abstract

In recent years, the use of experimental methodologies has emerged as a central means of evaluating international aid interventions. Today, proponents of randomized control trials (so-called *randomistas*) are among the most influential of development experts. This article examines the growth of this thought collective, analysing how uncertainty has become a central concern of development institutions. It demonstrates that transformations within the aid industry – including the influence of evidence-based policy, the economization of development, and the retreat from macro-planning – created the conditions of possibility for experimentation. Within this field, the randomistas adeptly pursued a variety of rhetorical, affective, methodological, and organizational strategies that emphasized the lack of credible knowledge within aid and the ability of experiments to rectify the situation. Importantly, they have insisted on the moral worth of experimentation; indeed, the experimental ethic has been proposed as the way to change the spirit of development. Through causal certitude, they propose to reduce human suffering. The rise of experimentation has not, however, eliminated accusations of uncertainty; rather, it has redistributed the means through which knowledge about development is considered credible.

Keywords: experimentation; uncertainty; development; aid; evidence.

Introduction

In 2011, Abhijit Banerjee and Esther Duflo published *Poor economics* to widespread praise. The MIT economists are pioneers in the use of randomized control trials (RCTs) within international development. Their book, which promised ‘a radical rethinking of the way to fight global poverty’, is premised on the use of experimentation to assess the effectiveness of international aid programs. They use RCTs to measure interventions on populations randomly assigned to facilitate comparison across groups. For example, in a population of 500 poor, young men, half might receive financial literacy training while the other half do not, with researchers tracing differences in years to come. In the past fifteen years, the use of experimentation by social scientists studying poverty and development in the global South has boomed. Today it is a crucial epistemic practice that has been adopted throughout the academy and by NGOs, the World Bank, and governments. The proponents of the method - whom some have called the

randomistas (e.g., Ravallion, 2009a) - are today among the most influential experts within the aid world.

The history and politics of experimentation within medical science have been well-documented (Marks, 1997; Epstein, 2009; Timmermans & Berg, 2003; Lakoff, 2006). Yet, scholars have paid little attention to this methodology as a mode of knowing within international development nor as a political rationality in the global South, a phenomenon Petryna (2009) calls ‘experimentality’ (but see Rottenburg, 2009b; Berndt, 2015). Neither those voices critical of the shift (of which there are some), nor the literature produced by proponents (of which there is much), chart the rise of the *randomistas* as a sociological and political phenomenon (cf. Teele, 2014). This paper analyzes the growth of experimentation as a shift in the politics of knowledge within the aid industry. *Randomistas* problematized settled practices and knowledge, turning ‘matters of fact’ into “matters of concern” (Latour, 2004) through their insistence on the pervasive *lack* of knowledge within international development circles.¹ They offered experimentation as the authoritative and feasible means of achieving certainty and, eventually, the reduction of global poverty. In doing so, they translated epistemological and ethical values into practice. Proponents have depicted RCTs as the key to conclusive knowledge about ‘what works’ in development through their engagements with multiple publics, including academics, aid workers, policymakers, and what Monika Krause (2014) calls the ‘donating public’. Yet in the wake of their success, an ongoing search for epistemic closure has, in turn, problematized RCTs, leading *randomistas* to search for ways to repair their favored method. The result is not the achievement of certainty but rather an extended ‘experimental system’ (Rheinberger, 2010) in which the *randomistas* became among the most authoritative voices in international aid.

Facticity has solidified as a field of political power in the contemporary world, and this article seeks to clarify how such a transformation occurred within international aid. To trace this shift, I follow the repertoires through which individuals have justified their approach and criticized alternatives. This orientation analyzes the situated moral and epistemological reasoning of actors, their narratives, ethos, and methods. To do so, I have surveyed academic literature, popular and social media, training handbooks, instructional videos, policy briefs, and promotional discourse. This recent history is further evidence that accusations of uncertainty can, in the right context, serve as potent means of reshuffling the frameworks and methodologies through which policy decisions are made and resources distributed. Yet, it equally gives pause to optimism about the ability of new methodologies to consolidate consensus over aid policy and practice.

I argue that these evaluative styles and strategies have been furthered by, first, transformations in the international aid landscape and, second, the consolidation of a network of allied organizations, techniques, and discourses. This paper first introduces this thought collective before turning to the conditions of possibility for the rise of the *randomistas*, including the emphasis on accountability, the economization of aid, and the retreat from macro planning. It

¹ On the social study of ignorance, see Proctor and Schiebinger (2008), High *et al.*, (2012), McGoey (2012), Gross and McGoey (2015).

then outlines the importance of approaching the randomistas as critical actors who have pursued strategies of legitimization ranging from the rhetorical and emotional to the methodological and organizational. The elective affinity between these conditions of possibility and legitimating strategies facilitated the expansion of RCTs as a means to reduce uncertainty. However, as the final section of the paper documents, indeterminacy has persisted within the aid world, occasioning ongoing debates and novel responses by the randomistas.

The randomista thought collective

The adoption of randomized control trials within development economics and the aid industry has been propelled by a version of what Fleck (1981 [1934]) called a ‘thought collective’. In his pioneering explanation, scientific facts emerge within a particular community’s conceptual and perceptual repertoires.² The *randomista thought collective* is united, first and foremost, not in a consensus about a series of facts but rather in agreement about methodology. They share disciplinary matrices rather than policy prescriptions (Breslau, 1998). This epistemic commitment orients adherents toward a style of inquiry that utilizes RCTs as the means to achieve knowledge that is useful for development policymakers. In the words of Dani Rodrik (2008), the randomistas “tend to be suspicious of claims to *ex ante* knowledge about what works and what does not work.” In the quest to conclusively know ‘what works’, the randomistas insist upon the virtues of experimentation but, in contrast to some criticism, the randomistas do not practice methodological monocropping; instead, as I detail below, RCTs have been the paradigmatic method for a broader transformation in development knowledge and practice, the ur-method around which complementary ones are assembled.

The randomista thought collective is a porous network of interrelated but distributed organizations and discourses. Particular individuals, styles of reasoning (Hacking, 1992), and ‘cognitive infrastructures’ (Hirschman & Berman, 2014) circulate within this transnational discourse community. As Mirowski (2014, p. 43) has suggested with regard to neoliberalism, the structure of the randomista thought collective might be conceived as something like a ‘Russian Doll’ - a core group of adherents surrounded by concentric, linked layers (see also Mirowski & Plehwe, 2009). The origin of the thought collective and its continuing centre of authority is the Abdul Latif Jameel Poverty Action Lab (J-PAL) at MIT. Founded in 2003 by Esther Duflo and Abhijit Banerjee, J-PAL facilitates and organizes RCTs, advocates for the method, and disseminates their findings.³ As of 2015 its affiliates conducted more than 600 RCTs, but its influence through legitimating and diffusing the method has been far greater. If RCTs were once a novelty in development, by 2015 they were a highly sought after staple of the field.

² For the related idea of ‘epistemic communities’ in international relations, see Haas (1992).

³ Both economists are well regarded and influential. Duflo, for example, has won the John Bates Clark Medal and a MacArthur fellowship, addressed the UN General Assembly, and held the youngest chair at the Collège de France (in ‘Knowledge against Poverty’).

A similar organization at Yale University, Innovations for Poverty Action, was established in 2002. These organizations work in a collaborative fashion, providing infrastructure, financing, and networks to facilitate RCTs, often conducted by doctoral students. These academic centres are in frequent partnership and exchange with other researchers, aid organizations, and governments. The thought collective has grown through funding, meetings, publishing, and implementation. Particularly important is the World Bank's Development Impact Evaluation (DIME) unit, the Center for Global Development (CGD), a Washington, DC think tank, and the International Initiative for Impact Evaluation (3ie). DIME was purposefully established in 2005 (and re-launched more ambitiously in 2009) to conduct and promote experimental evaluation of aid projects; by 2015 it had conducted at least 175 evaluations in 47 countries.⁴ CGD does not implement aid interventions but has served as a vocal proponent of evaluation, including experimentation (see CGD, 2006). 3ie, for its part, is a grant-making organization founded in 2008 to promote impact evaluation and systematic reviews of evidence. Since then, it has distributed more than US\$84 million in 200 grants in over 50 countries. Underlying the spread of the randomista thought collective are significant amounts of money and support, especially from the Bill & Melinda Gates Foundation, the William and Flora Hewlett Foundation, and the UK Department for International Development (DfID). Governments, too, have adopted the emphasis on evaluation, with Mexico a forerunner in the adoption of RCTs.

The use of RCTs provides access to certain resources, networks, and prestige. As one expands beyond the core sources of finance, authority, and ideas, the thought collective can include a dizzying array of initiatives and organizations. 3ie has more than 100 partner organizations and identifies 45 distinct repositories for impact evaluations (Mishra & Cameron, 2013). IPA has more than 400 research associates. J-PAL has offices on five continents. Indeed, it is an indicator of the very success of the original premise that it has been adopted or responded to by nearly all major stakeholders in the aid industry. Duflo has described herself as an 'institution builder' (Parker, 2010), and the randomistas are active proselytizers, publishing training materials and handbooks, creating new academic journals, offering online courses, and using blogs and social media to particular effect. Esther Duflo's TED Talk has been viewed more than 700,000 times and their work often appears in the media. *Poor economics*, Banerjee and Duflo's award-winning book, was reviewed positively in publications from *The New York Times* to *The Cleveland Plain Dealer*. The creation and circulation of this discourse cultivates the multiple publics with which they engage and expands the thought collective.

Since 2009, there has been a more concerted effort to shape government policy. The idea for J-PAL was to "engage decision-makers not just as experimental partners but as adopters of programs that have already been vetted" (Parker, 2010). Yet, they found that their evidence did not always translate into policy action. As Michael Kremer, a pioneer of development RCTs,

⁴ Other aid organizations have created initiatives to promote the use of experimentation and generalized the practice throughout their work. For example, USAID's Development Innovation Ventures promotes them and the US Millennium Challenge Corporation does so for 40 per cent of its projects (Gilbert, 2013).

noted, ‘While evidence can play a very important role for certain policy makers, it’s far from enough on its own’ (quoted in McMurtrie, 2014). In response, core individuals from the randomista thought collective have begun implementing aid interventions directly. Kremer co-founded Deworm the World which is today an initiative of Evidence Action, an NGO initially housed at Yale’s IPA that seeks to expand interventions determined effective by RCTs. Similarly, with the support of the Hewlett Foundation, behavioural economists associated with the thought collective formed ideas42, an entity ‘with the goal of using scientific insights to design innovative policies and products’.⁵

The changing landscape of the international aid industry

The randomistas have emerged and now operate within a broader disciplinary and institutional regime. If the latter half of this article attends to the ‘social life of methods’ (Savage, 2013), this section considers what Eyal (2013) calls the “background of practices” that are the ‘condition of possibility’ for expertise. It asks not who is authorized to do or say what, but rather what forms of life or shared orientations were necessary for the rise of the randomistas to be a reasonable possibility, let alone occur. Three are particularly salient: the field of evidence-based policy, the economization of aid, and the proliferation of NGOs.

Before turning to these, it is important to note that experimental assessments of social policies have a longer history, particularly in the United States. In 1969, for example, the psychologist Donald T. Campbell published a programmatic essay entitled ‘Reforms as experiments’. Social scientists, he claimed, should be the ‘methodological servants of an experimenting society’ (Campbell, 1973), enrolled to ‘try out new programs designed to cure specific social problems’ (Campbell, 1969). In the coming years, dozens of social services were evaluated experimentally, including a study of more than 1200 households in New Jersey to assess the effects of a negative income tax credit and an early assessment of the Head Start program (see Dehue, 2001; Bartholomé, 2005). Such histories, however, have rarely been discussed publicly within international aid networks. Instead, as I discuss below, when development economists and policymakers sought to promote social experimentation, they more frequently invoked a schematic history of medical experimentation, using the presumed authority of biomedicine as a justification for emulating it.

Evidence-based policymaking

As this special issue makes clear, ‘evidence-based policy’ is a discourse and practice reshaping the work of states and NGOs across the globe. The promise of evidence-based policy is the stage on which the randomistas’ arguments and activities take place. Indeed, many of the disputes in which randomistas engage are about how, exactly, aid practitioners can *best* implement evidence-based policies. Rarely do individuals who decry the predominance of RCTs question

⁵ See: <http://www.povertyactionlab.org/partners/ideas42>

the broader goal of evaluating development projects; instead, they seek an appropriate methodological balance in order to investigate “what works and what does not in the fight against poverty” (Ravallion, 2009a, p.5; see also Ravallion 2009b; Deaton, 2010). Yet, the randomista thought collective is relatively autonomous from, and therefore not reducible to, the broader evidence-based policy movement (but see Pearce & Raman, 2014). Randomistas coalesce in their own institutions and publish in their own venues. Importantly, they also understand themselves to be semi-autonomous, often critically engaging with other proponents of evidence-based policy by offering a more certain style of knowing.

There is a long history of ‘monitoring and evaluation’ (M&E) within international aid organizations. Such techniques serve multiple purposes. Like other ‘technologies of trust’ (Porter 1996), these audits serve as a means of ascertaining and governing ‘action at a distance’ (Power, 1997; Rose, 1999). There is a heightened emphasis on accountability within international aid because public revenue is spent at a remove from citizens, top-level management, and other constituencies (Rottenburg, 2009a; Jensen & Winthereik, 2013). Audits also function as a measure of worth. As Krause (2014) has detailed, aid organizations operate within a competitive field in which ‘agencies produce projects for a quasi-market in which donors are consumers’. Methods such as interviews and logframes facilitate the evaluation of aid and, crucially, the comparison of heterogeneous projects, indicating which should receive support. However, as Margarita Rayzberg (2014) has documented, the randomistas emerged in contrast to existing M&E approaches which they deemed less ‘rigorous’.

This was important because, around the turn of the century, there was heightened emphasis on improving the effectiveness of aid. For many observers, however, there simply was not enough evidence to know one way or the other if aid was effective. For example, the US Congress’s Meltzer Commission (2000) found that only 5-10 percent of World Bank projects were reviewed within 3-5 years. The Center for Global Development’s Evaluation Gap initiative documented, over a number of years, the paucity of studies of aid effectiveness. As its president, Nancy Birdsall (2006), said, ‘Without impact evaluations that are rigorous, independent, and thus credible we cannot know what programs work. We cannot even argue convincingly that foreign aid itself works’. By 2006, major aid organizations responded with a substantially expanded commitment to impact evaluation. And newer entrants like the Gates Foundation are particularly supportive (Gates, 2013).

Such dynamics are closely related to budgetary concerns, especially after 2008 when foreign aid budgets have been under further pressure. The randomistas frequently emphasize the importance of experimental evaluation of cost effectiveness. Governments, such as the United Kingdom, place considerable emphasis on using RCTs to achieve ‘value for money’ (Whitty & Dercon, 2013). Ruth Levine (2014) - who co-authored CGD’s Evaluation Gap report, advised USAID on evaluation policy, and subsequently directed the Hewlett Foundation’s development work - explains that evaluation can ‘lead to better use of money’ by aid organizations. J-PAL’s director **Rachel Glennerster and Harvard’s Michael Kremer (2011)** valorize RCTs for allowing funders ‘to design successful and cost-effective programs’.

This broader discourse enhanced the audibility of the randomistas. Their commitment to experimental evaluation was understood by others within the aid world as a more precise means of achieving their goals. RCTs displayed “generosity”, allowing others to incorporate them into their work (Eyal, 2013). By offering an answer that was deemed both effective and feasible, the randomistas became the most prominent experts within the energized field of evidence-based aid.

The economization of development

The rise of the randomistas was also facilitated by the existing influence of academic economics within key institutions in the aid industry. Organizations like the Gates Foundation, DfID, and especially the World Bank maintain close ties to neoclassical economics departments at universities and often hire or fund academic economists. Although new sources of funding have somewhat diminished the financial sway of the World Bank, it maintains its leading position as a source of intellectual authority and influence.

The heightened role of economists within the World Bank dates to the presidency of George Woods (1963-1968) who established the economics department and the influential position of chief economist. Under his tenure, the number of economists increased by 25 per cent. By 1991, 80 per cent of senior staff in the Policy, Research, and External Affairs departments were trained in economics or finance (Woods, 2000, p. 152). Economists replaced engineers, medics, geologists, and agronomists, whose services are now usually contracted when needed (Moore, 2007).

As Stein (2008) has demonstrated, the ties between universities and the World Bank accelerates the uptake of academic trends within aid policy. Within the discipline, methodological innovation and precise measurement are rewarded (Fourcade *et al.*, 2015). Evidence is understood as quantified, causal proof rather than data observed or experienced. The randomistas’ emphasis on ‘what works’ for policy interventions is indicative of the discipline’s commitment to ‘fixing the economy’ (Mitchell, 1998). It is a style of reasoning which pays ‘little attention to history’ (Fourcade *et al.*, 2015, p. 108) and seeks to both generalize and expand into new domains - often those traditionally associated with other disciplines (Fine & Milonakis, 2009).

However, some economists have been critics of the rise of randomized control trials (e.g., Deaton, 2010; Ravallion, 2009b; Rodrik, 2008). The reasons for the disputes differ, but most have been specific, often technical, and within the bounds of broader shared commitments - not least for improved knowledge.⁶ At stake for both the randomistas and these interlocutors is how

⁶ There exist at least two important exceptions. First is Basu’s (2014) call for non-quantifiable forms of knowing such as wisdom, reasoned intuition, and judgment. The other is Ravallion’s (2014) concern about the ethics of RCTs. As of 2015, both approaches sit at odds with the prevailing approaches of the randomistas.

best to reduce or eliminate ambiguities. Crucially, the randomistas and these critics with whom they engage operate within the bounds of mainstream, neoclassical economics. They do not challenge the disciplinary presuppositions.⁷ Orthodox economics is the taken-for-granted background on which the critique offered by randomistas has been deployed. It serves as an institution that stabilizes and fixes meaning, offering a grammar in which worth is debated. Where the randomistas do differ from the orthodoxy - such as the use of behavioral economics - they remain in good professional standing (Davis, 2013).

Within neoclassical economics, however, the 1990s were a time of change. What Angrist and Pischke (2010) call the ‘credibility revolution’ brought applied microeconomics to the fore. A newer, more determined approach to causality, research design, and interference was contrasted with more theoretical work in the 1970s and 1980s (Fourcade, *et al.*, 2015, p. 92). In the realm of development economics, this was particularly a critique of macroeconomic growth regression analyses which previously dominated debates about the effectiveness of aid. For proponents of RCTs, such methods were akin to ‘speculating on a grand scale’ (Banerjee & Duflo, 2011, pp. 3-4). They advocated moving past the ‘big questions’ of theory testing and towards specific policy questions (cf. Rodrik, 2008; Bates, 2007).

NGOification and the retreat from large-scale planning

The randomistas have, therefore, been associated with a scalar transformation in aid practice. In the early years of the Bretton Woods regime, international finance institutions and bilateral donor agencies focused predominantly on large-scale infrastructural investments. By the late 1960s, however, such initiatives were the subject of considerable critique. At the World Bank, Robert McNamara downplayed infrastructure development in favour of poverty reduction and ‘basic needs’. The 1973 US Foreign Assistance Act directed bilateral aid toward food, nutrition, health, and education (Stein, 2008, pp. 14-15). And a considerable network of formal institutions and individuals pursued ‘community development’ across the global South (Immerwahr, 2015). This approach was significantly furthered, first, by the discrediting of state planning by the late 1980s, and later by the UN’s Millennium Development Goals which served as a powerful means of channelling funding toward smaller interventions (Picciotto, 2014). By the late 1990s, the project-centred approach spurred the worldwide growth of non-governmental organizations.

The resulting arrangement, at both major institutions and smaller NGOs, means that much aid is funded and administered through individual projects, not sector- or country-wide schemes (Krause, 2014). RCTs were particularly well-suited to the down-scaling of aid interventions. Randomistas do not do not generally seek to assess the effectiveness of, for example, aid *writ large* or general budget support. The method is not suited for that. Rather, they

⁷ The historical consolidation of orthodox economics is both well-documented and frequently criticized for its positivism, methodological individualism, and mathematicization. See, for example, Mirowski (1991, 2002), Steinmetz (2005) and Milonakis and Fine (2008).

seek to assess the impact of particular programs and projects (e.g., the distribution of textbooks or bed nets).

Doing so requires closely working with implementers, sometimes even changing the intervention in order to make it fit the methodology (cf. MacKenzie, 2008). For the randomistas, working with NGOs is highly desirable. As Glennerster and Kremer (2011) write of the earliest aid RCTs, ‘NGOs, in contrast to governments, proved to be highly flexible and open to experimenting with new ideas’. Today, many NGOs seek experimental evaluation due to funders’ requirements, efforts to improve their work, and the desire to demonstrate effectiveness in a competitive field. The multiplicity of NGOs (in contrast to governments) means randomistas can find willing partners, even when they are interested in studies that may debunk the status quo.⁸

The NGOification of aid has necessitated new indicators. A previous era of development positioned economic growth and national income as the key measures of success. However, since at least 1990, when the Human Development Index was created to combine indicators of health and education with income measures, the aid world has been host to a profusion of indicators. These classifications, numbers, and ranks have emerged as a key mechanism of government (Davis *et al.*, 2012). RCTs, first through the evaluation of specific interventions and then through the amalgamation of many experiments, are a similar technique of knowing and governing. They enumerate not populations but specific groups of beneficiaries (cf. Foucault, 2009). They assess not ‘need’ but ‘impact’ (Krause, 2010).

The affordances of the experimental method - including intensive data-gathering across treatment and control groups - suit project-based aid work. It permits an audit deemed more accurate than alternatives, whether focus groups, interviews, or otherwise. And experimentation also displays a temporal affinity with the prevailing logics: the duration of both project funding and RCTs often align (a couple of years) and, furthermore, RCTs seem to offer an evaluation more quickly than other means (such as waiting for the next census or general household survey).⁹

If RCTs proved particularly compatible with small-scale interventions at the start, in recent years this is changing. As RCTs have entered the mainstream, and techniques for designing and managing trials have improved, the scope is expanding. One took place throughout the entirety of Andhra Pradesh (population nearly 50 million) while another, as part of a move into the global North, tested job training programs in half of the cities in France (*The*

⁸ In his profile of Duflo, Parker (2010) writes that she ‘had long wanted to use experimental methods to put microfinance to the test. As she saw it, there was little beyond anecdote to support claims that the technique had a special power to combat poverty, gender inequality, and ill health’. Only ‘after a lengthy search in an industry wary of subjecting itself to this kind of scrutiny’ did she find a microfinance institution willing to subject itself to an RCT.

⁹ This is not always the case: in some cases, the experiment is considered too short or begins too late, leading to questions about whether it truly captured a project’s effect. Randomistas have, in turn, sought to influence project design at an early stage, furthering the performativity of the method. For a critique of the short timeframe of RCTs, see Olofsgård (2012) and Woolcock (2009).

Economist, 2013). Such a trajectory fits with the ambitious narrative of the randomistas which seeks to accumulate general knowledge through multiple experiments.¹⁰

Experimentation *contra* indeterminacy

It was within this broader context that the randomista thought collective grew; however, there was nothing preordained about its emergence, and a fulsome understanding of the politics of experimental knowledge requires close attention to the manner in which proponents questioned the status quo and established agreement about the desirability of development RCTs. To do so, it is useful to follow the pragmatic sociology of critical capacity developed by Luc Boltanski (2011) and his collaborators (Boltanski & Thévenot, 2006; Boltanski & Thévenot 1999).¹¹ Their oeuvre attends to the manner in which actors justify, generalize, and confirm, or (alternatively) qualify, challenge, and critique, the situations in which they operate. They analyze the ethics and spirits which motivate critique and animate communities. Pragmatic sociology draws a distinction between ‘worlds’, which are regimes of social consensus and representation, and ‘reality’, which is always at least partially at odds with representational schemes. Critique operates to open up that gap, challenging and destabilizing worlds through ‘reality tests’. For Boltanski and Thévenot (2006), disputes take place within certain schema which they call ‘grammars of justification’. When successful, critique undermines the ‘semantic security’ that is indicative of stable institutions and practices; it insists that signs, words, or means of calculating are out of step with reality.

Such an interpretive approach highlights the moral commitments immanent in seemingly technocratic discourse. Similarly, within science studies, there is a long tradition of analyzing the literary, visual, and discursive production of scientific achievement (Shapin & Schaffer, 1985; Latour, 1987). McCloskey (1998) in particular has pioneered the study of economists’ representations, demonstrating the ways in which authority is performed and achieved through rhetoric. Credibility requires that economists - like other scientists (Hilgartner, 2000) - manage their narrative structures, frames, and evidence in ways that are compelling for particular audiences.

The benefit of such an approach is to foreground the disputes that emerge due to conflicts over these regimes of value and justice within the aid industry. The randomistas have engaged in two, interlinked critiques. First, they have disputed the reality of aid’s effectiveness. As I detail below, they justify their approach through the ultimate desire to improve the ways in which public policy is able to reduce poverty, mortality, and other forms of human suffering. Closely related to this is their denunciation of the prevailing representations of the effectiveness of aid. Within the justificatory regime of development aid, credible knowledge is highly valued by funders, implementers, and the public at large. RCTs are presented as a methodology that

¹⁰ For further discussion of this expansionary logic, see Blattman (2011), Ludwig *et al.* (2011), Banerjee and Duflo (2009) and Duflo and Kremer (2005).

¹¹ For discussions of pragmatic sociology, see Browne (2014), Bénatouïl (1999), Celikates (2006) and Wagner (1999). For an application to economics, see Davies (2014) and Boltanski and Chiapello (2006).

determines whether or not knowledge about aid is accurate. In this case, their targets for denunciation have been other ways of knowing. Through the skillful legitimation of experiments, they have enrolled adherent to their view that the existing manner of assigning worth was not a proper reflection of reality.

This is a critical, evaluative orientation that privileges empirical results obtained through a particular method. But the randomistas have not been merely denunciatory nor empirical; they have sought epistemic closure, not confoundment. They are *not* oriented toward critique for its own sake, and they do *not* cast suspicion on the very possibility of development nor the feasibility of aid *writ large*. They are reformists, not radical. Their critique demonstrates a differential between *what should be* and *what is*. It is sympathetic and constructive, which accounts for some of their success and some of the reactions to their success: because they are legible contributors, large portions of the aid industry recognize them as beneficial, but because their commitments do not depart from a larger status quo, they are also the subject of strident rebuttals by critics who see RCTs as reproducing an overly narrow economics (Reddy, 2012).

Between critique and confirmation: Legitimizing experimentation

If transformations within the aid industry made the rise of the randomistas possible, particular actions and strategies made it occur. Central to their success has been the cultivation of a consensus that RCTs are a compelling manner in which to reduce uncertainty about the effectiveness of aid. However, because their overarching themes and methodological commitments have been valorized to multiple publics – including academics, policymakers, aid workers, and the donating public – the significance of uncertainty has differed. In economics journals this takes the form of a careful research design to achieve valid causal results. For policymakers, the indeterminacies of accountability and cost-benefit analysis are emphasized. And when addressing less expert audiences, uncertainty is to be reduced because it impedes the effectiveness of aid and, therefore, contributes to human suffering. Such flexibility functions to enroll numerous supporters.

In each of these registers, experimentation is justified as a moral methodology. As Daston and Galison (2010, p. 36) argue, the nature of scientific objectivity has changed over time but is, fundamentally, an ethical issue because it involves the ‘suppression of some aspect of the self’. For the randomistas, RCTs have ‘epistemic virtue’ because they remake aid workers by removing subjective bias. But because they are engaged in public policy debates, the randomista thought collective adopts an even more urgent moral message. They exhort others to adopt the methodology not only to reform themselves but to transform the distribution of resources. The value-laden proposition of the randomistas is that more objective knowledge about causation (‘what works’) is a crucial factor for a host of humanistic goals (poverty alleviation, gender equality, and so forth). This section discusses how the randomistas have fashioned themselves and been understood by their audiences. To do so it examines the variety of legitimation strategies deployed by the thought collective.

For the randomista thought collective, RCTs demand a different type of aid researcher. ‘The great virtue’, Banerje’ (2007) writes, ‘is that they force us to venture inside the machine. To implement a proper evaluation, one has to know the exact details that define a program. And as economists think about them, they begin to build stories about them and get ideas about how to change them for the better’. RCTs reform economists from ‘sitting in your office downloading data and working regressions’ (Glennerster quoted in McMurtrie, 2014) to getting ‘on-the-ground experience [that] shows us the realities that might otherwise have been left out of our models’ (Glennerster & Kremer, 2011). Proponents value field experiments not only for offering better results but because it suits their dispositions and reveals new insights. Ian Parker (2010) reports that Esther Duflo is ‘effusive’ about fieldwork: ‘I love it, I love everything about it. It is the only way, when you work on development, to get an intuitive sense of how people really live their lives’.¹²

Intimacy, however, brings risks. Historically, guidelines within the aid industry have sought to minimize the interaction between implementers and evaluators to avoid any conflict of interest. RCTs, in contrast require close collaboration between evaluators and implementers (in order to ensure random assignment, for example). For proponents, the attendant risks of impropriety are eliminated by the methodology.¹³ It disciplines evaluators by minimizing discretion and the possibility of error. As the director of J-PAL writes, despite the close collaboration, ‘it is possible for randomized evaluations to provide independent or objective results’. She continues,

This is because, for the most part, the results of a randomized evaluation are what they are... [T]here is relatively little flexibility for the evaluator to run the analysis different ways to generate the outcome they want to see... [C]ompared to much other evaluation work carried out by development agencies, randomized evaluations provide results which are harder to manipulate and thus are reasonably objective... (Glennerster, 2013)

In this way, the experimental method is valuable because it enforces what Porter (1996) has called ‘procedural objectivity’, or the application of impersonal rules to subjective practice. This is particularly important for two, related reasons. In public policy, where transparency is valued, numbers are understood to reduce arbitrary decision-making. In applied research domains, which are often less prestigious and vulnerable to accusations of interestedness, objectivity is even more a virtue (Breslau, 1998).

Experimentation is also justified because it establishes causal relationships, removing potentially confounding variables.¹⁴ In the early days, many spoke of RCTs as the ‘gold standard’ for economics research, but today the most common way for randomistas to characterize their evidence is ‘rigorous’. This is because a randomized control trial, if properly

¹² For a critique of the underappreciated role that ‘intuition’ plays in popular development economics, see McGovern (2011).

¹³ For an excellent illustration of this, see Margarita Rayzberg’s forthcoming “Fairness in the Field.”

¹⁴ Breslau (1998) notes that it is a particularly Humean notion of mechanistic cause-and-effect. On the varieties of causation, see Hirschman and Reed (2014).

implemented, has ‘internal validity’, meaning the inference of a causal relationship between two variables is warranted.¹⁵ As John List (2009), a prominent experimental economist, argues, the fundamental challenge in the social sciences ‘is how to go beyond correlational analysis to provide insights on causation’. In the view of the randomistas, RCTs are uniquely capable of doing so because they “represent a mixture of control and realism usually not achieved in the laboratory or with uncontrolled data, permitting the analyst to address questions that heretofore were quite difficult to answer” (Levitt & List, 2009). Moreover, RCTs are promoted as self-evident, very nearly speaking for themselves: ‘the evidence is simple to interpret. The beauty of randomized evaluations is that the results are what they are’ (Banerjee *et al.*, 2007).

This is particularly important for policymaking. The randomistas promise causal certitude in their aspiration to provide useful science. Historically, program evaluation has lacked prestige in aid organizations, but the randomistas have overcome this by shifting from assessments of *past* projects to guides for *future* interventions. Their motto is not “what worked in that instance” but rather ‘what works’ generally. They are ‘policy entrepreneurs’ (Krugman, 1995), actively disseminating their findings as recommendations for funders and implementers, and often called upon as experts, collaborators, and consultants.

The acquisition of this authority has benefitted from the invocation of medicine as a model for international development. The adoption of RCTs is an aspect of what Davies (2013) calls a change in economics, ‘from aspiring to the status of physics to aspiring to that of biology’ (cf. Joffe, 2013, 2014). In the case at hand, economists seek to learn from medicine, emulate it, and use it as a means of justifying their work (cf. Mirowski, 1991). As Banerjee and Duflo (2011, p. 8) write, ‘the cleanest way to answer such questions [of causality] is to mimic the randomized trials that are used in medicine to evaluate the effectiveness of new drugs’. In Karlan and Appel’s (2012) book promoting aid experimentation, they argue that the RCT “has long been the gold standard throughout the sciences for determining effectiveness. To take an example, the Food and Drug Administration requires data from an RCT to warrant approval for new medicines. In general, if you need rigorous and systematic evidence of effectiveness on a large scale, you use an RCT to get it.” When discussed in more popular media, medicine is also the reference (e.g., *The Economist*, 2013).

It would be an overstatement to suggest the randomistas naively equate biomedicine and social policy; rather, a more diffuse and subtle process is at play. In the above cases, medicine is an explanation-at-hand but, given its presentation as an authoritative science, it also serves to legitimate the randomistas. This happens through mundane conventions, such as referring to aid projects as ‘treatments’, and major advocacy reports, like the Center for Global Development’s Evaluation Gap report (CGD, 2006). Authored by individuals with both medical and economics training, this influential document repeatedly drew an equivalence between medicine and economics. They write,

¹⁵ The qualification is significant. As an enormous literature on medical trials attests, properly designing and implementing an RCT is a considerable difficulty. These qualifications have rarely intruded on the rise of the randomistas.

No responsible physician would consider prescribing medications without properly evaluating their impact or potential side effects. Yet in social development programs, where large sums of money are spent to modify population behaviors, change economic livelihoods, and potentially alter cultures or family structure, no such standard has been adopted.

Continuing later, they declare that ‘The simple truth is that many well intentioned social programs are like promising medical treatments - we cannot really know if they do more good than harm until they are tested’. Of course, the assumption that medical experimentation offers obvious, unmistakable evidence has long been questioned by social scientific and medical observers alike (e.g., Will & Moreira 2010). What is notable in this case of methodological change, however, is the aphasia that occurs as experimentation crosses disciplinary boundaries.¹⁶

Justifying aid RCTs has involved overlooking other aspects of experimentation, too. For the adherents, the use of the experimental methodology is a “radical” innovation. This word is rarely defined but rather signifies as an iconoclastic break from business as usual.¹⁷ The randomistas self-fashion as an upstart group, fighting against the tide. J-PAL’s director recounts that ‘When we started, there was a huge amount of resistance and hostility in the development community’ (Glennester in Parker, 2010). Despite their enormous influence, they continue to depict themselves as peripheral. The influential evaluation proponent William Savedoff (2014), for example, notes that more are being done than ever before but says that

RCTs are marginal because only about 200 of them (my estimate based on the 3ie database) are being started in any given year on topics related to development programs. This is dwarfed by the thousands of evaluations being conducted using expert interviews, focus groups, non-purposive samples, and quasi-experimental methods.

Thus, for key proponents, the measure of critical success is the quantitative proportion of RCT evaluations, not the epistemic or discursive authority of the method.

Randomistas believe proper methods can reveal evidence unsullied by ideology, politics, and fads (cf. Latour, 1993). The economist Diane Coyle (2011) praises the randomistas for departing from a ‘landscape of development economics [that] has been scarred by ideological battle’. Another proponent bemoans the fact that “politics still drives most Western countries’ foreign development aid’ instead of ‘scientific and evidence-based tools for policymaking and

¹⁶ On historical ‘aphasia’, see Stoler (2011).

¹⁷ Randomistas have successfully portrayed their approach as novel despite a history of field experiments in economics since at least the 1920s and the use of experimental evaluation of social policies throughout the postwar era. More often than not, this history is effaced, but when it is not, the older generation of experiments is, paradoxically, deemed both too expensive and too limited. As J-PAL’s Glennester and Harvard’s Kremer (2011) write, ‘Randomized trials have been used to study social and economic conditions in the developed world for some time. In the 1970s the US government conducted large-scale evaluations of a negative income tax and of health insurance. But, while these evaluations were useful, they tended to be expensive one-offs, designed to measure the impact of a single policy with many components, making it difficult to learn in a cumulative way over time’.

priority setting' (Schmitt, 2014). The lack of certainty means that development economics 'has promoted a great many 'big-think' fads' (Glennerster & Kremer, 2011). Experimental evidence is portrayed as a rupture from these histories. The causal certitude revealed through RCTs makes for a common narrative structure: *people have typically thought X, but now we know it is actually Y*.¹⁸ To pick just one such proclamation, "Many economists believe those who most need a product are more likely to pay for it. They're wrong" (Glennerster & Kremer, 2011). Thus the randomistas public proposition is an epistemic iconoclasm, not only strategically positing ignorance but claiming to have rectified the deficit.

But, quite importantly, the rhetoric of the randomistas is not reducible to facticity and bravado.¹⁹ Indeed, there is often an emphasis on modesty and inquisitiveness. They speak of a 'quiet revolution' that improves things 'at the margin' (Banerjee & Duflo, 2011). Their approach is to see 'the fight against poverty... as a set of concrete problems that, once properly identified and understood, can be solved one at a time' (Banerjee & Duflo, 2011, p. 3). For those within DfID, experimentation is indicative of 'humility' and 'honesty' and explicitly set up in contrast to an alternative ethos of "using the social power of the 'expert' to imply we know the answer when we actually have no solid evidential basis for our opinion or prejudice" (Whitty & Dercon, 2013). Ruth Levine (2014) says that insisting upon evaluation is actually indicative of a rejection of 'dogmatism' by 'saying we aren't so sure about the effects of our actions, [and] we're open to surprise and to learning'. In doing so, randomistas position a method and its corresponding ethos in contrast to approaches based on *a priori* policy prescriptions and immunize themselves from allegations of technocratic hubris (e.g., Eyben & Roche, 2013).

This inquisitive modesty opens up new venues for their work and facilitates the expansion of the thought collective. For example, by positing that 'it's too soon to tell if behavioral economics can help a young woman in Uganda', the Hewlett Foundation links promissory expectations with a commitment to disinterested evaluation (Choi, 2014). Here, as elsewhere, the attribution of ignorance or uncertainty serves to instantiate a "matter of concern" (Latour, 2004). As McGoey (2009) suggests, uncertainty 'demands attention, debate, funding, and most crucially, experts to determine how the situation should be resolved... [T]he expert's insistence on the uncertainty of a situation is virtually unchallengeable, for expert uncertainty, unlike expert knowledge, is difficult to dispute'. The randomistas strategically assign uncertainty and therefore bolster their rise (McGoey 2012).

¹⁸ Consider a 2012 debate at NYU. Angus Deaton, an economist critical of RCTs and informed by the philosopher of science Nancy Cartwright, focused on the epistemological shortcomings of RCTs while Abhijit Banerjee instead detailed the new findings from recent RCTs, avoiding philosophical or methodological issues. See: <http://www.nyudri.org/events/annual-conference-2012-debates-in-development/deaton-v-banerjee/>

¹⁹ Amongst others, Fourcade *et al.* (2015) have documented the 'self-confidence' of economists. As they note, 'That confidence is perhaps the greatest achievement of the economics profession - but it is also its most vulnerable trait, its Achilles heel'. The randomistas, I would suggest, have a more complex performance, not least as a means to immunize themselves from critique.

But the randomistas do not remain in the domain of inquisitiveness alone. The applied nature of their research and the imperatives of policymaking encourage epistemic closure. This is in contrast to the role of experiments in other styles of economics. In her comparison of economic models and experiments, Mary Morgan (2005, p. 317) has argued that models may ‘surprise’ economists but ‘experimental results may be unexplainable within existing theory and so ‘confound’ the experimenter’. Development RCTs have a different trajectory, rarely confounding but rather understood as ‘clinching’, to use Cartwright’s (2007) phrase. The randomistas also avoid another pitfall: critique that remains in the realm of empirical denunciations of prevailing consensus risks dissolving into a form of relativist nihilism (Boltanski, 2011). Had RCTs been used as ‘gotcha’ methods, always disproving, the randomistas might have been seen as mere denouncers. Yet, by offering policy prescriptions they have avoided such a possibility. There is thus a strategic, dialogic pivoting between iconoclasm and modesty, critique and confirmation.

The methodology was also furthered through the adoption of a cautiously optimistic tone. Against ‘the rather melancholy view active in economics today’ (Banerjee & Duflo, 2011, p. 237), the randomistas offer hope. Duflo says this is ‘less depressing than the view that it is a big conspiracy against the poor. Name your favourite enemy- capitalism, corruption... Our view is easier. You think hard about the problems and you can solve them. That is why I feel generally a happy person, not at all discouraged’ (cited in Gapper, 2012). Such cautious optimism may seem to be an awkward bedfellow for their radical iconoclasm, but it is the multivocality of the randomistas which gives them their influence.

To fellow economists, they debate econometric methods. To politicians, they offer low-risk bang for your buck. To humanitarians of all stripes, they suggest progress is within reach. The discursive aspects discussed here - iconoclasm, modesty, inquisitiveness, optimism - are efforts to reconstitute the ethos of international aid. It is an approach often lauded by observers as surprising and sensible. *BusinessWeek* (2010) calls them ‘pragmatic rebels’. *The Economist* is a frequent exponent and popularizer, finding the approach ‘fascinating’ (2011a), ‘engrossing’ (2011b) and ‘more fruitful’ than ‘an animated fight over political profundities’ (2011c). A review of *Poor economics* is particularly illuminating for how influential media understand the thought collective. It is the ‘best [recent] book about the lives of the poor’ in large measure because they ‘take the poorest billion people as they find them. There is no wishful thinking. The attitude is straightforward and honest, occasionally painfully so. And some of the conclusions are surprising, even disconcerting’ (*The Economist*, 2011d).

It is in their most public engagements, however, that experimentation most emerges as a moral methodology. Many randomistas are motivated by a humanitarian ethos to end “distant suffering” (Boltanski, 1999). Duflo has said she felt ‘that what I should really do in my life is help the poor’ (quoted in Parker, 2010) and when addressing large audiences she foregrounds what the historian Thomas Haskell (1985) called ‘humanitarian sensibility’.²⁰ For example, her

²⁰ Parker (2010) describes Duflo’s motivation as ‘shaped by ‘Protestant left-wing Sunday school’ and by the international response to the Ethiopian famine of the mid-eighties (Band Aid, Live Aid). And it was

widely viewed TED Talk from 2010 opens by invoking the 200,000 deaths in the Haitian earthquake the month before. The presentation continues, juxtaposing stark imagery of poverty and disaster with schematic charts and statistics. She asserts that there is a ‘Haiti earthquake every eight days’ due to ‘entirely preventable causes’. The trouble is, aid in the moment is ‘not any better than the Medieval doctors and their leeches’. Duflo acknowledges that what Chouliaraki (2006) calls ‘the spectatorship of suffering’ is ‘a bit cheap’ but desirable for her purposes because ‘I’m already understated, and not very funny. I have to be a little in-your-face’ (Duflo, quoted in Parker, 2010).

The randomistas, in these public appeals, argue for a merger of methodologies and orientations. From a genealogy of technocratic planning, they apply quantitative indicators, cost-benefit analysis, and experimentation. From a genealogy of progressive politics, they seek human betterment, deploying an affective repertoire and creating “structures of feeling” (Williams, 1977). Evidence Action, the NGO created by the randomistas and discussed below, is a prime example of this. Its staff explain that they “are driven by passion” but seek to avoid being blinded by it: “we need to divorce ourselves from the products we are so passionate about, and focus on an empirical assessment of what the problem is we are trying to solve” (Evidence Action, 2014). This self-discipline and hybridity brings forth its own representational styles, forms, and content. The Evidence Action website (pictured), for example, displays the visual techniques of distant moral sentiments (anonymous African and South Asian children, in particular) with those of experimental rigor (charts and graphs that simplify and condense statistical and experimental analyses). It equally includes academic papers, policy briefs, and ‘colorful’ launch events with magicians, folks dances, and Sanskrit verses being chanted (Ramachandran, 2013).



shaped, too, by the work done by her mother, a doctor, who, from the late seventies onward, left her pediatric practice in Paris for a few weeks each year to treat child victims of war, first in Western Sahara and later in El Salvador and Rwanda. Duflo described her mother as ‘a generous human being to the point where it’s unnerving for the rest of us’.

Attention to the representations of the randomistas reveals that far from being constrained within a methodological straightjacket - as some critics have alleged - the thought collective actually displays a great variety of legitimating strategies. The variety of approaches - from the emotive to the scientific - are not understood by the proponents nor their audience as inconsistent; instead, the randomistas have grown by weaving together different registers of value and rationality (cf. Weber, 1978, pp. 24-26). If these strategies have successfully translated ideas about the virtue of certainty into institutional practice, they have not eliminated indeterminacy. Instead, they have fueled an expansion of the “search for the unequivocal” (Breslau, 1998) – a dynamic to which the final section of this article turns.

After experimentation: The persistence of uncertainty

In January 2015, the *American Economic Journal: Applied Economics*, a leading source of experimental economics, published six RCTs in a single issue. Each experiment analysed the impact of microcredit, in settings from Morocco to Mongolia. Arguing that prior evidence from microcredit’s proponents ‘failed to disentangle causation from correlation’, Banerjee and his colleagues concluded that microcredit was not capable of ‘transformative effects’. At best, these experiments proved it was ‘modestly positive’ with regard to indicators of health, education, women’s empowerment, or other topics, though some of the experiments found no significant changes. For some observers, the issue suggested ‘the final word on microcredit’ (Sandefur, 2015), but for others, knowledge was not certain, and “there are still many questions to be answered” (Ogden, 2015). If microcredit is any indicator, the capacity of RCTs to resolve indeterminacy about the effectiveness of aid was less than promised. Indeed, the rise of the randomistas has heralded less a closure of questions than a redistribution of the ways of answering them.²¹ As this final section shows, while the randomistas successfully questioned prevailing assumptions about how to assess aid, in more recent years, RCTs themselves have become the subject of uncertainty.

Ongoing concerns about the reliability of evaluative knowledge has led to a twofold transition. On the one hand, the “experimental system” (Rheinberger, 2010) has been expanded to include three new methods: experimental replication, trial registries, and systematic reviews. On the other hand, the randomista thought collective has grown to include aid organizations directly implementing projects based on experimental results. Both expansions follow from the normative presuppositions of the thought collective. The former seeks to reinforce the epistemic virtues of experimentation, minimizing uncertainty due to subjective bias or particular contexts. The latter is an effort to reduce human suffering by acting upon knowledge considered unequivocal and bypass the politicized delays of governments and aid institutions. That is, the

²¹ The question of ‘how experiments end’ (Galison, 1987) is foundational to science studies. In Collins’s (1985) influential explanation, scientific tests cannot prove conclusive without a broader set of social institutions and forms of life. Any given experimentation or replication is subject to a potentially infinite regress: ‘since experiment is a matter of skillful practice, it can never be clear whether a second experiment has been done sufficiently well to count as a check on the results of a first. Some further test is needed to test the quality of the experiment - and so forth’ (Collins, 1985, p. 2).

former are recognizably still driven by efforts at epistemic closure within the bounds of evidentiary tests while the latter reflect a more localized consensus about ‘what works’ and a moral urgency to act.

The extension of the thought collective demonstrates the heightened influence of its styles of reasoning, its techniques, and its practices within the international aid community. Experimentality - Petryna’s (2009) term for the expansion of experimental rationalities of governing - simultaneously redistributes the dominant ways of assigning worth within aid and who is capable of doing so. Some critics of the randomistas have emphasized a different ethos for aid work. Kaushik Basu (2014) insists that ‘we have to rely on intuition, common sense and judgment’, while Lant Pritchett (2014a) says ‘there is no shortcut around using judgment and wisdom’. Yet, no thought collective has emerged to promote wisdom, judgment, intuition, and common sense in aid; instead, randomistas can adopt such diffuse concepts as complementary to their existing practices (see Glennerster 2014; Blattman 2014). The inverse - insisting upon RCTs - cannot so easily be incorporated without expanding the thought collective’s influence and authority. In the discussion that follows, I document how a continued insistence on uncertainty has expanded aid’s experimentality.

Extending the experimental system: Replication, registries, and reviews

The most enduring critique of aid experimentation is the issue of ‘external validity’. Because the precision of an experiment comes at the cost of a narrow scope, the cost of an *internally* valid study is relative ignorance about its applicability elsewhere (Cartwright, 2007). The limitations are most often considered spatially (e.g., experimental evidence from Western Kenya may not illuminate a southern Indian case – or even elsewhere in Kenya). But there is also temporal particularity: RCTs are a form of historical documentation, telling you what resulted in one case. They do not offer sturdy grounds for projection into the future (Pritchett, 2014b).

Members of the thought collective acknowledge these limitations in technical fora but less commonly raise them when speaking to other publics. Indeed, the limitations of external validity sit uneasily with the desire for policy relevance. The slippage from *here* to *there* and from *ex post* evaluation to *ex ante* justification is implicit in the effort to produce useful social science – and critics have often asserted that the results of RCTs are *not* therefore unequivocal.

In response, proponents have sought to cultivate a version of ‘controlled decontextualization’ (Cooper & Waldby, 2014, p. 11). Through methodological, institutional, and discursive means they have sought to improve the generalizability of their findings. A key means of doing so is through experimental replication. For example, Innovations for Poverty Action experimentally evaluated the Ultra Poor Graduation program in six countries on more than 10,000 subjects, leading them to conclude that ‘a multifaceted approach to increasing income and well-being for the ultrapoor is sustainable and cost-effective’ (Banerjee *et al.*, 2015).²² The congruence of results in contexts as varied as Honduras and Ethiopia is interpreted as a strong basis on which to endorse the intervention.

²² Other examples of replication are discussed in Browne *et al.* (2014, p. 228).

Such significant research efforts respond to concerns that ‘the dappled world’, to use Cartwright’s (1999) felicitous phrase, is a source of uncertainty within international development. But efforts such as these are expensive and unwieldy, unlikely to become the norm. Randomistas are doubtful of the possibility to replicate one-off experiments in another location (so-called ‘external replication’) (Levitt & List, 2009, p. 14). Advocates of experimentation want funding bodies to encourage ‘replication in multiple settings and thus external validity on assessments of promising interventions’ (Birdsall & Perakis, 2011), yet this may not be feasible.

In contrast, ‘internal replications’ may be less costly, risky, and difficult. In these, the models and calculations of published studies are replicated. Proponents realize that existing disciplinary practices and incentives militate against internal replication and that prior efforts to do so have failed (Hamermesh, 2007, p. 723); however, they argue that it is important because it not only corrects errant results (McCullough & McKittrick, 2009), but also ‘provides incentives for the experimenter to collect data carefully’ (Levitt & List, 2009, p. 14).²³ Replication, therefore, is envisioned as a means of reaffirming the epistemic virtues of experimentation, and organizations like 3ie have recently begun funding, guiding, and conducting international replications.

As a technique to discipline scientific selves, calls for replication are being joined by efforts to create experimental trial registries in development economics. Trial registries permit randomistas to submit their plans for research and analysis prior to conducting the work. This similarly responds to suggestions of impropriety that make experimental results less trustworthy. For one, the competitive pressures of academic research encourage biases in what is published. The dissemination of null results is infrequent despite the regular airing of concerns and suggestions for reforms – including one for a *Journal of Failed Experiments* from the World Bank’s Chief Economist (Basu, 2014). Moreover, there is a concern that researchers are ‘fishing’, or adapting their “models and specifications in order to yield statistically significant results” (Brown, Cameron, and Wood, 2014). Such critics are attuned to the social dynamics of research and seek institutional and methodological means to transform them, shoring up the trust in experimentality.

For her part, Esther Duflo believes that ‘The FDA requires reporting results of any funded medical trial. Institutions of this type need to be developed for field experiments’. The thought collective’s methodological handbook asserts that ‘there is virtually no downside to registering that we are undertaking an evaluation, and there is an important public benefit, so this form of registration is strongly encouraged’ (Glennerster & Takavarasha, 2013, p. 375). In 2013, the American Economic Association launched a registry for social science experiments, and 3ie developed a related initiative to catalogue impact evaluations of international aid projects.

The standardization of experimental knowledge permits the commensuration and, therefore, comparability of results. Similarly, the growth of “systematic reviews” in international

²³ Deaton (2010, p. 424) documents the methodological difficulties of experimentation that he believes ‘undermine any claims to statistical or epistemic superiority’.

development is a means to gather disparate studies into an authoritative assessment of particular interventions.²⁴ Organizations like the World Bank, 3ie, and J-PAL create systematic reviews in order to use multiple sources to minimize the uncertainty associated with any individual researcher or location.²⁵ In the words of 3ie proponents, *systematic* reviews differ from conventional literature reviews because they have ‘a clear protocol for systematically searching defined databases over a defined time period, with transparent criteria for the inclusion or exclusion of studies, as well as the analysis and reporting of study findings’ (Waddington *et al.*, 2012, p. 360). Thus, as with other aspects of the experimental system, systematic reviews are appealing because their procedural objectivity is said to remove subjective bias and errant results.

From evidence to implementation

If replication, registries, and reviews have sought to enhance the epistemic virtues for which RCTs were promoted, in more recent years the randomista thought collective has expanded one layer further. Although the thought collective has always been concerned with applied research and useful science, it is important to emphasize that core members are today engaged directly in the implementation of aid projects. The most notable example of this is Evidence Action, an NGO established in 2013 and initially ‘incubated’ by Innovations for Poverty Action, the Yale-based hub for experimental evaluations of aid. Evidence Action focuses on ‘scaling’ ‘interventions whose efficacy is backed by substantial rigorous evidence’ (Waddington & Leach, 2014). They seek to provide ‘cost-effective impact for many’. The impetus for such a shift is the acknowledgment that ‘There is a gap between what research shows is effective in development and what happens in practice’. Evidence Action’s first initiative was to absorb Deworm the World, a children’s health initiative created by Michael Kremer. Today it also distributes chlorinated water dispensers in areas without reliable infrastructure, often partnering with governments in places like India, Kenya, and Ethiopia. In both cases, Evidence Action bases its work on experimental evaluations of these programs, including Kremer’s early evaluation of deworming and subsequent work by J-PAL and IPA on clean water dispensers. It aims to ‘bridge the gap between rigorous research and pilot programs on the one hand, and institutionalized programs on the other’.²⁶

If the randomista thought collective is visualized as a Matryoshka doll with J-PAL at the center, Evidence Action is perhaps the outermost nesting doll. Its recent formation represents the growth of the thought collective from a few economists engaged in academic research, to organizations suited for coordinating and translating such research, to now the actual aid organizations themselves. Earlier work sought to convince and enroll allies into the randomista

²⁴ Systematic reviews are most commonly associated with the biomedical work of the Cochrane initiative, which randomistas invoke as a model.

²⁵ Moreira (2007) discusses systematic reviews in medicine as a process of ‘disentanglement’ and ‘qualification’ of results.

²⁶ See: <http://www.evidenceaction.org/who-we-are/#vision-values>

worldview, but Evidence Action acts on behalf of a consensus within the thought collective about ‘what works’. As a semi-autonomous entity, it can move forward with its work rather than needing to enroll many others for each action. This status was made clear in a recent dispute (the so-called ‘worm wars’) where the study offered by Evidence Action to support deworming was questioned by a replication (Davey *et al.*, 2015). While a full analysis of the dispute is beyond the scope of this paper (but see Evans, 2015), what matters to emphasize is both the ongoing debates about the credibility of development knowledge and the way in which Evidence Action was able to continue its work despite the uncertainty. It is, then, an institutionalization of the randomistas’ rise.

Conclusion

Published in 1986, *An anthropological critique of development* long preceded the profusion of evidence-based policy and the randomistas. Yet in that volume - subtitled ‘the growth of ignorance’ - Mark Hobart (1986, p. 4) noted that ‘claims to knowledge and the attribution of ignorance are central themes to development and remain seriously under-studied’. In the intervening years, more attention has been turned to the politics of knowledge within development, yet Hobart’s insight about the importance of claims to knowledge or ignorance remains pointed. The rise of the randomistas represents perhaps the apogee of ‘the attribution of ignorance’. In the past 15 years, the nature of knowing and the state of knowledge have become matters of concern within an aid industry transformed by demands for accountability, the influence of economists, and the proliferation of NGOs. Adeptly legitimated by proponents, RCTs have filled this void but, as scholars of experimentation in medicine have documented, RCTs are incapable of securing certainty. Indeed, as McGoey (2010, p. 71) argues, it ‘is the very methodological weaknesses of RCTs that imbues them with the authority they hold: for to deny the reliability of a particular study, one must reach for more data, more studies, larger RCTs, in order to justify the validity of one’s objections’. In the case of the aid randomistas, it is a tacking back and forth – between promises of certainty and accusations of uncertainty; between registers of argumentation; and between audiences – that has helped reconstitute the spirit of international aid. Their repertoires of justification have certainly included their own uncertainties and evasions, but the history demonstrates they have successfully attributed uncertainty elsewhere. Furthermore – in a logic recalling Power’s (1997) theorization of the audit society – this lack of certitude has not called into question certainty *writ large* but served to justify further experimentation.

Attention to the dynamics of critique in international aid reveals in part how institutions like the World Bank and disciplines like economics achieve their enduring influence. While some criticisms remain unheard within the halls of the World Bank (or at the very least unengaged), in other ways, aid organizations are highly responsive to critique. In the case of the randomistas, this was facilitated due to the conditions of possibility and their legitimating strategies. But this incorporation of critique also functions to displace more radical voices, a dynamic that Boltanski (2011) suggests is crucial to the dominance of institutions whose

reflexivity can immunize them by reappropriating and modifying critique. As James Ferguson (2014) wryly notes, ‘The World Bank has always been a very articulate critic of positions that it held ten years earlier’. In part, the rise of the randomistas suggests, this is due to the moral orientations and grammars which permeate its regimes of justification. Such an approach differs from those who view technocracy as powerful due to its popular inaccessibility or its depoliticization by elites; such an understanding – for which Habermas’s (1985) concern for the colonization of the lifeworld by instrumental rationality and market forces is a touchstone – can only tell at most half the story. They too often miss the affective and ethical appeals immanent within technical domains. The randomistas show how the translation of an ethos into a thought collective can reorient methodology and practice—and with it the distribution of authority and resources. These changes deserve criticism of their own, but in addition to marshalling a critique, this paper has suggested we need a better understanding of the social, moral, and epistemological dynamics at play, without which, the criticism is likely to miss its mark.

Acknowledgments

This paper was substantially written while in residence at the Max Planck Institute for the History of Science and I am grateful to its director, Dagmar Schäfer, for welcoming me. At Michigan, Paul Edwards, Howard Stein, and Dan Hirschman were particularly helpful interlocutors. Emily Brownell and Sam Dinger read and commented on the entire paper. Conversations with Margarita Rayzberg, Luciana de Souza Leão, Jigar Bhatt, Fiona Achi, Diana Graizbord, and Zach Griffen were enlightening and reveal how much more there is to say on this topic. The four *Economy and Society* reviewers were astute, critical, and constructive. Finally, I am grateful to Ann Kelly and Linsey McGoey for organizing the special edition.

References

American Economic Journal: Applied Economics: Vol. 7 No. 1 (January 2015)

Angrist, J. & Pischke, J. (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. Working Paper 15794. National Bureau of Economic Research. Retrieved from <http://www.nber.org/papers/w15794>.

Banerjee, A. (2007). *Making aid work*. Cambridge, MA: MIT Press.

Banerjee, A. & Duflo, E. (2009). The experimental approach to development economics. *Annual Review of Economics* 1, 151-78.

Banerjee, A. & Duflo, E. (2011). *Poor economics: A radical rethinking of the way to fight global poverty*. New York, NY: Public Affairs.

Banerjee, A., Karlan, D., & Zinman, J. (2015). Six randomized evaluations of microcredit: Introduction and further steps. *American Economic Journal: Applied Economics*, 7 (1), 1-21.

- Banerjee, A., Duflo, E. Goldberg, N., Karlan, D., Osei, R., & Pariente, W.** (2015). "A multifaceted program causes lasting progress for the very poor: evidence from six countries." *Science* 348(6236).
- Bartholomé, Y.** (2005). Campbell, Donald T. *Encyclopedia of Social Measurement*, 239-244. Elsevier.
- Basu, K.** (2014). Randomisation, causality, and the role of reasoned intuition. *Oxford Development Studies*, 42 (4), 455-472.
- Bates, R.** (2007). Forum response: Making aid work. *Boston Review*.
- Bénatouïl, T.** (1999). A tale of two sociologies: The critical and the pragmatic stance in contemporary French sociology. *European Journal of Social Theory*, 2(3), 379–96.
- Berndt, C.** (2015). Behavioural economics, experimentalism, and the marketization of development. *Economy & Society* 44 (4), 567-591.
- Birdsall, N.** (2006). A major step forward on impact evaluation. Center for Global Development. Retrieved from <http://www.cgdev.org/article/major-step-forward-impact-evaluation>
- Birdsall, N. & Perakis, R.** (2011). "The New USAID Evaluation Policy Is Not Getting Nearly Enough Attention." *Center For Global Development*. Accessed October 28. <http://www.cgdev.org/blog/new-usaid-evaluation-policy-not-getting-nearly-enough-attention>
- Blattman, C.** (2011). Impact evaluation 3.0? Retrieved from <https://chrisblattman.com/2011/09/02/impact-evaluation-3-0/>
- Blattman, C.** (2014). The latest in faith-based development: Randomized control trials? *Chris Blattman Blog*. Retrieved from <http://chrisblattman.com/2014/03/11/the-latest-in-faith-based-development-randomized-control-trials/>
- Boltanski, L.** (1999). *Distant suffering: Morality, media, and politics*. Cambridge: Cambridge University Press.
- Boltanski, L.** (2011). *On critique: A sociology of emancipation*. London: Polity.
- Boltanski, L. & Thévenot, L.** (1999). The sociology of critical capacity. *European Journal of Social Theory* 2(3), 359-377.
- Boltanski, L. & Thévenot, L.** (2006). *On justification: Economies of worth*. Princeton, NJ.: Princeton University Press.
- Boltanski, L. & Chiapello, E.** (2007). *The new spirit of capitalism*. London: Verso.
- Bourdieu, P. (1990). *Homo Academicus*. Palo Alto, CA: Stanford University Press.
- Breslau, D.** (1998). *In search of the unequivocal: The political economy of measurement in US labor market policy*. Santa Barbara, CA: Praeger.
- Browne, A., Cameron, D. & Wood, B.** (2014). Quality evidence for policymaking: I'll believe it when I see the replication. *Journal of Development Effectiveness*, 6 (3), 215-235.
- Browne, C.** (2014). The institution of critique and the critique of institutions. *Thesis Eleven*, 124 (1), 20–52.
- BusinessWeek.** (2010). The pragmatic rebels. Retrieved from: <https://www.bloomberg.com/news/articles/2010-07-02/the-pragmatic-rebels>

- Campbell, D.T.** (1969). Reforms as experiments. *American Psychologist* 24(4), 409-429.
- Campbell, D.T.** (1973). The social scientist as methodological servant of the experimenting society. *Policy Studies Journal* 2(1), 72-75.
- Cartwright, N.** (1999). *The dappled world: A study of the boundaries of science*. Cambridge: Cambridge University Press.
- Cartwright, N.** (2007). Are RCTs the gold standard. *BioSocieties*, 2(1), 11-20.
- Celikates, R.** (2006). From critical social theory to a social theory of critique: On the critique of ideology after the pragmatic turn. *Constellations*, 13(1), 21-40.
- CGD.** (2006). *When will we ever learn? Improving lives through impact evaluation*. Washington, DC: CGD.
- Choi, H.** (2014). We could all use a little nudge. Work In Progress: The Hewlett Foundation Blog. Retrieved from <http://www.hewlett.org/blog/posts/we-could-all-use-little-nudge>
- Chouliaraki, L.** (2006). *The spectatorship of suffering*. London: Sage.
- Collins, H.M.** (1985). *Changing order: Replication and induction in scientific practice*. Chicago, IL: University of Chicago Press.
- Cooper, M. & Waldby, C.** (2014). *Clinical labor: Tissue donors and research subjects in the global bioeconomy*. Durham, NC: Duke University Press.
- Coyle, D.** (2011). Small changes, big results: Forum response. *Boston Review*.
- Daston, L. & Galison, P.** (2010). *Objectivity*. New York, NY: Zone Books.
- Davies, W.** (2014). *The limits of neoliberalism*. London: Sage.
- Davey, C., Aiken, A., Hayes, R. & Hargreaves, J.** (2015). Re-analysis of health and educational impacts of a school-based deworming programme in western Kenya: A statistical replication of a cluster quasi-randomized stepped wedge trial. *International Journal of Epidemiology*, 44 (4), 1581-1592.
- Davis, J.B.** (2013). Economics imperialism under the impact of psychology: The case of behavioral development economics. *Oeconomia*, 3(March), 119-138.
- Davis, K., Fisher, A., Kinsbury, B. & Merry, S.E.** (2012). *Governance by indicators: Global power through classification and rankings*. Oxford: Oxford University Press.
- Deaton, A.** (2010). Instruments, randomization, and learning about development. *Journal of Economic Literature*, 48(2), 424-455.
- Dehue, T.** (2001). Establishing the experimenting society: The historical origins of social experimentation according to the randomized controlled design. *The American Journal of Psychology*, 114 (2), 283-302.
- Duflo, E. & Kremer, M.** (2005). Use of randomization in the evaluation of development effectiveness. In G. Pitman, O. Feinstein & K. Ingram (Eds.). *Evaluating development effectiveness* (pp. 93-120). London: Routledge.
- The Economist.** (2011a). Health and unpicked low-hanging fruit. Retrieved from http://www.economist.com/blogs/freeexchange/2011/04/banerjee_and_duflo_0

- The Economist*. (2011b). Untying the knot. Retrieved from <http://www.economist.com/node/18584122>
- The Economist*. (2011c). Fiesta de los randomistas. Retrieved from http://www.economist.com/blogs/freeexchange/2011/04/banerjee_and_duflo
- The Economist*. (2011d). Remembering to respect the preferences of the poor.” Retrieved from http://www.economist.com/blogs/freeexchange/2011/04/banerjee_and_duflo_1
- The Economist*. (2013, December 14). Random harvest.
- Epstein, S.** (2009). *Inclusion: The politics of difference in medical research*. Chicago, IL: University of Chicago Press.
- Evans, D.** (2015). Worm wars: The anthology. *Development Impact*. Retrieved from <http://blogs.worldbank.org/impactevaluations/worm-wars-anthology>
- Eyal, G.** (2013). For a sociology of expertise: The social origins of the autism epidemic. *American Journal of Sociology*, 118(4), 863–907.
- Eyben, R. & Roche, C.** (2013). The political implications of evidence-based approaches. *From Poverty to Power*. Retrieved from <http://oxfamblogs.org/fp2p/the-political-implications-of-evidence-based-approaches-aka-start-of-this-weeks-wonkwar-on-the-results-agenda/>
- Ferguson, J.** (2014, June 10). Interview with James Ferguson. *Humanity Journal*.
- Fleck, L.** (1981[1934]). *Genesis and development of a scientific fact*. Chicago, IL: University of Chicago Press.
- Fine, B. & Milonakis, D.** (2009). *From economics imperialism to freakonomics: The shifting boundaries between economics and other social sciences*. London: Routledge.
- Foucault, M.** (2009). *Security, territory, population: Lectures at the College de France 1977-1978*. New York: Picador.
- Fourcade, M., Etienne, O., & Algan, Y.** (2015). The superiority of economists. *The Journal of Economic Perspectives*, 29 (1), 89–113.
- Haas, P.M.** (1992). Introduction: Epistemic communities and international policy coordination. *International Organization*, 46(1), 1-35.
- Habermas, J.** (1985). *Theory of communicative action*. Boston, MA: Beacon Press.
- Hacking, I.** (1992). Statistical language, statistical truth, and statistical reason: The self-authentication of a style of scientific reasoning. In E. McMullin (Ed.), *Social Dimensions of Science* (pp. 130- 57). Notre Dame, Indiana: University of Notre Dame Press.
- Hamermesh, D.S.** (2007). Viewpoint: Replication in economics. *Canadian Journal of Economics* 40 (3), 715-733.
- Haskell, T.** (1985). Capitalism and the origins of the humanitarian sensibility, Part 1. *American Historical Review*, 90(2), 339-361.
- High, C., Kelly, A. & Mair, J.** (2012). *The anthropology of ignorance: An ethnographic approach*. London: Palgrave Macmillan.
- Hilgartner, S.** (2000). *Science on stage: Expert advice as public drama*. Palo Alto, CA: Stanford University Press.

- Hirschman, D. & Berman, E.P.** (2014). Do economists make policies? On the political effects of economics. *Socio-Economic Review*, 12(4), 779-811.
- Hirschman, D. & Reed, I.** (2014). Formation stories and causality in sociology. *Sociological Theory*, 32(4), 259–82.
- Hobart, M.** (Ed.) (1986). *An anthropological critique of development: The growth of ignorance*. London: Routledge.
- Galison, P.** (1987). *How experiments end*. Chicago, IL: University of Chicago Press.
- Gapper, J.** (2012). Lunch with the FT: Esther Duflo. *Financial Times*. Retrieved from <http://www.ft.com/cms/s/2/81804a1a-6d08-11e1-ab1a-00144feab49a.html>
- Gates, B.** (2013). *Measuring progress: Annual letter 2013*. Gates Foundation. Retrieved from <http://www.gatesfoundation.org/Who-We-Are/Resources-and-Media/Annual-Letters-List/Annual-Letter-2013>
- Gilbert, N.** (2013). International aid projects come under the microscope. *Nature*. Retrieved from <http://www.nature.com/news/international-aid-projects-come-under-the-microscope-1.12268>
- Glennerster, R. & Kremer, M.** (2011). Small changes, big results. *Boston Review*. Retrieved from <http://www.bostonreview.net/glennersterner-kremer-changes-big-results>.
- Glennerster, R. & Takavarasha, K.** (2013). *Running randomized evaluations: A practical guide*. Princeton, NJ: Princeton University Press.
- Glennerster, R.** (2013). RCTs provide independence or objectivity through methodology. *Running Randomized Evaluations*. Retrieved from <http://runningres.com/blog/2013/12/3/rcts-provide-independence-or-objectivity-through-methodology>
- Glennerster, R.** (2014). Comparing cost-effectiveness across contexts. *Running Randomized Evaluations*. Retrieved from <http://runningres.com/blog/2014/3/11/comparing-cost-effectiveness-across-contexts>
- Gross, M. & McGoey, L.** (Eds.). (2015). *Routledge international handbook of ignorance studies*. London: Routledge.
- Immerwahr, D.** (2015). *Thinking small: The United States and the lure of community development*. Cambridge: Harvard University Press.
- Jensen, C.B. & Winthereik, B.R.** (2013). *Monitoring movements: Recursive partnerships and infrastructures*. Cambridge, MA: MIT Press.
- Joffe, M.** (2013). Teaching evidence-based economics. *Royal Economic Society*. <http://www.res.org.uk/view/art6Oct13Features.html>
- Joffe, M.** (2014). Can economics be evidence-based? *Royal Economic Society*. Retrieved from <http://www.res.org.uk/view/art4aApr14Features.html>
- Karlan, D. & Appel, J.** (2012). *More than good intentions: Improving the ways the world's poor borrow, save, farm, learn and stay healthy*. New York, NY: Plume.
- Krause, M.** (2010). Accounting for state intervention: The social histories of 'beneficiaries'. *Qualitative Sociology*, 33 (4), 533-547.

- Krause, M.** (2014). *The good project: Humanitarian relief NGOs and the fragmentation of reason*. Chicago, IL: University of Chicago Press.
- Krugman, P.** (1995). *Peddling prosperity: Economic sense and nonsense in the age of diminished expectations*. New York, NY: Norton.
- Lakoff, A.** (2006). *Pharmaceutical reason: Knowledge and value in global psychiatry*. Cambridge: Cambridge University Press.
- Latour, B.** (1987). *Science in action*. Cambridge, MA: Harvard University Press.
- Latour, B.** (1993). *We have never been modern*. Cambridge, MA: Harvard University Press.
- Latour, B.** (2004). Why has critique run out of steam? From matters of fact to matters of concern. *Critical Inquiry*, 30, 225-248.
- Levine, R.** (2014). What's the one really good reason not to evaluate? *Work in Progress*. Retrieved from <http://www.hewlett.org/blog/posts/friday-note-whats-one-really-good-reason-not-evaluate>
- Levitt, S.D. & List, J.** (2009). Field experiments in economics: The past, the present, and the future. *European Economic Review*, 53 (1), 1–18.
- List, J.** (2009). An introduction to field experiments in economics. *Journal of Economic Behavior & Organization* 70(3), 439-442.
- Ludwig, J., Kling, J. & Mullainathan, S.** (2011). Mechanism experiments and policy evaluations. *Journal of Economic Perspectives*, 25 (3), 17-38.
- MacKenzie, D.** (2008). *Material markets: How economic agents are constructed*. Oxford: Oxford University Press.
- Marks, H.** (2000). *The progress of experiment: Science and therapeutic reform in the United States, 1900-1990*. New York, NY: Cambridge University Press.
- McCloskey, D.N.** (1998). *The rhetoric of economics*. Madison, WI: University of Wisconsin Press.
- McCullough, B.D. & McKittrick, R.** (2009). *Check for numbers: The case for due diligence in policy formation*. Vancouver: Fraser Institute.
- McGoey, L.** (2009). Pharmaceutical controversies and the performative value of uncertainty. *Science as Culture*, 18(2), 151–64.
- McGoey, L.** (2010). Profitable failure: Antidepressant drugs and the triumph of flawed experiments'. *History of the Human Sciences*, 23 (1), 58–78.
- McGoey, L.** (2012). Strategic unknowns: Toward a sociology of ignorance. *Economy & Society*, 41(1), 1-16.
- McGovern, M.** (2011). Popular development economics: An anthropologist among the Mandarins. *Perspectives on Politics*, 9(2), 345-55.
- McMurtrie, B.** (2014, June 9). Poverty under the microscope. *Chronicle of Higher Education*.
- Milonakis, D. & Fine, B.** (2008). *From political economy to economics: Method, the social, and the historical in the evolution of economic theory*. London: Routledge.

- Mirowski, P.** (1991). *More heat than light: Economics as social physics, physics as nature's economics*. New York, NY: Cambridge University Press.
- Mirowski, P.** (2002). *Machine dreams: Economics becomes a cyborg science*. New York, NY: Cambridge University Press.
- Mirowski, P.** (2014). *Never let a serious crisis go to waste: How neoliberalism survived the financial meltdown*. New York, NY: Verso.
- Mirowski, P. & Plehwe, D.** (Eds.). (2009). *The road from Mont Pelerin: The making of the neoliberal thought collective*. Cambridge, MA: Harvard University Press.
- Mishra, A. & Cameron, T.** (2013). Register of impact evaluation published studies protocol. 3ie Working Paper. Retrieved from http://www.3ieimpact.org/media/filer_public/2013/05/07/rieips_protocol_2_1.pdf
- Mitchell, T.** (1998). Fixing the economy. *Cultural Studies*, 12(1), 82-101.
- Mitchell, T. (2002). *Rule of Experts: Egypt, Techno-Politics, Modernity*. Berkeley: University of California Press.
- Moore, M.** (2007). Response. *Making aid work*. Cambridge, MA: Boston Review.
- Moreira, T.** (2007). Entangled evidence: Knowledge making in systematic reviews in healthcare. *Sociology of Health & Illness*, 29 (2), 180-197.
- Morgan, M.** (2005). Experiments versus models: New phenomena, inference and surprise. *Journal of Economic Methodology*, 12(2), 317–329.
- Ogden, T.** (2015). Hopefully not the final word on microcredit. *Financial Access Initiative*. Accessed October 28. Retrieved from <http://www.financialaccess.org/blog/2015/8/5/hopefully-not-the-final-word-on-microcredit>.
- Olofsgård, A.** (2012). Why political short-sightedness and randomised control trials can be a deadly mix for aid effectiveness. *Vox*. Retrieved from <http://www.voxeu.org/article/deadly-mix-aid-effectiveness>
- Parker, I.** (2010). The poverty lab. *New Yorker*.
- Pearce, W. & Raman, S.** (2014). The new randomised controlled trials (RCT) movement in public policy: Challenges of epistemic governance. *Policy Sciences* 47 (4), 387-402.
- Petryna, A.** (2009). *When experiments travel: Clinical trials and the global search for human subjects*. Princeton, NJ: Princeton University Press.
- Picciotto, R.** (2014). Is impact evaluation evaluation? *European Journal of Development Research*, 26 (1), 31-38.
- Porter, T.** (1996). *Trust in numbers: The pursuit of objectivity in science and public life*. Princeton, NJ: Princeton University Press.
- Power, M.** (1997). *The audit society: Rituals of verification*. Oxford: Oxford University Press.
- Pritchett, L.** (2014a). Rigorous evidence isn't. *Building State Capacity*. Retrieved from <http://buildingstatecapability.com/2014/02/20/rigorous-evidence-isnt/>
- Pritchett, L.** (2014b). An homage to the randomistas on the occasion of J-PAL's 10th Anniversary: Development as a faith-based activity. Center for Global Development.

Retrieved from <http://www.cgdev.org/blog/homage-randomistas-occasion-j-pal-10th-anniversary-development-faith-based-activity>

Proctor, R. and Schiebinger, L. (2008). *Agnology: The making and unmaking of ignorance*. Palo Alto, CA: Stanford University Press.

Ramachandran, S. (2013). A colorful launch yet again in Delhi. Evidence Action. Retrieved from <http://www.evidenceaction.org/blog-full/a-colorful-launch-yet-again-in-delhi?rq=delhi%20launch>

Rayzberg, M. (2014, April). From impact assessment to impact evaluation. Presentation at *STGlobal*.

Rayzberg, M. (forthcoming). "Fairness in the Field: The Ethics of Resource Allocation in Randomized Controlled Field Experiments." *Science, Technology & Human Values*.

Ravallion, M. (2009a). 'Should the randomistas rule?' *The Economists' Voice*.

Ravallion, M. (2009b). Evaluation in the practice of development. *The World Bank Research Observer*, 24 (1), 29-53.

Ravallion, M. (2014). Taking ethical validity seriously. *Development Impact*. Retrieved from <http://blogs.worldbank.org/impacetevaluations/taking-ethical-validity-seriously>

Reddy, S. (2012). Randomise this! On *Poor economics*. *Review of Agrarian Studies* 2 (2), 60-73.

Rheinberger, H-J. (2010). *An epistemology of the concrete: Twentieth-century histories of life*. Palo Alto, CA: Stanford University Press.

Rodrik, D. (2008). The new development economics: We shall experiment, but how shall we learn? *HKS Working Paper No. RWP08-055*.

Rose, N. (1999). *Powers of freedom: Reframing political thought*. Cambridge: Cambridge University Press.

Rottenburg, R. (2009a). *Far-Fetched Facts: A Parable of Development Aid*. Cambridge: MIT Press.

Rottenburg, R. (2009b). Social and public experiments and new figurations of science and politics in postcolonial Africa. *Postcolonial Studies* 12 (4), 423-440.

Sandefur, J. (2015). The final word on microcredit? Retrieved from: <https://www.cgdev.org/blog/final-word-microcredit>

Savage, M. (2013). The 'Social Life of Methods': A Critical Introduction. *Theory, Culture & Society*, 30(4), 3-21.

Savedoff, W. (2014). End the evaluation wars: A Plea to shift from the abstract to the specific. 2015. *Center for Global Development*. Accessed October 28. Retrieved from <http://www.cgdev.org/blog/end-evaluation-wars-plea-shift-abstract-specific>

Schmitt, D. (2014). Donors should put evidence before politics and diplomacy. *SciDevNet*. Retrieved from <http://www.scidev.net/global/aid/opinion/donors-evidence-politics-diplomacy.html>

Shapin, S. & Schaffer, S. (1985). *Leviathan and the air-pump: Hobbes, Boyle, and the experimental life*. Princeton, NJ: Princeton University Press.

- Stein, H.** (2008). *Beyond the World Bank agenda: An institutional approach to development*. Chicago, IL: University of Chicago Press.
- Steinmetz, G. (Ed.). (2005). *The politics of method in the human sciences: Positivism and its epistemological others*. Durham, NC: Duke University Press.
- Stoler, A.** (2011). Colonial Aphasia: Race and disabled histories in France. *Public Culture*, 23(1), 121-156.
- Timmermans, S. & Berg, M.** (2003). *The gold standard: The challenge of evidence-based medicine and standardization in health care*. Philadelphia, PA: Temple University Press.
- Teele, D.** (2014). *Field experiments and their critics*. New Haven, CT: Yale University Press
- Waddington, H. et al.** (2012). How to do a good systematic review of effects in international development: A tool kit. *Journal of Development Effectiveness*, 4(3), 359-387.
- Waddington, H. & Leach, B.** (2014). How much evidence is enough for action? The value of systematic reviews of the totality of evidence. *Evidence Matters*. Retrieved from <http://blogs.3ieimpact.org/how-much-evidence-is-enough-for-action/>
- Wagner, P.** (1999). After justification repertoires of evaluation and the sociology of modernity. *European Journal of Social Theory*, 2(3), 341–57.
- Weber, M.** (1978). *Economy & Society: An outline of interpretive sociology*. Berkeley, CA: University of California Press.
- Whitty, C. & Dercon, S.** (2013). The evidence debate continues. *From Poverty to Power*. Retrieved from <http://oxfamblogs.org/fp2p/the-evidence-debate-continues-chris-whitty-and-stefan-dercon-respond/>
- Will, C. & Moreira, T.** (Eds.). (2010). *Medical proofs, social experiments: Clinical trials in shifting contexts*. Farnham: Ashgate.
- Williams, R.** (1977). *Marxism and literature*. Oxford: Oxford University Press.
- Woods, N.** (2000). The challenges of multilateralism and governance. In C. Gilbert & D.Vines (Ed.), *The World Bank: Structure and policies* (pp.132-158). Cambridge: Cambridge University Press.
- Woolcock, M.** (2009). Toward a plurality of methods in project evaluation: a contextualised approach to understanding impact trajectories and efficacy. *Journal of Development Effectiveness*, 1(1), 1-14.

Kevin P. Donovan is in the programmes in Anthropology & History and Science, Technology & Society at the University of Michigan. His current research examines the reformulation of sovereignty and value in the wake of empire in East Africa. In addition, he studies the cultural politics of corporations and the material politics of infrastructure in sub-Saharan Africa. More information is available at <http://kevinpdonovan.com/>