# Evidence in action:
# an anthropology of global poverty alleviation efforts

Fiona Gedeon Achi

Department of Anthropology, Faculty of Arts
McGill University
Montréal, QC, Canada

August 2019
A thesis submitted to McGill University in partial fulfillment of the requirements
of the degree of Doctor of Philosophy

# TABLE OF CONTENTS

# Acknowledgements

The end of the dissertation writing process, which consists of many solitary hours of work, offers a particular vantage point for reflecting on the many relations without which this PhD project would not have existed, let alone be completed. I find that the gratitude I owe to people is often too difficult to summarize in words.

I am filled with sincere appreciation for all the institutions that opened their doors so that I could conduct fieldwork and learn about their compelling work. I thank the team at the Center for Effective Global Action in Berkeley; the team of Evidence Action in both Nairobi and around Kisumu; the J-PAL staff and the NGO Educate Girls staff in India; my interviewees at J-PAL Global in Boston and GiveWell in San Francisco; the organizing and teaching team during the J-PAL executive training in Denmark; the members of Elimu Impact Evaluation Center in Kenya, in particular its director Dr. Matthieu Chemin at McGill University; and several governmental officials in Kenya. While I cannot cite them for reasons of anonymity, many people working at these field sites have become friends who have accompanied me in multiple ways throughout this PhD work. At McGill University, Shelley Clark was so kind to introduce me to some researchers conducting randomized evaluation, thus helping me secure my first fieldwork site in India. In Davis, California, where I lived some months before undertaking fieldwork, I was lucky to be kindly welcomed and supported by the Department of Agricultural and Resource Economics and the Department of Anthropology at the University of California at Davis. Thank you in particular to Michael Carter for introducing me to the Center for Effective Global Action. While doing fieldwork, I was lucky to encounter enduring friendships: thank you Mareike Decker, Anna Dubuis, Steve Gathe, Sammy Gebre, and Ashley Wilson.

For continuous intellectual guidance, inspiration, and support, I am very grateful to my supervisor Tobias Rees. My two committee members Stephen Collier and Setrag Manoukian also provided significant intellectual guidance, especially as I prepared the bibliographic essays and the proposal for this PhD work. Thank you also to Vincanne Adams, Vololona Rabeharisoa, and Kaushik Sunder Rajan for listening and exchanging about my research.

Many friends and colleagues have shaped this dissertation by providing feedback at various stages – they have made thinking more enjoyable for me. Members of the "Writing

To friends who saw me walk the PhD route, I am deeply grateful to you all for infinite reasons: Nassima Abdelghafour, Bilge Beril Aksoy, Alessio Bacci, Roberta Badiali, Sophia Bamert, Elettra Bietti, Linda Boukhris, Ramazan Cansever, Kim Colopinto, Anaïs Couasnon, Laurianne Coutier, Kim Colopinto, Céleste Fournier, Andrew Friedman, Marcos Ginestra, Soizic Jean-Baptiste, Tamar Kaneh-Shalit, Ebru Kayaalp, Handan Kekec Cansever, Baker Kiggundu, Sam Jang, Cosima Jend, Silvia Lami, Émilie Liabeuf, Alan Marie, Marc Moraine, Marie-Christine Naffah, Maissam Nimer, Souad Osseiran, Bekir Başar Özer, Alix Pahaut, Marco Quaranta, Nicole Rigillo, Ishani Saraf, Akiva Shalit-Kaneh, Quentin Simons, Raphaëlle Selliez-Vandernotte, William F. Stafford, Molly Tenzer, Abdullah Terzioğlu, Mariana Tzoneva, Duygu Yeral, Maria-Giovanna (Magi) Ziccardi, and Victoria Zurita.

I feel immense gratitude for the care and kindness I received from my multiple families: la "tribu Achi", but also Giacomo, les Stoeffler-Puault, les Vouriot-de Matos, and Shani's "Turkish" family: Fernanda, Mert, E.J, and Mustafa. To Quentin and Shani (and pezzettino), I owe a life I could have never imagined. I dedicate this thesis to Monica, my mother, smiling at the thought of what she might have said when reading it.

*What is a good life? What is its opposite? These are questions to which no two men will give the same answers.*

-Salman Rushdie, *The Golden House*

# Abstract

Based on fourteen months of fieldwork among those who evaluate anti-poverty interventions and work to transform "successful" interventions into policy and programs "at scale" around the world, my dissertation examines the practice of *evidence-based development*. Since the early 2000s, an influential group of economists and development practitioners insists that global poverty alleviation efforts should be grounded in scientific evidence about "what works" to reduce poverty and propose randomized evaluations as the best instrument for identifying effective solutions worldwide. A method that originated in agricultural sciences and medicine, randomized evaluations are a specific form of experiment that consist of dividing randomly the target population into treatment and control groups in order to measure the impact of an intervention. For instance, I have studied a project in India that compared educational outcomes between adolescent girls that received an empowerment intervention and those who did not.

In a context where the eradication of "poverty in all its forms everywhere" has been positioned as the top global priority by the United Nations' Sustainable Development Goals, to be reached by 2030, and a substantial debate persists over the relevance of international aid, evidence-based development has taken shape as a significant movement tying together new research and policy institutes, funding agencies, NGOs, and governments across the planet. Following people and projects across Kenya, India, and the USA, this dissertation aims to document the imaginations of a better world made of ideas but also of centrally constituted through repeated practices and technologies that sustain evidence-based development.

In the first chapter, I investigate how my interlocutors situate their work within older and contemporary development efforts, with a focus on understanding what might give randomized evaluation evidence such actionable force to guide decision-making. The remaining chapters are all concerned with showing how evidence-based development seeks to impact millions of lives around the world either by extending a "proven" water sanitation initiative in Western Kenya (Chapter 2), by partnering with governments to reach the highest level of scale across countries (Chapter 3), or by devising frameworks to promote the generalizability of randomized evaluation's findings across "contexts" (Chapter 4). Across all these case studies, this dissertation aims to make apparent the multiple (political) projects and imaginations that evolve alongside evidence-based development and condition both its possibilities and its limits to shape policy efforts around the world. I conclude with an actor-oriented reflection on the difficulty of assessing to what extent, and exactly in which ways, evidence-based development reconfigures funding priorities and policy agendas worldwide.

# Abrégé

S'appuyant sur quatorze mois de terrain parmi ceux qui évaluent des interventions contre la pauvreté et s'efforcent à transformer les interventions « efficaces » en des politiques et programmes « à grande échelle » autour du monde, cette thèse examine la pratique du *développement fondé sur les preuves*. Depuis les années 2000, un groupe influent d'économistes et de praticiens du développement préconise que les efforts mondiaux pour la réduction de la pauvreté devraient être basés sur des données scientifiques illuminant « ce qui fonctionne » et présente les expérimentations aléatoires comme le meilleur instrument pour identifier les solutions efficaces à travers le monde. Une méthode qui trouve ses origines dans les sciences agronomiques et la médecine, les expérimentations aléatoires consistent à diviser aléatoirement la population cible en groupe « de traitement » et groupe « de contrôle », afin de mesurer l'impact d'une intervention. Par exemple, j'ai étudié un projet en Inde qui comparait les résultats éducatifs de femmes adolescentes qui recevaient un programme d'autonomisation avec ceux de femmes adolescentes qui ne recevaient pas le programme.

Alors que l'élimination de la « pauvreté sous toutes ses formes partout » est aujourd'hui positionnée comme la priorité majeure des Objectifs de Développement Durables établis par les Nations Unies pour être atteints d'ici 2030, et tandis qu'un important débat subsiste autour de la pertinence de l'aide internationale, le développement fondé sur les preuves est devenu un mouvement considérable qui rassemble des nouveaux instituts de recherche et politiques publiques, des organismes de financements, des ONG, et des gouvernements autour du monde. Étudiant des projets à travers le Kenya, l'Inde, et les États-Unis, cette thèse vise à documenter les imaginations d'un monde meilleur – faites d'idées mais aussi constituées par des pratiques et des technologies répétées – qui soutiennent le développement fondé sur les preuves.

Dans un premier chapitre, j'analyse comment mes interlocuteurs situent leur travail vis-à-vis des efforts de développement (à la fois passés et contemporains) dans le but principal de comprendre ce qui donne aux données produites par les expérimentations aléatoires leur force d'action. Les chapitres restants ont pour objectif de montrer différentes façons dont le développement fondé sur les preuves cherche à toucher des millions de vie à travers le monde, que ce soit en implémentant une intervention « prouvée efficace » pour assainir l'eau dans l'ouest du Kenya (Chapitre 2), en collaborant avec des gouvernements pour atteindre la plus grande « ampleur d'échelle » à travers les pays (Chapitre 3), ou en développant des cadres pour promouvoir la « généralisabilité », au travers de différents « contextes », des résultats obtenus

par expérimentations aléatoires (Chapitre 4). Dans toutes ces études de cas, cette thèse cherche à rendre visible les multiples projets (politiques) et imaginations qui évoluent aux côtés du développement fondé sur les preuves et qui conditionnent à la fois ses possibilités et ses limites d'action sur les politiques (publiques) autour du monde. Cette thèse se conclue avec une réflexion, guidée par les questionnements de mes interlocuteurs, sur la difficulté d'estimer dans quelle mesure, et exactement de quelles manières, le développement fondé sur les preuves reconfigure les priorités de financement et les agendas politiques à l'échelle mondiale.

# Introduction

In March 2014, I began fieldwork on a development program in northern India. In many ways, my fieldwork days resembled those described in many ethnographies of development I had read before leaving for India. I woke up before dawn and travelled long distances in a 4x4 vehicle on uneven roads between the regional office of the NGO and the scattered villages where an educational development program was taking place. I spent long hours waiting and observing while the NGO staff delivered training and activities to program recipients or met with local officials. While I accompanied NGO field coordinators on their monitoring routines, I took notes of implementation issues such as when an orphan boy was included in a school parliament designed to comprise only girls to enhance adolescent girls' school attendance and improve their life skills and sense of empowerment. At first sight, the development program I followed looked "uncannily similar in many of [its] daily practices" to other development programs routinely deployed around the world to address issues ranging from health and education to financial inclusion and agricultural productivity (Gupta 2012: 273)

However I went to northern India not to learn directly about this NGO-ran educational development program but instead to study the randomized evaluation of the program. During my fieldwork months, the program was being evaluated through a randomized controlled trial (RCT) set up by a team of economists in the United States (USA) and field researchers in India who were external to the NGO. Over the previous year, thirty schools had been chosen in the district: half of the schools had been randomly allocated to receive the program of "girls' parliament" while the other half did not receive the program and served as a control group. The research team aimed to understand the causal impact of the program (e.g. did it have a positive effect on girls' self-perception?) by comparing the outcomes among both groups. My fieldwork therefore mostly consisted in following the RCT research team as it visited both control and treatment schools, and also administered questionnaires and surveys to students in order to

gather data about school attendance and life aspirations. Unless one had been informed explicitly about this research project or shown the list of schools included in the study, the architecture of the randomized controlled trial was invisible to the naked eye.

The tying knot of this dissertation is the RCT – this elusive object that is hard to see and whose boundaries are difficult to define. The RCT I followed in northern India is one in thousands of randomized evaluations of development programs conducted in over 80 countries around the world in the last two decades by research teams based in major universities or international organizations. Yet this dissertation is not only about the RCT, but more broadly about the larger phenomenon that I have come to term *evidence-based development* and into which the RCT was my first empirical entry point[1]. I conceptualize evidence-based development as a polymorphous project to reduce poverty in the global South by generating evidence about the effectiveness of development programs through RCTs and then using this evidence to shape policies and programs at scale around the world. And just like the architecture of the RCT I followed in India was superimposed on the girls' parliament educational program, evidence-based development is a project that simultaneously takes shape within and hovers above current global poverty alleviation efforts.

In *When Will We Ever learn? Closing the Evaluating Gap*, an influential group of economists and development experts decried that "after decades in which development agencies have disbursed billions of dollars for social programs, and developing country governments and nongovernmental organizations have spent hundreds of billions more, it is deeply disappointing to recognize that we know relatively little about the net impact of most of these social programs" (Savedoff, et al. 2006: 1). In the aftermath of this report, randomized evaluations (RCTs) emerged in the 2000s as a vastly popular and exported tool to assess the success and failures of the manifold poverty alleviation projects devised by NGOs,

---

[1] I return to explaining the term as both an actor's and an analytical category below.

governments, philanthropies, and international organizations around the globe[2]. Over the last decades, individual-level poverty has taken shape first as an urgent problem requiring action and imaginations beyond governmental intervention, and second as a complex issue whose exact manifestations, causes, and solutions largely divide the opinions of development practitioners.

The turn of the twenty-first century in fact witnessed the proliferation of institutions, programs, funding agencies, methods, innovative solutions, and collaborations committed to alleviating the plight of the 'world's poorest', mobilizing a vast array of non-state actors. In this perspective, development – or global poverty alleviation – comes to the fore not as a project to care for the Kenyan economy or for Cambodians, but as a project to care for the poor at a global scale. The target of poverty alleviation has emerged as a global group of people, known for instance as the "bottom billion", whose membership bypasses national and cultural affiliations (Collier 2007). While the urgency of alleviating individual-level poverty makes today solid consensus, many disagreements persist within academic and policy circles about how to understand and address poverty (Green 2006; Rist 2008; World Bank 2015). With regards to identifying the best ways to address poverty, debates and unanswered questions include, among others: what works to improve long-term educational outcomes? Are cash-transfers or microcredit loans better suited to lift people out of poverty? How can we reduce corruption and leakage in large-scale, governmental programs to provide subsidies or services to the poor? Are governmental interventions or NGOs more effective in improving key poverty issues?

---

[2] RCT is not a direct abbreviation of randomized evaluation, but rather of "randomized controlled trial" that is more directly associated with the medical world. My interlocutors within evidence-based development employed randomized evaluation and RCT interchangeably, often also using randomized experiments and field experiments. I follow their usage in this dissertation.

In this conjuncture, starting at the beginning of the 2000s, RCTs have been presented by its early advocates – academics and development practitioners in universities, NGOs, and some international organizations and governments across the planet – as an inescapable overarching solution for poverty alleviation: that which will allow understanding which anti-poverty interventions work, for what ends, and why, by following the causal effects of programs on their recipients. One can hypothesize that this interest in data-driven interventions, showcasing evidence of impact, was spurred by a multiplicity of factors including frustration with past development efforts, the adoption by the United Nations of the Millennium Development Goals and the focus on individual-level poverty as an overarching development goal, the shrinking budgets for social services, and the concurrent entrance of private funders with high demands on accountability and efficiency (Adams and Rees 2014; Kapur, et al. 1997; Kelly and McGoey 2018).

While the following chapters also tell about accounting logics, cost-effectiveness analysis, and econometric methods, this dissertation is not primarily an examination of how the ubiquity of "metrics" and quantitative indicators are transforming priorities and resource allocation around the world, a crucial endeavor undertaken by several scholars before me (Adams 2016; Best 2017a, b; Biehl and Petryna 2013; Jerven 2013; Merry 2016). This dissertation rather documents evidence-based development not only as a field dominated by "narrow specialists" or technocrats but, borrowing from the words of economist Albert Otto Hirschman in *The Rise and Decline of Development Economics*, as a practice initiated by "a group of social scientists (…) impelled by a vision of a better world" (1981: 21). What are the concrete technologies, this dissertation asks, that practitioners of evidence-based development (academic researchers, but also NGOs and international organizations, data analysts, field coordinators, funding bodies, and even government representatives) devise to enact their visions of a better world?

Importantly however, my work is not circumscribed to understanding the worldviews and narrative strategies of the "randomistas", the expression now widely used by commentators to designate the thought collective advocating for the application of RCTs to "reduce poverty by ensuring that policy is informed by scientific evidence"[3]. The appellation "randomistas", whether used by critics or apologists of RCTs in development, equally fixates and simplifies people and realities rather than illuminates them. Focusing on randomistas amounts to reducing evidence-based development to a discursive field dominated by academics in leading institutions in North America and Europe when it is in fact a heterogenous phenomenon not only made of people including, as mentioned above, project managers, research assistants, field coordinators, engineers, project recipients and government officials in and from the global South, but also constituted by both tangible and intangible things such as research designs, surveys, software, slippery roads, hard work, policy and programs, hopes, and collaboration with governments.

In each chapter of this dissertation, I aim to show how these different things enact, but also centrally constitute and condition, the imaginations of a better world that sustain evidence-based development by looking in turn at: the physical and causal architecture of the RCT to identify "better lives"; the kinds of hard and repeated work involved in maintaining an evidence-based water sanitation program; the very specific ways in which the government of Kenya is enrolled to scale up pilot projects; and how "contextualization" is practiced by my interlocutors when promoting the generalizability of randomized evaluation results. My goal, overall, was to follow imaginations as they are being made in the minds and words of people I encountered, but crucially also as they are being shaped – both reconfigured and limited – in the daily work of those who practice evidence-based development.

---

[3] https://www.povertyactionlab.org/, accessed July 12, 2019

**Situating evidence-based development**

To some readers, the term evidence-based development might evoke the more general phrase of evidence-based policy or the field of evidence-based medicine. A brief overview of how evidence-based development both resembles and differs from these other two phenomena is a good way to further explain the object of my investigation. It is important to note that the term evidence-based development is not a widely used term within global poverty alleviation efforts and as such should be taken in this dissertation as an analytical rather than as an actor's category, one that nevertheless refers to a well-established empirical reality which I describe in this section[4]. My interlocutors in fact often use the expression "evidence-based policy" to refer to their current work of collaborating with governments to find successful evidence-based programs to improve the lives of poor people. In so doing, they position their work, and the institutions carrying it out, squarely within the larger agenda of evidence-based policy. Scholars have placed the beginnings of evidence-based policy as a self-identified movement in the United Kingdom in the late 1990s, during the New Labour Government under the leadership of Prime Minister Tony Blair, that informed policy-making by way of rigorous, objective evidence, often quantitative in nature (Pawson 2006).

---

[4] Some might find "development" a strange choice of term given that my interlocutors are primarily interested in alleviating individual-level poverty in the global South, rather than concerned directly with international development as a project to modernize and generate productive national economies (Collier, et al. 2017; Mitchell 2002). The term of evidence-based global poverty alleviation might therefore seem more fitting, but its length hinders rather than enhances understanding. Furthermore, while not ubiquitous, the term "evidence-based development" has been used by my interlocutors in interviews or media communication (see for instance https://www.povertyactionlab.org/fr/news/evidence-based-development-trend-whose-time-has-come, accessed July 18, 2019). Finally, I privilege the term development because it keeps alive both 1) the heterogeneity of this evidence-based movement happening through non-governmental organizations but also in new units within bilateral organizations such as USAID (see https://www.usaid.gov/documents/1870/strengthening-evidence-based-development-five-years-better-evaluation-practice-usaid, accessed August 4, 2019), and 2) the ambiguity around what exactly are its goals, especially when RCT advocates partner tightly with governments on the design and scale-up of projects, something I discuss later in this introduction and in Chapter 3.

Of course, one might question the phrase *evidence-based policy*, which seems to imply that until the late 1990s, policy decisions were precisely *not* based on evidence, but rather on other sets of factors such as ideologies, beliefs, presuppositions, theories, "guesswork" or "good intentions"[5]. Rather than a turn to *evidence* proper, some scholars have acutely noted, evidence-based policy amounts to the inclusion of certain types of knowledge (quantitative indicators, statistics, numbers and metrics) and the exclusion of other kinds of knowledge, deemed as not rigorous and often anecdotal, in the process of managing people and things around the world (Adams 2013, 2016; Colvin 2016). It remains true nevertheless that evidence-based policy is a historically specific, albeit diffused, agenda that takes shape around the question "*What works*?" and builds on the logics of evidence-based medicine (Bemme 2018; Solesbury 2001).

Evidence-based medicine, crystallized at the end of the 1980s, is an approach to medical practice that shifts the locus of decision-making from individual physicians using selected pieces of research to a systematic production and utilization of clinical evidence, based around the randomized controlled trial of medical treatments and systematic reviews of available evidence. As anthropologist Adriana Petryna (2009, 2010) has documented in *When Experiments Travel*, the logics of evidence-based medicine today pervade the very lucrative global pharmaceutical industry and its relentless search for human subjects to conduct clinical trials around the world, increasingly in low-income settings. In evidence-based medicine circles, the name of British doctor Archie Cochrane is often cited as one of the founding figures of the practice who promoted the importance of RCTs for gaining medical knowledge in his 1972 treatise titled *Effectiveness and Efficiency*. The name Cochrane was subsequently re-used in 1993 when the Cochrane Collaboration was established in the United Kingdom as an

---

[5] See Esther Duflo's words in https://www.ted.com/talks/esther_duflo_social_experiments_to_fight_poverty (accessed August 4, 2019), who talks of taking the "guesswork" out of policy-making with randomization and (Karlan and Appel 2011)'s who advocate for "more than good intentions".

international collaborative of health practitioners dedicated to summarizing "the best evidence" in medical research.

Interestingly, in 1999, the Campbell Collaboration was established as a sister organization to the Cochrane collaboration with a mission to provide systematic reviews and evidence synthesis not about medical research this time, but about social interventions in order to promote positive social and economic change[6]. Like for the Cochrane collaboration, the eponym name chosen for the organization's title was an homage, and Campbell in this case referred Donald T. Campbell, an American social scientist who promoted "an experimental approach to social reform" in the United States and other "modern nations" starting from the late 1960s (Campbell 1969: 469). Campbell was writing his ideas which promoted an experimental view to society-making at approximately the same time that a limited number of field randomized experiments of social policies were being conducted in the United States. These include a series of experiments around a taxation scheme known as the negative tax income from 1968-1982 (see Orcutt & Orcutt 1968) and also the RAND health insurance medical experiment that focused on studying the health care costs and people's associated behaviors between 1971 and 1986[7]. These experiments are well-known today and often cited as precursors to the movement of evidence-based policy. One can cite the Coalition for Evidence-Based Policy based in the USA and the Alliance for Useful Evidence in the UK as two recent influential organizations, composed of academics, policy-makers, and funders, that seek to increase government effectiveness using rigorous evidence about "what works"[8].

If I have termed the phenomenon I study evidence-based *development* to distinguish it from evidence-based *policy*, this is for three main reasons. The first reason, certainly the most obvious and apparent, is to indicate that I study institutions and projects whose main topical

[6] https://campbellcollaboration.org/about-campbell.html, accessed July 17, 2019
[7] https://www.rand.org/health-care/projects/hie.html, accessed July 17, 2019
[8] http://coalition4evidence.org/mission-activities/, accessed July 17, 2019

interest is improving the lives of the world's poorest (a global population that cuts across national boundaries) and whose geographical focus is therefore concentrated in evaluating programs and policies in the global South. In fact, the rise of evidence-based development has been accompanied the creation of new institutes in the 2000s to promote evaluations of development programs, with the ultimate mission of reducing poverty through more effective policies. A major starting point for evidence-based development was the founding of the Abdul Jameel Poverty Action Lab (J-PAL) in 2003 at the MIT, the first research and policy center founded by three development economists specifically to support the use of randomized evaluations for poverty reduction research. Other such centers were formed within existing international organizations, such as the Development Impact Evaluation Unit (DIME) created in 2005 at the World Bank.

Overall, the institutional and programmatic sites in which evidence-based development is practiced are either organizations headed primarily by people whose professional lives have been concerned with international development and extreme poverty in the Global South (development economists, staff from NGOs, international organizations, philanthropies, etc.), or structures, including ministries, located in developing countries. In the later part of the 2000s, J-PAL opened regional offices in Europe and North America as a way to promote the use of randomized evaluations in guiding social programs also in "rich countries". While my research project is not oblivious to this evolution, its focus is restricted to an investigation of evidence-based development as it applied to the alleviation of poverty in "developing countries". This is both because my study of RCTs was routed through an interest in the past and present of international development efforts, which has its own history and trajectory of poverty reduction efforts in the developing world.

Secondly, when I talk of evidence-based development, I focus on those organizations committed to randomized evaluations over and beyond other methods for producing knowledge

and modes for influencing social change. While evidence-based policy as a movement privileges quantitative evidence, it nevertheless relies on a panoply of research designs, which are based around different assumptions about how to derive causal claims about the effectiveness of programs. Similarly, development organizations, practitioners, and researchers are today using various *impact evaluation* methodologies to evaluate the causal effects of development programs beyond the RCT. These research designs and econometric methods are summarized in *Impact Evaluation in Practice* (Gertler, et al. 2011), a World Bank publication that promotes evidence-based policy-making in international development. Nevertheless, the organizations and projects I followed during fieldwork were primarily grounded on the randomized evaluation as source of knowledge and authority[9].

As an example, early in my PhD work in September 2013, I attended a week-long training in Denmark led by J-PAL about evaluating social programs. While a small portion of the course was dedicated to presenting diverse impact evaluation methods that are available, the main focus was rather on explaining to the audience why randomization (into treatment and control groups) was preferable for properly measuring causal effects of a program. The training taught the audience how to conduct randomized evaluations of development programs and made us practice, in small teams, how to design and present a potential RCT. In my work, I chose to focus on randomized evaluations because, for better or worse, it has acquired in the last decade the aura of being the "gold standard" of impact evaluation, with implications for publication decisions and funding allocation. And it is true that for those development economists who are engaged regularly in impact evaluation, RCTs remain the preferred method

---

[9] These are mostly J-PAL (already mentioned), and its sister organizations such as Evidence Action and Innovations for Poverty Actions discussed later in the Introduction, but also a charity "ranker" GiveWell. The Center for Effective Global Action (CEGA) at Berkeley where I also conducted fieldwork had a wider focus, including research around measurement tools used in impact evaluations and transparency in research. Nevertheless, when I was there, the leadership of CEGA held RCT as the highest standard of evaluation for obtaining rigorous evidence, and the Center was partnering with J-PAL on a few initiatives, including some funded by the Gates Foundation, to ran randomized evaluations of agricultural initiatives, which I discuss in Chapters 1 and 4.

(over other quasi-experimental or non-experimental designs) to evaluate a program whenever possible. My intention is of course not to contribute to enhancing the uniqueness of RCTs, but rather to investigate a compelling topic on a whole series of debates and practices that revolve around their rise to prominence, especially regarding how RCT-focused organizations engage in policy work in developing countries.

This is where, however, it becomes crucial to separate organizational mission and individual researchers. On the one hand, J-PAL, its sister organization Innovations for Poverty Action (IPA), but also the Center for Effective Global Action (CEGA) at the University of California Berkeley, are organized around large networks of affiliated researchers in universities across the world. Most of these affiliates, in addition to having used RCTs to evaluate development programs, also conduct other kinds of research that rely on different methodologies. This is why, as I already alluded to earlier, it does not make sense to speak of "randomistas" because most people involved in randomized evaluations are not *solely* evaluating randomized interventions, with perhaps the exception of a few researchers who direct those centers[10].

On the other hand, J-PAL and IPA only promote and disseminate results of *randomized evaluations* as part of their mission. For these organizations, RCTs are a threshold of quality and a criterion for inclusion to appear on their database or even for collaborating with them as partners, whether NGOs or governments. While J-PAL core staff often gladly recognize that if a program *cannot* be randomized, it is of course much preferable to evaluate it using quasi-experimental studies rather than *not* evaluate it *at all*, the organization does not include such studies when it publishes policy briefs or conducts advocacy efforts. Given this practice, the influential group of economists heading J-PAL and related organizations have been accused of

---

[10] Indeed, "randomistas" is not an actor's category, used by my interlocutors to describe their own work, but rather a term used by commentators of evidence-based development, whether the in the press or by scholars.

distorting knowledge about poverty alleviation by suggesting that only evidence about what can be randomized holds epistemological and hence actionable value[11].

The third and perhaps most important reason for relying on the term 'evidence-based development' is to complexify its direct association with governments and national plans that a reliance on the word "policy" indexes. In fact, different institutions carrying out evidence-based development exhibit distinct relationships when working with governments, and even within the same institution, these relationships have changed over the last two decades. As geographers Webber and Prouse (2018) argue, until recently, NGOs were the preferred partner institutions for J-PAL, IPA, and Evidence Action (a related organization described in greater details in the "Field sites" section below). Until the 2010s, J-PAL and its sister organizations predominantly evaluated programs ran by NGOs in distinct countries and then explored paths to successfully scale up these programs by employing routes that could involve (or do away with) governmental means.

To recount their genesis, my interlocutors in the J-PAL nexus always referred to Michael Kremer, a professor of economics at Harvard, who in the mid-1990s partnered with an NGO in Kenya to lead one of the first randomized evaluations of development programs (in this case concerning educational initiatives including the distribution of textbooks)[12]. J-PAL members, but also staff members at the World Bank, have suggested that the approach pioneered by Kremer was instrumental in leading RCTs to take-off and in fomenting the movement for evidence-based development. Nevertheless, starting in early 1998, a team working at the International Food Policy Research Institute, an international center based in Washington D.C., had designed and implemented the randomized evaluation of PROGRESA, a program of conditional cash transfers introduced by the federal government of Mexico as a

---

[11] See the section below "Evidence-based development and its critics" for further details on this line of criticisms and its authors.
[12] See https://www.youtube.com/watch?v=YGL6hPgpmDE, accessed July 18, 2019

major rural anti-poverty program[13]. The efforts around the evaluation of PROGRESA and especially those of one of the team members, Paul Gertler, who subsequently served as Chief Economist at the World Bank, are often credited as another potential line of genesis for the movement of evidence-based development. The example of PROGRESA was set to represent a major precedent for the creation of the Development Impact Evaluation Unit at the World Bank in 2005, an initiative bringing research teams to evaluate specifically *governmental* programs this time, given that the World Bank's core work is mostly done through collaboration with national governments.

Interestingly, J-PAL and its sister organizations are increasingly eager to partner with governments and to do so in a more upstream fashion as it pertains to the pathway from the production of evidence to the implementation of projects at scale. They have created several initiatives whose goal is to include governments in the process of co-designing policies that are relevant to national priorities *before* they are evaluated through RCTs. Because my fieldwork was concentrated around organizations within the J-PAL / Evidence Action nexus, my dissertation, especially Chapter 3, reflects on the evolution of their working relationships with governments.

Today, the commitment to randomized evaluations and impact evaluation more generally has become powerful within development efforts. Within economics as an academic discipline, the results of randomized evaluation now represent an important share of publications in top economics journals; graduate students working under the supervision of the founders of CEGA and J-PAL are generally the preferred candidates for hire at the highest-ranked universities around the world. In 2017, Esther Duflo, one of the founders of J-PAL, was invited to give the keynote address at the American Economic Association Annual Meeting. During my fieldwork, there were rather serious rumors that this group of academics, promoting

---

[13] For a paper reporting on the evaluation of PROGRESA, see for instance Gertler (2004).

randomized evaluations, had been contenders of the 2015 Nobel Prize of Economics, which was in the end won by Angus Deaton. Over the last years, the Development Impact Evaluation Unit and other philanthropic funders have significantly increased their funding for randomized evaluations and other impact evaluations of their projects.

Despite this, the percentage of the development assistance budget spent on randomized evaluations, as opposed to financing projects, remains very low. On the policy side, J-PAL and its sister organizations have established numerous collaborations with governments in the global South, including to deliver training about evidence-based development or the signed Memoranda of Understanding to initiate evaluations of specific government programs. Yet, it should be noted that it is extremely difficult, if not impossible, to determine both the nature and the size of the impact of evidence-based development on anti-poverty policy and programs at scale around the world, something I return to in the Conclusion of this dissertation.

**Evidence-based development and its critics**

Along with gaining in popularity among key development funders, the movement for evidence-based development has also attracted an important number of critics whose perspectives I found important to summarize here so that the reader can refer to them as she sees fit when she engages with the arguments made in the following chapters[14]. The criticisms aimed at the application of randomized evaluations to solve poverty around the world can be summarized into four main categories.

The first line of criticism, often embodied by the economist and Nobel prize winner Angus Deaton, is concerned with what is called in econometric terms *internal validity.* Deaton puts into question the claims often advanced by the advocates of evidence-based development

---

[14] A significant number of critical figures and arguments emerging from within development economics itself were brought together in a 2014 edited volume titled *Field Experiments and Their Critics*, which has inspired the title of this section.

that the RCT research design is superior to any other for assessing causal relationships. While the RCT design allows to reduce statistical biases significantly, critics argue, one can never be certain to have eliminated all biases with randomization in such a way that you might attribute the causal impacts on your intervention (e.g. free midday meals in school) when they are in fact created by an "omitted variable" (e.g. people in the treatment groups live closer to the road which leads to the school)[15].

Furthermore, even once the RCT evidence has been collected, analysis and regressions must be performed, implying econometric choices, which impinge on the result. This was at the heart of what had been called the "worm wars", an important controversy that took place in 2015 when a research team tried to replicate the results of a seminal RCT conducted in the early 2000s and according to which there were positive links between mass deworming of parasitic infections in schools and improved educational outcomes in Western Kenya[16]. If different analytical methods changed the results, what did this say about the RCT as a knowledge producing instrument? But this controversy was ultimately a debate about the replicability of quantitative *research* more broadly, not only of randomized evaluation, and spurred the creation of methods to promote greater research transparency.

The second set of criticisms concerns the *external validity* of randomized evaluations, that is the extent to which the research findings (of an RCT) are generalizable beyond the study setting and are therefore relevant for informing policy at scale. For instance, could the identification of a positive impact of cash transfer programs in Morocco potentially form the ground for policy decisions in Mali? The ways in which those who practice evidence-based development both generalize and contextualize RCT results is the subject matter of Chapter 4,

---

[15] While RCTs are implemented with the very objective of neutralizing the risk of "omitted variable", Deaton (2009) argues that RCTs can never fully balance the infinity of factors that may affect a program's outcome.

[16] This RCT and the subsequent programs that were scaled-up as a result of it are addressed empirically in Chapters 1 and 3. For the original RCT, see Miguel and Kremer (2004) and for the articles that spurred the "worm wars" see Aiken, et al. (2015) and Davey, et al. (2015).

premised on the recognition that the way in which generalization is being drawn has significant consequences for what programs are scaled-up around the world and who receives services and how.

Martin Ravallion, former director of the research department at the World Bank, is one of the most out-spoken figures for the third kind of criticism, already alluded here above, which argues that an overreliance on randomized evaluation as the "gold standard" for conducting research about poverty is likely to distort knowledge given that "this tool is only feasible for a non-random subset of the topics that matter"[17]. In other words, some development interventions, for instance the construction of large dams, might not be amenable to being randomized and allocated only to the treatment as opposed to the control groups. If, therefore, as the critique goes, funders and organizations exclusively or primarily rely on RCT evidence as a source of information, this might end up significantly changing development priorities and the kinds of development programs put in place (or excluded) around the world. In his tribune "Should the Randomistas Rule?", Ravallion deplores that centers such as J-PAL "have put their preferred methods ahead of the questions that emerge from our knowledge gaps", when "in reality a rich set of non-experimental tools is available to the issue of concern to policy-makers" (2009: 2-3). Ravallion (2009: 5) welcomes the emphasis put on obtaining "better knowledge" about development effectiveness but insists that "the important task of investigating what works and what does not in the fight against poverty cannot be monopolized by one method."

The last category, which rallies a broader audience (from development practitioners, to anthropologists, but also the press and the lay public) is about the ethics of randomized evaluation: whether it is ethical to collect evidence about the impact of a program by granting the benefits of an intervention to a treatment group and withholding them from the control

---

[17] https://blogs.worldbank.org/impactevaluations/six-questions-martin-ravallion?CID=WBW_AL_BlogNotification_EN_EXT, accessed July 19, 2019

group. Rejoinders here include that such a criticism assumes that development programs are beneficial when they might actually hurt people, but also that when a program will be phased out or limited in scope (perhaps because of financial constraints), choosing who among those equally in need will receive the program via lottery – or randomization – might not seem any less ethical than selection by means of other approaches, such as regional or ethnic favoritism[18].

Of course, the criticism around the ethics of randomized evaluation assumes also a more radical version which questions the moral base of experimenting on people, especially in low income countries, for the sake of knowledge and government (but also potentially profits in the case of medical clinical trials). Some scholars of global health have underscored how research on medical drugs or poverty and intervention are often intertwined, suggesting that contemporary Africa is subjected to a widespread regime of "experimentality" in which "the boundaries between governmentality and scientific investigation are porous" (Geissler 2013: 14). Several anthropologists have pointed to the imperial logics which might underpin these "experimental" projects – whether in medicine or development – as vulnerable populations in postcolonial countries are constituted into subjects of trials by NGOs and international organizations in ways that are reminiscent of how colonial territories were constituted into laboratories of colonial power and administration (Geissler 2013; Geissler and Molyneux 2017; Rottenburg 2009b).

In its conceptual orientation, my doctoral research recognizes the importance of attending to how current poverty alleviation efforts might show continuities with earlier political projects and their legacies, not least of which the technologies deployed in "colonial aspirations toward empire" such as mapping, counting, censing, and transposing (Adams 2014). Yet, it also takes seriously recent calls for an anthropological engagement that conceives

---

[18] For a sociological account that take the ethics of resource allocation in randomized evaluations as a direct object of empirical investigation, see Rayzberg (2019).

development and global health as endeavors that, like anthropology, seek to navigate complex worlds through a multitude of logics, instruments, and aspirations (Ferguson 2010; McKay 2012; Venkatesan and Yarrow 2014). In so doing, the point is not "to deconstruct" in the sense of assessing the validity of evidence-based development against other poverty alleviation efforts. Rather, my intention is to highlight that distinct anti-poverty projects configure and address the issue of poverty differently, with consequences for policy-making and thus for present and future livelihoods.

**Field sites**

Studying the entire arch of evidence-based development, from evidence-production in randomized evaluations to the transformation of programs at scale around the world, required a multi-sited fieldwork across several organizations and countries. While research took place in different sites, the structure of my fieldwork was not random but shaped by an initial interest in following the trajectory of the Jameel Poverty Action Lab (J-PAL) as the first and most influential center promoting evidence-based development and randomized evaluations of anti-poverty projects. In addition, my empirical interest in J-PAL was underpinned by my willingness to study how non-state actors partake in policy-making around the world today. My fieldwork therefore brought me across many institutions and locations in the nexus composed by J-PAL and its sister organizations, rather than organizations housed in existing international organizations such as the Development Impact Evaluation Unit at the World Bank.

Though I tried several times to be granted access to J-PAL global offices in Boston to carry out longer-term fieldwork, I never obtained permission. I was however able to visit their offices once in September 2015 to carry two interviews with staff members, one with the Associate Director of Training and one with a Senior Policy Manager. Nevertheless, fieldwork

led me to several cities and countries including Denmark, India, Boston, Washington D.C., California, Nairobi, and other regions of Kenya, and to interact with people and projects who had been there almost since the inception of J-PAL. While my fieldwork brought me to government offices to meet with staff from funding agencies, and also with program recipients, my primary focus was on those organizations whose central mission is to promote and practice evidence-based development. If I spent a considerable portion of my fieldwork in Kenya, as the foci of my chapters reflect, this research was not an empirical investigation into social policy-making in Kenya. Rather, Kenya served as an illustration of how global poverty alleviation efforts are implemented in concrete programs. For the sake of clarity however, I will detail in turn my three main field sites. There, my inclusion into daily activities and events was dictated by a combination of what my interlocutors were eager to share and what they refused to grant me access to. My position as anthropologist in a field organized around economic expertise triggered a significant range of reactions from my interlocutors. There were those whose greeted me with explicit apprehension, those who interacted with me only with a hidden reticence, while many others expressed warm welcomes and a real curiosity about my research.

1) From March to June 2014, I shadowed the randomized evaluation of a girl's empowerment project conducted in 30 schools of a rural district in Rajasthan. At stake in this field site was to understand how a randomized evaluation is concretely unfolding on the ground, including design choices, navigating implementation obstacles, and the daily routine of travelling across research sites. The study was funded by J-PAL, ran by a team of two principal investigators from a US university in the Midwest and the International Food Policy Research Institute (one of the investigators was a J-PAL affiliate), and relied on a team of research assistants and field coordinators hired by J-PAL South Asia in India. All the coordinating activities and preparation were housed in the district main town in the office of

the NGO whose program was being evaluated. I followed the end of the evaluation when the end-line data collection process was happening and surveys were being administered to students across schools. In this field site, participant-observation also meant that I became a member of the team who led the qualitative component of the study and conducted focus groups in the schools with boys and girls about their daily lives and aspirations[19].

2) The Center for Effective Global Action at the University of California-Berkeley (CEGA) nurtures very tight links with J-PAL, even managing funding pools together. CEGA was founded by Edward Miguel, a close collaborator of Michael Kremer, and together they authored a randomized evaluation study in western Kenya published in 2004 as "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities", one of the most cited papers in economics and widely credited for having launched the movement for evidence-based development. I spent four months at CEGA's headquarters from September to December 2015 during which time I was granted a desk in one of the cubicles of the Center's main room in Giannini Hall on the campus of UC Berkeley.

The time I spent on campus at UC Berkeley gave me the opportunity to conduct dozens of interviews with researchers who had carried out randomized evaluations, including Paul Gertler who was involved in the 1990s evaluation of PROGRESA, the Mexican government program of conditional cash transfers mentioned above. I also learned about how the Center managed RCT portfolios, disseminated results through briefs and workshops, and constituted its own mission through logical models. In between staff meetings, interviews, departmental seminars, and conferences, I attended some undergraduate lectures on microeconomics, chatted with the Center's staff members, and transcribed field notes.

---

[19] I obtained all necessary approvals to inhabit both research roles, as an anthropologist of randomized evaluations and a qualitative researcher on the randomized evaluations. In addition, my interlocutors knew about my research and my primary goal for carrying out fieldwork there.

3) I spent eight months between January and August 2016 following the work of Evidence Action in Kenya. Evidence Action is a non-governmental organization whose mission is to transform anti-poverty interventions identified as successful through randomized evaluations into programs "at scale" through various pathways, with and without governmental means. Institutionally, Evidence Action emerges out of the non-profit initiative "Deworm the World", which used to be housed under IPA and mentioned here above (a sister organization to J-PAL with offices around the world to implement RCTs). "Deworm the World" was created to carry out advocacy efforts and technical assistance to governments to implement national programs of school-based deworming to reduce parasitic worm infections in schools and improve schooling outcomes. As an idea Deworm the World grew out of the paper published by Edward Miguel and Michael Kremer reporting on the randomized evaluation of a deworming study in Western Kenya, as mentioned above.

In 2013, Evidence Action was created with the mandate to continue the work of "Deworm the World" but also to expand its portfolio and to support the implementation of other development programs at scale around the world. Evidence Action has offices and activities in several countries including Malawi, Uganda, Vietnam, Ethiopia, and Bangladesh, but the biggest offices remain its headquarters in Washington D.C. and its Kenyan branch, with two main regional offices in Nairobi and the western city of Kisumu. In addition to learning through interviews about more recent programs tested at scale by the organizations, I shadowed its two oldest initiatives in Kenya, which are both said to reach millions of people around the country.

The first program is the National School-Based Deworming Programme (NSBDP), a Kenyan government program that benefits from substantial technical assistance from the Deworm the World team from Evidence Action. Studying the NSBDP brought me to travel across to country to attend internal weekly logistical meetings, government county planning

meetings, training sessions for government officials, and also to observe Deworming Day, the public release of the results of the deworming program, and the launch of the National Plan for the Control of Neglected Tropical Diseases.

The second program I shadowed was the Dispensers for Safe Water (DSW) which operates in western Kenya and also through similar models in Malawi and Uganda. It is a program that works to install chlorine dispensers across water sources in rural areas so that households can disinfect their water upon collection. In addition to interviewing the Directors of the program, I followed the operations of a sub-regional office in Kakamega through team meetings, trainings and meetings with government officials, trainings of villages at water sources, and dispenser monitoring and repairs. Studying both the NSBDP and the DSW allowed me to learn about two very different models of scale: one which enrolls the national infrastructure of education and health and one which precisely seeks to provide safe water where such public infrastructure and pipes are missing.

In addition to these three well-demarcated field sites, my research led me to visit people and projects in the nexus of J-PAL at other diverse locations. Overall, I conducted about 45 open-ended interviews, including with the directors and staff members of the Kenyan office of IPA, with Directors of Vision 2030 (the overseeing organ for Kenya's national development program), a program operations specialist at Development Innovation Ventures (a USAID program to fund, test, and scale new solutions to "critical" challenges affecting millions around the world), and with a research analyst and an outreach associate at GiveWell in San Francisco (an organization that ranks NGOs and philanthropies to advise donors on best-giving practices and that has consistently ranked Deworm the World among its top-ranking charities).

In addition, my first step to conducting participant-observation of evidence-based development was to attend, in September 2013, a week-long training on impact evaluations and randomized evaluations led by J-PAL, taught by professors of economics from European

universities and affiliates of J-PAL and attended by philanthropists, academics, and government officials from around the world. I also attended JPAL's 10th anniversary celebration on the campus of MIT in December 2013, visited the Elimu Impact Evaluation Centre housed in rural central Kenya and directed by Dr. Matthieu Chemin (professor at McGill University and operated daily by a mix of McGill students and Kenyan staff), and interviewed a team member of the Busara Center for Behavioral Economics in Nairobi, which conducts lab experiments to identify evidence-based solutions for responding to pressing social and economic issues.

*A note on pseudonyms:* I did not change the names of the organizations where I conducted fieldwork because their programs are well-known even in the media, but I changed the names of individuals, unless they are considered public figures of evidence-based development.

**Conceptual framing and guiding questions**

This dissertation aims to document how poverty is constituted as an object of knowledge and a field of intervention through the technologies of evidence-based development, which certainly includes the RCT, but also outgrow research practices in the field of policy-making, for instance when "scaling-up" programs in partnerships with governments. Underpinning this empirical interest is a broader preoccupation with following how imaginations of a better world – at a global scale – are concretely enacted. This focus takes inspiration from anthropological work that has called for studying imagination not only as a set of ideals and discourses, but as constituted through repeated practices to bring into existence "the good" as that which is beyond "the already there" (Robbins 2013: 453; see also Redfield 2012; Sneath, et al. 2009). Of course, the intention is not to claim that evidence-based development does succeed in bringing about a better world. Whether it fails or not is a question

that certainly cannot be answered by a simple yes or no. While fieldwork conjured an abundant set of thematic questions, it also triggered an important ongoing reflection regarding the epistemological and moral bearings of evidence-based development. I hope this reflection is manifested in the ways I expose arguments in the following chapters, by trying to practice a scholarship inspired by scholars such as de Laet and Mol (2000) and Peter Redfield (2016), that does not define itself primarily by proof and certainty, but through a nuanced analysis eager to adapt its conclusions in light of current events and further learning.

In my first days at the Center for Effective Global on the campus at the University of California at Berkeley, one of the Center's directors, a professor of economics, shared with me that most of his encounters with anthropologists or qualitative researchers had been colored by the impression that the latter found economic practice to be "evil" and one that "hid behind numbers". He was expressing reservations and worries about my fieldwork at the Center, about what I would write in my dissertation and what impact my doctoral work would have on the Center's work and its people. "If you are part of this discipline that expects you to say that development and economics is evil", he kept probing, then "how can you do otherwise if you want to be published and have an academic career?" This dissertation therefore has been animated by the challenge to take seriously and generously the work of my interlocutors, without falling solely on either side of dismissal or apology, and trying to keep alive hopes and aspirations, differences and distinctions, as much as incoherencies and contradictions. In this endeavor, I was helped by anthropologist Lionel Robbin's words: "It is not that imaginings of the good cannot sometimes be set aside in practice or put to use in ideological projects that support the continued existence of structures of violence and suffering, but if we assume that ideals always and only get either ignored or deployed in nefarious ways, then the anthropology of the good can never get off the ground" (2013: 457).

In his book *Poverty and the Quest for Life* centered around an ethnography of rural Rajasthan, Bhrigupati Singh (2015: 9) seeks "to define the quality of life somewhat differently from economists" and beyond minimal indicators such as hunger, caloric intake, educational level or income. By contrast, I want to define life with economists, not only in their terms and not exclusively through their work since evidence-based development is a movement that calls upon far more than economic knowledge. Yet, I want to define life *along* with them, by following their randomized evaluations, the ways in which they disseminate results, and the programs and policy they put in place, along with the help of multiple other actors involved in evidence-based development. What it means today to pose global poverty as an economic problem is a complicated question and one that guides my research endeavor.

To further present the conceptual framing of this dissertation and the literature it is indebted to, my research interest in evidence-based development can be broken down into through three related – and equally significant –research concerns: the economic, poverty, and the global.

1) The economic: It is essential to recognize that economic tools are not the only possible entry point for studying global poverty alleviation. Many organizations and concerned scholars propose other modes of attending to both poverty and global health issues that might require rethinking political systems and inequality (Roy, et al. 2016). If I have chosen to focus on microeconomic tools (the RCT and its adjacent economic methods and theories) to empirically investigate poverty alleviation efforts in the global South, it is not because I attribute to them superior epistemological or ethical worth. However, the critical role played by economic thinking and practice in fostering conceptions of poverty (and health) has been highlighted by several scholars and deserves rigorous examination (Adams 2016; Bashford 2014b; Mitchell 2005; Rees 2014b; Rodgers 2011a). In the history of international

development at least, poverty has often been posed as an economic problem in various guises requiring economic solutions.

The World Bank, which was to help the reconstruction of national economies after World War II, turned to developing poor economies after this period. In the 1970s it established individual-level poverty as a development goal, and since the 1950s the World Bank has been dominated by economists (Kapur, et al. 1997; Staples 2006). And while sometimes multidisciplinary, the teams funded by major donors (USAID, Gates Foundation, etc.) to implement and evaluate anti-poverty projects today are overwhelming headed by economists. Poverty is thus made into a problem that can be solved using the tools of micro and macroeconomics. In other words, the world and humans are to be best understood and intervened upon using economic analysis: field experiments, economic theories and models, large quantitative datasets, statistical regressions, and their policy implications. Through microeconomic investigation, randomized evaluations seek to determine which solutions are most effective in producing desired anti-poverty outcomes, which vary depending on how one defines poverty. Here, economics is applied to the work of finding the best ways to improve the living conditions of the poorest.

Interestingly, the global concern with individual-level poverty took shape in a period when the general trend shifted toward microeconomics with its globalized idea of the market and its myriad rational actors. As discussed in more detailed below in the "poverty" sub-section, early international development strategy was concerned with national economic growth in developing countries. In the 1970s however, through a series of political events and intellectual changes, the grounding assumptions of Keynesian economics – imperfect markets and necessary state intervention – were cast into serious doubt. Macroeconomics was rethought on microeconomic foundations, with maximizing individuals and self-regulating markets at its center (Rodgers 2011b: 63-8). It can only be hypothesized that the tools of microeconomics to

study and change human actions (including its heavy use of computer modeling) now both available and dominant in academic and policy circles, were well suited to tackle the issue of individual-level poverty and probably helped in establishing it as central global issue well into the twenty-first century.

Concurrently, evaluating development projects was made 'more feasible' in recent years because the increasing focus on poverty as a condition of people through the tool of microeconomics has had the consequence of reducing the "scale" of projects (e.g. regional instead of national) and of targeting individual outcomes. For instance, it is easier to follow the effects of a cash transfer program on key dimensions of livelihoods than of a long-term project to construct roads across the country, which is characteristic of earlier nation-wide, development projects. Second, evaluating projects on individual-level outcomes (such as health or education) implies a lighter process of data collection which, unlike tracking measures of economic growth, does not require full-fledged national statistics offices.

In following the "work of economics" and its techniques to manage people and things (Mitchell 2005), the conceptual approach of this dissertation bears resemblance to and takes cues from the scholarship that has been concerned with "the performativity of economics" to emphasize that economics does not only represent the world, but that it is significantly involved in making it, guiding political projects and enacting the future (Çalışkan and Callon 2010; Callon 1998, 2007; Escobar 1994; Fourcade 2009; MacKenzie, et al. 2007; Mitchell 2008). Like Hannah Appel's call in her article "Toward an ethnography of the national economy" (2017), this dissertation pays attention to the *incompleteness* of performativity by documenting not only the power of evidence-based development, but also the resistances and limitations it encounters, for instance when my interlocutors are challenged on the generalizability of their RCT findings or when they implement a program through governmental means and must account for the fact that governments hold different plans for the future. When following the

scale-up of programs found successful through RCTs, one of my aims is to make apparent the multiple (political) projects that "may work alongside each other", enrolling similar technologies such as national infrastructures, but that may also work "against each other, making visible independent, incongruent imaginings" (Sneath, et al. 2009: 26).

My approach however departs from this literature in one important way. In fact, scholars studying "the performativity of economics" have been predominantly focused on studying how economics constitutes a particular object – namely the *national* economy – as an object of government. Here, the "economization of life" engendered by those economic practices "value and govern life in terms of their ability to foster the macroeconomy of the nation-state, such as life's ability to contribute to the gross domestic product (GDP) of the nation-state" (Murphy 2017: 6). Instead, my interlocutors are simultaneously concerned with much smaller-scale and large-scale objects. They intervene at the level of the individual, not in the name of the national economy, but in the name of alleviating the burden of poverty worldwide – in other words, of reducing the number of people who live under extreme conditions of poverty around the world. Boosting economic growth, or national plans, might be an indirect product of their work, especially insofar as evidence-based development increasingly enrolls governments and their priorities, but it definitely is not the central aim of their work. Instead, their target is a global group of people (one might say a global population), whose living conditions measured in income, health standard, education, agricultural productivity, empowerment at the individual level must be improved. That tension and the gap between the governing of national populations and of global ones suffuses the chapters of this dissertation, and I return to them when I present my research concern on "the global" below.

2) Poverty: A key conceptual grounding of this research project is to inquire into poverty itself. Rather than assuming a fixed concept "poverty" and looking at how efforts to reduce it have changed over time, this study focuses on the co-constitution of both: how ways

of understanding poverty are transformed in tandem with ways to solve it. First, I draw here on a wide number of scholars, from history to economics and international relations, that have detailed how in the history of international development and beyond, the type of poverty that receives attention at different times in history and in different places is not the same (Beaudoin 2007; Brown 2002; Geremek 1994; Gutton 1971; Himmelfarb 1984a, b; O'Connor 2009; Ravallion 2013). Despite the diverse range of interests, a common theme runs through many of these works: the crucial shift(s) that occurred in the late nineteenth century around the economic roles attributed to the poor and the root causes of poverty in the industrial nations of Europe and North America.

Whereas until the late nineteenth century, poverty was understood as a moral problem caused by the bad behaviors and the ill nature of the poor, and at the same time a source of cheap labor (Donzelot 1984; Ferguson 2015; Ravallion 2013), this changed gradually at the end of the nineteenth century when poverty started to be explained through social causes. Several disparate, yet related, factors are said to have contributed to such a change in perceptions: the rise of socialism, debates over schooling and using education as a tool to reduce unemployment, and the rise of the social sciences. In particular, the elaboration and application of the social survey in rural and urban regions of Europe permitted to uncover that poverty was the outcome of external economic and social causes, rather than of inferior personalities. First, if poverty was not the result of human nature and failings, it was not an inevitable condition. Second, the new statistical dimension to the condition of poverty nourished the belief that it could be solved through scientific and systematic means (large programs of insurance, state-led health support, a dynamic economy, etc.). In the industrial nations, at the start of the twentieth century, social progress started to be measured against the incidence and prevalence of severe poverty.

While these historical shifts indeed help explain the beginnings of a growing awareness and concern for the category of the poor in the richer nations of the world, it would still be a long way before economists and development practitioners became concerned with alleviating the material and other kinds of hardship of the poor within developing countries. There is indeed a very distinct trajectory of poverty reduction efforts in the developing world which must be attended to. My research about evidence-based development is therefore heavily indebted to scholars who have studied the history of international development (Best 2013; Escobar 1994; Ferguson 1990; Finnemore 1997; Kapur, et al. 1997; Li 2007; Mitchell 2002; Seers 1969; Toye 1987). The genealogy traced by these scholars emphasizes how international development, born in the aftermath of World War II as colonial countries were achieving independence, was initially a project to modernize national economies in "underdeveloped" countries, wherein "poverty was a condition of countries, not of individual human beings" as the goal was the economic "gap between rich and poor countries" (Arndt 1981, 1987).

An important reorientation happened in the early 1970s in the aftermath of deep skepticism about the economic growth paradigm triggered by the widespread recognition that the benefits of economic growth had not 'trickled-down' to the poor, as evidenced by a wave of empirical studies on income distribution in Latin America or poverty in South East Asia (Fishlow 1972; Pearson 1970). It is within this context of contestations that, in 1973, the World Bank, under the presidency of Robert McNamara, proposed the reduction of absolute poverty as an additional development goal alongside economic growth. Absolute poverty was defined as the condition of those forty percent of the peoples in the developing countries earning under 30 cents per day and who live in "conditions of malnutrition, illiteracy, and squalor… that no set of statistics can begin to describe" (McNamara 1973b: 8). Poverty was thus presented as a situation that affect people equally across multiple national boundaries. In that sense, the figure of the absolute poor generated a picture of poverty as international category, with comparable

dimensions in many different countries. Most importantly, Martha Finnemore insists, McNamara's effort to make poverty alleviation a part of development "internationalized responsibility for the world's poor" (Finnemore 1997: 220).

The enthusiasm for the plight of the (absolute) poor was only short-lived and marginal in the 1970s, overshadowed by the more consensual target of economic growth as the 1980s turned toward structural adjustment programs and the bulk of macroeconomic policies promoted to solve economic crises and deficits (Toye 1987). A sustained focus for individual-level poverty only reemerged in the late 1990s in the wake of the realization that the decade of structural adjustment programs had not translated into major poverty reduction in developing countries. Several international agencies recognized the shortcomings of adjustment policies, stating that "[their] burden of adjustment policies often fall disproportionately on the poor" (Cornia, et al. 1987: 7). It is within this international context marked by stark contrasts that the issue of individual poverty was granted a new role within the development project. Moving beyond its status as an intermediate step toward reaching economic growth, or at best a complementary aim, individual-level, extreme poverty alleviation became in the 1990s the main developmental goal.

In parallel, once it had become an overarching development goal, poverty was reconfigured in important ways (Kapur, et al. 1997: 375-6). In the 1990s, the metric of absolute poverty, developed by the World Bank, came under important criticisms from many development circles which argued that improvement in human livelihoods cannot be equated to reducing material scarcity. In this view, income poverty is only one dimension among many others that make up poverty, such as lack of health, political representation, and gender inequality (e.g. Sen 1999). In other words, a closer attention to individual-level poverty complicated its object. In the 2000s indeed, the fight against poverty lost its character as a straightforward endeavor of tackling the problems of the extreme poor (those living on less

41

than 1$ per day) to being about "ending poverty in its all its forms everywhere" (post-2015 Sustainable Development Agenda). Like international development earlier, worldwide individual-level poverty became a programmatic and imaginative category, asking what it means to lift someone out of poverty and structuring collective imaginations of a better world.

The last decades have witnessed the proliferation of small-scale interventions, targeting the "world's poorest" in manifold places and poverty in all its forms (Van de Walle, et al. 1995). Instances of microcredit structures, insurance and cash transfers programs, or humanitarian design in the form of water filtration bottles have been put in place around the world and have been extensively studied by anthropologists across South America, Asia, and Africa (Collier 2013; Elyachar 2005; Ferguson 2015; Redfield 2016; Roy 2010). Poverty is thus addressed on a global scale through a multiplicity of exportable (or at least exported) solutions to improve individual-level outcomes (income, health, education, etc.), that are now being evaluated using randomized evaluations around the world.

In fact, theories about poverty regularly undergo revisions, such that poverty does not always impact the same groups of people (do we focus on urban or rural poverty?), is not always only about material misery (but also mental conditions), has changing figures of the poor at its center (delinquent or victim, rational or irrational actors), and is hence not similarly experienced. For this research it matters less whether evidence-based development manages to target the "true" manifestations and causes of poverty; rather, what matters is that "imaginations" about poverty and its concurrent solutions (made of ideals but also concrete technologies) have real consequences, bringing forward concrete programs and policies to address it.

3) The global: While situated within the history of international development, worldwide poverty alleviation has outgrown it in various ways, as previously alluded to. Far from being a project that falls only to governments and their initiatives, a vast array of non-

state actors has joined the responsibility to care for the world's poor (Bornstein and Redfield 2011; Fassin 2012; Feldman and Ticktin 2010; Ferguson 2010; McKay 2012; Redfield 2013; Rees 2014a; Rigillo 2014; Ticktin 2011). Large and small NGOs, private philanthropies and universities, think-tanks, the media, multinational corporations, all "seek to actively involve ordinary men and women in the struggle against global poverty" (Roy 2010: 10). Scholar Ananya Roy documents:

> Global philanthropic foundations, global justice campaigns, and global non-governmental organizations are leading the fight against poverty, the largest of them commanding resources, power, and influence that far exceed the scope of most nation-states (ibid).

There is in other words the existence of both a "global conscience" and a global mobilization for poverty alleviation that escapes the international, in the sense that its forces and scopes far exceed the purview of national governments and their resources (Roy 2010). What has changed in calling it global has to do with "for whom" poverty is today a problem and a responsibility. In a recent article in fact, anthropologist James Ferguson asks, "what happens when the key implementers of poverty policy are not national states at all, but transnational NGOs, or private transnational foundations?" (2010: 169). It is important to mention here that several anthropologists have documented similar frameworks of responsibility and action in talking about the shift from international to global health (Adams 2016; Brown, et al. 2006; Rees 2014a). My understanding of the global in this study is thus heavily indebted to these ethnographic and conceptual investigations into global health and global poverty[20].

---

[20] It would be very difficult to draw a neat demarcation between global health projects and anti-poverty interventions. As I previously alluded to, different poverty experts work with distinct definitions of poverty (alleviation), some of which include improving health outcomes. As a consequence, many projects targeted at the world's poorest are health-based projects (immunization, distribution of bed nets, healthcare access, etc.). A vast number of randomized experiments evaluate the effectiveness of those projects on outcomes ranging from health to future earned incomes. Investigating global poverty hence necessarily means treading upon field sites that also belong to the question of global health and entering in dialogue with anthropologists of global health.

My interest in the global is therefore routed through this global responsibility for poverty, which enrolls non-state actors, such as the organizations leading evidence-based development among whom I conducted fieldwork. Today, anti-poverty projects emerge around collaborations that span multiple sectors, professions, and countries, and that cannot be easily categorized as public or private, academic or operational, field-based or digital, nor bound to one specific geographical region, let alone a government, as eloquently summarized by Vincanne Adams and Tobias Rees:

The new alignments or global health today frequently make it difficult to clearly state where the NGO ends and the state begins, or where public institutions end and pharmaceutical industries begin. As a result, we have moved beyond the normal constitutive terms of our political vocabulary: here society, there the market; here the social good, there capitalism; here the private, there the public (2014: 1).

It is precisely the hybrid nature of development programs and the ambiguity of the goals at stake in distinct interventions (biopolitical aims? national plans? reaching millions?) that this dissertation has been concerned with. At the same time that my approach is grounded in the significance of non-state actors for development efforts, it is also intent on drawing a specific portrayal of the global, one that neither dispenses nor hovers above the nation-state, but that mobilizes the apparatus of the state in important ways and often even depends on it.

In this light, evidence-based development seems to straddle what are taken as two common axes for the management of people and things: on the one hand, the nation-state form which supposedly should care exhaustively but only for a restricted citizen population; and on the other hand, global actors that address a global population – what one may call humanity (Rees 2014a) – but with vertical, circumscribed interventions, which often harbor the language of investment and calculation. How might evidence-based development be drawing ways of managing worldwide poverty that come "after the social" and the logics of provision based on social citizenship, and yet that are perhaps not so radically "decoupled from the nation-state" (Ferguson 2010: 168; but also Ferguson 2013; 2015)?

A second axis of my concern for the global is about the scope of the global as an object of intervention whereby the projects I study seek to reach millions of people, often targeting people across national boundaries[21]. Here, the aim is to examine the political and epistemological implications of a project such as evidence-based development that aspires to find, apply, and design solutions to poverty with potentially global relevance. This means that, like in Chapter 4, I also detail how my interlocutors devise frameworks to qualify the generalizability of RCT results, as they face the challenge of how a world divided into nation-states and "contexts" both limits and transforms findings and programs. At stake here is an interrogation of what it means to imagine both the general and the singular, the universal and the cultural. In different ways therefore, this dissertation is an inquiry into the production of anthropological knowledge as much as of economic practice. This orientation is not particularly new if one remembers James Ferguson's injunction in "Anthropology and its evil twin" that "to be critical of the concept of development requires, at the same time, a critical reevaluation of the constitution of the discipline of anthropology itself" (1997: 170).

**Chapter summaries**

Chapter 1 "The Better Life: On the Quality of Life and that of Evidence" is an actor-oriented account of how evidence-based development departs from earlier approaches for addressing poverty in the global South. Ethnographically, the impetus for this chapter was the willingness to understand what my interlocutors meant when they repeatedly described their work on poverty alleviation as "sector agnostic" and guided instead by a "filter of evidence".

---

[21] I prefer to talk of "millions of people" rather than rely on the term "global population" in this dissertation when talking about visions and aims that are not directly about citizens and national populations. Population as a concept is not necessarily restricted to *national* populations given that the authority establishing equivalences (in the forms of statistics, census, enumeration and registration), and thus engineering populations made from comparable subjects, can be divorced from the state (See Bashford 2014a; Curtis 2002; Murphy 2017). Nevertheless, population is a concept whose emergence has been strongly associated with that of the nation-state, as well as with questions of sovereignty and biopolitics, most notably in the scholarship of Michel Foucault.

The chapter is therefore divided into two main parts. First, I detail what a "sector specific" approach to development is in the eyes of my interlocutors as a segue into a broader reflection of how poverty has been intervened upon in the past fifty years of development efforts. Here, I pay attention to portray evidence-based development through its differences but also to show how it evolves within already existing theories of poverty.

Second, I suggest that the specificity and appeal of randomized evaluations used to guide policy-making relies on making visible "the counterfactual" as the possibility of a different life. In this way, I argue that the RCT must be understood not only as a tool to figure out "what works" and to enroll in cost-effectiveness analysis of development solutions. Instead, I present the RCT as a research design that in its comparison between treatment and control groups identifies and sketches how to enact what I conceptualize as *the better life*: improvements in the quality of life which, for my interlocutors, bear the mark of causal certitude and hence a strong actionable force. I end the chapter with a brief examination of how evidence-based development, as a project to shape funding decisions, is practiced as a combination of RCT evidence and other, yet not "sector specific", factors.

Chapters 2 and 3 both attend to how interventions initially identified as effective through randomized evaluations are transformed into larger programs in Kenya. Both chapters follow how the NGO Evidence Action seeks to reach millions of people across national boundaries but focus on very different pathways to scale. As briefly alluded to above, Chapter 2 documents a program to provide safe water where public networked infrastructure is missing, whereas the program studied in Chapter 3 precisely relies on pre-existing national infrastructural systems.

Chapter 2 follows a sanitation program and the daily work necessary to fulfill the aspiration of granting access to chlorine dispensers for water disinfection to millions of people. In this chapter I ask: how do we understand ambitions to scale in places that are beyond the

reach of public infrastructures for living, such as water or electricity? I propose to conceptualize the sanitation program, reliant on a light, transportable chlorine dispenser technology, as an "infrastructure of efforts" to underscore how the viability of the program is dependent on long journeys, physical work, repeated maintenance, and planning. I suggest that a focus on "infrastructure of efforts" also calls for studying aspirations to scale as a *process of reach* that must account for changing topographies, unexpected issues, and different sensibilities, rather than engaging with scaling up as a project of replication and mere standardization. By offering *reach* as an additional analytic category to *scale* or *scalability*, I do not mean to flatten hierarchies and ignore power or inequality. I wish however to bring attention to how expansion and repetition might be generative, and sometimes life-enabling, when they sustain welfare provision in (off-grid) infrastructures that provide basic services.

Chapter 3 focuses on the Kenyan National-School-Based Deworming Programme (NSDBP) to explore the roles granted to the state in the project of evidence-based development. In the first part it describes, somewhat playfully, how the government and its operational capacity are enrolled as instruments of scale in visions of poverty alleviation whose target exceeds a national population. Second, I develop the metaphor of *enmeshment* to further analyze the collaborations around the NSBDP, specifically to underscore that the boundaries between the aims of the Kenyan state on the one hand, and that of the NGO Evidence Action on the other, cannot be so easily established. This chapter offers an instantiation in which global plans crucially depend on state resources and where non-state actors must frame their work through state priorities if they want their programs to be implemented. At stake in this chapter is to offer a complex alternative to social science writings that emphasize how global non-governmental projects erode, dispense, or radically marginalize the sovereignty of political states.

Chapter 4 prolongs the interest in scaling and impacting millions of lives around the world by looking at how my interlocutors debate the "external validity" of their RCT findings, in other words how they generalize beyond the initial study setting. In fact, my interlocutors acknowledged that in their policy work – for instance in their efforts to disseminate results or promote the adoption of a program – governments or other stakeholders always raised up the issue of the generalizability of findings beyond one region or one country. This chapter therefore looks at the frameworks used by my interlocutors to generalize findings and expand programs considered effective. I explore how my interlocutors engage in "contextualizing practices" as they qualify the generalizability of their results and adapt programs to new (political) contexts, while justifying their hopes that results might have application across several contexts. By weaving together different ethnographic vignettes, I look at multiple and potentially equally valid ways in which similarity and difference can be produced to represent and act on the world. In analyzing the contextualizing practices of my interlocutors, I simultaneously interrogate the concept of context as it has been deployed in anthropological scholarship and question the purview of invoking context as a major critical move to repair accounts that are seen to universalize wrongly. Without advocating inattentive globality or circumscribed singularity, the intention is to make visible difference (cultural or national) as a product that must be reckoned with rather than assumed as a given.

In the Conclusion, I return to the contrast between the enthusiasm of my interlocutors around how the RCT measures causal impact with certitude and the more difficult task, pointed out to me also by my interlocutors, of assessing the policy impact of evidence-based development around the world. I conclude this dissertation with a reflection on the power of evidence – both in its presence and its absence – within evidence-based development, but also beyond.

At the end of an interview during fieldwork, an economist encouraged me to write in a way that economists could understand my arguments, insisting that there is no complicated idea that cannot be written in a simple form. That day, I was tempted to tell him that perhaps econometric equations and mathematical signs in economics articles could then be transformed into explanations using prose. Yet, his words stuck with me and I hope to have written in a way that could be understood by all those among whom I have conducted fieldwork. May they learn about anthropological curiosity in this way just like I learned about their imaginations of a better world.

# Chapter 1 - The better life: On the quality of life and that of evidence

**A filter of evidence, or sector agnosticism**

The fourth Kenya National Deworming Day took place on February 25, 2016. On this day, teachers in schools across Kenya distributed deworming tablets to millions of children aged 2-14 and recorded the quantity of drugs used on paper forms. They did so as part of the National School-Based Deworming Programme that falls under the Kenyan Ministries of Health and Education, funded by external donors, and supported in its implementation by the NGO Evidence Action, whose work I had come to Kenya to learn about. The program aims to treat school-aged children in Kenya infected with parasitic worms to enhance health, cognitive, and educational outcomes. The main parasitic worms targeted by the program are soil-transmitted helminths and schistosomiasis which are organisms that live and feed in the human stomach and transmitted by being in contact with infected waters or soils, or by drinking or eating infected water and food.

I observed the day's proceedings at the Kalani public school located in Kisumu, the main city of western Kenya, along with Lucas, a field coordinator, Floriane, a monitor, and Lucy, the data collection manager, all three from the Kisumu office of Evidence Action. We follow Floriane as she monitors the distribution of deworming pills in classrooms both downstairs and upstairs. Students are sent to wash their hands; they form lines in front of the water taps in the courtyard, and some come back with water on their faces. One designated teacher goes around the classrooms, handing the deworming tablet with a spoon to every student in turn. Teachers take note of those children whose parents have asked, through a letter they received in the previous days from the school, not to deworm their children, either because they do not want them to be, or because their child has already been dewormed. We leave as

Floriane is conducting interviews with the kindergarten teachers to know their perspective on the deworming program.

A few hours later, I fly back to Nairobi with the Kenya Deworm the World team (a group of about ten managers, coordinators, and project assistants working for Evidence Action) and the "Ethiopian Delegation", who all went together to Nandi, east of Kisumu, to observe Deworming Day. That week in February, about ten government officials from Ethiopia had traveled to Kenya to learn about Evidence Action's model of technical assistance to national deworming programs. In January 2015, Ethiopia had announced the launch of its Deworming Program Targeting Children, becoming the third country – after Kenya and India – to provide deworming treatment to school-aged children as part of national programs to improve health and educational outcomes. Evidence Action supports the implementation of the national deworming program in all three locations. The organization is also its key instigator, drawing a list of countries where deworming prevalence is high and the infrastructure exists to support a school-based deworming program, looking for donors and partners on the ground, and then advocating for its significance at various levels of government.

In the weeks that preceded the Deworming Day, the feverishness that naturally accompanies the culmination of many months of hard work was palpable in the Nairobi office of Evidence Action. The Deworm the World weekly meeting, two Mondays earlier, lasted over two hours as team members went around the table providing updates and plans about following-up on the transportation of drugs and deliveries to pharmacies, the training of district officials and teachers, program budgets and monitoring activities, as well as the presentation to the Ethiopian delegation. Amid that sustained agitation, I was very happy to have managed to schedule a first meeting with Eshan, Director of Evidence Action for the Africa Region. On the scheduled day however, I arrive almost late to our meeting.

Overhearing me talking to his assistant, Eshan greets me into his office. We sit on the table with two armchairs close to the window which overlooks the parking lot of the compound. Spontaneously, he details how his education and work in India, and then the United States, brought him to what he does today. I ask him about his main responsibilities in his role as Director of Evidence Action for the Africa Region. He oversees the delivery of programs at large scale, he says. This includes making sure, Eshan says, that numbers are met, for instance that 5.2 million school-aged children receive deworming treatment in Kenya on Deworming Day, or that over 3 million people are granted access to chlorine dispensers to disinfect water across Kenya, Uganda, and Malawi as part of the Dispensers for Safe Water program. But numbers, coverage and reach, he insists, are only terms in an equation in which program quality and cost are also important considerations.

"Why does Evidence Action focus its work on scaling up these two programs – the Deworm the World Initiative and Dispensers for Safe Water?", I ask Eshan.

Eshan answers by prefacing that this is a question for which I should triangulate answers, gather bits and pieces of information here and there during my stay.

"In my perspective, Eshan explains, "this was because there was robust evidence. For deworming, there is an RCT [randomized controlled trial] which started in the 1990s. There was a good amount of evidence combined with investment from donors."

Eshan was referring to a randomized evaluation study conducted by Edward Miguel and Michael Kremer, two economists from leading American universities. In 2004, their paper reported the results of an intervention in Kenya where deworming pills were randomly assigned to some "treatment" schools starting in 1998[1]. The data analysis found that mass deworming had significant positive impacts on health and educational outcomes for children in treated

---

[1] These were schools that met the prevalence cut-off recommended by the World Health Organization. For more details, see the paper by Miguel and Kremer (2004) "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities".

schools, but also estimated that it had positive spillover effects on school attendance in some untreated children as well, because of reduced parasite load which lowered transmission of the parasitic worms. The causal effects of the deworming intervention were calculated by comparing outcomes between the kids in the treatment group against those in the "control" group, composed of schools that did not receive the deworming pills intervention[2]. The paper's conclusion argued for the cost-effectiveness of mass deworming relative to other programs aimed at improving school participation, such as textbook provision, training for teachers, school subsidies, or incentives for teachers based on students' test scores and dropout rates[3]. Subsequent research on deworming, based on following-up the cohort of children from the initial randomized field experiment, concluded that mass deworming treatment had long-term impacts, as those who had been treated as children had higher education and earnings even a decade later[4].

"For Dispensers for Safe Water", Eshan continues, "there is less robust evidence because the [randomized evaluation] study was conducted only over three years, not a time series like deworming, but everyone thought it was a good idea: researchers, donors, etc."[5]

So," Eshan recapitulates, "there is the evidence, and the interest from donors, government, partners."

Eshan also reminds me that these two programs, Deworm the World and Dispensers for Safe Water, in some ways pre-existed Evidence Action. Initially, they were institutionally housed under Innovations for Poverty Action, an organization that evaluates "effective solutions to global poverty problems" around the world through its many project managers,

---

[2] Within development circles, this study by Edward Miguel and Michael Kremer is widely credited for having started the movement to apply randomized evaluations to the problem of global poverty, as discussed in more details in the Introduction of this dissertation.
[3] The programs which served as comparison were also ran by the same NGO based in western Kenya as the deworming intervention and were also evaluated by randomized evaluations.
[4] See Baird, et al. (2016) "Worms at Work: Long-run Impacts of a Child Health Investment".
[5] See Ahuja, et al. (2010) "Providing Safe Water: Evidence from Randomized Evaluations".

field coordinators, surveyors, and data analysts which bring together NGOs, governments, and academics around worldwide collaborations[6]. The two programs became part of Evidence Action's portfolio in 2013, when the latter NGO was launched to take "research findings" from Innovations for Poverty Action and focus *specifically* on the *implementation* side: launching, running, and supporting such programs at scale.

"These two programs", Eshan underscores, "have influenced how Evidence Action looks at other, new programs to take up as part of its portfolio of work".

We have been talking about forty minutes when I ask Eshan whether Evidence Action has a perspective on what the key problems of poverty are. Eshan lists current programs of the organization, both those implemented at large scales, such as the deworming program and the dispensers for safe water program, and those still being tested for their "potential to scale". He underscores that these programs span a wide range of issues including seasonal migration, remedial education, sanitation to prevent disease and improve child development, and HIV prevention for girls. "So", he settles, "the cross-cutting theme is evidence".

By this, Eshan means that the common element to all the programs in Evidence Action's portfolio is that they are "backed" by evidence of intervention effectiveness, obtained by conducting randomized evaluations. Eshan also adds that my question is timely because there is an internal debate going on within the organization about whether "we should be sector-specific or sector-agnostic". I inquire about what the arguments are for or against for each of these choices. Eshan underscores that he can only give his opinion but cannot answer on behalf of Evidence Action, given that nothing has been decided yet. He explains that being sector-specific (e.g. focusing only on projects about sanitation) adds an additional filter on top of the filter of evidence already there to determine which projects can be taken up by Evidence Action, and it is not obvious whether this is desirable.

---

[6] From IPA's motto, https://www.poverty-action.org/, accessed May 15, 2019

Yet, *evidence* worked as a strategic *filter* because the organization *only* implements programs for which there is "rigorous evidence" of their positive impact, such as the study by Miguel and Kremer which reported that deworming led to a significant improvement in educational outcomes. If a program is not backed by this rigorous evidence mostly derived from randomized evaluations, it would never be considered by Evidence Action. By being sector-agnostic, as it had been so far for the first three years of its existence, Evidence Action could consider interventions in any thematic area, or what Eshan called a *sector*, whether it be health, education, agriculture, financial inclusion, and more, insofar as there was acceptable evidence of its impact. By opposition, being sector-specific would entail that Evidence Action would only do work in one sector, let us say sanitation for instance, and in that sense the sector "sanitation" would be an *additional filter* on top of that of evidence. In that scenario, Evidence Action would only run programs *about sanitation*, and specifically programs about sanitation *backed by rigorous evidence of their effectiveness* at increasing different sanitation outcomes (e.g. diarrheal diseases; condition of toilet facilities).

"The upside of being sector-specific", Eshan concludes, "is to be specialized in one sector."

Our meeting ends on this sentence, which sounds like a tautology. Perhaps Eshan was aiming to suggest that there are also advantages to be gained in being sector-specific, such as having extensive knowledge of pressing sanitation issues in Kenya and beyond. But this would also entail choosing one sector over others, a choice that Eshan's words seemed to be saying that it is not an easy one to make.

**Development failures, uncertainties, and agnosticism**

In my fieldwork about evidence-based development, Eshan was not the only one to underscore that his outlook on poverty alleviation was sector-agnostic to foreground instead

the importance of evidence as filter. The deworming study discussed with Eshan is indeed one in about a thousand of randomized field experiments (often abbreviated RCTs) that have evaluated the causal effects of both non-governmental and governmental interventions around the world to find "what works to reduce poverty". As shown in Figure 1 below, as a research design RCTs work by randomly dividing the population into two groups (one which receives the intervention and one which does not) and comparing their outcomes. For instance, another randomized evaluation might measure the rate of savings between households which are granted cash transfers and those which are not.
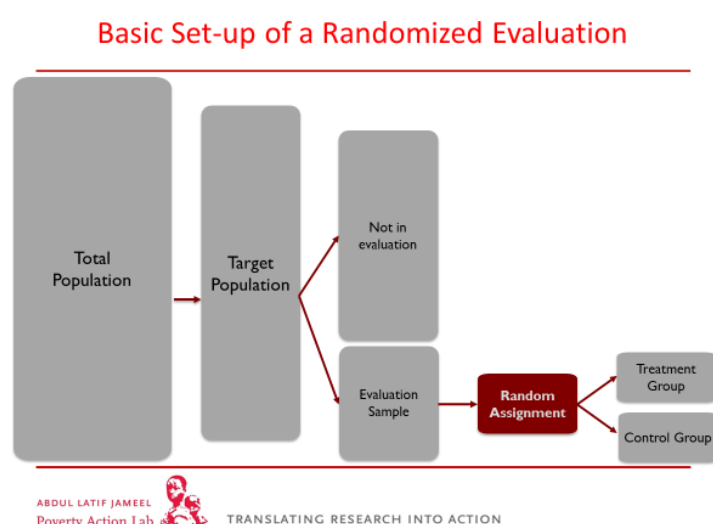


**Figure 1: Basic Set-up of a Randomized Evaluation (Source: *Why Randomize?* Slide from J-PAL executive training in Denmark, September 2014)**

Far from being just "an academic exercise" however, the RCT movement in development has consolidated in the last fifteen years into a nexus made of various international centers that take policy-making and *policy change* as central to their mission (Ogden 2016: xxi & xxvii). These include the two organizations cited above: Evidence Action, whose mandate is to scale evidence-based and cost-effective programs to reduce the burden of poverty for millions of people[7]; and Innovations for Poverty Action which has about twenty

---

[7] https://www.evidenceaction.org, accessed May 20, 2019

country offices conducting RCTs worldwide. In addition, one can cite the Jameel Poverty Action Lab at MIT, a global network of researchers and policy staff working to reduce poverty by ensuring that it is informed by RCT-produced evidence; the Center for Effective Global Action, based at the University of California-Berkeley, which funds, launches, and disseminates RCTs and their results; the Development Impact Evaluation Unit (DIME) at the World Bank; Development Innovation Ventures, the United States Agency for Internal Development's incubator program to identify and test "creative solutions to any global development challenge"[8]; or the International Initiative for Impact Evaluation, which supports the generation and effective use of high-quality evidence to inform decision-making and improve the lives of people living in poverty[9].

The policy work of these organizations exists under various guises, from the simple "policy brief" summarizing insights and recommendations to delivering training workshops to government officials about RCTs. It also includes holding sessions within ministries to share key lessons in a specific sector such as education or working with government agencies to identify "innovative" solutions to a particular issue (e.g. petty corruption) and evaluate through RCTs new government programs that involve these innovative solutions. This policy work furthermore encompasses lobbying and supporting governments to implement large-scale programs backed by evidence such as the Kenya National School-Based Deworming Programme with which I opened this chapter.

My interlocutors working in evidence-based development organizations often stated – timidly, with pride, even urgency for some – that their approach was agnostic, or alternatively ecumenical or not ideological. In other words, they asserted that their work was not defined by

---

[8] https://www.usaid.gov/div, accessed May 20, 2019
[9] See https://www.3ieimpact.org/about-us, accessed May 28, 2019. As detailed in the Introduction of the dissertation, I conducted fieldwork and interviews with Evidence Action, Innovations for Poverty Action, the Jameel Poverty Action Lab, the Center for Effective Global Action, GiveWell, and Development Innovation Ventures. For further details about my field sites, see the Introduction of this Dissertation.

a focus on agriculture, health, education, finance, gender, governance, or any combination of these development sectors. While they admitted that the purpose of their work was to reduce poverty around the world, they told me that they were agnostic because they had no preconceptions about important sectors, issues, and programs, about things that work and how to implement them, even at times describing their work in contrast to other organizations which, in their view, hold such specific ideas. I contend that my interlocutors used this image of *agnosticism* as a way to demarcate their approach from other –sometimes earlier—forms of practicing development and what they saw as their failures[10].

Scholars have pointed out that evidence-based development took shape against a backdrop of perceived "past development failures" at different levels (Best 2017b; Kapur, et al. 1997; Kelly and McGoey 2018). Historian Jacqueline Best (2017a) has documented a general sense of paralysis beginning in the late 1990s emerging from the World Bank, as both former state-led and market-dominated approaches to development were treated as failures. In the early 2010s, for instance, the value of microcredit projects was put into question as default rates skyrocketed and the media documented "suicide epidemics" among those who could not repay their loans in India (Roy 2012). Development institutions and experts tried to design new ways forward, such as evidence-based development framed as a "third way" beyond both "neoclassical market-oriented policies" and "traditional development aid and [its] large-scale state interventions (Berndt and Boeckler 2016: 22-23).

Anthropologist Kevin Donovan has recently discussed how a consensus emerged within the aid industry that views RCTs as "a compelling manner in which to reduce

---

[10] While religiously connoted, the idea of being sector-agnostic has an investment logic to it: the idea that one will follow the best solutions to poverty problems whichever sector in which one finds them. Doing some research, I found that two bloggers attribute the coining of the term to Ralph Smith, the vice president of the Annie E. Casey Foundation, a philanthropy focused on building better lives for disadvantages children (e.g. https://thesectoragnostic.wordpress.com/about/, accessed May 28, 2019). Ralph Smith used it somewhat differently from my interlocutors to mean that the foundation can succeed only to the extent that it is willing to pursue solutions wherever it finds them, in the public, private, or social sector.

uncertainty about the effectiveness of aid" (2018: 37). Donovan shows that experimentation is justified by those who apply randomized evaluations to the problem of worldwide poverty "as a moral methodology" which provides epistemological and pragmatic certainty for knowing the world and how to improve it (ibid). The methodology is moral, in this sense, because it is engaged in a project to end human suffering through causal certitude and to change the distribution of resources (Donovan 2018: 45). These writings underscore how the rise of experiment, evaluation, and calculation in international development must importantly be read through different valences of (un)certainty, ambiguities, ignorance, failures, and redefinitions.

How did my interlocutors articulate the necessity of their work within this recognized uncertainty around whether decades of foreign aid had translated into success and better livelihoods? How did their approach brought forth new pairings of certainty and uncertainty to validate their reliance on the randomized evaluation as a precondition to successful global poverty alleviation? In this chapter, I delve into multiple moments of self-representations in which my interlocutors underscore the agnosticism of their work to portray the specificity of evidence-based development as imagined by those who practice it. While my interlocutors might at times overemphasize the differences between evidence-based development and other kinds of development approaches, I nevertheless believe that there is much to be learned about the history of international development and the current reconfiguration of global poverty alleviation efforts by paying serious heed to their claims about the unique nature of their work.

My intention is not to comment on the moral worth of evidence-based development, neither to praise it nor to discredit it. I am interested in weaving together this actor-centered, field-based, "history of the present"[11] (Foucault 1977) because it can help us appreciate how

---

[11] I use the term "history of the present" to emphasize how this chapter is marked by a "critical concern to understand the present" rather than "a historical concern to understand the past", one that was also prompted by a feeling of puzzlement, and sometimes uneasiness on my part, about the practice of evidence-based development (Garland 2014: 373). In addition, my interlocutors' portrayal of their work often points to "an aleatory path of descent and emergence that suggests the contingency of the present and the openness of the future" (Garland

evidence-based development might differ from both earlier and contemporary quantification projects carefully studied by other scholars: for instance, about the emergence of statistics in Europe (Desrosières 2002), the production and use of economic development data by national offices of statistics in Africa (Jerven 2013), the circulation of global health metrics (Adams 2016), or techniques for generating global indicators on human rights (Merry 2016). While analytic attention within the social sciences has been more geared toward the dichotomy between qualitative and quantitative knowledge (Colvin 2016), often to show their incommensurability, I insist that it is important to examine the different perceived qualities of quantitative knowledge itself. In my analysis, distinct projects to quantitively measure and calculate are built on epistemologies and ontologies that can be very diverse from one another.

I therefore attend to how my interlocutors purport to be rewriting sites and modes of interventions, thereby enacting particular ways of establishing truth, value, and commensurability[12]. In so doing, I focus less on the randomized evaluation *per se*, as a regime of evidence production with corporate and financial logics brought to bear on development efforts, this undefined endeavor that polarizes consciences and imaginations. I rather take evidence-based development as its own kind of full-fledged project of global poverty alleviation, one importantly with a specific understanding of evidence at its center, to explore what kinds of implication this might have for how poverty will be managed around the world in coming years.

This chapter documents the *quality* of the evidence claimed to be produced by randomized evaluations, such that this evidence becomes itself an index of a certain improvement in the quality of life[13]. Such an endeavor sketches the comparison at the heart of

---

2014: 372), gesturing to struggles, alliances, and repurposing, while themselves drawing the limits of current institutions and practices of development.

[12] I am grateful to my colleague Julianne Yip for this formulation.

[13] As stated in the Introduction, my use of the phrase "quality of life" is indebted to Bhrigupati Singh's work *Poverty and the Quest for Life* (2015). In his ethnography of rural Rajasthan, Singh (2015: 9) seeks to define "quality of life differently from economists" beyond, he states, minimally adequate levels in indicators such as

RCT design (between treatment and control groups), which I conceptualize as the *better life*, as a useful heuristic device to make sense of the tensions in evidence-based development: seemingly sector agnostic, and yet deeply engaged in shaping policy-making and funding decisions. I argue that the identification of the *better life* serves indirectly as a sort of answer, even if unfinished and likely insufficient, to the conundrum of having to establish priorities, compare goals, and weigh outcomes, that comes unavoidably when practicing development as an NGO, funder, state body or philanthropist[14]. In paying close attention to the RCT design, I am informed by anthropologists who have examined how in their "very design" technologies of development, whether portable water filters or prepaid meters to measure electricity consumption, "are scripted with, and come to reflect, specific ethico-political projects, targets, and expectations" (von Schnitzler 2013: 672 but see also Redfield 2012).

Importantly, my aim is not to map radical ruptures or epochal breaks but instead to describe a moment of making-up and reflection (Rabinow 2008). If you recall, Eshan had told me that there was ongoing internal debate within Evidence Action about whether the organization should be sector-agnostic or sector-specific. This shows that evidence-based development cannot be studied as a stabilized phenomenon with a singular history. Here, I also take cues from historian Beatrice Cherrier (2015) who has warned against relying exclusively on "backward" histories in which economists reconstruct their evaluation practices as a bounded field with straightforward origins (e.g. Leigh 2018). In conversation with other scholars of evidence-based development I attend to how my interlocutors themselves recognize continuities with older ways of practicing development, as they encounter similar struggles and face similar doubts to those faced by organizations they wish to demarcate themselves from

---

caloric intake or hunger. Here I precisely seek to follow what "quality of life" might mean for those economists who seek to address poverty worldwide through randomized evaluations.

[14] I thank Nassima Abdelghafour for helping me better articulate this idea through the remark she made (at a workshop in Edinburgh) that evidence-based development enfolds two different kinds of evaluation: the RCT as scientific evaluation and policy-making as normative/political evaluation.

(Abdelghafour 2017; Bhatt 2011, 2017, 2018; de Souza Leão and Eyal 2019; Donovan 2018; Webber and Prouse 2018). In the story I trace therefore, tensions and incoherence in my interlocutors' outlook are not an empirical issue to be concealed, nor it is to be taken as a sign that the project of evidence-based development must be debunked. Instead, I take tensions as a conceptual entry-point because following them allows to sketch a nuanced and complex understanding of how evidence-based development hopes and aims to redress what it sees as (past) failures in the short history of international development.

Before analyzing the quality of the evidence advocated by proponents of randomized evaluation, I provide an examination of what my interlocutors meant by a sector-specific approach to development and what it reveals in terms of both how poverty has been apprehended and intervened upon and how the reduction of poverty has been measured since individual-level poverty became a central component of the development agenda in the 1970s. I then discuss how evidence-based development seemed both to outgrow and to evolve within this sector-oriented approach to global poverty alleviation, but importantly how it proposes evidence – in the form of the embodied counterfactual – as a "new", principal sector. I conclude the chapter by a short exploration of how evidence-based development, as a project to influence policy-making and funding decisions, is practiced as a pairing of RCT evidence with a constellation of other, yet not sector-specific, considerations.

**A history of sectors**

On two separate occasions and unprompted by my questionings, interlocutors distanced their work from that of the Gates Foundation to give contours to the agnosticism of their work. The Gates Foundation is a private philanthropic association started in 2000 by Bill Gates, the founder of Microsoft, and his wife Melinda Gates whose official mission statement today is to address "extreme poverty and poor health in developing countries, and the failures of

America's education system"[15]. My interlocutors saw the Gates Foundation as holding preconceptions about what it takes to improve the lives of the poor, that is mostly working in global health.

The first occasion was in October 2016 in California during a staff meeting at the Center of Effective Global Action (CEGA), the research and policy institute founded by Edward Miguel, one of the two principal investigators of the deworming RCT study. That day, Edward was giving a presentation about the "new" framing of CEGA. Edward showed us a visual graph detailing CEGA's logical model steps: 1) identify big problems, such as hunger and child mortality; 2) find innovative solutions and measure them; and 3) finally, outcomes: poverty reduction. Chiara – one of the Center's team members – interjects with a "what about economic growth?", stating that she has been asked already by partners and donors in the development field how CEGA's work contributes to economic growth. Edward says that CEGA cares about different outcomes. He brings in the example of the Gates Foundation whose framing, he says, is ultimately about health.

"[The Gates Foundation] funds different kinds of projects but what matters for them in the end is health", he adds.

On the other hand, Edward explains, outcomes cross-reinforcing each other, that's how CEGA sees it, [good] health reinforces income for instance. Edward concludes that CEGA has an ecumenical view.

Chiara follows-up: could we say CEGA's framework is about improving livelihoods? Edward approves, even though perhaps not fully, that one could phrase what CEGA aims at in terms of improving living standards or quality of life. Edward adds: "I'm fine saying that CEGA takes a Senesque view on this [from economist Amartya Sen], the notion that poverty

---

is multidimensional. Maybe we can agree on six, seven outcomes". But, to conclude he insists that "we've never had an ideological view on this", distancing the work of CEGA from other NGOs including the Gates Foundation, which might have specific ideas about which issues (or sectors) matter most when it comes to poverty and its alleviation.

The second time this occurred, I was at my desk in my room in Nairobi and on Skype with Will, at the time the new CEO of Evidence Action. From his office in Washington D.C., Will is unpacking what he meant when he just told me that Evidence Action was part of a counter-culture in aid.

"Most people in development will say *the problem is…* and they then search for solutions to that problem which they see as important. This is because", Will explains, "99% of the development field feels more comfortable being sector specific. Sector and country specific."

It is early March 2016, a few weeks after my meeting with Eshan, so this mention of being *sector-specific* grabs my attention.

"For instance," Will continues promptly, "if you suggest to Bill Gates that transportation safety might be a promising avenue, he'll say no because it is not on the list of the problems he thinks are important. Most of the development field is led by people who know the problems. Gates, for instance, says *here's the problem*, such as polio [disease], and then runs a grant competition to find the best solutions. This is a huge improvement already because at least they don't assume the solutions, but still they pick a problem. The Global Innovation Fund[16] or Evidence Action instead have no sectors and can thus explore transportation safety [for instance]."[17]

---

[16] The Global Innovation Fund is a charity investment fund to support social innovation backed by rigorous evidence, generated via randomized evaluations or other strategies to identify causal impact. See https://globalinnovation.fund/apply/steps/am-i-a-good-fit-for-the-global-innovation-fund/, accessed May 20, 2019.

[17] When I asked Jeff more details on transportation safety as a possible sector of intervention, he justified it in the following way: "a) road deaths are a bigger deal than most diseases, b) most transport safety falls to road building

Sector is a very common, if loosely defined, term in international development circles to refer to the different thematic areas that have long been heralded as being constitutive of global poverty alleviation. Large international or bilateral development agencies, such as the World Bank, the United States Agency for International Development, or the UK Department for International Development, which provide loans and grants to support development projects around the world, divide their work into sectors. The tab "what we do" on their respective websites lists sectors as varied as education, water and sanitation, gender equality and women's empowerment, financial inclusion, global health, agriculture and food security, democracy and human rights, or infrastructures. Taken together, these thematic sectors are reckoned to represent the main global poverty problems, and each must be tackled to contribute to successful poverty alleviation around the world. These sectors are also echoed in the Sustainable Development Agenda, the collection of 17 goals established by the United Nations to "ensure that all people enjoy peace and prosperity"[18] and that include: no hunger and to achieve food security, good health, quality education, gender equality, clean water and sanitation, and more. According to this outlook, poverty is a composite category, made up of deprivations in these different sectors (Webber and Prouse 2018: 183).

When Will said that 99% of the development field feels more comfortable being sector-specific citing the Gates Foundation as an example, he was referring to the fact that an important part of NGOs, philanthropies, charities, consortiums, for-profit social enterprises implementing anti-poverty programs around the world usually focus only on one sector at a time. The Gates Foundation in fact focuses on two main sectors in developing countries: health and extreme poverty because it "think[s] they are the biggest barriers that prevent people from

---

ministries, c) most donors focus on diseases not road fatalities. So something like this falls into the cracks between ministries."

[18] https://www.undp.org/content/undp/en/home/sustainable-development-goals.html, accessed May 22, 2019.

making the most of their lives" [19]. The Gates Foundation's website states that it addresses extreme poverty by developing "new techniques to help farmers in developing countries grow more food and earn more money". As such, the Gates Foundation effectively focuses primarily on health and agricultural productivity.

Similarly, in India in 2014, I followed a randomized evaluation led by researchers affiliated with the Jameel Poverty Action Lab who assessed a school-based program to empower young adolescent girls ran by the Indian NGO Educate Girls. As its name indicates, the NGO focuses on girls' education, an issue at the crossroads between two widespread development sectors: education and gender equality. Through its different programs, Educate Girls sought to increase the number of girls who go to school and enhance the quality of their education. I first realized during my preliminary fieldwork in India that *poverty* itself is rarely explicitly at stake in the many interventions assessed by randomized evaluations, because it is never a directly measured category. Almost every day I went to several middle schools where the surveys were administered, along with the research team. I had come to the randomized evaluation field site expecting to encounter 'poverty' and had not found it. By this, I do not mean the graphic poverty of a child begging in the street. But evidence-based development, after all, aims to find effective solutions to reduce "poverty".

This observation led me to realize that *poverty* was a category absent from all stages of the randomization evaluation, from its design to the implementation, never used as a descriptive or analytical tool by the principal investigators, the research team on the ground, or the NGO workers. There were discussions about gender disparities, empowerment, lack of education, and life aspirations, which stood as proxies – whether explicitly or implicitly – for poverty reduction. For instance, the many documents that were circulated among the team members

---

[19]   https://www.gatesfoundation.org/Who-We-Are/General-Information/Letter-from-Bill-and-Melinda-Gates, accessed May 21, 2019.

and the meetings I attended often tackled the questions of how to best capture the causal impacts of the empowerment program, which methodological tools to use, and what outcome categories to include (e.g. life aspirations? change in daily schedules to measure time spent on homework?).

Yet, there was no explicit mention of how empowering women translated into less poverty, whether one conceptualizes less poverty as higher income or uses a broader conception of welfare made up of a combination of income and other improved factors, such as education, health, and resilience to risk, in a more multidimensional framework (Alkire and Foster 2011; Sen 1999). The mechanisms linking these dimensions, such as gender equality, women's empowerment, or life aspirations, to a decrease in *poverty* were rarely discussed, even though you might remember that Edward had explicitly underlined that CEGA's ultimate goal was poverty reduction, and the first and overarching Sustainable Development Goal is the end of poverty around the world. I argue that this is because the theory of change leading from women's empowerment to poverty reduction was actually there in the background, most often assumed[20].

In the case of the project I followed in Rajasthan, poverty was conceived to result primarily from women's lack of control over household resources because of severe gender inequalities. There, the poor came into view as adolescent girls who needed to be empowered within a framework that sees investments in human capital, mostly education and health, as the silver bullet to fight global poverty and stimulate growth[21]. But this is only one of the many theories on poverty and development that have emerged over the last fifty years and are given

---

[20] See the work of my colleague Doerte Bemme (2019) for an ethnographic and in-depth discussion of the theory of change as a flexible and explicit epistemic device for those who address global mental health.

[21] See for instance Nike's girl effect https://news.nike.com/news/nike-foundation-launches-new-girleffectorg or https://www.theguardian.com/opportunity-international-roundtables/2017/oct/04/global-poverty-child-marriage-education-girls, both accessed May 22, 2019.

concrete form through NGO programs, government projects, philanthropies or for-profit private initiatives.

As described in the Introduction, it is important to note that the poverty *of the individual* has not always been an explicit aim of international development efforts. As several authors have shown, throughout the 1950s and 1960s the main development concern was rather the gap between rich and poor countries, to be solved via economic growth (Arndt 1981; Kapur, et al. 1997). There, poverty was a category that referred mainly to national economies to differentiate between poor and rich countries. In the 1970s, the World Bank significantly changed this by adding to its development agenda the reduction of absolute poverty, "a condition of life so degrading as to insult human dignity and yet a condition of life so common as to be the lot of some 40% of the peoples of the developing countries" (McNamara 1973a: 7). Numerically, absolute poverty was established as the condition of those forty percent of the peoples in developing countries earning under 30 cents per day (McNamara 1973a: 10). This institutionalization of absolute poverty into the development agenda transformed poverty into a concept that emphasized individual livelihoods (Finnemore 1997; Roy 2010). It also foregrounded individual-level poverty, alongside growth, as central preoccupation of development efforts (Best 2013). Most significantly for my argument, the 1970s' turn toward individual-level poverty spurred a multiplicity of theories about poverty, both about how to define it and how to reduce it, giving rise to an abundance of projects that have sought to solve poverty over the last fifty years around the world.

While today's development agenda often uses an income-based poverty line to track the evolution of poverty levels, the individual earning less than 1.90$ /day is not the *direct* object of intervention of most development projects. At randomized evaluation field sites therefore, through the programs they assess for their causal impacts, we encounter different types of "poors", already embedded in specific 'theories' about the problem of poverty: young

adolescent girls taken to be subjects of gender inequality for empowerment programs; farmers whose agricultural productivity needs to be improved in interventions designed to increase fertilizer use such as those funded by the Gates Foundation; or still the poor as natural but cash constrained entrepreneur in microcredit projects. What works to reduce poverty is thus measured through its nested elements, through programs about education, housing, food security, or financial inclusion[22]. Sociologist Nassima Abdelghafour (2017) has also pointed to how RCTs often evaluate interventions that are already black boxed and subscribe to common practices of development and common theories of how to solve poverty, such as improving access to formal banking or intensifying in agriculture.

As several scholars have argued, projects to help the poor are importantly embedded within economic theories and political philosophies about the world at large, beyond the specific issue of poverty (Elyachar 2005; Ferguson 2013, 2015; Roy 2010). For instance, the agriculture and food security sector has roots in a conception of poverty as low rural productivity which took shape in the 1960s within a post-World War II world focused on national production and unemployment and where states were central driving actors (McNamara 1973b). As for the financial inclusion sector, it relies on the figure of the poor as entrepreneur which gained center stage in the late 1980s in the wake of a world globalized around the idea of self-regulating market where microeconomic actors freely exchange (Rodgers 2011a). And the gender equality and women's empowerment sector identified as a target for solving poverty could be situated at the intersection of twentieth century feminism and neoliberal ideas about the self as entrepreneur (Bashford 2014a; Foucault 2004; Murphy 2017).

---

[22] My colleague Doerte Bemme (2018) makes a similar argument in her dissertation by showing that the project for Global Mental Health brings forward a moving understanding of the human as an aggregate human who can be constantly assembled, disassembled, and reassembled amongst different projects.

This analysis brings into sharp focus the ambiguous role played by the issue of poverty within the history of international development. The existence of poverty is that against which this or that economic or political view about the world must measure itself. Poor people are matter out of place, troubling reminders of the limits of contemporary economic policies and politics (Douglas 1966). The lives of the poor are sites where dominant ideas about the economy, the market, and how best to govern are failing. Beyond its positioning as moral imperative, alleviating poverty worldwide thus emerges as a project through which existing theories about the world can be challenged, refined, solidified, and extended to turn the poor into a figure that can also take part in economic and political projects of the time, explicit or implicit. Hence, for instance, the global movement started in the late 1990s to "make markets work for the poor" by correcting market failures via microfinance or cash transfers.

I contend that by telling me that the Gates Foundation, like many other development organizations, was sector specific, Will meant to gesture to these presupposed theories of poverty (e.g. poverty is predominantly the result of low rural productivity). Likewise, when Edward had repeated that CEGA had an ecumenical view on what were the main poverty issues, he was suggesting that significant theories of change (e.g. greater fertilizer use leads to better yields which will translate into higher income which means increased welfare) were based on preconceptions, or in his words were "ideological", and had to be evaluated through randomized evaluations, demonstrated via rigorous evidence, rather than taken at face value. In their decision to be sector-agnostic, Eshan, Edward, Will, and several of my other interlocutors were indirectly challenging these competing theories of poverty and their embedded worldviews. Poverty had always already been pre-defined, my interlocutors' outlook suggested, by people and organizations that claimed to know which sectors were most crucial for poverty alleviation efforts and how change would be brought about.

Websites of organizations committed to evidence-based development in fact insisted that major donors and foundations choose their sectors (or causes) for reasons that are entirely from "the heart", for instance because of "personal experience with a problem"[23]. As an illustrative case from my fieldwork, Lia Fernald, professor of public health at University of California-Berkeley, explained her first research focus on nutrition in developing countries by evoking an episode she experienced of her childhood. She was thirteen, she told me, when there were the first famines in Ethiopia, and she wondered: "how can there be children who die of not enough food?" It is not that my interlocutors would ever deny that health or nutrition are key welfare services to provide to people worldwide. Rather, they suggested that by being sector-agnostic, one could be open to discovering other avenues – such as transportation safety – a quick example provided by Will, that have potentially great(er) impacts on lifting people out of poverty and increasing their overall wellbeing. As I show in the next section, it was not always clear however to what extent evidence-based development concretely moved away from a sector-oriented approach to development and how deeply it put to the test widespread postulates about poverty and their accompanying theories of change.

**Measuring welfare?**

In my conversation with him, Will said that "they pick a problem" but "at least they don't assume the solutions" when describing how the Gates Foundation works. Instead, the Gates Foundation run competitions and funding rounds to find out what would be the best solutions to enhancing health outcomes and improving agricultural productivity in the development world. In fact, during my fieldwork, CEGA and the Jameel Poverty Action Lab (J-PAL), two evidence-based development centers, had been collaborating since 2009 to

---

[23] https://blog.givewell.org/2012/05/02/strategic-cause-selection/ and https://www.effectivealtruism.org/articles/introduction-to-effective-altruism/, both accessed May 21, 2019.

administer and carry out the Agricultural Technology Adoption Initiative (ATAI) funded by the Gates Foundation. The goal of ATAI is to understand what helps and prevents farmers "to adopt technologies and access markets, in a manner that increases their profitability and improves their welfare". At the time of writing, ATAI has funded fifty-one randomized evaluations across Africa and South Asia to assess the causal impacts of multiple agricultural development strategies on farmers' livelihoods[24].

In November 2015, I attended an event to disseminate the policy lessons derived from all the ATAI research projects conducted in the past six years by professors based in leading universities across the world. Chiara, the project manager in charge of the presentation, said that the Board of J-PAL and CEGA affiliates leading ATAI recognize the importance of studying [technology] adoption "because, of course, there can't be any impact if people don't take up [the technologies, such as fertilizer or specific varieties of crops]". She also underscored however that "we also want to know beyond adoption: we want to know whether these things adopted are good for farmers, such as their nutrition or welfare". When I asked her, Chiara confirmed that ATAI's research so far had produced more evidence about adoption than impact, but that in ATAI 2.0, the continuation of the first six years of collaboration and the second phase of the project, they hoped to encourage impact studies of those technologies and things that are found to be well adopted by farmers.

In presenting this vignette about ATAI, I wish to show how my interlocutors practicing evidence-based development simultaneously entrench *and* unsettle a sector-specific approach to global poverty alleviation and its concurrent theories of poverty and change. By drawing a

---

[24] https://www.atai-research.org/about/, accessed May 21, 2019. Fieldwork interviews offered interesting perspectives on the story behind ATAI, which brought together Michael Kremer, one of first economists to have applied randomized evaluations to the question of international development, and Rajiv Shah (then at the Gates Foundation, but who has since become a USAID administrator and is now the president of the Rockefeller Foundation) around their interests in behavioral questions, such as: Why are people not getting immunized?, Why aren't farmers investing in varieties of seed that will protect from drought or excess rainfall? or, Why aren't people making choices that will benefit them?

difference between impact and impact studies, the ATAI board was explicitly underscoring that it is not enough to know about adoption of agricultural technologies. Chiara has emphasized that it is important to know whether this adoption of agricultural technologies translates into increased welfare. My interlocutors were indicating that the causal pathways between agricultural technology adoption, increased agricultural yields, and reduced poverty could not be taken for granted. By requesting more impact studies, my interlocutors were willing to put into question that *agriculture* was necessarily a promising avenue – or sector – of global poverty alleviation. After all, what if impact studies revealed that agricultural technology adoption did not result in increased welfare?

At the same time however, the policy lessons document drafted by ATAI discussed at length the findings derived by randomized evaluations around the determinants and barriers to technology adoption. For instance, the policy insights suggested that insurance against weather shocks made it worthwhile for farmers to invest in farming technologies. It also talked about training interventions enrolling trainers from within the farmers' network to increase the likelihood that they follow advice and adopt the technology. In that sense, positive impact on adoption was framed as a marker of success. Was there not here at play some "technology fetishism" (Redfield 2016: 175), perhaps the product of the fact the randomized evaluations conducted as part of ATAI are funded by The Gates Foundation, an organization started by a tech giant, such as Bill Gates?

I recurrently encountered this focus on measuring adoption in my fieldwork. Many research papers and policy documents based on findings derived from randomized evaluations reported on how different programs affected levels of adoption or usage for different types of services, ranging from fertilizer, immunization, bed nets, insurance, to chlorine dispensers to

sanitize water[25]. In journal articles, policy reports, but also across conferences and other dissemination events, high adoption rates were often taken as a sign of a successful and promising development project.

Earlier in my fieldwork, Liam, a professor of economics at the University of California-Berkeley, had called this phenomenon "the Gates Foundation problem", making again visible the Gates Foundation as a particular way of practicing development against which to differentiate oneself. According to Liam, "[in the 2000s] the Gates Foundation came along, established this 'engineered' list of problems, giving one million here and one million there for these problems. And development economists did not have a lot of money at the time – so it worked." What Liam means is that development economists were granted funds to study programs funded by the Gates Foundation which were in specific sectors, such as education or health. Economists therefore started producing analysis about program $x$ to improve education or program $y$ to increase health, but less on the *last step*, in other words the impact on welfare or poverty itself, however one defines it, due to a lack of income or more broadly due to an aggregate set of outcomes embedded in the notion that poverty is multidimensional, as alluded to by Edward above.

To draw on a famous example, in a 2010 TED talk about "Social Experiments to Fight Poverty", Esther Duflo, co-founder of J-PAL and one of the first proponents of RCTs in development, discussed three programs "which work" to alleviate poverty. Duflo presented a program to incentivize people to immunize their children by giving them bags of lentils when they visit the clinic and an intervention that tested different pricing mechanisms for their impact on bed nets usage, showing that people who get them for free use them at least as much as those who buy them. Again, Duflo framed these two interventions as successful *poverty*

---

[25] For instance in regards to immunization, see Banerjee, et al. (2010), for bed nets take-up see Cohen and Dupas (2010), and for fertilizer adoption see Duflo, et al. (2011).

*reduction* strategies. To be clear, I am not hereby trying to debunk these assumptions – some of which might make intuitive sense – nor questioning the fact that one might provide health and educational services as an end itself. What I seek to rather emphasize is that Duflo, and many of my interlocutors, sometimes turn *means toward* poverty alleviation (e.g. measures of technology adoption) into *ends* of poverty alleviation (e.g. impact on broader welfare). By equating *adoption* of services with poverty reduction, Duflo and other interlocutors implicitly subscribe to a sector-oriented approach to development and their embedded worldviews. For instance, agricultural technology adoption will make farmers more productive and wealthier and thus increase their overall wellbeing. Individuals who have slept under bed nets or have been immunized, by being healthier, will fare overall better in life.

Here again, however, it is worth underscoring how difficult it is to categorize the work of my interlocutors. In her well-known TED talk, Esther Duflo also presented the deworming study with which I opened this chapter, in which it was argued that providing treatment to children to reduce parasitic worm infection increased school attendance. Duflo framed school-based deworming as a successful anti-poverty program, *because* it enhanced school participation[26], without necessarily looking further down the line whether increased school attendance transforms into longer and better education which then translated into increased welfare over a life time. While Duflo does not make mention of it in her TED talk, if you recall, subsequent research on deworming, which followed the cohort of children from the initial randomized field experiment and published in 2011, concluded that mass deworming treatment had long-term impacts, in that those who had been treated as children were found to have higher education and earnings even a decade later.

---

[26] And it does so with little cost, being cost-effective. I return later in the chapter on the question of cost-effectiveness, especially because there have been important debates around the ethics of cost-effectiveness as a guiding principle of policy, insofar as it imports financial logic into welfare service provision. I argue nevertheless that cost-effectiveness as a consideration is subordinated to that of showing proof effectiveness and hence to the main focus of evidence-based development, namely evidence.

The follow-up deworming study therefore makes visible a longer pathway of causal impacts: deworming treatment → not having worms → higher school participation → improved education → higher earnings → increased welfare. It stops at higher earnings and does not directly address the issue of how higher earnings link to overall greater well-being, but it already *measures* an important stretch of the theory of change. On the contrary, the first study only measured the first steps until improved education, leaving *assumed* the steps further down the line. Yet, what the follow-up study suggests is that one cannot *presuppose* that higher school participation will induce long-term life changes but must instead track and trace it by following the cohort of people who received deworming treatment at a younger age. One can therefore reckon with the follow-up deworming study as a challenge to a sector-oriented approach to development which assumes, rather than empirically assesses, these theories of change.

I want to provide two additional short case-studies to show the ambivalence of my interlocutors' work, as it appears in turn as sector-agnostic, scrutinizing established sectors and their embedded theories of change or proposing new possible sectors, and at other times as reaffirming these theories of change, without aiming to demonstrate them empirically. One of programs that Evidence Action was bringing at large scale while I was in Kenya was G-United, in collaboration with the Kenya Ministry of Education. Started in 2014, the G-United program is organized around recent university graduates deployed around the country to help with learning in schools by providing remedial education. The G-United program is part of a broader initiative headed by J-PAL called "Teaching at the Right Level" and heralded since the mid-2000s as a very effective approach to improve educational outcomes. It is a "pedagogical approach" which consists of organizing learning groups by learning levels rather than age or grade[27].

---

[27] https://www.povertyactionlab.org/case-study/teaching-right-level-improve-learning, accessed May 23, 2019.

The initial two randomized experiments, whose results were published in 2007, assessed a remedial education program by an Indian NGO called Pratham (Banerjee, et al. 2007). Since then, other similar interventions have been evaluated in India and across several African countries. Recently, the Teaching at the Right Level initiative led by J-PAL has been one of the recipients of a US $80 million grant provided by Co-Impact, a global collaborative focusing on what they call "systems change, focused on improving the lives of millions of people around the world", and who list the Bill & Melinda Gates Foundation and the Rockefeller Foundation as their donors [28].

The new Teaching at the Right Level website presents the approach as "an evidence-backed educational approach that helps children develop basic reading and mathematics skills, *opening doors to a brighter future*."[29] In other words, while Teaching at the Right Level is presented by J-PAL and its partners as a successful solution of poverty, it is an intervention that remains resolutely within a usual sector of development, namely education, and its embedded theory of development change, which assumes that improved education translates into poverty reduction, or to use the website's word, that it "opens doors to a brighter future".

By way of a contrast, in 2015 appeared in the *American Economic Journal: Applied Economics* an article titled "Six Randomized Evaluations of Microcredit: Introduction and Further Steps". Among the authors were Abhijit Banerjee, also co-founder of J-PAL, and Dean Karlan, founder of Innovations for Poverty Action mentioned earlier, an organization composed of twenty offices worldwide that conduct randomized evaluations of global poverty problems and disseminate their results. The paper reported that the six randomized evaluations of microcredit program, conducted in different countries, all reported "a consistent pattern of modestly positive, but not transformative effects" (Banerjee, et al. 2015b: 1).

---

[28] https://co-impact.org/, accessed May 23, 2019
[29] https://www.teachingattherightlevel.org/, accessed May 23, 2019

As you might know, microcredit – as part of a focus on financial inclusion – is the provision of small loans to people who do not usually have access to mainstream credit services to allow them to start small businesses. Microcredit emphasizes an idea of the poor as natural entrepreneurs who, if given financial resources in the form of loans, will set up productive businesses. In the early 2000s, microcredit polarized debates between those who saw it as a magic bullet and those who were more doubtful about its potential for poverty reduction. In 2006, Mohammed Yunus and the Grameen Bank which he started in 1983 to provide microcredit services were awarded the Nobel Peace Prize "for their efforts to create social and economic development from below"[30].

Findings from the six randomized evaluations conducted in Bosnia-Herzegovina, Ethiopia, India, Mexico, Mongolia, and Morocco, reported that there was no evidence of the impact of microcredit on income, while acknowledging that "increased income is essential to poverty reduction" (Banerjee, et al. 2015b: 12). The paper also reported findings about consumption, because "it is a widely used proxy for living standards", finding no evidence of an increase in household consumption, but some "potentially important if not transformative effects" on "durable consumption", such as assets (e.g. cows) or durable goods (e.g. a washing machine) (ibid). Overall, the analysis was taken to show that "microcredit doesn't live up to the promise of transforming the lives of the poor"[31] as it had been heralded by the key institutions promoting it but also by public figures such as the Prime Minister of Bangladesh who claimed that the microcredit campaign was one of the most effective ways to fight exploitation and poverty.

Of course, this study by Banerjee and colleagues was neither the only, nor the first, piece of research to inquire into the concrete effects of microcredit (e.g: Kabeer 2005; Karim

---

[30] https://www.nobelprize.org/prizes/peace/2006/summary/, accessed May 23, 2019
[31] https://www.povertyactionlab.org/sites/default/files/2015.01.22-Microcredit-EurekAlert!.pdf, accessed May 23, 2019

2011). Yet, it had important repercussions within the development world for putting into question microfinance as a secure sector toward global poverty alleviation. Commentators suggested that the liquidity offer by microcredit might be useful for households in meeting "a wide range of needs rather than solely boosting business income" (Cull and Morduch 2017; Morduch 2017). Overall, the evidence from these randomized evaluations about microcredit has contributed to eroding the theory of the poor as a natural entrepreneur embedded within the microcredit movement. Instead, in their widely read book *Poor Economics*, Banerjee and his colleague Esther Duflo (2011) have advanced an image of the poor as a reluctant entrepreneur who opens a business because no other options were available, who knows that she might never save enough to make it profitable, and who rather wishes that her children find salaried jobs[32].

I have juxtaposed these different examples of randomized evaluations and the programs they assessed, such as Teaching at the Right Level and the microcredit, to show how evidence-based development sometimes puts into question certain taken-for-granted theories of the poor and poverty – e.g. the natural entrepreneur – while at other times it rather evolves within and takes on the assumptions of a sector-specific approach to development, such as when it presents remedial education as a successful poverty alleviation strategy leaving presupposed – rather than tested – the notion that better education eventually leads to less poverty in the form of improved overall welfare. On its website, J-PAL sorts its evidence and policy insights by sectors, alphabetically ordered in tabs ranging from agriculture to political economy and governance. In that sense, RCTs are proposing to evaluate what works best to reduce poverty,

---

[32] Doubts around microcredit might have contributed to the rise of cash transfer programs around the world, as governments and big development agencies like the World Bank are putting in place schemes to provide supplemental income in the form of cash transfers to vulnerable households as a way to increase income, consumption, and durable investment, and ultimately promote poverty reduction. Cash transfers offer their own anthropologies of poverty, especially in their declination as unconditional or conditional (e.g. conditioned to sending your child to school) as well as their own political imaginations, which have been significantly studied by James Ferguson (2015) in his book *Give a Man a Fish: Reflections on the New Politics of Distribution*.

but also to re-inscribe their results within development sectors, which are not readily comparable.

"Most people in development will say *the problem is…*and then they search for solutions to the problem which they see as important", Will had told me, as mentioned above, to underscore that most development institutions were sector-specific, unlike the organization Evidence Action which he headed. Taking Will's word as a cue, and the general insistence of my interlocutors to explain their work in contrast as grounding in a sector-agnostic or ecumenical approach, I have sought to understand global poverty alleviation efforts by being attentive to this reality of sectors as an organizing logic in the history of development. My analysis has revealed that development efforts have apprehended poverty as a composite reality, often addressed through its nested elements (or sectors) – education, health, agriculture, or microfinance – whereby the links between tackling these sectors and the *last steps*, in other words poverty reduction, are not always explicitly addressed or measured by randomized evaluations because they are often presupposed, even at times by my interlocutors.

When deciding which programs to take on and to implement at large scale, Eshan in the opening vignette had said that Evidence Action uses a *filter of evidence* instead of a filter of sector. When one practices development in a sector-specific way, one looks – like the Gates Foundation – for the best solutions to improving a reduced list of issues such as health or agricultural productivity. If, on the other hand, an organization aims to be sector-agnostic, as both Will and Eshan had purported to be, one could cut across and encompass all the different theories of poverty. The cross-cutting theme of his organization, Eshan had explained, instead of it being a sector, such as sanitation or education, was *evidence*. Now that I have examined the filter of sector, I turn to examining how a filter of evidence bears on the practice of development among those who see randomized evaluations as a precondition to successful global poverty alleviation.

**Evidence as "new sector", or the embodied counterfactual**

"But", I asked Will during that same Skype conversation mentioned above, "if there is no sector, if one removes *"the problem is…."* part of the equation, what becomes the criterion to adjudicate between two projects in terms of their success? Against what background are different solutions assessed?"

I was talking with Will about how the organization Evidence Action, which he directed, made choices about which development programs to take up as part of its portfolio and to implement at large scale, as I had done earlier with Eshan when I asked him why Evidence Action focused on scaling-up the Deworm the World Initiative and the Dispensers for Safe Water program in several countries around the world.

"Well, if you want", Will concluded, "my sector is a *cost-effective* program *that works* and *can be scaled*."

When Will talks of a program "that works", he is indirectly making reference, just like Eshan, to a filter of evidence. In this section, I play close attention to the quality of the evidence at the heart of randomized evaluations and why it has such a strong actionable power for my interlocutors. What is this evidence which leads to replace a focus on sectors and priorities with a filter of "robust evidence" as a strong justification for action? When I talk about the quality of evidence, I do not refer to quality as a measure of value or good standard. Instead, I seek to make visible the distinctive characteristics of RCT evidence as its proponents understand it to show how it might differ from other kinds of quantification projects, even within international development, with possibly important implications for collective practices about what kids of futures to envision.

In his reply, Will had mentioned two other considerations: cost-effectiveness and the ability to scale. In fact, all development programs implemented at scale or tested as promising

interventions for scale by Evidence Action had three things in common: a preoccupation for evidence of causal impact, an unfailing regard for cost-effectiveness, and a concern with scalability. The notion that randomized evaluations serve as important guides to allocate limited resources toward poverty alleviation has repeatedly appeared throughout my fieldwork. Silicon Valley investors "who want to make sure they spend their money well" also figure as key players in Will's explanation. In the beginning of this chapter, you might recall, I had underscored that the RCT deworming study in Western Kenya argued for the cost-effectiveness of mass deworming relative to other programs aimed at improving school participation, such as textbook provision or training for teachers. For donors, results of randomized evaluations can form the ground for cost-effectiveness analysis (is it more efficient to train teachers or reduce class-size to increase test-scores?) and determine which programs have impact at relatively low cost. Yet the uniqueness and extensive visibility of randomized evaluation in today's international development world emerge primarily from their focus on identifying causality (Barbu and White 2006; Duflo, et al. 2007; World Bank 2015).

In turn, scale is linked to cost-effectiveness, because the more your program is effective, the bigger the number of recipients you can have with the same budget. Nonetheless, the focus on scale is not reducible to cost-effectiveness but must be understood in light of stated aims to tackle issues affecting millions of people worldwide. Chapter 2 and 3 of this dissertation address the issue of scale within evidence-based development, looking at how scalability is both imagined and infrastructurally enacted in two distinct development programs supported by Evidence Action. I return briefly to the criterion of cost-effectiveness in the last section of this chapter when I looked at a cost-effectiveness model used by GiveWell, a non-profit organization that ranks global development programs to suggest the best opportunities for donors and guide their donation decisions.

Nevertheless, I contend that the relevance of RCTs for those who design, implement, and scale their results cannot be captured exclusively in financial terms such as a return on investments or cost-effective analysis[33]. What I find compelling is that during fieldwork my interlocutors always presented the existence of RCT evidence as the first element that justified policy action, although of course they also discussed other (equally) important motivations (including cost-effectiveness and scale). By examining how my interlocutors describe that RCT evidence successfully captures causal impacts, I foreground other kinds of relevance of evidence-based development beyond accounting logics to see if they might tell a different story about the envisioned possibilities of evidence-based poverty alleviation.

In general, scholars of evidence-based development have discussed how "when development economists and policymakers" seek to promote randomized evaluations, they "frequently invoke[d] a schematic history of medical experimentation, using the presumed authority of biomedicine as justification for emulating it" (Donovan 2018: 31; but also Jatteau 2013). As anthropologists Adrian Petryna (2007) and S. Lochlann Jain (2010a) have discussed in their research, clinical trials have become a central component in our hope for medical progress. Clinical trials are medical experiments that follow the same methodology as randomized evaluations for poverty alleviation, namely the randomized controlled trial. It assesses the effectiveness of a drug by comparing between a treatment group (that receives the drug in question) and a control group (that does not receive the drug but instead a placebo or an older version of the drug).

One my interlocutors, an economics professor who conducts randomized evaluations of agricultural technologies in India argued, along with historians of economics (Cherrier 2017;

---

[33] A significant debate emerged in recent years among development practitioners after the publication of an article (Banerjee, et al. 2015a) in which it was reported that the results of randomized evaluations of a project called "Graduation Program" to lift people out of poverty was very effective across several countries, but at a very high cost given that the program included multiple components including productive asset grants, trainings, and temporary cash consumption support.

Mirowski 1992), that economics evolved from explicitly borrowing the methods of physics toward enrolling the tools of medicine, as both fields use the randomized controlled trial to produce evidence and truth[34]. Going further, I explore here the quality of RCT evidence that makes it so powerful that it is taken up in today's global medicine but also in today's fight against global poverty. I argue that the answer is to be found by zooming in on the treatment and control group comparisons, which makes visible a counterfactual analysis, or the possibility of a different and better life. Let me unpack.

By evidence that a program works, Will and Eshan do not mean numbers and quantitative indicators, or at least not *any kind* of number. All the evidence-based development centers where I conducted fieldwork, such as J-PAL, Evidence Action, but also CEGA in California, all report on, disseminate results of, and lobby for those programs for which there is, as they see it, *rigorous* evidence about *the counterfactual*, namely knowledge about *what would have happened in the absence of the program*. How would have people fared had they not received the intervention? As illustrated in Figure 2 below, the impact of a program, or evidence that it works, is reckoned to be the difference between what happens with the program and what would happen in the absence of the program, as was explained to me by a professor of economics.

---

[34] Whether medicine was truly the field that first applied the randomized controlled trial is somewhat beside the point, but some identify the statistician Ronald Fisher and his agricultural plot experiments as the father of the modern randomized controlled trial as it is practiced today (see Fisher 1935).
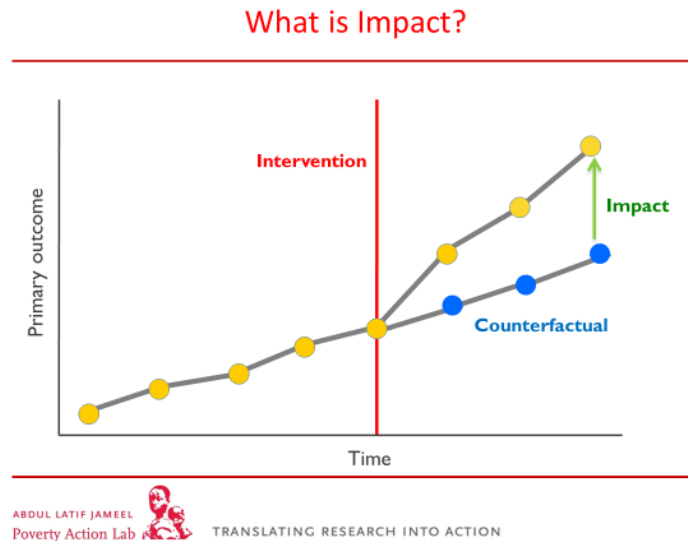
**Figure 2: What is Impact? (Source: *What is Evaluation?* Slide from J-PAL executive training in Denmark, September 2014)**

Back at CEGA in the fall of 2015, Louise, who works to promote transparency in research, had in fact expressed a concern that would repeatedly come up during my fieldwork: "people mixed quantitative data with right away proving causality". We talked at length about the fact that a large, representative sample does not mean that you can right away identify and explain causal mechanisms; rather, "you are still only left with correlations".[35] By this, Louise meant that you can observe a positive relationship between ice cream sales and incidents of drowning, but it is not that ice cream sales cause people to drown, but rather that *both* happen in the summer, when it is hot outside. I am here giving this example as an explanatory illustration, but one can think about other unpleasant correlations, like the ones that link skin color to unemployment or religion to criminality.

Randomized evaluations, in this perspective, are perceived to be a marker of quality. "RCT is the most reliable method for counterfactual analysis, which is required to make a credible claim about causal relationships", Dilara, then executive director of CEGA, told me

---

[35] See also (Ogden 2016: xv) who explains the rise of randomized evaluations in development in light of "the oft-repeated dictum: 'correlation does not equal causation". See Judea Pearl in (Pearl and Mackenzie 2018: 12) who states that the first thing students learn in statistics is that "correlation is not causation".

when I asked her about what the Center understood by rigorous evidence. In statistical terms, randomizing into treatment and control groups allows us to generate two identical populations, which differ systematically in no other way than the treatment itself – people in the treatment group receive for instance microcredit and those in the control group do not – thus making it possible for us to trace back any observed difference between the two groups to the impact of the microcredit intervention itself[36]. As I was often explained during fieldwork, one randomizes to erase out the possibility that other confounding factors explain the impacts observed.

To explain the particularity of randomized evaluations, my interlocutors sometimes contrasted them with other evaluation methods used to assess development programs. At an executive training program on randomized evaluations in which I participated in Denmark in 2013, led by J-PAL and attended by philanthropists, NGOs, government officials, private sector employees, and academics, David Luco, professor of economics at the Paris School of Economics, introduced several other impact evaluation methods including pre-post, simple difference, difference-in-differences, and statistical matching. With the pre-post method, Dr. Luco told the audience, we would compare the test scores of a class before and after they receive an educational program. Let us say the average test score was 25% before the program and 51% after the educational program was implemented in the school. This represents a 26% increase over the year. But, Dr. Luco insisted, "we cannot know whether or how much of the increase can be attributed to the educational program" rather than to other factors. Maybe the test they used to assess children's reading skills was not the same both times, and was easier the second time around, which would explain the higher test scores? Maybe this reflects a time

---

[36] The unit of randomization can differ, such as in the deworming study where schools, rather than people were the unit of randomization. This procedure is called clustered-randomization, where groups of subjects are randomized, rather than individual subjects, through their belonging to a school or village for instance. Clustered-randomization requires more participants to obtain the same statistical power and to draw conclusive findings, statistically speaking, as researchers have to take into account that children in the same school (for instance) are more likely to be similar than those across different schools.

trend and the fact that children would have improved their reading abilities anyway over the year?

The simple difference method measures the difference between program participants and non-participants after the educational program is completed. But if we did not randomize program participants in the first place, can I compare the two groups? It might be that the test scores of those enrolled in the educational program are 10 percentage points higher than those not enrolled, but this could be due to the fact that the best students were chosen, or chose themselves, to take part in the program. "So", says Dr. Luco, "this doesn't tell me anything about the impact of the program. It works *only* if non-participants are exactly equal to participants". I will end with Dr. Luco's discussion of the regression analysis method. In this case, you compare children who have the same pre-test score, the same age, the same gender, the same parents' income, and some other observed characteristics. "If these are the only variables that affect the test scores and are correlated with the treatment", Dr. Luco says, "your estimate of the program impact can be said to be accurate. But often, the estimate is biased, because I am capturing the effect of a factor I did not control for, which I confound with the impact", such as for instance distance to the school.

In contrast, according to my interlocutors, randomly allocating people into treatment and control groups avoids that people who receive microfinance are chosen because they are more motivated, more educated, or more knowledgeable about credit and loans, which are more likely to lead to higher income or other increased welfare outcomes in this group, regardless of the microcredit program. Thus, one can avoid attributing any improvement (or decline) in outcomes incorrectly to the intervention when in fact some other element may have caused it. What Dr. Luco sought to emphasize to his audience is that in all previous examples, such as the simple difference method or the regression analysis, "you have to argue that the group is as good as the counterfactual." By contrast, Luco added, "with randomized evaluation, the groups

are statistically identical before the program", because participants are randomly allocated into treatment and control groups (as shown in Figure 3 below). If you therefore measure a difference between the control group and the treatment group at the end of the program, my interlocutors underscored, you can attribute this to the impact of your program (whether it is a positive or negative impact).
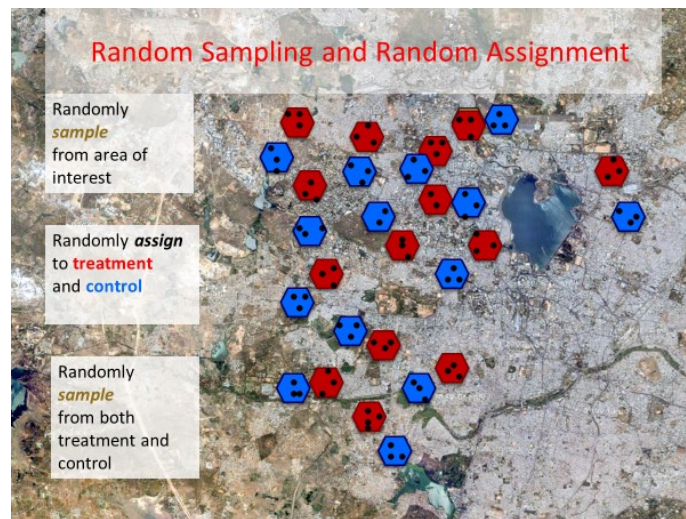


**Figure 3: Random Sampling and Random Assignment (Source:** *What is Evaluation?*
**Slide from J-PAL executive training in Denmark, September 2014)**

The day before, during that same executive training, Mattias Kirac, professor of economics at Aarhus University (Denmark), asked the audience: "How do you control for how many people would have been immunized, even in the absence of your program to encourage immunization?" In other words Dr. Kirac was asking: how do you make sure that you are observing the impact of your program and not just of people's decisions regardless of the program? To disentangle this, Dr. Kirac told us: you need the counterfactual, which he called *the construction of what did not happen.* I allude to Dr. Kirac's words, because I want to foreground how the counterfactual is *a construction*. As Dr. Luco put it, the counterfactual is *the state of the world* that participants would have experienced in the absence of the intervention. But this alternative state of the world is illusory precisely because no one can simultaneously have received and *not* received, for example, the microcredit program provided

by the Grameen Bank. Put differently, if you want to capture the impact of a program you ran into a problem, Dr. Luco said, "because the counterfactual cannot be observed. Instead, we need to "mimic" or construct the counterfactual".

In this view, randomized experiments create a reliable control group that acts as an incarnated proxy for an impossible counterfactual, given that no one can live twice, and serve as their own comparison. I cannot, over the same period, under the same weather conditions, have both *received* and *not received* the 500$ microcredit loans provided to me by the Grameen Bank. I either did or did not. Dr. Luco himself repeated to the audience of the executive training that "randomized evaluations are *a way to artificially create the counterfactual*." RCTs are therefore, importantly, a way to trick time, a sort of magical empiricism[37]: to find artificial ways to escape the unidimensional, linear direction of time as we conceive of it, where there is no going back in time and observing the same things again, tweaking only whether you receive or not the program to observe its effects. The control group provides, according to my interlocutors, the best opportunity to know how people would have fared had they not received the program.

Of course, the preoccupation with capturing causality is hardly new in economics. The entire field of econometrics, understood as statistics applied to economic problems, was developed over the course of the twentieth century to derive causal relationships in the absence of experimental data (obtained by randomizing into treatment and control) for example on consumer behavior (Wooldridge 2006). I am not hereby endorsing randomized evaluations, attributing to them epistemic superiority to know cause and effect relationships over other econometric methods, or even other modes of knowledge, whether scholarly or otherwise. I am neither interested in assessing the quality of evidence produced through RCTs, whether in theory or in practice when it is collected in the field, for there will always be people in favor,

---

[37] I owe my colleague Nicole Rigillo for this formulation of magical empiricism.

others pointing limits, across all kinds of knowledge production. I however want to insist on the RCT design as a distinct way of ascertaining causality because it provides *physicality* to the concept of the counterfactual in the form of an embodied control group, made of people that represent an alternative state of the world, to assess how people would have fared in the absence of the intervention (see Figure 4).



**Figure 4: Students in the treatment group who have received an educational program as part of a randomized evaluation in northern India answer a survey about their life aspirations (Photo by Fiona Gedeon Achi)**

Scholars in the social sciences have described counterfactual thinking as the possibility to imagine alternative worlds, or as the word itself suggests, as what is *contrary to facts* (Ehrenstein and Muniesa 2013). Counterfactual analysis is also a method used in the field of history, although not embraced by all historians, but discussed at length in Geoffrey Hawthorn's *Plausible Worlds* (1993). It uses thought experiments such as "What if the railroad had never been invented?" or "What if humans had not crossed the Bering Strait into the Americas?" to derive the importance of these events to world history. What kind of world would we live in today if they had not happened?

In *The Book of Why*, a recent book about causation, computer scientist Judea Pearl and co-author Dana Mackenzie (2018: 16) define counterfactual reasoning as that which deals with

"what-ifs". Pearl, citing philosopher David Lewis, says that counterfactuals have been seen as fundamental to causality since the second half of the eighteenth century. Pearl cites Hume's definition of causation as an illustration: "We may define a cause to be an object, followed by another, and where all the objects similar to the first are followed by objects similar to the second. *Or in other words where, if the first object had not been, the second never had existed*." (2018: 25, emphasis added). While I am less interested in Pearl's claims that counterfactuals are the building blocks of moral behavior as well as scientific thought, I cite Hume's definition because it summons two related yet different ways of understanding causality, one grounded in regularity (the cause is regularly followed by the effect), and the other one in the counterfactual (if the first object had not been...). I similarly seek to foreground the counterfactual as a central component of the power that randomized evaluations have today in shaping development economics and policy evaluation today.

In her research about cancer in North America, anthropologist S. Lochlann Jain (2013) talks of the counterfactual as the possibility for a different life. That other life is one without disease. What if she, and others, had gone for preventive screening? How would have things turned out then? The counterfactual remains as the haunting picture of a potential better and longer life (Jain 2013: 63). It has a hold on us, even in our most insignificant, daily experiences. The impossibility of the counterfactual is that thing which our friend implicitly gestures to when she tells us: "Stop regretting, there's nothing you can do about it now" or "There's no point in regretting, you don't know how things would have turned out otherwise". It reminds us that all we have is this one life. It reminds us that all we have left is to come. That we cannot change the past but perhaps influence the future. Through her ethnography, Jain (2010b: 230) documents the randomized controlled trial and its artificially-generated, embodied counterfactual as functioning through a logic of war. It is a research method, she argues, in

which "two groups compete, one wins", and in the case of medical research, there is nothing patients can do to change the outcome. Really, it all depends on the drug they are assigned.

In Jain's account, the counterfactual comparison is a matter of life and death. I argue that for my interlocutors the counterfactual comparison indexes in the broadest sense a *better life* for those populations they are intervening upon, or those they are trying to help in terms of education, health, increased income, or something else. If I use the term *better life*, it is to highlight its normative value, or at least the ability to guide action. I contend that the comparison between treatment and control groups in the randomized evaluation, one could say a technical comparison, bears an important actionable force for those organizations which promote evidence-based development, including Evidence Action.


**The better life**

To elucidate the circumscribed but nonetheless important argument I wish to draw, it is useful to start by contrasting the notion of *what works* as used by my interlocutors and the category of the *better life* which I crafted as an analytical category. I suggest that there is a consensus about what is meant by finding *what works* when conducting a randomized evaluation (Deaton and Cartwright 2018). We might contend that RCTs are incapable of capturing causal effects well or argue that they can only measure some interventions and ignore others, or that in reality implemented RCTs are much messier than the methodology is in theory. The fact remains nonetheless that an intervention *that works* when assessed through an RCT experiment usually refers to one that has been shown to bring positive effects on a certain set of outcomes of interests (tests scores, immunization, fertilizer use, etc.) in the treatment group as compared to the control group. In that sense, mass deworming, in the published article by Michael Kremer and Edward Miguel, was found to be effective at increasing school attendance.

When in the field, I saw that positive impacts reported by RCTs were often conceptualized as a program worth bringing forth into the policy arena, however elusive that image of the policy arena and its policy-maker is[38]. Eshan had listed "robust evidence", you might remember, as one of the key reasons for scaling-up mass deworming into a national program, all the while stipulating that his outlook was sector-agnostic. What mattered for him therefore could not be that evidence was *about* health and educational outcomes, but rather that it was RCT evidence, one that offers a comparison between outcomes in a treatment versus a control group. This was evidence, one could write, with its qualifier truncated (evidence *about* or *of*), bereft of sectors, which on its own stood as an evaluative standard. Evidence acquired *by itself* an actionable valence. This could include writing policy insight briefs about how the evidence "challenges the status quo" and "shifts global thinking" about social policies[39]; organizing meetings to encourage governments to take up the evidence; collaborating with governments or NGOs to do randomized evaluations *at scale* of that program to confirm the thrust of the evidence; or in the case of Evidence Action, supporting (technically and financially) governments, whether in Kenya, Ethiopia, or India, to put in place a national School-based Deworming Program but also finding donors and resources to scale up the chlorine dispenser for safe water initiative across different countries such as Kenya, Malawi, and Uganda, as Eshan had mentioned to me.

The RCT's ability to make claims about comparisons between treatment and control (e.g. deworming improves school attendance) holds a political valence. In 2007, it allowed a group of researchers to advocate the importance of mass deworming to a Young Global Leaders Education Task Force at the World Economic Forum Annual Meeting in Davos. The Task

---

[38] Sometimes an RCT or a set of RCTs reported not about a successful intervention to be replicated or scaled, but about some lesson about human behaviors, like around incentives (e.g. for immunization) that could be translated into elements to include to design effective policies. See also chapter 4 for a longer discussion about how RCT findings are generalized.

[39] https://www.povertyactionlab.org/evidence-to-policy/shifting-global-thinking, accessed May 28, 2019

Force subsequently launched the organization Deworm the World, which brought together development practitioners and government officials to lobby for the uptake of deworming into national public health programs, starting with Kenya in 2009. Indeed, RCT advocates propose *the better life* (i.e. the comparison between control and treatment groups) as (one of) the most compelling measures of success that should guide action.

Put differently, *what works* signals *what should be done.* And what works has this potency in important part because, I contend, its identification relies on a very telling, physical, visual tool: that embodied counterfactual that makes visible not only a simple descriptive contrast between two groups, but the possibility of what I have termed a *better life* – in this case improved education for those who have taken deworming treatment. For instance, just below in Figure 5 is a slide used during the executive training I attended to explain the strength of randomized evaluations to evaluate program impacts. You can notice the use of pictures of villagers to represent visually the treatment and control group, but also the use of simple schematics, helping the audience to understand how randomized evaluations and the evidence they report on should be interpreted[40].

---

[40] Interestingly, at a logistics meeting I attended in Kenya to train county officials to implement the national Deworming Day, the visual slides used made reference to the randomized evaluation of deworming in the first slides but did not use visual tools to inform its audience, perhaps because the presentation was happening in a context where the Kenya National School-Based Deworming Program was already taking place in the last few years, rather than in a context of advocacy efforts.
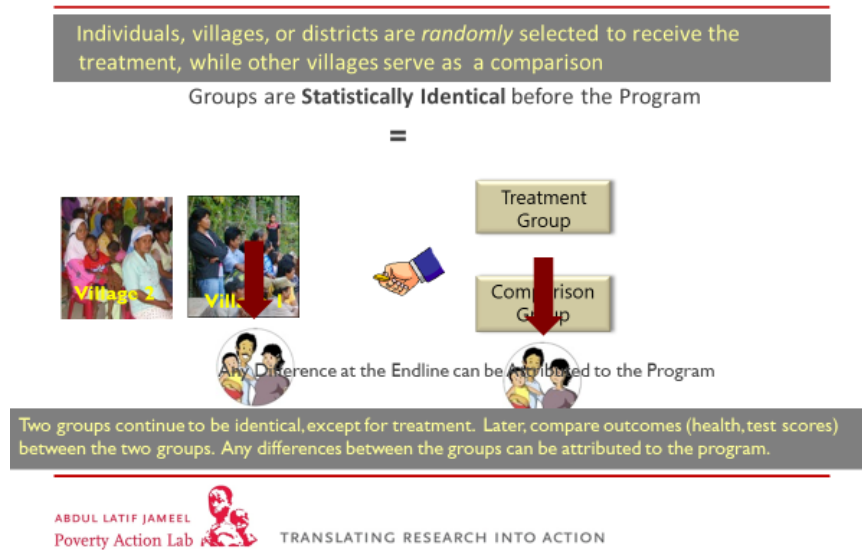
**Figure 5: Randomized Evaluations (Source: *Why Randomize?* Slide from J-PAL executive training in Denmark, September 2014)**

In "The Rise of the Randomistas", Kevin Donovan (2018: 39) discusses how RCTs are promoted as self-evident, speaking for themselves, because the evidence is so simple to interpret. With their telling illustrations of causal certitude, Donovan states, those who promote randomized evaluations do not only purport to assess "*past* projects" but to offer "guides for *future* interventions", not only to show "what worked in that instance", but rather "what works" generally (ibid)[41].

This *better life* is not measured against a set definition of the good life, but in comparison to an artificial, embodied, counterfactual of myself, or my family, or my village, or my city. Within evidence-based development, the improved life achieved is a difference between two versions: with and without the intervention, whether it is regarding education, agriculture, health, or anything else. Implicitly, RCTs are seen to generate a sort of a double-present, which might signal the contours of a *better life*, one not fully located in the future, and

---

[41] I am here interested in the translation of evidence into *future* projects and in influencing policy-making. This chapter is concerned primarily with the conceptions of temporality and time at the heart of evidence-based development. For questions that have to do with generalizability of findings across "contexts", see Chapter 4.

yet one that one now knows how to bring forth for more people, through deworming, cash transfer, or any other project found effective through randomized evaluation[42].

In his article "Bioexpectations", anthropologist Peter Redfield (2012: 180) explores what (neoliberal, biopolitical) techniques of care emerge when expectations that people should live "even under extreme conditions of crisis, neglect, and poverty" are combined with "doubts about the state capacity to safeguard populations"[43]. "The result", he contends, is "a set of technologies built around minimalist forms of care", such as food provision or clean drinking water, provided by humanitarian organizations but also private corporations, which seek to "save the present through auto-empowerment at bodily level, whatever that might mean for the future", based on realist, rather than utopian visions, in a "second-best world"[44] (ibid). While my field sites exhibited these minimalist, individual-level forms of care, that have an orientation toward present lives, such as with deworming or household-level sanitation, the future is never out of sight.

Through the counterfactual, through the artificial creation of a double-present, the future is brought into the present. What randomized evaluations suggest is that, it might be too late for some, but for someone, somewhere, such as the children of today, a different life – another life – is still a possibility. Deworming will alter the course of that child's life who is likely to earn more as an adult; free chlorine for water sanitation will keep that second baby from dying. It requires the right intervention. Regardless of the sector or issue it tackles, the life will be *better* (even if not fully good). Evidence-based development is a project moved principally by the certainty of a comparative, incremental change, what I have called the *better*

---

[42] On the question of "scale" of the intervention, and the fact that most RCTs have concentrated on evaluating little development devices (Collier, et al. 2017) and cannot actually assess bigger structural questions, see the Introduction to the dissertation. I also believe that this remark is not confined to evidence-based development but reflects a more general trend in development alleviation toward micro interventions.

[43] For an examination of how evidence-based development is situated in complex relations vis-à-vis the state's project and its capacity to provide for populations, see Chapter 3 of this dissertation.

[44] When making reference to the second-best world, Peter Redfield cites Fiona Terry's (2002) *Condemned to Repeat? The Paradox of Humanitarian Action*.

*life*. While evidence-based development is an heterogeneous project to find what works to solve poverty around the world, I argue that my interlocutors were often not directly measuring, nor addressing and achieving, an ultimate reduction in poverty, whether defined as increased income or by several other dimensions. The *better life* is not defined like the good life by several dimensions of what makes it, but is rather characterized by being based on evidence about the counterfactual in the form of randomized evaluation.

In their edited volume titled *The World of Indicators,* anthropologists Richard Rottenburg and Sally Engle Merry (2015: 9) argue that "a crucial difference between large-scale interventions typically understood to be the hallmark of modernity and newer neoliberal experiments to design evidence-based melioristic interventions is the way they conceive of time". In their view, development interventions of modernity were based on heroic narratives of progress, in which the future could be known and shaped *in the longue durée.* By contrast, evidence-based interventions evolve within understandings of the future as "unknown and risky". "Accordingly," Rottenburg and Engle Merry write, "interventions only envisage the next step beyond what they have done, rather than a long trajectory of improvement" (ibid), what Peter Redfield (2016: 175) has also termed "small, incremental approaches to social change", premised on "the modest but urgent desire to 'make a difference' in an often intractable world". I believe that my field sites, and the way my interlocutors practiced evidence-based development, gesture to this cautious orientation toward the future and possibilities for improvement. Deworming, sanitation, or agricultural technology adoption might be narrow definitions of the good life, but perhaps ones that try to make up in certainty and immediacy the ambition they lack in scope[45].

---

[45] And that also, just as importantly, tries to make up in scale for this lack in scope, as discussed in Chapters 2 and 3 of this dissertation.

In her 2010 TED talk "Social experiments to fight poverty", Esther Duflo, founder of J-PAL and now Nobel Laureate in Economic Sciences, had said: "How do you know what would have happened without the aid? Maybe it would have been much worse, or maybe it would have been better. We have no idea. We don't know what the counterfactual is. There's only one Africa." So when I talk of the *better life,* far from me to suggest that RCTs are bringing about improvements which are not transformative. Instead, I seek to underline how RCTs can be read as one proposal for causal certitude in a world riddled with both epistemic and moral ambiguities. Let me unpack further through the emergence of the RCT movement in development, as recounted by Valery, Director of Global Innovation at Evidence Action, and the person who has probably been at organization for the longest time.

Do you know the story of how it all started?", Valery asks.

I first met Valery as I overheard her instructing in a cheerful voice the Deworm the World team on how to record some video clips to promote the work of Evidence Action. "Show the documents, the boxes of deworming drugs, and the black bags with names of schools so that people can understand what we do", she told them. Valery had lived in Kenya for over twenty years and first started working with Michael Kremer, the second principal investigator of the deworming project, when he was still one of the few development economists running randomized experiments.

"People sometimes overlook", Valery tells me, "that randomized evaluations started with Michael Kremer", now professor of economics at Harvard.

Michael Kremer, Valery related, had lived in Kenya during his undergraduate studies and then worked for an organization called WorldTeach. Through this, he got to meet Chris Tucker who was with International Christelijk Steunfonds (ICS), at the time a Dutch Christian

charitable organization that headed different educational programs[46]. Years later, after Kremer obtained his PhD, he went back to Kenya on a vacation and met a few of his friends, including Tucker. Kremer suggested to Tucker that they select 14 schools and allocated the ICS' sponsorship program to sponsor children to go to schools to seven schools. So with Andrew Minso, a colleague of Kremer, they started partnering with ICS to run a lot of randomized evaluations out in Busia in Western Kenya. That was in the mid-1990s.

Around 2004, Valery says, ICS got a new board and they did not approve of Kremer running all these experiments. Larry, who now works for Evidence Action but joined ICS around that time, also mentioned these tensions in the organization's leadership whose members felt that too much of its activities were now turned toward research in the form of randomized evaluations. The board thought it was awful to do experiments with children, Valery explains. ICS wanted to help children. "But how would they *know* if they were helping them?", she ponders. Valery also says that, in all fairness, perhaps no one took the time to explain RCTs well to the operational staff at ICS. She also insists that she means this with a lot of understanding considering that she too initially moved to Kenya from the United States to launch an NGO. The ICS projects might or might not have helped. But maybe the organization *did not want to know* [whether they helped or not], Valery adds. Valery reminds me that ICS had all these initiatives, such as giving out free uniforms, which, through randomized evaluations, were found *not* to be effective. And while Valery does not express it, I catch myself thinking: or the projects perhaps made things worse.

A few weeks later, in fact, in another meeting where I asked her about the importance of causality in the work of Evidence Action, Valery confided that "part of why I think is that we were allowed to do things without evidence is because we thought intentions were

---

[46] The organization has changed names several times and since 2011 operates under the name Investing in Children and their Societies.

enough"[47]. Here, Valery is alluding to the fact that for several decades, the impact of development schemes were not regularly evaluated, whether through randomized evaluations or any other methodology mentioned above, such as pre- and post- comparison (Kapur, et al. 1997). Until the late 1990s, most evaluations focused on output rather than impact, in other words results that the program most directly and visibly generated, such as: the number of schools in which uniforms were distributed; the number of bridges built in the country; the number of households that the food supplement program reached. A focus on outputs, it is argued by Karen and other interlocutors, does not tell anything about whether intervention A has (positively or negatively) impacted poverty alleviation outcomes (e.g. increase in school enrollment, economic growth, future earnings).

Intentions and "unintended effects" echo the arguments made in James Ferguson's *Anti-politics Machine* (1990), a seminal work whose approach has influenced an important part of the anthropology of development in the last twenty years. Certainly, at least some of the "negative effects" of development work that Valery points to are of a different order than those documented by Ferguson (1990: 87-8), whereby development, a universal form of expertise which "knows development" rather than local realities, transforms political problems (e.g. political subjugation) into technical problems (e.g. lack of education or access to credit) requiring exportable, simple solutions. Valery focuses on individual people, and her perspective on their well-being does not explicitly include political existence, at least not in the sense understood by Ferguson and other anthropologists, as self-determination, whether at national or smaller scales. Nevertheless, it seems noteworthy that I have heard about the

---

[47] The point for me is not to delve into the issue of "intentions" at the heart of international development efforts, which Valery implies were "good ones", even if the effects were negative. Others, including several scholars, might suggest that the "real" stakes of development work are not to help countries or people but economic dominance or political stability. On the question of whether development is a post-colonial project, and therefore an illegitimate endeavor, I feel ill-equipped to speak authoritatively but would tend to call, with other authors, for more nuanced accounts, which make room for uncertainty (e.g. Rottenburg 2009a).

mismatch between intentions and consequences emphasized not only in anthropological literature but in the midst of my fieldwork as well[48].

Surely ICS, like many other development initiatives, had its own ways of tracking the success of its projects. It might have counted the number of uniforms and textbooks that were distributed in schools to provide an idea of how many lives it had impacted. It might have garnered survey data to document how children were feeling after receiving uniforms. It might even have collected test scores across a sample of villages. And yet, Valery had wondered, "maybe the organization did not want to know". Knowing here meant generating a randomized control group – an embodied counterfactual – as the best approximation of causal impacts. In this portrayal, ICS could be said to address poverty *without knowing*.

And this, I read into Valery's tale, seemed an illegitimate approach to poverty alleviation. In fact, precisely because Evidence Action's work is about *development*, "which is about changing outcomes, not reducing 'something' right now, but having lasting effects", Valery had insisted, "causality and the counterfactual is something we should care about". "If your approach is merely humanitarian for instance", Valery had contrasted, such as distributing food or providing shelter to relieve temporary suffering, "the burden of proof is much lower", meaning you are not required to link your intervention to long-term (positive) effects, but simply to provide items to sustain immediate lifesaving relief efforts.

This is a small vignette but one that illustrates that randomized evaluations are also positioned by my interlocutors in reaction to a vision of a past in which NGOs, governments, and international organizations were perceived to put in place projects that triggered futures whose quality was unknown, because the effects of interventions were not evaluated, or, if they were, not with the right tools. These could be better, unchanged, but also poorer futures,

---

[48] See the book *More Than Good Intentions: How a New Economics is Helping to Solve Global Poverty* (2011) authored by the founder of IPA Dean Karlan and his field researcher colleague Jacob Appel on randomized controlled trials for research about poverty alleviation.

whether individual or collective. In a recent article, Kevin Donovan similarly discusses how "randomistas" have justified the need for randomized evaluations as a first step toward successful poverty alleviation "through their insistence on the pervasive *lack* of knowledge within international development circles" (2018: 28).

Here I wish to make visible that RCTs might be understood not only as a tool of efficiency and accountability in an unchanging landscape of development, or for generating simply yet another type of statistical indicator, but also as one specific *project* of global poverty alleviation. It is a project that sidesteps the ends of development to instead foreground an agnosticism about what are the most pressing poverty problems. Evidence-based development equates the lack of a certain form of knowledge – that of the counterfactual in the form of a control group– with a failed vision of development. In this framework, assessing causal pathways emerges as a precondition for any (possible) and durable improvement in the quality of life, regardless of how one defines it.

This analysis summons an important tension in my fieldwork between sector-agnosticism on the one hand, as in the refusal to define what the most important poverty issues are, and a lack of *methodological agnosticism* on the other hand, which insists on randomized evaluations (and sometimes related research designs) as the best way to identify causal impacts, to know what works to reduce poverty, and to reach effective anti-poverty policy around the world[49]. Social scientists have commented on how this methodological 'dogmatism', whereby evidence or "rigorous evidence" becomes synonyms with RCTs, might trump all other kinds of knowledge and experiences in the process of informing development agendas (Drèze 2018; Jatteau 2017; Ravallion 2009).

---

[49] When this question arose in fieldwork, I was often reminded that many RCT designs include qualitative data to better understand the "mechanisms" behind causal impacts, to show a bridge, rather than a segmentation, between RCTs and other modes of knowledge. The extent to which this connection is an actual reality requires further investigation.

In recent work, historian Jacqueline Best (2017a) warns that the ubiquitous focus on evaluation and calculability within international development risks, in fact, to undermine the redefinitions of development that occurred in the 2000s, which proposed broader and messier development objectives that linked together economic, political, and social spheres. "The question of what counts as success, Best writes, is both highly technical—involving questions of evaluation and calculation—and normative—raising the question of what we value enough to define as success" (2017a). Nonetheless, she claims, international development today requires that everyone plays the game of calculability, whether to prove program impact to relevant governments, to demonstrate cost-effectiveness to donors, to respond to demands of accountability from international organizations, or to determine adequate pricing schemes in for-profit social enterprises. The result is that practitioners (might) end up consistently prioritizing the technical, "design[ing] their policies so they are as to measure as possible, distorting development objectives to make them appear calculable". The consequence is a proliferation of more measurable, or 'testable' interventions (such as mass deworming or cash transfers), concerned with individual level outcomes (nutritional status, life skills' learning, and livestock possessions).

Conversely, within development circles, some voices are rather contesting this idea that policy-making is taken over by scientific evidence. Some economists have underscored how impact evaluations and research in general always play a minor role in policy-making, given that "policies (whether policies of government or of NGOs) are created through complex and opaque actions influenced by politics, capability, capacity, resource constraints, history and many other factors" (Ogden 2016: xxvii). Here, the process of policy-making and the process of knowledge production are painted as being separated from, rather than imbricated with, one another. This can even be read as a critique pointing to the limits of evidence-based

development which claims to influence policy change, but that "does not have a reasonable path to actually affecting policies and programs" in significant ways (ibid)[50].

I experienced a similar line of questioning when my fieldwork led me to visit government offices and attend meetings. It was not always clear whether governmental officials made the difference between RCT evidence and other types of research that supported diverse national development programs put in place in partnership with various international organizations or NGOs. My argument therefore is *not* that evidence-based development replaces political and ethical comparisons that require value judgements (such as deciding sectors to prioritize, where to allocate resources and funds, what constitutes good goals) by a technical comparison between treatment and control at the heart of the RCT design.

Instead, I want to highlight that this better life, which carries its own normative valence, is at the root of the tension I traced at the onset of this chapter: my informants claiming to hold no preconceptions, to be sector-agnostic or ecumenical, all the while being engaged in policy change. One may not know fundamentally, or in any case may refuse to state, whether it is more important to tackle health or education. Nonetheless, insofar as one has robust evidence produced by RCTs of the benefits of deworming on educational outcomes, one feels justified to seek to scale up that intervention. Of course, in practice, the transformation of deworming study into the actual Kenya School-Based Deworming Programme enrolled much more than evidence, including donors' interests and the collaboration of ministries[51]. The preoccupation with the *better life* becomes one small, yet significant, element in an intricate set of judgements, choices, values and objectives, efforts and trade-offs, interests, limits, and constraints to guide development efforts around the world. The equivalence between *what works* and *what to do* is far from complete, as I describe in greater detail in the next section.

---

[50] I return to this question of the power of RCT evidence to shape policy in the Conclusion of the dissertation by looking at the distinction between evidence of impact and impact of evidence.

[51] For a longer discussion on advocacy efforts that contributed to establishing the Kenya National School-Based Deworming, see Chapter 3.

**Multidimensional filters**

When I had asked Valery how Evidence Action chose which new programs to take on as part of its portfolio amidst all those interventions that had been reported to be successful via randomized evaluations, Valery had told me: "you're asking me, how do we compare apple and oranges? Education versus a program on another issue for instance…"

Valery had said that there were various techniques out there, making reference to QALY and DALY, which respectively mean "quality-adjusted life year" and "disability adjusted life year", two measures generally used in global health to quantify the relative burden of different diseases from mortality and morbidity and to calculate the value of diverse medical interventions to enhance both the quantity and quality of life[52]. The invocation of medical indicators to adjudicate between the potential of development interventions deserves examination in its own right, and this has been done by other anthropologists (Adams 2016; Rock 2000). What interests me here is that Valery continued this conversation by stating: "But it's very complicated. So there might end up being a multidimensional filter, including donors' interests".

As you might recall, Evidence Action was in fact devising a systematic organizational strategy to guide its future work. Eshan had told me that until now, the organization had been primarily guided by a filter of evidence, debating whether it should also become sector-specific, adding another filter on top of that of evidence, for instance by choosing a focus on sanitation or education. Now Valery was talking of a *multidimensional* filter – or in other words, the idea that the organization would be guided by several concerns, including donors' interest - which Eshan had also alluded to when stating that for the scale-up of deworming, "there was a good amount of evidence combined with investment from donors". Will also had already added cost-

---

[52] https://www.who.int/healthinfo/global_burden_disease/metrics_daly/en/, accessed May 28, 2019

effectiveness and scale to the preoccupation alongside evidence as relevant parameters to identify which programs to focus on for Evidence Action. Similarly, Valery shared that there were some filters already: the program had to be big and it had to be on poverty. Other NGOs might "slice it differently", she added, some focusing on Kenya, some only on programs about water, others only on collaborating with governments. Sara and other interlocutors had also pointed out how the short history of Evidence Action (started in 2013) had been marked by an "opportunistic way" of working.

Sara was heading the expansion of the G-United program in Kenya mentioned above, a more recent program of Evidence Action, which consists in pairing recent university graduates with schools to provide remedial services to students. Before, Sara said, Evidence Action had only two flagship programs: the Deworm the World and the Dispensers for Safe Water. Then, the question for the organization became, Sara told me, "What is the next program? What fish to catch?" Sarah emphasized that this process of selection was done very informally at first. There was the fortuitous convergence of "good evidence" and "good partners" which helped bring all these programs to scale. In contrast, Sara says, Evidence Action was moving toward "doing things more systematically", for instance, through "call for results". Sara was pondering out loud what the call for results should include: some thematic areas? But also, she wondered, "should we be calling only for evidence or also for specific partners, which are based in the region, who already do development work, but would like, but don't know how, to implement evidence-based programs?"

I am weaving together these discussions about the future strategy of Evidence Action because they pry open the question of how evidence-based development is routinely enacted. What is RCT evidence paired with in decisions about which programs to fund, to prioritize, and to implement? It must be noted that even within the field of development, the importance of this question has started to be recognized. For instance, development economist Jean Drèze

recently wrote in a blog post on evidence-based policy in India: "It seems to me that the relation between evidence and policy needs further thought. Evidence is about facts, policy is a political decision – in what sense is it to be 'based' on evidence?"[53]

I provide a last anecdote centered on GiveWell, a non-profit organization which ranks global development charities to suggest best-giving opportunities which are "evidence-backed, thoroughly vetted, and underfunded"[54]. GiveWell is part of the Effective Altruism movement, a project motivated by the question of how we can use resources to best help others. GiveWell is founded on the recognition that most people want to help reduce poverty but figuring out *what* helps is a very difficult process, from the larger challenge of deciding which problem to focus on to the more specific issue of which organization to dedicate oneself to, whether one is an individual donor, a development practitioner, a wealthy tech mogul wanting to start her own philanthropic foundation, or even a government or a large international organization, such as the World Bank.

In May 2016, I visited the head office of GiveWell in San Francisco. Celia, Outreach Associate, Ruya, Research Analyst, and I sat in a small and bright conference room on a top floor of the building. We look over onto grid-like streets and perpendicular offices and stores, that continue all the way into the docks lining up the San Francisco Bay. That year, the top three charities recommended by GiveWell included Evidence Action's Deworm the World, the Schistosomiasis Control Initiative (also working on the issue of parasitic worm infections), and the Against Malaria Foundation. GiveWell used four analytical criteria when researching and ranking individual organizations: evidence of effectiveness, cost-effectiveness, transparency, and room for more funding (i.e. "what will additional funds enable and what is the value of these activities?"). What I had come to ask the GiveWell team is how they compared charities

[53] https://www.ideasforindia.in/topics/miscellany/evidence-policy-and-politics.html, accessed May 28, 2019
[54] https://www.givewell.org/, accessed May 28, 2019

that worked on very different issues (e.g. malaria, worms, nutrition, transportation safety) in terms of how much they improved lives.

Celia makes references to their cost-effectiveness model, which is downloadable from their website. Both Celia and Ruya tell me that the cost-effectiveness model includes both fixed, objective elements such as, for instance, the prevalence of malaria around the world or the treatment effect of deworming, that are compiled through reviews of the literature. On the other hand, "one of the harder or most subjective [aspects] is how we compare the different outputs of the programs", Ruya says.

"The Against Malaria Foundation, their main effect we look is they save children under 5. Deworm the World: their main outcome is they increased income later in life. So how do you compare these lives saved to increased income? It's hard", she confesses.

We all laugh, out of a feeling of shared challenge, but also simultaneously of discomfort at this somewhat unfitting reality.

"We want people to put in their own value-judgments because we don't have other components for trade-offs", Ruya continues.

Celia and Ruya show me that the cost-effectiveness model includes an Excel line where one can enter "how many years of healthy life would you be willing to give up to get a year of extra income", what they call a subjective input. They explain that every staff member will enter their input to these questions and that ultimately, they try to compare different organizations and their interventions around a number that says "cost per equivalent life saved", a metric GiveWell uses to compare interventions with different outcomes, such as income improvements and averting death[55]. They even mention their worry that the process is

---

[55] While this measure of "cost per equivalent lives saved" can bring to mind Disability-Adjusted Life Year's logics discussed of reducing all interventions to the same denominator to quantify different burdens of morbidity and disease, GiveWell is explicitly unsatisfied with DALY as a comparative measure to determine the relative importance of issues because it reduces everything to the same unit of total years of life saved without differentiating among different outcomes (e.g. adult lives saved, cases of extreme misery averted, amounts of

not fully independent, because you can see other people's estimates, so all members of GiveWell could be anchored to each other in their judgement, and thus give similar values for these *subjective* inputs. In other words, GiveWell included staff values in their cost-effectiveness model, which you can see under the "Parameters" tab sheet in the picture of the cost-effective model below (Figure 6).



**Figure 6: GiveWell cost-effectiveness analysis –November 2016 (Source: https://docs.google.com/spreadsheets/d/148Z3fA0cz39CgeyQcaR5LNJuzsX0ZmcM3h3gUyg_pWg/edit#gid=2064365103)**

This conversation with GiveWell makes explicit the difficult questions that confront evidence-based development today, especially when one wants to think about the allocation of priorities and the distribution of resources *across* sectors, to use Eshan's words. Of course, there are innumerable RCTs conducted on a multitude of interventions, and not all interventions that are identified as having positive impacts are turned into more lasting or bigger programs, or are even discussed in the policy arena, an obvious state of the world that shows that, of

---

income increased, etc.). See: https://blog.givewell.org/2008/08/22/dalys-and-disagreement/, accessed May 28, 2019.

course, other evaluative standards are at stake beyond RCT evidence, its counterfactual and *the*

*better life*. Nevertheless, some organizations and people within evidence-based development

are more willing to engage with reality than others. Searching for GiveWell's current cost-

effectiveness model (they update it each year), I discovered that in 2017 the organization

attributed a name for these subjective parameters, talking about the importance of assigning

*moral weights* to cost-effectiveness analysis, as you can see in GiveWell's cost-effectiveness

model dated for this year (Figure 7 below).
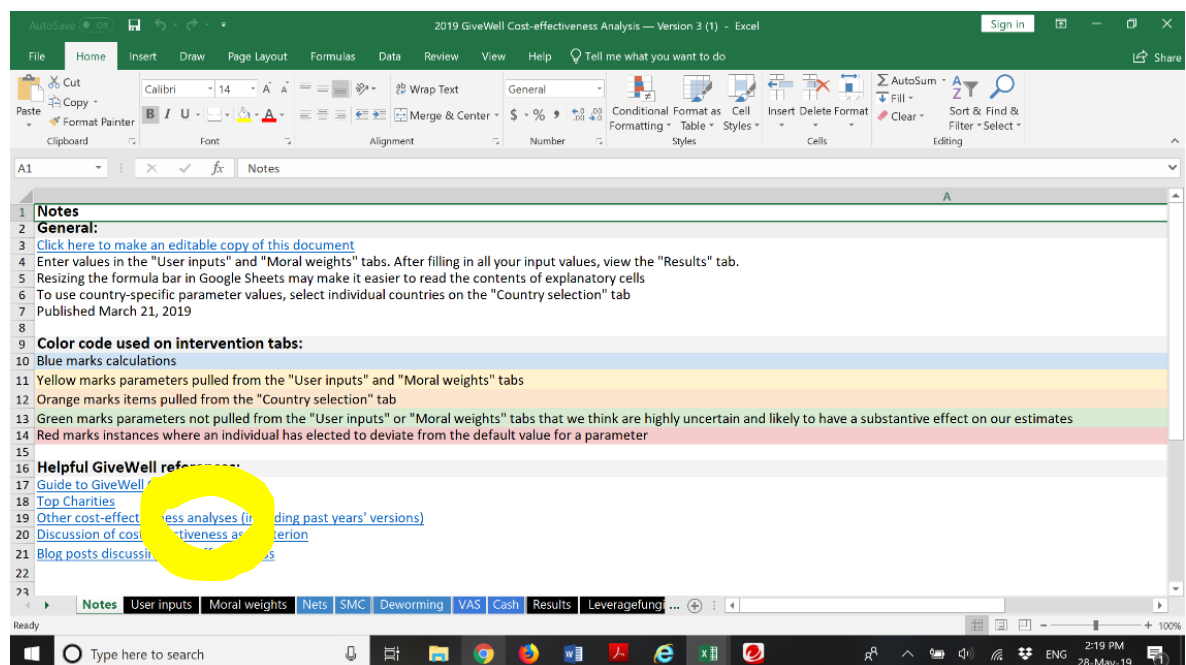


**Figure 7: GiveWell Cost-effectiveness Analysis –Version (Source:**
**https://docs.google.com/spreadsheets/d/1Y8gy8kHZwu40KuoE9GO_EWZS0x4ZMp3GX2_unD**
**Mtgyk/edit#gid=2064365103)**

In a blog post, GiveWell argued:

> However, when it comes to making a decision about where to donate or how cost-
> effective one charity is relative to another, one also needs to consider moral
> questions such as:
> •Valuing health vs. income: How much do I value averting the death of a 2-year-
> old relative to doubling the income of an extremely poor household?
> •"Age-weighting": How much do I value averting the death of a 2-year-old relative
> to averting the death of a 30-year-old?

Setting moral weights (incorporating the value judgments in the bullet points above, or similar, by giving them numeric value in our cost-effectiveness analyses) is uncomfortable and challenging. (…) However, answering these questions is unavoidable. Anyone deciding to donate to one charity over another is implicitly using moral weights, even if that person is not explicitly engaging with them[56].

On the same webpage, GiveWell foregrounds what they see as the impossibility of escaping moral choices in the form of attributing moral weights to different outcomes. Specifically, they cite a quote from an interview with Esther Duflo, co-founder of J-PAL and one of the development economists considered to have launched the RCT movement in development. I had just read that interview with Duflo in the book *Experimental Conversations* (Ogden 2016) a week before landing on GiveWell's website and had highlighted that same passage in orange which contains that excerpt. The book is a collection of interviews with key figures in the development sphere, engaged in various ways with evidence-based development, whether as academics or donors, critics or collaborators, working in institutions in the United States or elsewhere.

The interviewer, economist (and supporter of RCTs) Tim Ogden had asked: "Where do you feel the balance of evidence is on giving people money and letting them choose, or knowing there are some things they are underinvesting in and so we should give them, say, a bed net?" To which, Duflo's reply reads:

The problem is that no one has done that *comparison.* We know they are not buying bed nets with their GiveDirectly money (a program of cash transfers). I guess I don't know that for a fact, but it's knowable. *I think it's the type of question that will be frankly difficult to address because you're required to value benefits across sectors. So it depends what your objective is.* If your objective is to control malaria, you can do an experiment where you give money and you give people bed nets or you do something else and you can see how many people sleep under a bed net at the end of the day, and what's the number of malaria infections you've averted. So that's a question that's well defined. *Whether it's better to give cash or give bed nets would require you to make a judgement about what is the importance of making people*

[56] https://www.givewell.org/how-we-work/our-criteria/cost-effectiveness/comparing-moral-weights, accessed May 28, 2019

*healthy versus having them buy a roof. And that's one I'm not prepared to make* (Ogden 2016: 17, emphasis added).

It is an interesting coincidence that Duflo's words should catch my attention and as well GiveWell's when we were both writing about evidence-based development and thinking particularly about comparisons, not across treatment and control groups this time, but across sectors and goals, that involve value judgements or what GiveWell termed "moral weights", highlighting that these decisions were "uncomfortable" but also "unavoidable". One can sense a similar cautionary tale in the address given by Ruth Levine, well-known development economist and now director of Global Development for the Hewlett Foundation, at the closing session of 'Evidence Day' in September 2017 at the U.S. Agency for International Development:

As a society we are routinely asked to make trade-offs between conditions that are not intrinsically comparable: the well-being of a baby versus an elderly person, the welfare of those of us living today versus our descendants two generations hence. We have to make these choices around which there are no objective guides or guardrails, and individuals and groups draw on profound beliefs, whether from religion, historical precedent, philosophical tenets, or other sources. (…) Empirical analysis is not a substitute for the value judgments that inform a theory of justice in any society. But empirical analysis is an essential complement to those value judgments, helping to turn the 'what we believe' into the 'what we do.' [57]

My direct interlocutors were closer to Esther Duflo than to GiveWell's outlook, insisting, even in the face of a question that asked them to make choices, that they did not wish to make the choice, suggesting that this choice was not necessary or essential to their policy work[58]. Duflo might hold private ideas about the relative "importance of making people healthy versus having them buy a roof", but when interviewed in her capacity as Jameel Professor of Poverty Alleviation and Economic Development at MIT, she said that she was

---

[57] https://hewlett.org/moral-case-evidence-policymaking/, accessed May 28, 2019.
[58] As some of my readers have pointed out, there seems to be an infinite regression to the question of being sector-specific, which casts doubt on the possibility of ever being agnostic. Indeed, sector-specificism still seems to persist upstream in the selection of programs to evaluate even if one refuses to state priorities of development. Academics are making a choice, or judgement, about which projects to generate evidence about. How do they decide which areas to generate evidence about?

"not prepared to make" that choice. Edward, when pressed by Chiara to define what the end goal of the Center for Effective Global Action is (economic growth? improving livelihoods?), had repeated that the Center held a non-ideological, ecumenical view, with no preconceptions. When I asked him whether Evidence Action had a perspective on what were the most pressing poverty issues, Eshan answered in the negative, stating that the organization's cross-cutting theme was rather evidence.

At the same time, Eshan had also insisted, that the organization was debating whether to become sector-agnostic, leaving open the possibility that its staff might face in the near future the difficulty of choosing one issue over others. Eshan's words and the inclusion of "moral weights" as part of GiveWell's work sketch a third way. Pairing explicitly a filter of evidence with moral weights sketches a practice of poverty alleviation that is neither sector-specific *à la* Gates Foundation, by turning a few limited sectors into exclusive priorities, but nor is it totally sector-agnostic, by outsourcing value judgements and choice-making to others, whether these others are NGOs from which to demarcate one's approach or donors, as if those decisions were external to – rather than implicated in – one's work of evidence-based development.

One might argue, just like Idris, an economics professor I met at the University of California-Berkeley, that to interrupt this agnosticism is the role of policy-makers. "The idea is that I, as the academic, produce the evidence and policy-makers will take it up", Idris told me as we were drinking coffee on a terrace at a campus restaurant. Idris suggests a world with a clear division of labor between the entities that produce the evidence and those who use the evidence. There would be therefore a sort of gigantic database of evidence generated through randomized evaluations that policy-makers can consult to orient their decision-making, whether the policy-maker works in the education ministry of a country, is the director of a major international organization, or is a young entrepreneur launching their own nonprofit

development venture and deciding just which "sector" and approach to focus on. The reality of what I have encountered during fieldwork, however, is much less neat, blurring instead the boundaries between a strict figure of the scientist and that amorphous one of the policy-maker.

**Conclusion: ultimate ends, direct proof?**

This chapter examined a project of global poverty alleviation that characterizes itself by refuting to hold preconceptions about what constitute the most pressing development issues while taking them as central to its mission to influence policy-making. To an approach that understands development in terms of sectors (agriculture, health, education, etc.), my interlocutors propose an alternative practice of development centered around randomized evaluation, taken as a solution to finding out "what works" to solve poverty and guiding successful anti-poverty efforts around the world. In this chapter I argued that in this process the *better life*, in the form of a comparison between treatment and control groups at the heart of RCT design, becomes itself a measure of success to guide action, among many other factors guiding policy-making (e.g. funding, advocacy, political buy-in, chance encounters). This analysis draws attention to the ambiguity of poverty alleviation and development as projects. What are the priorities of poverty alleviation, and importantly for *whom*, are questions for which there are no clear answers. This ambiguity does not only exist in reports, scholarly work, development indexes, or development agendas, such as the Sustainable Development Goals; it is also palpable in the work of those who are embedded in development practice on a daily basis.

Earlier scholars have underscored the unfeasibility of finding "optimal solutions" to social problems, given that "there's no theory that tells us how out what might be considered

a societally best-state" (Rittel and Webber 1973: 168). In more evocative language when one thinks about randomized experiments, political economist John Stuart Mill (2009 [1879]: 10), in his famous treatise on *Utilitarianism*, had eloquently stated: "Questions of ultimate ends are not amenable to direct proof. Whatever can be proved to be good, must be so by being shown to be a means to something admitted to be good without proof. The medical art is proved to be good, by its conducing to health; but how is it possible to prove that health is good?". There is unavoidable uncertainty about what the good (development) is in the same way that there is no empirical way to ascertain what (sectors) matters most.

It is in this context that RCTs provide us with a limit case to think about the limits and anxieties inherent in a project to better the world. Not only just *how* to do it, but *what* exactly is this better world? What is not measured is not only what *cannot* be measured given evaluation methodologies (i.e. the effect of a hydropower plant), but also what cannot be *known* (i.e. the final parameters of a good, worthy life), at least in a human condition, haunted by secularism, empiricism, technicism, globalization, and cosmopolitanism. "How to be normative", asked anthropologists Marianne de Laet and Annemarie Mol (2000: 253) in a well-known article on Zimbabwean water infrastructure, "if there is no single, self-evident standpoint to speak from?" Can we read into evidence-based development's promise of causal certainty, and the *better life* implicitly sketched by RCT design and their incarnated counterfactuals, a certain recognition of the failure to ever close off the issue of the normative?

# Chapter 2 - Of energies and sweat: infrastructure of efforts, reaching millions, and safe water in Kenya

## Reaching human beings

This morning, Oscar is in Arala, a village in western Kenya reached through a bumpy dirt road on which the dust turns the adjacent forest trees brown. The local administration stakeholder meeting has gathered chiefs and assistant chiefs, representatives of the national government at the village level in Kenya. Oscar, in his capacity of head field officer for the NGO Evidence Action, is presenting the sanitation program that his organization is implementing in the district: the Dispensers for Safe Water program. Oscar has explained how the goal of the program is to provide households with access to chlorine dispensers, which can be used to disinfect drinking water upon collection at a water source.

"When households are handed chlorine bottles, they don't use them.", Oscar explains, implicitly referencing some recent research on domestic water access[1]. "So that's why we put dispensers close to the water source to act as reminder to disinfect your water".

Oscar has also shown the audience, using a real dispenser to illustrate his words, how the dispenser works: if you turn open the valve, it will release 3 mL of chlorine to treat 20 L of water (see Figure 8 below).

---

[1] See Ahuja, et al. (2010), available at
https://www.annualreviews.org/doi/full/10.1146/annurev.resource.012809.103919, accessed August 5, 2019

**Figure 8: Chlorine dispenser and jerry can containers**
**© Evidence Action**

Now he is addressing questions from the audience that mostly deal with technical, yet important, details. "How long can you keep the chlorinated water before it risks being contaminated again?"[2] What if you have only smaller jerrycans, let's say 10 liters, especially if a child must fetch the water?"[3] "What are the dangers of overdosing on chlorine?"[4]

"The next question we have received is", Oscar summarizes, "in a given area, where to station the chlorine dispenser if the water sources are far from each other?"

Given its finite resources, human and financial, the Dispensers for Safe Water program cannot install chlorine dispensers at every water point of the regions in which it operates. The question is therefore how to choose, within a designated area, where to install the available dispenser. Oscar proceeds to clarify that the staff at Evidence Action tries to choose the water point with the most people living in its vicinity. Once it has identified one, the organization

---

[2] Answer: *It depends on the vessel you use. Upon treatment, the water stays safe for three days in a jerrycan because chlorine evaporates. In a clay pot, however, the water should be changed after 24 hours because it is more exposed.*

[3] Answer: *One can fetch 10 liters, put the chlorine dosage, then fetch another 10 liters, and mix them together to obtain 20 liters. For a 5 liters jerrycan, this would mean fetching water 4 times.*

[4] Answer: *Chlorine is bitter; has a tough smell; it can provoke stomach discomfort and rambling, but it won't cause death. It's advised to wait 30 minutes between the moments the water is chlorinated and its consumption.*

does a verification to confirm the water point is viable (e.g. it is not dry) and is indeed being used.

"We give a minimum number of people that should be associated with the water point but no maximum.", Oscar insists. In summer 2016, at the time of this meeting, the number of people associated with one dispenser was on average 123. "Well", he states as a sort of overall conclusion, "our intention is to reach as many human beings as possible". At that point, in the first half of 2016, 2.3 million people were said to be served by 18,500 chlorine dispensers installed on the ground in Kenya[5].

**Reaching and repairing dispensers**

I realize how hard it is to reach people when, two days later, I accompany some employees of Evidence Action to repair two chlorine dispensers. The trunk is packed with concrete bags, black rubber poles on which to fix a chlorine dispenser, the blue plastic dispenser casing, and boxes containing chlorine. On usual days, James carries all this heavy material on his little motorbike down uneven and at times very slippery dirt roads. But today, as Oscar and I are accompanying him on his work routine, we travel by car.

We are leaving from Kakamega town, where the main Evidence Action office for that county is located. We are headed toward Ninoto, about forty minutes north-west, to reinstall a dispenser damaged by a cow. James and I share the rear seat. The driver Jacob drives the rented car. Next to him sits Oscar. Five minutes into our drive, we stop at James's house to collect some tools he had left there and that he will need to fix the dispenser.

James is a middle-aged and affable man whose personality is soft-hearted as well. A plumber by profession, he currently works as field assistant on the Kenya Dispensers for Safe Water program led by Evidence Action. In this role, James oversees that dispensers, within his

---

[5] If you multiply 18,500 by 123 (the average number of people served by 1 dispenser), you obtain 2.3 million.

area of responsibility, provide a constant supply of chlorine. This means that he is in charge of delivering new supplies of chlorine at waterpoints, but also of checking that dispensers are functional and repairing them when needed.

Before reaching Ninoto, we make another stop in Lilu market town, where James has his office between a small supermarket and temporary stalls of cereals installed by women on the sidewalk. There is a desk in the office, but it serves primarily as storage to keep log sheets, dispenser parts, and at least 500 white jerrycans containing chlorine. We pick-up a bag of cement that must weigh more than thirty kilograms.

The car comes to a stop in a circle of mud houses. A woman wearing a bright pink scarf on her head and an older-looking man greet us as they lead us toward the waterpoint where the broken chlorine dispenser is stationed. They are the "promoters", volunteers elected by the village to be responsible for adding chlorine before it runs out and for reminding people how to use the dispenser. Two men help James carry down the material to fix the chlorine dispenser, including the cement and the concrete. The path goes abruptly downhill, on very muddy terrain (see Figure 9). I try hard not to fall while I look at the men carrying very heavy loads on their backs. Behind me, women follow with young babies in their arms, not in the least seeming to lose their balance. I imagine people, sometimes even young children as I have seen them at other water sources, coming uphill with twenty liters of water single-handedly. I had learned from Oscar the previous day that this characterizes the Kakamega region – the fact that water points are downhill, often at the end of very steep slopes, and when it rains, these get extremely slushy.

**Figure 9: On the way to Ninoto's dispenser (Photo by Fiona Gedeon Achi)**

James must dig a round, narrow, yet deep hole in the uneven ground to stick in the rubber black pole that is to hold the dispenser casing. He mixes sand, concrete, and cement on the edge of the stream and fills the pole with mortar to fix the dispenser firmly in place (Figure 10). It takes James forty minutes to complete this physically toiling work, at the end of which his entire face is covered with sweat drops. He gladly welcomes the water bottle we hand him.



**Figure 10: Installing a new dispenser in Ninoto (Photo by Fiona Gedeon Achi)**

This day of my fieldwork in Kenya alludes to the difficult journeys undertaken by people to access dispensers, whether it is to repair them, monitor their workings, or for households to get chlorine. Here, reach is a substantial achievement, one in which minds, bodies, sweat, car, dirt, water, chlorine dispensers, and slippery paths mingle. Maintenance, planning and repair work figure centrally in this story to make sure that chlorine dispensers are well-functioning and do not run out of chlorine.

At the stakeholder meeting in Arala, when Oscar talked about "reaching as many human beings as possible", he was describing the willingness to make the Dispensers for Safe Water program accessible to the greatest number. By detailing this episode to repair a chlorine dispenser, I aim to draw attention to the chains of people and their energies that must be conjured to sustain the chlorine dispensers for just a few dozen people in the villages of Ninoto and Atani which I visited. Taken side by side, these two vignettes show us a few of the many labors that are needed to actualize Oscar's aspiration. They underscore that program success does not end only by installing 18,5000 dispensers across the region. Rather, reaching more than two million people involves at its core repeated acts of maintenance, such as travelling to ensure chlorine supply, accessing waterpoints to repair broken dispensers, training households about water sanitation, and re-designing parts of the dispensers to avoid deterioration.

This chapter attends to the kinds of work ensure the continuity of the Dispensers for Safe Water program (DSW) in western Kenya, including monitoring, training, repair, modification, and innovation to address daily occurrences, such as when the movements of a cow destroy a dispenser or, as will be discussed below, when people break dispensers on purpose. Building up on the anthropology of global health and development and the recent work in infrastructure studies, this chapter contributes to broader reflections around the forms, logics, and ethics of welfare provision, particularly where basic services are not provided

directly by the state but instead by a vast range of international and non-governmental actors (Ferguson 2010; Foucault 2004; Redfield 2012; Rees 2014a).

**Neither piped, nor light infrastructure**

By focusing on the mundane yet repeated work enrolled in both expanding to millions of people and sustaining through time a water sanitation initiative, I wish to make three points.

First, I want to trouble the binary divisions between technical fixes versus adequate social or political solutions on the one hand, and between outside imposition versus local agency on the other hand, that are commonplace in analyses about global health and international development efforts. Some anthropologists insist on the limits and dangers of these kinds of interventions that provide technical, circumscribed fixes to problems of a structural nature, which really should be addressed by developing large-scale, state-regulated public infrastructures (Biesel 2014; Pfeiffer 2013, 2014). These logics are thus reckoned to give rise to prepackaged solutions replicated around the world without much regard for cultural, geographical, or political particularities. In this perspective, the current rationalities of global health and development, often deemed as neoliberal (and post-colonial), are eroding more complex ways for understanding and providing for people's well-being.

The Dispensers for Safe Water program, ran by the non-governmental organization Evidence Action and funded in large part by the United States Agency for International Development, as attested by the USAID logo that adorns every chlorine dispenser, is evidently not a state-managed, centrally-planned, networked infrastructure of water provision, operating through pipes, pumps, and valves. It is rather an off-grid infrastructure, reliant on a simple and light technology, in the form of a blue plastic container installed independently at tens of thousands of water points throughout Kenya, but also as we see briefly later in Malawi and Uganda. Put differently, this sanitation program can ignite long-held and valid anxieties about

both the efficacy and concrete consequences of the international development apparatus. It can be read as a proposal to solve unequal access to reliable drinking water, a politically-fraught issue, with a standardized technical solution for disinfecting water (Escobar 1994; Ferguson 1990; Li 2007).

At the same time however, this development intervention requires constant and hard work to reach, maintain, and operate the infrastructure made of chlorine dispensers, but also of people, their endurance and choices, of bumpy roads and muddy ways, of motorbikes and legs for walking, of regional and sub-regional offices to stock chlorine refills, and of the mental work to learn about dispenser breakdowns and anticipate lack of chlorine. Whereas I insist on the hardworking nature of the work of replacing a chlorine dispenser, in their analysis of the Zimbabwean water Bush Pump, Marianne de Laet and Annemarie Mol (2000: 240) have emphasized the ease of fixing the pump, because the spare parts are locally available, because the repair only requires simple tools and few working hands. Both de Laet and Mol, and Peter Redfield (2016), in their analysis respectively of the Zimbabwean water Bush Pump and of personal water filtration systems, have foregrounded the mobility and flexibility of these water technologies in contrast to heavier, less portable water provision solutions such as piped systems.

Yet, even de Laet and Mol (2000: 245-246) discuss how the maintenance of the pump, its working condition and its cleanliness, depends on engagement on the part of the community in which it is installed, or else it falls into dereliction. While the chlorine dispensers for safe water share features among the technologies discussed by de Laet, Mol, and Redfield, in terms of it being a simple technology composed only of a limited number of components that do not require specialized knowledge to handle,  I nevertheless wish to make visible how even such a simple technology requires long journeys and sustained energies to remain operational. To draw attention to the long chains of people, things, labors, and relations that must be assembled

to bring something as basic as disinfected water to millions of people, I conceptualize the Dispensers for Safe Water program as an *infrastructure of efforts*[6].

By choosing the term of *infrastructure*, I follow historian William Rankin (2009) who has argued that the concept of infrastructure precisely emerged along debates about fostering economic development in postcolonial countries following the Second World War. Rankin writes that "infrastructure [became] a label for the technical-political systems required for growth and modernity", such as roads or electrical power, and thus embody (our own) ideas about development and sociopolitical modernity (2009: 61). In a recent journal issue about development in Latin America, anthropologists Kregg Hetherington and Jeremy M. Campbell (2014: 193) have similarly underlined how infrastructures are always about the future, as infrastructural projects are entangled with promise, hope, fear, possibility. Studying infrastructure therefore is also about examining orientations toward the future and how it is enacted.

In pairing *infrastructure* with the term *effort*, I take cues from scholars studying infrastructural realities in the global south, both state-provided and "off-grid", who show how infrastructures are centrally constituted of *relations*. This means not only formal ones regulated by contracts and legal transactions between users, workers, and providers of infrastructure, but also polymorphous relations that rely on persuasion, trust, motivation, good will, or patronage (Anand 2011; Harvey and Knox 2016; Howe, et al. 2016; Jensen 2017; Simone 2004; Street 2014b; von Schnitzler 2013). This attention to "infrastructure as relations" at the same time spotlights how access to basic services is predicated on a repeated work of self-engagement, monitoring, repair, and improvement.

---

[6] I owe this insightful formulation to my colleagues Nassima Abdelghafour and Adam Fleischmann. In her current doctoral work about evidence-based development, Abdelghabour (n.d.) examines the "living infrastructure" needed to turn villages into experimental sites for knowing what works for poverty alleviation.

With this analytical focus, I aim to depict global health and development programs not foremost as the product of replicated technologies (whether chlorine dispensers, solar lamps, microcredit programs, or vaccines) but significantly as heavy, thick, and even sticky infrastructures, embedding people's energies as well as their sweat and dirt, hopes and disappointments. In this reading, the aspiration to reach millions of people with one's program is made possible by these *infrastructures of efforts*, which render more ambiguous the distinction between what is technical and political, and between what is local (and pertains to the local community) and what is global (and not local).

Second, I wish to illuminate how ambitions to scale must reckon with changing topographies, different sensibilities and habits, and unexpected events. I suggest that the concept of *infrastructure of efforts* allows a specific reading about ambitions to scale at the heart of most projects to provide services to vast populations. Put differently, I underscore that the work of maintenance and repair at the center of the DSW program is never routinized (Orr 1999; Street 2014b). On the contrary, it necessitates adaptation, tinkering, re-design, and innovation as regional managers, field assistants, and technicians regularly face changing circumstances, unknown issues, and search for resolutions. How does one examine ambitions to scale, I ask, while witnessing a chlorine dispenser repair by a river in Atani village? What can slippery paths, broken dispensers, plastic casing melting, and absent volunteers illuminate about global ambitions? Reaching millions of people appears in this light as the product of collective involvement and achievement, which is always iterative, uncertain, and unfinished.

Taking erosion, breakdowns, and repairs as the starting points to understand this water sanitation initiative in Kenya forces an analytical shift away from an anthropology of "scale" toward an anthropology of "reach" (Jackson 2014; Jackson, et al. 2012; Moore 2008)[7]. In my fieldwork, my interlocutors at Evidence Action used the terms of scale and reach

---

[7] I thank Tobias Rees for helping me articulate this conceptual difference.

interchangeably. Empirically, their ambitions to reach millions of people with chlorine dispensers or scale-up the DSW program amounted to the same reality: that of their daily work as team members of the NGO. In contrast, this chapter argues that the concept of reach offers a distinct perspective on projects that seek to expand across people and places. The concept of *infrastructure of efforts* conveys efforts to *reach* people (efforts which must be constantly renewed) rather than the attempt to scale *a* program. Here, the vertical and top-down metaphor of *scale* is supplemented with a more multidirectional and collaborative metaphor of *reach*. The reference to collaboration is not to index agreement and success but instead to emphasize complex chains of labors that can result in success or failure, that can give rise to projects we value or condemn[8]. This is the challenge for understanding replication, not first and foremost as the inclination to expand without obstacles (Tsing 2012, 2015), but also in other terms, for instance in my fieldwork, as the yearning to reach as many human beings as possible.

In examining how programs expand and reach people, I build up on recent scholarship that underscores the importance to look at projects with global ambitions, not purely as homogenizing and standardizing, but also as processes that explicitly contend with unanticipated events and accommodate a diversity of motivations and outcomes (Bemme 2018; Carr and Fisher 2016; Carr and Lempert 2016; Ehrenstein and Neyland 2018; Kelty 2017; Moore 2008; Strathern 1995b, 2004). If we focus only on the dangers and detrimental consequences in past and current projects that aspire to scale, we risk forgetting that there are many *reasons* to scale. Scaling to provide drinking water to millions of people is a project that in many ways, I contend, is incommensurable with scaling for capitalistic profit-seeking, such as in the case of multinational corporations.

---

[8] Steven Jackson, scholar of media technology and infrastructure, reminds us not to "heroize" repair, which is not always directed toward "noble ends" and may function as much toward democratic ends and anti-humanist ones. To illustrate his point, he cites the recuperative strategies by which the Nazi war atrocities were sustained (Jackson 2014: 10).

Third and finally, this chapter invites careful consideration of sameness and repetition by weaving together both aspirations and efforts to reach millions of people as I encountered them in my fieldwork. As anthropologists E. Summerson Carr and Michael Lempert (2016: 19) write in their edited volume on scale, "there is a tendency to discern something dehumanizing—even violent—about scale, perhaps because of its association with measurement and ordination, on the one hand, and vertical power arrangements, on the other." In anthropological scholarship, scaling, expanding, replicating often denotes hierarchies, domination, homogenization, and standardization. This chapter seeks to foreground how sameness and repetition can also be life-enabling and generative, even a privilege when it takes the form of access to basic services at scale. Central to this line of questioning is to open-up further explorations of the complicated relationships between the universal and the particular, the global and the local, and hence about the value of different forms of welfare provision in today's world.

While I focus in this chapter on the Dispensers for Safe Water program, it is important to note that it is not the only development program of its kind in East Africa. Rather, the DSW is part of a larger initiative, led by an ecosystem of non-governmental research and policy institutes, to ground anti-poverty policy around the world in "scientific evidence". Researchers and partners identify effective solutions to various poverty problems (health, lack of education, or low income) through field-based, statistical evaluations of such programs, often at small, pilot scale. Programs identified as effective and worthwhile are then "scaled up" to target millions of people. This initiative within global poverty alleviation efforts, with roots in the early 2000s, is often referred to as "evidence-based development"[9]. It polarizes opinions in development circles, from those who take it as a welcomed solution to many decades of

---

[9] For further details about the evidence used to derive policy implications, see Chapter 1. Central to these reviews and decision-making is the randomized evaluation. The statistical assumptions on which the methodology is based, as well as how it is made into a herald of rigor and arbiter of value by its proponents, are discussed in Chapter 1 of this dissertation.

uncertainty around the usefulness of foreign aid to those who see it as a reductive epistemological and pragmatic practice likely to hinder rather than enhance anti-poverty efforts around the world, due to its simplistic statistical and anthropological assumptions (see Chapter 1).

In this chapter, I sidestep discussions around the direct implications of evidence-based development for knowing and managing poverty around the world – questions that are taken up in different forms in Chapter 1, 3, and 4 of this dissertation. In this chapter, I focus instead on the *implementation* side of the DSW water sanitation program to speak more broadly about the forms of welfare expectations that emerge when getting or giving access to such things as water, electricity, and health for populations that, for diverse and intricate reasons, are "beyond the reach of state infrastructures for living" (Redfield 2012: 178). As anthropologists have documented, these "substitute" infrastructures and technologies include single-disease interventions (vaccine, diagnostic tool, or treatment), microcredit or cash transfer programs, nutrition supplements, solar lamp products, or prepaid meters to access water or electricity (e.g. Elyachar 2010; Ferguson 2015; Karim 2011; von Schnitzler 2013). Scholars have been divided on how to conceptualize these attempts on the part of international organizations, large philanthropies, smaller NGOs, but also private corporations to fill in the "gaps", through technologies aimed at narrowed goals and limited timelines, often at the individual or the household level, such as the DSW program with which this chapter is concerned.

While they do not disagree with an "infrastructure ideal" of universal government infrastructure provision, several authors offer a complex portrayal of this non-governmental management and funding for "the social" and for basic services (Collier, et al. 2017). They point to the widespread failures of states to provide for their citizens, sometimes because of lack of resources and financial capacity, but in other instances also because of corrupt and nepotistic governments, more interested in engaging in graft and grabbing and fomenting ethnic

conflicts than extending physical and social infrastructures into the country (Ferguson 2010; Rees 2014a). Some of these scholars also question the desirability of "life on the grid" from an environmental and ecological perspective, suggesting that state-based, central, networked infrastructures are very energy-intensive and thus potentially unsustainable, in addition to being prone to pirating or hijacking (Gupta 2015; Street 2014b).

This view also underscores that these extra-governmental and restricted initiatives are primarily, and explicitly, concerned with "meeting the most urgent needs of poor people in extreme circumstances", and not generally claiming to make-up for governmental services (Redfield 2012: 158). These are (poor) people who might not be able to wait for the many years of political negotiation and construction required for a network of water pipes to be built throughout western Kenya, and whose needs therefore might be calling for other kinds of political imaginations. This conceptual lens takes seriously the moral impetus that underlies such "beyond the state" efforts, while also underscoring that they combine with other interests such as corporate investments or scientific research, showing the need for careful, and even perhaps adaptive, conclusions.

In line with this approach, the remainder of this chapter brings together the opening vignettes mentioned above along with further ethnographic details about the work of Evidence Action, the NGO that runs the DSW program, to explore what it means to aspire to reach millions with a non-state, off-grid water sanitation initiative in rural Kenya. After discussing Evidence Action's central and explicit concern with scaling-up programs, I provide a detailed examination of the work of maintenance, repair, and innovation carried by those actors involved in the DSW program to foreground it as an infrastructure of efforts. In the last section, I detail how a focus on infrastructure of efforts prompts an examination of ambitions to scale at the heart of international development and global health in terms of an anthropology of

*reach*, rather than the usual interest in *scale* as central category. I begin with a brief overview of access to water, in ways that are relevant to contemporary Kenya.

## Safe water provision in Kenya



**Figure 11: Fetching water in Atani (Photo by Fiona Gedeon Achi)**

At the end of the day of fieldwork with which I opened this chapter, as James had finished repairing another dispenser broken down by "drunk people" in Atani village, a little girl, likely eight years of age, came down the muddy track to fetch water from the stream, a 5-liter jerrycan in each of her hands (Figure 11). Thinking aloud, I emitted awe at the strength of this little girl. James turned to me and asked: are there many people in Europe who have water taps? That day, I was left wondering how to answer that it is rather the exception that people *do not* have taps. If in my daily life I had almost always taken the existence of a water tap for granted, in that moment, through my conversation with James, that banal object – the most apparent sign of a vast system that manages to bring running water into people's homes – emerged as an exception and the sign of privilege.

Of course, a vast literature has underscored how infrastructure tends to become invisible to its users, becoming apparent and obstructive only upon breakdown (Edwards 2017; Star

1999). These authors importantly underscore that this "invisibility of infrastructure" is situated, often the result of living in urban and wealthy environments. It is only when centrally-operated, grid or network-like infrastructure exists, and that it works well, that it can become invisible to its users who rely on it for their daily life, without much forethought. Given this, the anthropologist Brian Larkin has insisted that we talk of a "range of invisibilities" because "what is background for one person is daily object of concern for another" (2013: 336).

Many residing in large European or North American cities do not think much about the physical network of pipes and valves required to move around water, nor about the ecosystem of planners and workers that is in place to operate it. But for those who cannot turn the tap to get running water, or for those who have only intermittent service, getting water to drink – but also to wash oneself, one's clothes, one's house – becomes the result of some work of planning, and often a very physically demanding and time-consuming task predominantly carried out by young women (Crow and Odaba 2010; Koolwal and Van de Walle 2013; UN Women 2018; United Nations 2015).

In their history of water supply and sanitation in Kenya, civil engineer Ezekiel Nyangeri Nyanchaga and historian Kenneth S. Ombongi (2007) argue that the development of piped water in the country followed patterns of colonial development and the exigencies of the colonial state. In their view, the late nineteenth century construction of the British-controlled Uganda Railway, linking the interior of Uganda to the coastal town of Mombasa, was the first project for which piped water supply was used. The first piped water systems were therefore built and managed by the Railway company and reached the bigger towns on the territory such as Mombasa, Nairobi, or Kisumu, in which the colonial administrators resided. Until then, the authors write, "no attempts were made to bring water to the people; rather it was the person who went after the water" (Nyanchaga and Ombongi 2007: 273). They describe how people went to collect water from springs or rivers and transported them back to their homes in all

kind of receptacles ranging from "clay pots to buckets made from hides" and how, in arid regions, people would have to frequently move in search of water sources (ibid).

After Kenya's independence in 1963, government discussions and documentation posited that water would become a "social service" that the government was responsible for providing to its citizens, either free of charge or at a subsidized low cost. With the aim of reducing the incidence of diseases and increasing health outcomes, "effort was directed towards establishing water supply facilities so as to reach as many people as possible", but without much regard for the sustainability and maintenance of the facilities being put in place (Nyanchaga and Ombongi 2007: 283). As public institutions have been unable to provide widespread and reliable access to running (and drinking) water, a vast portion of the urban population relies today on the private sector to obtain this service, whereby water emerges as a commercial good. Slum dwellers are most affected by this situation, as they become subjected to informal water vendors to get their daily needs of water. Often dubbed by NGO reports and newspapers as the "water mafia", these informal companies often control access to the few water pipes present in the neighborhood, delivering water of changing quality at high price to women who have queued for hours[10].

In fact, access to water and water quality, while intimately linked, are not synonymous. Under a shaded tree along a stream, while attending a community training for the DSW program, I learned that clean water (*maji safi* in Kiswahili) is not necessarily safe water (*maji salama*). The former simply means transparent water, which seems clear to the eyes, while the second term refers to water that has been treated for drinking consumption. There are many ways in which water points can be contaminated: detergent residues due to washing clothes, rain that bring contaminants, animals drinking from the source, or open defecation. The World Health Organization reports that worldwide at least 2 billion people use drinking water sources

---

[10] See http://news.trust.org//item/20141126085900-3p4ar, accessed August 10, 2019

that are regularly contaminated and estimates that, globally, contaminated water causes about 500,000 to die of diarrheal disease each year[11].

While the recent 2010 Constitution of Kenya recognizes access to safe water as a basic human right, in practice this remains very limited. United Nations figures suggest that, in 2014, 38% of the Kenyan population was not served by improved drinking-water access, namely sources that are protected from outside contamination. Of these unserved populations, about 90% live in rural areas, indicating that most investments in water infrastructure have taken place in urban or peri-urban areas (World Health Organization 2015). In Kenya alone, the issue of lack of safe water access affects about 18 million people. Given this, the NGO Evidence Action sees its Dispensers for Safe Water program as an effective alternative to provide disinfected water to millions of households, in two main ways (Ahuja, et al. 2010).

First, Evidence Action aims to close what it sees as a gap left open by more mainstream sanitation programs. The latter have focused on "access to water", by building wells for instance, "but that is not enough", states the website of the organization. "Water may be relatively safe to drink at the well but does not stay safe in transit or when it is stored at homes."[12] In contrast, chlorinating water upon collection keeps it safe for at least a day (depending on the container used to store it). As Oscar explained in the vignette above, chlorine dispensers located close to water sources act as reminders for people to disinfect their water each time they pick some.

Second, the DSW program is predicated on making access to chlorine free of charge in the regions where it works. When I first met him upon my arrival in Nairobi in February 2016, Leo, then Deputy Director of Global Safe Water at Evidence Action, had made a reference to a study titled *The Price is Wrong*, a 2011 review of ten evaluations studying how price changes

---

[11] https://www.who.int/news-room/fact-sheets/detail/drinking-water, accessed January 21, 2019

[12] https://www.evidenceaction.org/dispensersforsafewater/#the-problem, accessed January 23, 2019

affect usage of health and education products among poor households (Abdul Latif Jameel Poverty Action Lab 2011). "When they have to pay for chlorine", Leo insisted, "the take-up is only 7-10% but when given for free, it's over 50%". Started in Kenya in 2012, the DSW is currently said to "reach 4 million" across Kenya, Malawi, and Uganda[13].

I avoid discussing here the choice of chlorinating water over other options for disinfection. It is fair to say that most water treatment processes, including in big cities in so-called developed countries, involve chlorination at some point in the chain, especially to avoid that water is contaminated during its transportation along the system of pipes. In addition, Evidence Action states that chlorine dispensers avert carbon emissions because people do not need to boil water to disinfect it[14]. When giving training to villages or government officials, the DSW team would ask people about the different methods to treat water. Examples cited included personal water filters, commercial water treatment options, but also boiling. The field officer giving the training would emphasize the risk of recontamination after boiling, unless the water is be stored in a closed plastic jerrycan.

One could also argue that different methods spread differently the burden of water sanitation. While a public municipal water system and its maintenance falls within the jurisdiction of the state, boiling water to sterilize it is entirely the responsibility of the household and must be done at home. In contrast, by acting as reminders at water points, chlorine dispensers might relieve some of that mental burden of remembering, as well as reduce the time spent boiling water and the financial cost of purchasing fuel (such as wood) used to boil the water. As to the scientific research concerning the virtue of boiling as opposed to chlorinating to kill microorganisms, the evidence remains mixed. Regularly, the DSW team

---

[13] https://www.evidenceaction.org/blog-full/evolution-of-dispensers, accessed January 21, 2019
[14] In his article on the LifeStraw, a water filtration device to purify water sold by a corporation to meet the needs for safe water in poor contexts, Peter Redfield (2016: 168) explains that the corporation used a similar rhetoric to advocate for its product, stating that by purifying water instead of boiling it, "the decline in wood burning would reduce local carbon production".

faces questions and challenges regarding the safety and trustworthiness of chlorine, such as when Oscar was asked by a government official during the meeting with which I open this chapter: "What are the dangers of overdosing on chlorine?"

As Nikhil Anand beautifully writes in "Hydraulic Publics", water infrastructures take form around "publics for whom water is a matter of life, and a matter *for* life" (2016). Water is a substance that knows no substitute and on which all human life depends (Crow and Odaba 2010; Fennell 2016). It might therefore strike as an odd choice for studying ambitions to scale as no one would disagree with the value of bringing water, in replicable form, to millions of people. One could argue, borrowing de Laet and Mol's formulation (2000), that some methods of water provisions might prove "easier to love", such as the Zimbabwean Bush Pump they analyze because as it delivers clean water to the most remote corners of the country, it also helps strengthen communities and build the nation, all the while being the product of a modest inventor who never aimed to patent it, thus allowing the Bush Pump to spread as a public, government-sponsored technology. "Uncomfortably close to [a technical fix], and yet also untouchably close to life", the DSW program proves hard, to use anthropologist Peter Redfield's words, "to [fully] embrace or [simply] reject" (2016: 160). It is precisely because water is a "limit case" that it is particularly interesting to be reckoned with. It illuminates that even a need as fundamental as access to safe water depends, as has been described in the case of the DSW program, on solid chains between people, terrains, things, work, and energies.

Informed by scholars who have studied water provision in various forms and places, this chapter underscores all the labors needed to make available something as basic and fundamental as safe drinking water, whether in its piped, running version, or at individual water points throughout a region (Anand 2011; de Laet and Mol 2000; Fennell 2016; Kaplan 2011; Redfield 2016). It is therefore my contention that this specific case can offer insights into how to analyze aspirations to expand embedded in other aims and motivations, whether or not they

are readily likable and worthy of support. The NGO Evidence Action's interest in scale in fact exceeds the DSW program and is a defining feature of their organizational mission. More broadly, "scaling up" programs evaluated and found to be effective is a central concern for all research and policy institutions involved in evidence-based development[15]. The next section turns to examining how team members of Evidence Action understand their work as having an impact at scale.

## The million scale

Asides from Oscar, many others during my fieldwork have emphasized a care for extending Evidence Action's programs to millions of people. Some have even informed me that they chose to work for Evidence Action because of the ambitious targets laid out in the job announcements they responded to. "What joy knowing that what I do today is going to impact lives, and a lot of them. 5 million children sounded like a challenge", Siya tells me she said during her hiring interview. Siya is a manager of the Kenya Deworm the World team, another program supported by Evidence Action that provides deworming treatment to treat parasitic infections in schools across Kenya. Siya's colleague, Sheila, also cited the scale of the deworming program as one of the reasons that made her want to join its team as director for Kenya[16].

The NGO Evidence Action was launched in 2013 specifically for the purpose of identifying and supporting programs that can operate "at scale". Grounding this is the idea that an anti-poverty solution "that works" in a study setting is not necessarily going to stay effective

---

[15]For more information, see the Jameel Poverty Action Lab's recent release about "Pathways to Scale" (https://www.povertyactionlab.org/evidence-to-policy) and Innovation for Poverty Action's new strategic ambition up to 2025 (https://www.poverty-action.org/sites/default/files/publications/IPA-2025-Strategic-Ambition-Brochure-Final-Web.pdf), both accessed August 5, 2019.

[16] For more details about the Deworm the World program that involves collaboration with Kenyan ministries, see Chapter 3. I use it as a case study to examine the relationship with the state, and the state itself as a project, in visions and plans where non-state entities take responsibilities historically associated with the state and whose targets exceed a given national population.

when implemented at "large scale". Here, scale has to be at least 1 million people, pondered Iris, then Director of Global Communications and Advocacy at Evidence Action, during our interview. "It cannot be a hundred, thousands, not even a hundred thousand", she emphasized. "Over 1 million as a minimum viable scale". For each of their programs, success is measured in terms of individual livelihoods and expressed in millions of people reached, within and across national borders.

On the website of Evidence Action, one can read the organization "only invest(s) in delivering services that have the potential to reach millions of people"[17]. This criterion – called *scalability* by Evidence Action – is among the four guiding principles used by the organization to determine which programs to test, keep, take-up, or discard at each stage of its work, along with the counterfactual criterion and the cost-effectiveness criterion, both discussed in Chapter 1. For Evidence Action, scalability – the potential to reach millions – includes two considerations.

First, interventions should be scalable: those impossible to scale are not retained by the organization. "If, for any reason,", the organization's website states, "we find we (or other implementing partners) cannot deliver a particular intervention at scale, we exit it from our pipeline"[18]. Of course, what is possible or impossible is of a pragmatic nature, rather than a utopian one – reasons that likely have to do with financial or logistical capacity. As discussed in Chapter 1, cost-effectiveness is an important criterion in Evidence Action's strategy and might tip the balance toward projects that are more "minimalist" in nature and aim, targeting immediate needs, rather than more "holistic" programs aimed at restructuring entire educational systems. What most interests me here however is that the definition of scalability heralded by Evidence Action also stipulates that interventions should tackle a "real need faced

---

[17] https://www.evidenceaction.org/beta/#the-beta-pipeline, accessed January 22, 2019
[18] https://www.evidenceaction.org/beta#the-beta-pipeline, accessed June 6, 2019

by millions of people". Or, in the words of Valery, Director of Global Innovation and Beta at Evidence Action, "we're explicitly saying that we're interested in programs that improve lives of millions".

I am sitting on the patio of the Nairobi office of Evidence Action and Valery and I are talking over the phone about scale-ups and scale as central in the strategy of Evidence Action. "That doesn't mean that there aren't very important programs", Valery insists, "that affect lives of smaller number of people". She mentions the example of trachoma, a parasitic infection that can cause blindness and is categorized as a neglected tropical disease by the World Health Organization. "Trachoma is devastating", Valery asserts, "but it's a problem that affects a lower number of people, so it's not what we do. And there are many other issues like that"[19]. A program might be scalable in terms of its logistical components, or because it is highly cost-effective, but if it addresses a need that affects a number of people much lower in its scale of magnitude than the "million", it will not be the focus of Evidence Action.

Put differently, the "million scale" is fundamental to the work carried by Evidence Action in the following way. It is an end goal; reaching millions of people is a marker of success but also a prerequisite, or a motivation for their work. Interventions that do not tackle an issue said to affect millions of people worldwide are not taken into the organization's portfolio in the first place[20]. The million scale *as a prerequisite* is a collection of individuals in need, albeit conceptualized as such through a development apparatus, and that can thus be targeted by many different interventions (Gupta 2012). These are, for instance, the 18 million Kenyan residents unserved by adequate water infrastructure and who could be granted access to drinking water

---

[19] Valery was comparing trachoma to parasitic worm infections because of Evidence Action's Deworm the World initiative. Whether Valery provides an accurate description of trachoma and its prevalence is not what matters: she could have provided many other examples to illustrate her point.

[20] While my interlocutors never made reference to a specific school of thought to justify their orientation toward the million as "horizon" of scale, my colleagues Nicole Rigillo and Julianne Yip have noted that this orientation resembles key tenets of utilitarianism, most famously exposed by Jeremy Bentham (1970 [1789]) and John Stuart Mill (2009 [1879]), as a moral philosophy that seeks the greatest happiness (or good) for the greatest number.

through the DSW, but also other kinds of initiatives, whether state-run or not. The million scale *as marker of success* is a recipient of a specific program, in our case the DSW program. In my fieldwork therefore, I encountered *scalability* as an explicit actor's category. It is that aspiration to put in place a program capable of reaching one or more million people, for two main reasons: because it targets a need faced by millions and because it can be operationalized for millions of people.

Anna Tsing is widely reckoned as the scholar who introduced an explicit focus on scalability and ambitions to scale within anthropology. According to her, the term "scalability" has its origins in the business world (Tsing 2012: 508). "Scalability in business", Tsing writes, "is the ability of a firm to expand without changing the nature of what it does" (ibid). In her work, scalability is an analytical category she uses to make sense of varied yet interlinked human endeavors. The expectation of scaling up characterizes such enterprises and epochs as colonialism, capitalism, globalization, science, or still yet international development (Tsing 2012). "Like business", Tsing argues, "development was supposed to scale up. The World Bank only funded village projects if they were already scalable; that is, if they could be spread to other villages without changing project elements." (2012: 508).

In general, Tsing summarizes in her last book *The Mushroom at the End of the World*, the quality of scalability is "the ability to make projects extend without changing their framing assumptions" (2015: 38). This is not an ordinary feature of nature, Tsing insists, but on the contrary, one that requires a lot of work to enact. Tsing often returns to the example of colonial plantation and the work of isolation, self-containment, reproduction, and equivalence it involved: "exterminate local people and plants; prepare now-empty, unclaimed land; and bring in exotic and isolated labor and crops for production" (2015: 39). Tsing also talks of anthropologist Clifford Geertz who, when sponsored to study mechanisms of economic development in the 1950s, reported that Indonesian firms relied almost exclusively on personal

relations between buyers and traders, rather than replicable protocols and practices. This state of things, in Geertz's reporting (1963), hindered economic development given that navigating such relationships required time and energy at the expanse of fast firm expansion. In light of this, Tsing (2015: 38) argues that scalable projects allow for "smooth expansion" by "banishing meaningful diversity, that is diversity that might change things" and making place rather for standardization, regulation, and universalization.

In the last decade, various scholars from anthropology, sociology, and geography have explored how the empirical realities they study are nuancing, and perhaps challenging Tsing's view of scalability as inherently problematic and dangerous. Of course, Tsing's writings provide much needed cautionary tales against often taken-for-granted and accepted contemporary realities, such as inequality, violence, overconsumption, and dire labor conditions. However, Ehrenstein and Neyland point out, for instance, that while "for Tsing…it seems that scale work is doomed to fail because it ignores or erases the diversity of our world", some scalable projects and their expansion are defined by the necessity to be "alert to the possibilities of change" (2018: 76). Authors Carr and Lempert (2016: 19) insist that there might be things we would all readily value that are produced through scalar practices: morality, communication, or welfare provision. The point is not to draw an exhaustive list of scalable projects, and sort the good from the bad ones, for anyway what is valuable or not is a situated endeavor.

Instead, the intention of this chapter is to move away from solidified epistemological and political assumptions about scale and toward a careful analysis of the "scalar practices of actors" to understand how and why people and organizations seek to replicate and expand actions, programs, things, or sentiments (Moore 2008: 221; but see also Carr and Fisher 2016). To open up the concept of scale is also to question other entrenched hierarchical metaphors through which the social sciences understand the world (global/local, macro/micro,

individual/society), as if each element of the pair in these dichotomies were a self-contained and ontologically powerful domain. By recognizing that scalability is a process before it is an end result, this chapter suggests that the boundaries between what are global or local ambitions are actually blurry, perhaps even elusive.

Anna Tsing (2005) underscores that smooth expansion often remains at an imaginative level because universal projects often encounters *frictions* – forces, often local in nature, that resist their scaling ambitions. Thomas Yarrow (2006: 295) has written that while "Tsing is aware of the problems of an analysis that assumes the distinctiveness of the local and the global", her work at times seems to re-entrench the very distinction she wishes to put into question. With a shift to infrastructures of efforts and reach as analytics, described in detail in the following sections, I wish to offer another alternative to trouble, and perhaps do away with the global versus local dichotomy. Here issues or unexpected events in the DSW program are not elements to resist a global plan that seeks to dispense with diversity. Rather, unknowns and challenges are constitutive of the program, as the program changes while confronting and incorporating them, albeit not always in ways that would draw consensus. Specifically, this chapter emphasizes the contingency of the efforts at the heart of the Dispensers for Safe Water program to make chlorine dispensers accessible to millions of people. In thinking through the concepts of efforts and reach, it becomes more difficult, I hope, to disentangle between local and global sources, between sources of resistance and forms of power. Reach aims to muddy these conceptual boundaries by bringing everything in one frame of analysis.

In the next section, we return to western Kenya to follow how the DSW team tries to sustain the continuity of their water sanitation program, through training, monitoring, maintenance, repair, discussions, and innovation. I wish to underscore how these different forms of labors account for different topographies, different ways of working, and different sensibilities. I then conclude by proposing that scalar practices in projects to provide basic

services to populations unserved by state infrastructures might be better understood as practices of *reach*, to emphasize the generative and life-enabling potential of some aspirations to scale.

**Infrastructure of efforts**

I opened this chapter with a vignette that focused on the chains of people and energies that must be conjured to sustain the chlorine dispensers for just a few dozen people in the scattered hamlets of Ninoto and Atani. Whether some people break dispensers on purpose in sign of protest (as we will see below), or like in the two vignettes above, because of a sudden movement by a cow quenching their thirst and who accidentally tipped the light equipment over, or some inebriated people who walked into the dispenser casing without realizing it, the NGO is frequently confronted with broken chlorine dispensers in need of repair. I detailed this tale of damaged dispensers and their mending to foreground how reaching people through the DSW program, let alone millions of them, must be reckoned as a difficult achievement that includes toiling tasks such as walking down steep slippery paths, carrying bags of concrete and mixing cement, building holes in uneven ground, or single-handedly fetching over 10 liters of water.

After dispensers have been installed and training conducted by field officers from the regional office of Evidence Action, a volunteer "promoter" is designated to encourage people to use the dispenser, refill the dispenser with chlorine, and request further chlorine supplies. During the process in which a person is elected to become a promoter, I learned that the eligibility criteria for this position, in addition to owning a cell phone, knowing how to read and write, and using this particular water source, include the capacity to walk an extra mile to teach people about chlorine. This further underscores how physical activity figures centrally in the ability to keep the DSW program running as walking and moving around is what links people in sparsely populated settings, where water points are rather scarce. In the case of the

volunteer promoter, the execution of this tiring labor depends on their good will, as their role is not formalized by any legally binding obligation (Street 2014b).

In emphasizing relations and labors as constitutive of "off-grid" infrastructures that provide basic services, I draw from recent scholarship on infrastructural realities in the global south (Abdelghafour n.d.; Howe, et al. 2016; Jensen 2017; Simone 2004; Street 2014a; von Schnitzler 2013). In her research about the rollout of rapid diagnostic tests for malaria in Papua New Guinea, anthropologist Alice Street (2014b) illustrates how conflicts between youths in the region resulted in the interruption of the delivery of these drugs to the Begasin Health Centre, which is only reachable through a 16 km hike in the forest. In his work on Mumbai's water infrastructure, anthropologist Nikhil Anand (2011) shows how getting access to piped water requires constant care and efforts on the part of city residents, especially when pressuring city officials to recognize the importance and legitimacy of their claims. Overall, these authors emphasize how the working of such infrastructural grids far outgrow contractual and regulated relations, or even reproducible elements; rather, they constantly require negotiation and endurance.

I insist that reach is a substantial achievement to introduce the concept of *infrastructure of efforts,* whereby global health and development programs are not only constituted by replicated frames and technologies, but also, just as importantly, involve heavy, burdensome, exhausting, and difficult actions and processes. In this reading, the ambition to reach millions of people – or the feature of scalability – is only made possible by these infrastructures of efforts that centrally depend on repeated maintenance and repair. In the rest of this section, I wish to show that besides being tiring and hard, the kinds of work at the core of these *infrastructures of efforts* must constantly contend with variable circumstances and unexpected events that require updating the parameters of the program.

In July 2016, on the day of Eid, which celebrates the end of Ramadan, I went to walk in the Kakamega forest with some employees of Evidence Action. On our way back, Oscar insisted that we make a detour to see the "crying rock", a very tall boulder that can remind the viewer of a weeping person. Oscar knows that I have a particular affection for these boulders decorating the hills of the Kakamega landscape. To me, it seems as if they have been added there, almost as a magical landing, after shapes and colors were already given to this region. As we drive away after our visit to the crying rock, Oscar and his colleagues intensely discuss, as they put it, about a "vandalized" dispenser. I remember, like them, having seen pieces of dark blue plastic as we walked through the household courtyards on the way to the crying rock. "How come people vandalize chlorine dispensers?" I ask.

They tell me that some people believe that chlorine is a tool for family planning. "They think we want to control their fertility", Oscar specifies, as if worried I did not fully understand.

The following morning, I go to the main regional office of Evidence Action in Kakamega town to attend the weekly staff meeting. In the front yard of the small compound the motorbikes used by field officers to go to the field are parked. There are stocks of chlorine lining-up the first floor below the staircase and also hundreds of chlorine refills in a room upstairs arranged in well-ordered rows of eightpacks of jerrycans. The weekly staff meeting opens with discussions on the challenges faced by the team in the past week, including the issue of dispensers being vandalized. I learn that a few years back, the same worries circulated around mosquito nets when they were being rolled out by different organizations throughout the region. To those reporting rumors linking chlorine and family planning, some Evidence Action team members respond:

"Look this dispenser has been installed three years ago: haven't you had children since then?"

On that same Friday morning, the challenges being discussed also entailed: field officers finding chlorine in the promoter's house but not in the dispenser; a dispenser near a stream that is not used (should it be removed or relocated to make accessibility more convenient? Why do some water points have less users? Is it too often dry or flooded?); the possibility of having two volunteer promoters for each village and if that would be detrimental to the DSW program since both would rely on one another to carry out the tasks. One field assistant reported that in a village people preferred the option of one promoter. Oliver, in his capacity as head field officer, makes sure to underscore that there are differences between communities: some prefer one promoter, others, two. "A special case where people said one promoter is better cannot come to be blanketed for the whole of Kakamega", he reminds them.

The weekly meeting also serves as an opportunity to test and reinforce the team's knowledge of the program. Mark, the regional coordinator, asks Dilara, who is in charge of conducting surveys to monitor usage and perceptions of the DSW program: "What do you do if chlorine is missing from dispensers when you visit the water points?" Dilara answers that she will tell the promoter so he or she can add chlorine. Mark acquiesces Dilara's answers but also completes the information she provided in front of the attendee by stating that the community service assistant, as the employee of Evidence Action in charge of that district should also be informed so "he can keep an eye of that dispenser". Out loud, Mark also considers whether Dilara could add chlorine or replace parts of malfunctioning dispensers when doing the monitoring evaluation, undecided as to whether doing so would constitute a conflict of interest given that Dilara is only supposed to be an auditor of the DSW program.

This weekly staff meeting, asides from raising awareness about the persistent association between chlorine dispensers and fear of family planning, is an effort on the part of the team to tackle singular occurrences, whether about empty or unused dispensers, and debate about the proper ways in which to address them. In talking about humanitarian interventions in

the global south, Peter Redfield writes that one must be careful not to see them, drawing on Richard Rottenburg's formulation (2009b), as "therapeutic domination", because "humans and nonhumans rarely prove fully obedient in practice" (2016: 175). Instead, Redfield claims, adjustments and back and forth steps are the hallmark of contemporary forms of interventions which are no longer oriented toward distinct and smooth phases of planning and action (2016: 176). Similarly, witnessing how the Evidence team went through different challenges felt like listening to unfinished conversations, for instance around the question of whether one or two promoters is a more favorable option, precisely because a team member highlights that there is no generally applicable solution. Often therefore, decisions are not taken but rather questions are left hanging to be taken up in future reflections and subsequent meetings[21].

In addition to a consideration for changing circumstances and new issues, this vignette also foregrounds that the DSW program necessitates a significant knowledge of the region and its inherent variability. One needs to identify which water points are dry too many months in the year or not used, for whichever reason, to avoid mistakenly placing a dispenser there. Similarly, program presentations and trainings include appreciation of how households store water, some in clay pots, other in plastic jerrycans. In Arala, if you recall, Oscar had said to government officials that the amount of time the water stays safe upon depends on the vessel used to store water.

---

[21] This orientation toward planning and action as perpetually repeated and centrally involving redesign and adaption is also reminiscent of Esther Duflo's keynote address at the 2017 American Economic Association Annual Meeting, titled "The Economist as Plumber". In the talk, Duflo presented the contemporary role of the economist –especially the one involved with policy, development, and working with governments– as that of a plumber (rather than that of engineers or even physicist). Duflo was arguing that the economist's job is no longer to find out grand theories or give general prescriptions, but in the challenge to influence the real world, the economist must engage with details –including knowing how households in the global South get access to basic services and how "to tinker" and adapt programs so that they fit a particular set of setting and needs (see also Chapter 4 on how evidence-based development accounts for contexts and necessary adaptations to answer contextual demands). Duflo said: "the economist-plumber needs to persistently experiment and repeat the cycle of trying something out, observing, tinkering, trying again", referring not only to the evaluation of programs, but also the tinkering involved in implementing "effective" solutions (2017: 16).

"The water stays safe for three days in a jerrycan because chlorine evaporates. In a clay pot, however, the water should be changed after 24 hours because it is more exposed", he detailed to the assembly.

In light of this vignette, I wish to ask: what enables a program to reach millions of people? On the one hand, the DSW program embeds several reproducible and replicated elements. In Malawi, Uganda, and Kenya, the three countries where the program operates, one finds the same chlorine dispensers installed tens of thousands times; at all these water points, one can meet volunteer promoters in charge of filling the dispenser with chlorine when it runs out; one can attend the same community engagement meeting to explain how the dispensers works and choose the "village promoter"; one can spot a guy on a motorcycle delivering chlorine supplies at each water source.

And yet, the operational implementation of the DSW requires a consistent reassessment of at least some of the project's framing assumptions, to use Tsing's words (2015:38). It means having to account for promoters who do not fulfill their roles, to negotiate with people breaking dispensers intentionally or not, to update one's knowledge about the region's water resources, to face questions about the dangers of chlorination, and to determine collectively how to approach these arising situations. The labors at the core of the DSW program thus call for dealing with situations that escape the more routine and habitual maintenance work of walking slippery paths to monitor chlorine supplies, of carrying cements and building holes to repair dispensers, and of carrying heavy receptacles when fetching water. In this reading, thinking about *infrastructure of efforts* suggests that off-grid efforts to provide access to basic services embed not only heavy work but also ones that require energies and ability to attend to unknown issues and search for resolutions.

In a very different setting, the anthropologist Antina von Schnitzler (2013) also shows how tinkering and innovation are central to infrastructures providing basic services at scale.

Her fieldwork research in South Africa examines the pre-paid meter now widely used in poorer urban areas which automatically disconnects users from access to water or electricity in case of non-payment. As residents find ways to bypass the meter, by disabling it or at times even fully breaking it, city officials and engineers look for ways to make the technology "smarter" and more unbreakable. Preferred innovations, von Schnitzler specifically shows, are not those that offer "global solutions" embedding generalized ideas about a "third world" customer base, with predictable behavior patterns and a certain level of trustworthiness. Rather, South African utility delegates favor meter products whose design elements take into account the short yet intense history of prepaid meter protesting and tampering in the country (von Schnitzler 2013: 684-686)

Technical redesign features also in the DSW program. While chlorine dispensers were still regularly in need of repair, engineers and technicians redesigned this technology over the years to make it sturdier and more durable, especially to withstand wear and weather conditions. Originally in metal, the dispenser casing was redesigned in plastic to avoid rust. The first kind of plastic, however, would become deformed under the sun over time and so the latest version of the dispenser at the time of my fieldwork was made of a "better" plastic to avoid such an issue. Overall, this analysis offers conscious thoughts as well as adaptation and innovation as an alternative focus to replication and standardization in understanding international development and its ambitions of scalability (Jackson 2014). Reaching millions is characterized by reproducibility and routine, but also fundamentally by the need to anticipate, to bear with difference, and to revise one's way of doing things, and not only when things breakdown, but weekly –and constantly. When scaling up a program, one has to account for different topographies, different ways of working, different sensibilities, and even different institutional realities.

It is my first week in Kenya and I am just starting to learn about the program ran by Evidence Action. To clearly illustrate his explanation, Leo, then Director of the DSW program for Africa, draws a schematic map on the white board, made of one square surrounded by round dots.
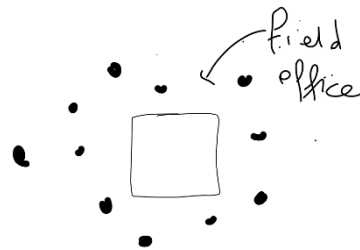


"The square, says Leo, "is a field office" (like the one I would later visit in Kakamega) "and the dots represent the villages and communities served by that field office". Leo adds that a person leaves from the office a motorcycle and then goes around the villages to deliver the chlorine. One of these guys on the motorcycle is James, the field employee I accompanied as he repaired dispensers.

"Is that guy on the motorcycle hired by Evidence Action?", I ask Leo.

In Kenya, yes, replies Leo. In the Malawi iteration of the DSW program, the guy on a motorcycle delivering chlorine is actually a government employee.

"In Malawi, it was great because the government was eager to partner with us. In that country, instead of field offices operated by Evidence Action, we rely on government house centers. Action delivers the chlorine there and then someone from the center delivers the chlorine to the communities.", Leo adds.
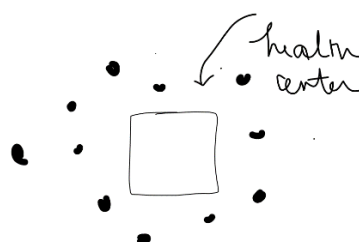
As he speaks, Leo writes *health center* next to the square on the map which was earlier a field office when he explained the implementation of the DSW program in Kenya. I remembered thinking, while listening to Leo, that it was interesting how a schematic drawing can serve to represent the space in different locations. That drawing, I contend, provides (quite literally) a good illustration of what happens when programs, at least the one I followed in fieldwork, scale. Some things stay the same – and the same map can be used to talk about the program in Kenya as well as in Malawi. Yet others change as planners on the ground reckon with changing circumstances for after all, Leo had to cross out the legend "field office" and replace it with "health centers".

In an effort to problematize the too often taken-for-granted antagonisms between global abstraction and local complexity, anthropologist Doerte Bemme (2018, 2019) underscores that those who scale up their programs within global health are acutely aware of the necessity to adapt their services or products to changing circumstances, or else they might simply 'not work'. In other words, scaling up actually mobilizes flexibility inasmuch as it produces uniformity. In so doing, Bemme provocatively suggests that contingency – and hence also diversity – might be features of the universal and the global, and not only of the local and of that which is not, or cannot, be scaled.

Along similar lines, Christopher Kelty (2017) discusses a form of "scalability" that is between "universalism and hyper-specificity" by examining the Participatory Development Toolkit, a briefcase filled with set images, cards, and games designed to draw people into discussing their experience and problems, and to offer possible solutions, in an attempt at an inclusive, bottom-up, participatory development. Kelty shows how understanding the kit exhaustively as a standardized device would be missing the point, because while the toolkit is a device for decontextualizing, "the tools are not automatic; a toolkit implies the existence of

a skilled tool user as well" (2017). "A toolkit", Kelty therefore concludes, "sits somewhere between an imagination of a context-specific, autonomous, and self-guided development without any facilitation on the one hand; and on the other, the large-scale, universal, automatic spread of one-size-fits-all solutions everywhere" (2017).

As Vera Ehreinstein and Daniel Neyland (2018: 77) argue when detailing a vaccine consortium in Africa, sometimes scaling up interventions is what produces the necessity to both attend to and incorporate diversity, as well as to modify one's program to be successful across distinct situations and people. My field research lent support to such a complex picture of scalability, one that weaves expansion with adaptability and reach with transformation. In presenting the DSW program as an *infrastructure of efforts*, reaching millions of people appears as the product of a collective achievement that is always iterative, uncertain, unfinished, and open to dissension. Taking these claims seriously forces, I contend, a conceptual shift away from doing an anthropology of scale toward ethnographically documenting *reach*, with important implications for how anthropology engages with projects with expansive and global ambitions, whether international development or national welfare initiatives.

**Anthropology of reach**

As category of analysis, reach enrolls a particular reading of projects that seek to expand across people and places. Reach is the result of infrastructures of efforts, sustained again and again, to give access to water or another basic service to millions of people[22]. Reach is therefore constituted by replicable elements, such as a light blue plastic dispenser or motorcycles to move around. But reach is also importantly made up of hard work, an awareness of variability and

---

[22] Chapter 3 of this dissertation discusses how reach in evidence-based development is also increasingly achieved through partnerships with governments and the reliance on state institutions.

the unexpected, and hence the willingness and capacity to adapt, redesign, and modify as I intended to instantiate with several vignettes above.

While I have so far provided lists of replicable and nonreplicable elements, as if these could be easily separated, a focus on reach troubles this dichotomy as technologies, things, people, procedures, tinkering, motorbikes, energies, and slippery paths mingle. Reach as a heuristic device puts into question conceptual binaries such as standardization/diversity, but also local/global or self-contained/collaborative. Reach emerges as a productive category specifically by studying explicit ambitions to scale to bring the DSW program to millions of people across Kenya (and neighboring countries). Thinking with reach, one cannot disentangle what fails and say that it is because they are reproducible and geographically or culturally insensitive elements of a project. Through the case of the DSW program, I have tried to underline that in the collective achievement of reach, what works well and smoothly, as opposed to what is broken, needs repair and adaptation, or is being resisted to, cannot be predicted in advance and can really be any component of the long chain of people and things surrounding chlorine dispensers.

The idea of reach, especially if taken in its verb form, is evocative to study what sustains off-grid infrastructures for living because it centrally conveys the idea of *effort*. To reach something or someone means to arrive at, to get as far as possible, and often to extend one arm out or to stretch one's limits to manage to reach. Reach also highlights slippery paths, like the ones treaded by James and me, that render movement and work difficult, perhaps even impossible at times. Sometimes one fails to reach what is intended to be reached, such as a chlorine dispenser because fuel runs out or there is a storm, or a program target because something or someone gets in the way, an authoritative decision was taken somewhere and the funding discontinued, or simply due to a lack of resolution. Perhaps one does not manage to bring disinfected water to millions of households, because new pathogens still reach their

disinfected water and re-contaminate it, because people may decide not to use the dispensers and chlorinate their water, or even because vandalized dispensers indicate a certain failure to take sufficiently into account existing worries about family planning.

To offer reach as a conceptual alternative to scale therefore does not mean to advance collective achievement as predicated on perpetual agreement or success. Geographer Adam Moore (Moore 2008: 219) has argued that, within the field of geography, network and scale have become the dominant conceptual devices and are often presented as describing mutually exclusive and contradictory phenomena. Where scalar practices and arrangements denote hierarchically and vertically-ordered forms of life, the network is opposed to denote empirical, epistemological, and political realities of a more democratic and equitable nature. Reach instead allows for a multidimensional metaphor, one that neither fully flattens nor fully assumed the nested vertical level. Importantly indeed, an attention to reach does not mean ignoring power relations, inequalities, disagreements, forces, resistances, and even mistakes and failures. By recognizing the hardship and inventiveness of our actors' labors, an anthropological focus on reach hopes nevertheless to render more difficult finalized conclusions and definitive judgements about the value of diverse forms of welfare provision.

By emphasizing a turn toward reach as an analytic category, I also aim to address an empirical and conceptual confusion present with scale as a heuristic. Anna Tsing has provided a substantial description of her interest in scale, which she approaches in the verb form. "To scale well", Tsing wrote "is to develop the quality called scalability, that is, the ability to expand — and expand, and expand — without rethinking basic elements" (2012: 505). As several recent scholarly initiatives have tried to qualify Tsing's claims, these have worked with very broad –and in my view loose– definitions of scale. These empirical investigations importantly show that scaling practices can "spawn a sense of intimacy and an ethic of interrelatedness" at the same time as they serve projects to discriminate and silence. In this

view, important kinds of relationships are the product of scaling projects "which can transport us across space and time, introduce us to countless other actors, and dock us to any number of shores", like the possibility of accessing essential services (Carr and Fisher 2016: 153).

Authors who issue these conclusions and who claim to study contemporary scaling practices examine realities that are incredibly diverse, such as a 200-ton Japanese concrete dock (the product of tsunami debris) washed up on the Oregon shore (Carr and Fisher 2016), but also language politics in Hungary (Gal 2016), or the court system of a Polynesian island (Philips 2016). Yet, these diverse topics are all conceptualized as scale or scaling. What does "scale" refer to here, and does this focus remain conceptually productive? In those instances, it seems as if scaling is equated, if not with hierarchy, then with making things "bigger". But if scale is simply bigness, then can one not find scale *everywhere* as a potential object of analysis? And if everything can be explored through the heuristic of scale, what remains of its conceptual thrust? In my reading, this confusion calls for devising new analytic devices to differentiate between various scaling projects.

Studying aspirations to scale as *processes of reach* conjures the difficult journeys undertaken to both provide and access basic services, but also, centrally, the repeated work of maintenance and repair upon which these off-grid infrastructures fundamentally rely. This analytical lens joins together the care that is given to material technologies, like when James repairs broken chlorine dispensers, and the importance attributed to the idea of a collective life where everyone has access to basic services such as safe water. Here, the burdens of maintaining the DSW program merge with the ethical responsibility to provide for the urgent needs of those millions of people living without access to state infrastructures bringing water, electricity, or transportation. Building up freely from Steven Jackson's piece on *Rethinking Repair* (2014)*,* I wish to emphasize how an ethics of mutual responsibility can be used to talk

of large technological systems as they emerge primarily as invested with moral concerns and fractious hopes for a good life.

I asked throughout this chapter: How does one think about ambitions to scale while witnessing a chlorine dispenser repair by a river in Atani village? One will catch sight of that little girl, likely eight years of age, who is coming down a muddy track to fetch water, a 10-liter jerrycan in each of her hands. She needs water to cook, to wash herself and her family's clothes, to clean the house she lives in. Just like billions of other humans and infinite other kinds of beings, whose lives ultimately depend on water. In examining the labors that enable scalability as encountered in my fieldwork, and thus focusing on *reach* rather than *scalability*, I also aimed to unearth the invisibility of the grid-like infrastructures from which one, as a scholar, tends to receive basic services. I contend that taking for granted certain "already scaled" infrastructures constitutes the condition of possibility for social scientific analyses that decry technical fixes and uphold such a strong division between the political and the technical, as if the distinction between the two were obvious.

Standing on that slippery path in Kenya to reach water, I saw in the defense of the *diverse* against standardized technologies, generalized services, repeated actions, or replicated projects, also a problematic rejection of the simple, the ordinary, the common, the average, and perhaps the necessary. When one opens the tap to get drinkable water, rather than "go to the water" to fetch it, one might forget that running water was itself made into a scalable project. When one studies ambitions to scale-up as the effort to *reach* millions (rather than merely scale a project (capitalism) or a program (international development), one might come to think of sameness not only as deadly or authoritarian, but also as necessary and transformative. In that perspective, sustaining our shared biologies might be foregrounded again as an impetus as important and worthy of preserving (or fostering) as cultural diversity and that rather than a given should be reckoned with for what it is: an achievement. Differently put, might an

anthropology attentive to the dynamics of reach request from us that we reconsider an intellectual genealogy that divides between bare life (mere biological needs sustained by humanitarian and development initiatives) and the good life (conceived as the product of variable, unscaled existences) (Agamben 1998; Fassin 2007; Malkki 2002; Reid 2010).

**Conclusion: on the unfinished**

In this chapter, I have drawn attention to the many forms of labor that suture the continuity and expansion of a water initiative in rural Kenya. Taking seriously the NGO's aspiration to reach millions of people with its program, I explored how this program is made scalable, foregrounding that the work of repair, maintenance, but also re-design and adaptation, figure centrally in that endeavor. Of course, this story provides only a limited view of the Dispensers of Safe Water, which involve much larger horizons, both spatially and temporally, than the ones I was able to cover during my fieldwork research. Nevertheless, I think that its emphasis helps appreciate the DSW program as a thick, heavy, burdensome, and always improved upon off-grid infrastructure. I contend that by focusing on the fact that something so often taken-for-granted as service is made possible by great chains of people and things, and the links their work and efforts trace, that we gain both broader and deeper perspectives on the global projects that animate global development efforts today around the world.

The achievements of scalability that I studied are always iterative, uncertain, and unfinished. I suggest that conceptualizing the DSW as an infrastructure of efforts and presenting efforts to scale it up to millions of people through the analytic of reach helps appreciate those tentative and adaptive realities. As anthropologist Nikhil Anand (2011: 558) wrote before, the intention is not to be clever with native categories (such as *efforts* and *reach*) that one finds in fieldwork through the labors and the aspirations of our interlocutors and merely squeeze them until one has exhausted their power of analogies. Instead, those two

concepts – *infrastructure of efforts* and *an anthropology of reach* – help to appreciate the changing and complex realities that are knowingly faced by those who in the first place wish to replicate a problem at a very large scale. Our efforts, as I alluded above, can serve us as often as they can fail us, in that we cannot repair all the dispensers that need it, or we might forget to bring necessary chlorine supplies, or we might be unable or unwilling to resolve protests against the DSW program, and hence might not be able to reach millions of people and grant them access to the opportunity of disinfected water.

By focusing on the mundane yet repeated work enrolled in both expanding to millions of people and sustaining through time a water sanitation initiative, I wished to trouble conceptual binaries widely used to analyze international development and global health. I showed that, if one takes seriously those whose lives revolve around the DSW program, the opposition between band-aid technical solutions and real political change does not hold, unless one withdraws any capacity for reflection, imagination, and action from the very people one studies. In presenting scaling up a program as a process of reach that must account for different topographies, ways of working, and sensibilities, I have also aimed to problematize vertical understandings of the world, premised on an evident distinction between global imposition and local variation, scalability and cultural diversity, homogenization and difference. In this other reading, sameness is not inimical, but generative and constitutive of diversity, insofar as the provision of basic services at scale promises a good and multifaceted existence.

An attention to ambitions to scale as *reach* enabled by *infrastructure of efforts* does not propose a naïve and laudatory alternative to critical approaches to globalization and international development. Rather, it mimics – here in a metaphorical way – the unfinished and adaptive features of the empirical realities it explored to propose conclusions that are unstable. By this, I do not mean that I am advancing claims that I take lightly, just like my interlocutors did not take lightly their work of managing and caring for an initiative about safe water.

157

Nevertheless, and in conversation with other anthropologists who have advocated for the value of ambivalence in guiding inquiry, this chapter is also indirectly an effort to propose a scholarship whose value and thrust grounds not foremost in authoritative certainty but rather in its combination of careful analysis always open to the possibility of its own limits on the one hand, and its need to revise its positions on the other hand (Redfield 2016; Stewart 2008). As I was writing this chapter, I became aware of a recent follow-up study that questions the effectiveness of the DSW program in averting deaths from waterborne diseases, citing the need to take better into account the combined routes of people and pathogens in their homes (Null, et al. 2018). How might these findings change the shape of the DSW program in the years to come, and possibly the very ideas traced in this chapter?

# Chapter 3 - What is and what is not of the state? Partnership as enmeshment in Kenya

**State's absences**

In March 2016, I attended the "Release of Year 3 Results" for the last deworming round of the Kenya National School-Based Deworming Programme (NSBDP) in a primary school of Kirinyaga county in central Kenya. Bringing together the Kenya Ministries of Health and Education, the Kenya Medical Research Institute, non-profit organizations, donors, and pharmaceutical companies, the NSBDP is an effort to treat millions of children affected by parasitic worms in schools across the country, with the overall goal of improving their health and educational outcomes. To reach this event aimed at publicly disseminating the results that the NSPDP produced during the previous year, I travelled with employees from the NGO Evidence Action and a manager of the Child and Adolescent Health Unit of the Ministry of Health. Upon entering the school premises, we were greeted by the rest of Evidence Action's Deworm the World team who had arrived a day early to finalize the event's organization.



**Figure 12: Setting-up Release of Year 3 Results in a school in Kirinyaga (Photo by Fiona Gedeon Achi)**

White tents had been set up in the sandy school yard to shade chairs and tables where the invited audience, including students' parents, would sit. On the left and right, schoolchildren wait on the grassy lawn strewn with tall lean trees. A banner, portraying the picture of two small children holding deworming pills, indicates that this event is about the National School-Based Deworming Programme. "*Kwa Afya Na Elimu Bora, Tuangamize Minyoo*!" says the slogan on the banner: "For health and good education, destroy worms!". The event is scheduled to start at 9am but at 10am, the Master of Ceremony and the County Governor have not yet arrived. Following the arrival of the Governor's representatives at around 11:30am, the event finally starts with dances and songs performed by the town's scout organization.

After a morning of speeches delivered by a long list of school employees, government officials, staff from Evidence Action, and donor agencies' representatives, the event is closed by the symbolic unveiling of the Year 3 Programme Report. Ceremonially, one of the speakers cuts the white ribbon and pulls up the blue fabric that hides the billboard symbolizing the Report, thus revealing one figure – 6.1 million – which refers to the number of children reached by the national deworming program in 2014-2015. On the way back, someone jokes that the governor could afford to arrive late or not show up all, because five hundred children are not five hundred voters, alluding to the students from the Kirinyaga primary school who have been recipients of the NSBDP in the last years.

This was not the only time during my fieldwork that a public event was delayed because of a senior government official's tardiness or absence. A month earlier, the County Director of Health arrived hours late to a meeting for planning the logistics of the upcoming round of the NSBDP in Makueni county, which was attended by about forty local and national-level representatives from the Ministry of Health, others from the Ministry of Education, and some managers from Evidence Action. In June 2016, the ceremony for the Launch of the Second

Kenya National Strategic Plan for Control of Neglected Tropical Diseases started late as the audience waited in a Nairobi conference hall for the arrival of the Cabinet Secretary for the Ministry of Health who ended up not coming[1].

Relying on the joke about the governor mentioned above, it might be tempting to transform these episodes of absence and delays into an icon of the state's indifference or weakness, but my informants offered also alternative interpretations. In the car back to the office of Evidence Action, we discussed the absence of the Cabinet Secretary of Health with the team from Evidence Action. Elijah, budget coordinator for the deworming program at Evidence Action, said that most people think these ministers are snobs when in fact they are extremely busy. To illustrate his point, Elijah reminded his colleagues: "you meet with them and at 6pm they still haven't had time to eat and you can see it on their face"[2]. Elijah's words foreground that government officials, in Kenya just like anywhere else, have intense and conflicting schedules, especially when they are Ministers or Governors. Not showing up is here taken to be the outcome of the inevitable busyness that comes with the task of administering people and things within a national territory. When explained through the demanding work life of governmental managers, episodes of absence acquire a prosaic quality that nuances certain social science narratives about African states as incapable, failed, or stunted (along with Brown 2015; Ferguson 2010; Mahajan 2017; Mann 2015; Redfield 2012).

The vignettes above speak to the contradictory roles granted to the Kenyan government by the global poverty alleviation NGOs with whom I conducted fieldwork: simultaneously reliable and unreliable, functional and dysfunctional, present and absent. Why, this chapter asks, despite this ambivalent portrayal of the state, NGOs such as Evidence Action are eager

---

[1] This is the equivalent to the Minister of Health in Kenya.

[2] While this could be taken as an image of NGOs and other development actors being a big strain on the government as its officials cannot attend all partnership meetings, I want to emphasize that the absence of the Cabinet Secretary of Health was noticed at the Launch of the Second Kenya National Strategic Plan for Control of Neglected Tropical Diseases, which was a governmental event to inaugurate a new set of national policies related to the elimination of neglected tropical diseases in Kenya.

to partner with governments when implementing their programs? Why were the different evidence-based development organizations I worked with always discussing and sharing their experience around successfully "engaging developing countries' governments"[3]? Sidestepping the vague explanation that state capacities are the obvious choice to route one's non-governmental program for reasons of legitimacy and sustainability, I seek the answers to this question in another type of the state's absence. Reading over my field notes, I realized that so many of my interlocutors talked to me about the *government*, whether to praise it or loath it, whether as a blurry shadow or personified civil servants, whether as closed partner or elusive political candidate. Almost no one however had invoked the Kenyan *state*, or even the Indian state when I was conducting preliminary fieldwork in Rajasthan.

This could be unsurprising given the well documented rise of global neoliberalization and marketization in recent decades that pushed back the state as the main provider of services, both in actuality and in people's expectations (Ferguson 2006a; Hearn 1998). It could also have to do with the fact that evidence-based development, the movement to reduce poverty by ensuring that policy is informed by "scientific evidence" and my topic of research, functions through an Americanized world view, in which the state has never played a central role in providing welfare, unlike in other European countries such as France and Germany (Collier 2016; Gupta 2012: 301)[4]. Without denying these claims, I argue that this exclusive use of the term "government" across my field sites reveals another interesting empirical puzzle. It gestures to an instrumental relation to the apparatus of the state, one in which the fundamental goals of the NGO and that of the state might not be shared, but where operational capacities converge to put in place a scheme such as the Kenya National School-Based Deworming

---

[3] See for instance: https://www.eaglobal.org/talks/engaging-developing-country-governments/, accessed June 12, 2019

[4] For a detailed discussion of the nature of the evidence at the heart of the rising movement known as evidence-based development, see Chapter 1. By exploring relations and collaboration with the government, the current chapter prolongs an examination of the infrastructural and institutional components critical to achieving scale and reach for NGOs, already addressed in Chapter 2.

Programme. My interlocutors yearned to partner with the Kenyan government, but also other national governments, to gain access to the state apparatus and its institution that allows the wide distribution of a deworming intervention across a national territory. It is the enrollment of a system to govern a state – i.e. the Kenyan *government* – into a program that is not *only* about the traditional aims of the state, since Evidence Action seeks to deworm millions of children worldwide rather than solely enhance the health of a national population.

By drawing on eight months of participant-observation shadowing the Deworm the World team at Evidence Action and more limited interviews conducted with Kenyan government officials, this chapter shows how working with the government to deliver deworming treatment to children across schools in Kenya is reckoned to be as necessary and enabling as it can be limiting and constraining. The analysis of this chapter decenters the naturalness of the nation-state as the taken for granted political form. It documents how, from the perspective of Evidence Action, the government emerges predominantly as an instrument of scale, rather than as a system mandated to represent and execute a state's sovereign authority to administer a specific population within defined borders (Redfield 2012: 157). At the same time however, by foregrounding how the state and NGOs need each other's operational capacities to implement a functional program, this chapter shows non-governmental interventions as more than neocolonial enterprises that erode national sovereignty (Biesel 2014; Hanlon 1991; Hearn 1998, 2007; Nguyen 2010; Pfeiffer 2013). Through a detailed examination of the Kenya NSBDP, I document how the project of Evidence Action and that of the Kenyan state mutually implicate each other in ways that cannot be easily captured and even less easily explained away by either one being subsumed by the other. Overall, I argue that the collaborations between non-governmental organizations committed to evidence-based

development and national governments importantly bring forth political sovereignty under several different guises[5].

In this way, this chapter provides an indirect ethnography of the state to understand current transformations in the role of the African state, both imagined and actual, that moves beyond theorizations premised on the inadequate and vanishing state, or on the ontologizing of its superior legitimacy (Gupta 1995; Trouillot 2001). It offers a conceptualization "from the margins" because I encountered the state only indirectly in my fieldwork which was focused on the work of non-governmental organizations on poverty (Das and Poole 2004a). My view from the periphery is of course of a very different nature than that initially intended by Veena Das and Deborah Poole when they proposed "a radical rethinking of the state" from the margins, whereby the object of study was primarily marginal populations that the state attempts to manage, transform, or pacify (2004a: 7). Importantly, however, my chapter builds on these and other authors to approach the state as an "incomplete project" (Das and Poole 2004a: 7), an "unfinished project" (Mann 2015: 9) and also a "composite reality" (Gupta 2012: 46) that shares with other non-state entities the responsibility of administering people and things within its territory (also Elyachar 2005)[6].

In conversation with other scholars, this chapter interrogates how state sovereignty must be rethought of in light of the emergence and increasing influence in the last decades of non-governmental actors to address health and other social issues in African countries (Bemme 2018; Brown 2015; Elyachar 2005; Fadaak 2019; Ferguson 2010; Ferguson and Mansbach

---

[5] This chapter is primarily concerned with what Radkhika V. Mongia (2007: 395) has called the "internal" dimension of sovereignty, namely the authority of the state to adjudicate over a fixed territory as well as a definable population, rather than with the "external" dimension of sovereignty which is about the recognition by other nation-states in the international order and the potential infringements by other states into the state's sovereignty.

[6] In a personal communication, Stephen Collier emphasized how paying attention to the distinction between the government and the state is also useful for thinking about "*distributed sovereignty* where one doesn't actually think that there is a singular "state" that is going to be the agent of a scale-up for instance" (November 16, 2016).

2007; Rees 2014a)[7]. I speak of the nation-state here primarily understood as a modern biopolitical project: a state concerned with enhancing (and controlling) the life of its population (Foucault 2004)[8]. This is a different emphasis than that pointed out by Julie Hearn (2007: 1095), according to whom "given the close working relationship with both northern and southern states, the extent to which NGOs are 'non-governmental' is a source of much debate." The issue in this chapter is not so much whether Evidence Action is truly non-governmental, but rather how close partnerships between NGOs and governments might ask us to think differently about the nation-state as a project.

This chapter therefore contributes to an important literature that has sought to decenter the state as the main political problem to ask instead "what is government?"; or in other words, how are we being governed today (Ewald 2016; Ferguson 2010; Ferguson and Gupta 2002; Foucault 2004; Mann 2015; Murphy 2017)? When I ask "what is government?", I refer to a different notion of government from the one used until now to invoke the government of Kenya. In this case, "government" is not the system to administer a political community, but the process of governing, or what Foucault has called "arts of governing" and the "conduct of conduct", which include management and control techniques not limited to state politics alone (Ewald 2016; Foucault 1982). To inquire into what government is, instead of asking "what is the state?", means also to recognize that national populations are managed by interventions, actors, ideas, and actions that cannot be reduced to those of the state and its apparatus.

---

[7] I am grateful to Hannah Brown (2015) and Gregory Mann (2015) for helping me realize that that contemporary global health partnerships in Kenya must be understood as something that occurred at a particular historical conjuncture that is neither within the framework of the new post-colonial state, full of vision at the height of its sovereignty, nor within a conceptualization of the state as it was thought of at the apex of structural adjustments in the 1990s "when NGOs were supported over and above the state because the state was viewed as bloated and incapable of properly delivering resources" (Brown 2015: 352). In contrast, Veena Das and Deborah Poole (2004b) tellingly discussed how for classical theorists of the European state form, such as Thomas Hobbes, Jean-Jacques Rousseau, or later Max Weber, the problem of the state was that its boundaries were always threatened by forces of nature that did not conform to the state's rationality.

[8] Of course, it remains a question as to who and what (interests) does the state really serve (the public interest? of which public? private, nepotist interests?), especially when the state is considered to be a composite entity made of multiple of institutions and programs.

Nevertheless, by asking "what is government?" rather than using the terms "governmentality" or "governance" preferred by some authors (e.g. Brown 2015; Ferguson and Gupta 2002), the intention is to underline that in this chapter I am concerned with the consequences of transnational forms of government precisely in regards to a state's political sovereignty[9].

After a condensed description of the NSBDP, I propose to conceptualize the partnership between Evidence Action and the government of Kenya in terms of an enmeshment to bring into view how hard it is to tell apart what is and what is not of the state in this deworming program[10]. First, I document this enmeshment by showing how, from the perspective of Evidence Action, the government emerges as an instrument of scale. I show how this scale is achieved through a set of public infrastructures to highlight the contradictory roles granted to state resources as both necessary and yet insufficient. While in this case enmeshment is about operational capacity, I move beyond a focus on shared technologies of government in the second part of this chapter to show that the boundaries between the aims of government of the state and those of the NGOs cannot be so easily established. This demands exploring how we are being governed by forces beyond the state today without relying exclusively on a narrative about the "straightforward marginalization of the state" (Mahajan 2017: 1; but see also Collier 2011; Ferguson 2015; Mann 2015).

**Enmeshment, or in the shadows of the government**

---

[9] I rely on the adjective "transnational", foregrounded by Gupta and Ferguson's (2002) seminal article on *Spatializing States*, to denote processes of governing that go beyond the state. Others have also proposed the term "global", stressing that the global is not necessarily that which dispenses with the state, but also accommodates it in configurations that are unexpected (Tobias Rees, May 2019 in personal communication). This chapter similarly aims to show that global projects –in the sense of those led by NGOs and having at stake goals that are global in scale– do not always hover above the state, but centrally necessitate it. Nevertheless, the term transnational remains more widely used in the literature, while the term "global" is also more centrally an actor's category such as in "global health" or "global poverty". I thus used the term transnational here for reasons that have to do with clarity.

[10] In their reader on the anthropology of the state, Aradhana Sharma and Akhil Gupta underscore that the distinction between state and civil society, often assumed to be universal and natural, might not capture postcolonial realities "where the boundaries between state and non-state realms are blurred" (2006: 166).

A few days after attending the Release of Year 3 (2014-2015) Results for the Kenya National School-Based Deworming Programme, I was given a copy of the Year 3 Programme Report, co-authored by the Kenyan Ministry of Health and the Kenyan Ministry of Education. Reading on, one learns that "the goal of the programme is to remove parasitic worms as a public health problem for children in Kenya" (Kenya Ministry of Education and Ministry of Health 2016: 5)[11]. The program targets both soil-transmitted helminths and schistosomiasis, parasites nested in the intestines of people and transmitted via contaminated water or food sources. Taking place in schools across "at-risk areas endemic for parasitic worms", the program targets children aged 2 to 14 years, whether or not they are enrolled in school, and aims to treat at least five million Kenyan children each year during five years between 2012-2017. The report adds that: "Regularly providing deworming tablets to children through schools is a proven cost-effective and safe treatment strategy due to the readily available, extensive and sustained educational infrastructure." (ibid). The NSBPD, the report also details, "is a government programme implemented by the Ministry of Education, Science, and Technology and the Ministry of Health" (ibid). The report also thanks its donors and "recognizes Evidence Action for their technical assistance and fiscal management of the programme" (2016: 2).

The funding came from two philanthropic donors, one focused on improving the lives of children and the other on ending neglected tropical diseases, dedicated specifically to the NSBDP. Geared toward sustaining five years (2012-2017) of the national program, the funds are in principle the government's responsibility. Yet, the funds were actually routed through the NGO Evidence Action for funding allocation and budget management, in part because of the daunting, even incapacitating, public procurement system. The procurement system in

---

[11] The prevalence for both these infections is reported to have dropped in the 2012-2015 period (Message from Dr. Jackson Kioko, Ag. Director of Medical Services-Ministry of Health in Y3 Report).

Kenya operates through elaborate systems of competitive bidding (supposedly) designed to avoid corruption (Wrong 2009). In practice, as I was often told, the system makes it very difficult to operate budgets through government means because the timelines are *just* so long that they turn interventions into irrelevant ones. "When the money comes to the government", said Adrian from the Ministry of Health, echoing many others during my fieldwork, "the process of getting that money is so tedious and difficult, it is much and faster with NGOs, it makes it easier to run a program" [12].

The Deworm the World team at Evidence Action, who manages the funds and logistics of the country-wide NSBDP on behalf of the government of Kenya, is made-up of about ten employees. In the Nairobi office, Damien – in charge of the implementation cascade – spends hours every day on the phone trying to reach schools, government offices, and health centers to schedule meetings, trainings, transport of drugs, and the printing of monitoring forms to collect data about the program. As the budget coordinator, Elijah sends hundreds of SMS messages and emails to remind teachers, headmasters, county directors of health and education to provide their expenditure receipts. The country program director Sheila attends bi-monthly meetings with the program steering committee to review "high-level, strategic directions, and ways forward for the program". As program managers, Siya, Marc, and George go around the country to lead dozens of meetings to organize the details of the NSBDP in different regions, to train teachers to implement the annual deworming day in their schools, and to teach local officials to use the monitoring forms to report on the program's unfolding, including the number of children who took the deworming drug in different schools across Kenya.

---

[12] There were in addition worries about the misappropriation of funds that seem to have been founded. In the early days of the NSPDB, which started in 2008, the program was funded through a pool fund granted by external donors to the government of Kenya to finance, among other programs, its national deworming program. There were concerns about how the pool fund was handled by the government officials in charge of it, resulting in legal indictments and in the temporary interruption of the program for a few years before starting again in 2011.

The government, on the other hand, provides what Siya has called "in-kind contribution". These include man-hours of all the government staff members who take part in the program, including Adrian and Elizabeth, representatives from the Ministry of Health and also of Education respectively, who travelled along with us to lead meetings and trainings with county officials during my fieldwork months. Siya also lists "infrastructures" as part of these in-kind contributions from the government, "such as the schools where deworming takes places". Produced by two big pharmaceutical companies, "the drugs come through the WHO donation program", Siya says, "but the transport and storage of drugs is also government-led."

To anyone who has worked in international development, this National School-Based Deworming Programme might strike them as being a typical collaboration between, on the one hand, a developing country government variably involved in the initiative at stake, and a group of experts on the other hand, from international organizations or NGOs, who act as technical assistants and through funds, energies, and insistence support the program. To me however, this configuration came together as a puzzle, one pieced slowly and imperfectly through fieldwork. I had gone to Kenya stating that "I am going to study the Kenya Deworm the World program ran by Evidence Action". I left Kenya, eight months later, feeling like I had learned ethnographically about the Kenya National School-Based Deworming Programme – a program where the boundaries between what is of the state and what is of the NGO are at least blurry, if not entirely impossible to tell apart as suggested by the program description I drew above. As an anthropologist, had I been studying a national program embedded in the Kenyan state's project or shadowing a non-governmental program emerging from non-state and global initiatives? Were the goals of the NSPDB to be understood, using the Ministries' Report discussed above, in terms of enhancing the public health of a circumscribed population? Or should the aims of the NSPDB be examined in light of Evidence Action's mandate to "deworm the world", which extended beyond Kenya's national borders to intervene in other countries

as well? This left me wanting to examine further what it meant, in the parameters of my fieldwork, to be 'of the state' or 'outside of it'.

To capture the difficulty to provide a square answer to these questions, I propose to analyze the collaboration between Evidence Action and the Kenyan Ministries of Health and of Education through the image of *enmeshment*[13]. The metaphor of enmeshment is not meant to signal equal powers but serves to move beyond the lens of "technical assistance" which is state-centric and implicitly naturalizes the state as the legitimate authority to govern. By proposing enmeshment as an analytical tool, this chapter follows a growing anthropological interest to "attend to the ways that 'global health partnerships' are described by parties who are differently positioned" because it might reveal distinct expectations, metaphors, and ontologies (Taylor 2018: 4; but see also Brown and Green 2017; Crane 2010; Rottenburg 2009a). However, by proposing enmeshment instead of relying on the actor's category of "partnership", this chapter seeks to bring into focus not only what separates the different partners in their appreciation of the same intervention but is also "attuned toward modalities of coming together" (Brown 2015: 350) as well (Mahajan 2017). The notion of enmeshment sheds light on how Evidence Action works with Kenyan state's capacities to meet its aim of delivering deworming treatment to millions of children in Kenya, keeping in mind that the aims of the government might be a global ambition that exceeds the Kenyan territory[14].

---

[13] The verb "enmesh" is used in the preface of Das and Poole's (2004a) edited volume *Anthropology in the Margins of the State* to describe the relationships between "local worlds" and "the state" to denote that the state is both an overwhelming power and a source of local desires. I had used this term to conceptualize what I saw in fieldwork before discovering Das and Poole's usage, but theirs highlighted the need to reiterate that enmeshment does not mean equal standing. I also do not use the term of "hybrid" used by others (Biesel 2014; Redfield 2012) because it refers predominantly to a multiple configuration, which brings in public-public private partnership for funding and management but does not directly address the question of how plans and visions might be enmeshed as I tackle the subject in the later section of this chapter.

[14] In his compelling discussion of personal water filtration systems developed for African countries by the corporation Vestergaard Frandsen, Peter Redfield (2016: 173) discusses how the corporation engages the Kenyan government as partner, even though its endeavor is never framed as one of nation-building. Redfield presents there the term "partnership" as a common trope of global health, "in which contractual equality can mask both asymmetrical power relations and an absence of fixed obligations".

An analysis of transnational government through enmeshment asks us to ponder carefully about – and even sheds productive confusion on – what the aims, terrains, and actors of government are. In her famous book *The Will to Improve*, anthropologist Tania Li conjures Michel Foucault to contend that to govern is "to seek not one dogmatic goal, 'but a whole series of specific finalities'" (2007: 9). These finalities, Li insists, might be incompatible and borne out by different powerful entities such as politicians, missionaries, or commercial lobbies, to underscore that there was "no unitary purpose" to colonial rule in the case of colonial Indonesia about which she writes (ibid). Situated in independent, post-colonial Kenya, my chapter nonetheless attempts to follow these diverse finalities. In the NSDPB, what and who is being governed, and by whom and how? What becomes of states' designs under these circumstances? What kinds of understanding of political sovereignty are therefore foregrounded?

**Scale = evidence\*government**

"Another thing I should mention is that increasingly we recognize that, you know, you can't just say: *ok, go government of X, scale-up deworming*. [The governments] need a lot of technical support in all of the mechanics of doing that, and it's something I don't know much about at all because it's about working inside of a government and thinking about regulations and approvals and budgets and line items and just really detailed stuff about technically implementing a program at large scale in government, especially a highly, highly bureaucratic government like India…". These cautionary words were issued by Robert, Senior Policy Manager at the Jameel Poverty Action Lab (J-PAL), at the beginning of my fieldwork when I visited his office in Boston in September 2015, several months before I was to start fieldwork in Kenya and learn specifically about the "mechanics of working with a government".

Robert was explaining that in order to achieve lasting change, it was not enough for himself and his colleagues to produce convincing evidence about the positive impacts of mass school-based deworming on both health and school attendance, and then advise governments to transform this pilot intervention into larger programs in their countries. Instead, Robert's words indicated that his organization is increasingly acknowledging the need to work closely with governments to ensure that policy is informed by scientific evidence. In fact, the work of Evidence Action must be situated in a broader movement of evidence-based development that brings together several research and policy partners such as J-PAL and Innovations for Poverty Action, all headquartered in the USA but with offices and operations around the world[15]. As mentioned in previous chapters, what animates the work of those NGOs is a willingness to disrupt a certain perceived status-quo in the current global policy landscape to ensure that resources are directed toward funding and implementing programs backed-up by field-based evidence drawn from randomized evaluations that can be scaled to millions.

I write *increasingly acknowledging the need to work with governments* because, as has been documented by geographers Sophie Webber and Carolyn Prouse, in the early 2010s advocates of evidence-based development at J-PAL positioned "NGOs as superior partner organizations to state-based entities, [as] more flexible", more reliable, and also acceptant of the randomized evaluation methodology to evaluate anti-poverty interventions (2018: 176).

On December 7, 2013, I attended the tenth anniversary of J-PAL that involved a public celebration attended by more than a 1,000 people from across the planet, seated in an auditorium featuring talks, videos, stories, and panel discussions from academics, NGOs staff, field researchers, and policy-makers. One speech was dedicated to "The Origin and Evolution of Randomized Evaluations in Development" and was led by Michael Kremer, professor of

---

[15] For further details about this nexus of organizations, see the Introduction of this dissertation, particularly the section Field sites.

economics at Harvard University, who recounted how he first had the idea, in the late 1990s, to implement randomized evaluations in Kenya in collaboration with a Dutch NGO called ICS to evaluate their educational programs and how this initial work continued in India when a colleague, Abhijit Banerjee, asked Kremer in 1996 to come to India to explore the possibility of conducting randomized evaluations with the NGO Seva Mandir. Kremer further detailed how his approach to randomized evaluations, centered around partnering with NGO programs, offered an important edge in terms of costs and opportunities over other institutions of development, such as the World Bank's Development Impact Evaluation unit, that evaluate government programs. Similar to what I encountered in the field, Webber and Prouse contend that "only recently have [J-PAL] researchers begun to recognize the importance of working with governments for large-scale antipoverty impact" (2018: 176).

In the last decades, anthropologists have increasingly studied entities that provide welfare services for populations alongside and beyond the state: international organizations, big philanthropies, humanitarian bodies, corporations, and NGOs (Escobar 1994; Ferguson 2015; Li 2007; Rigillo 2014). While several of these actors, such as the Gates Foundation or the World Bank, emerged explicitly to remediate for the failure or absence the state (Biesel 2014; Feldman and Ticktin 2010; Mahajan 2017; Redfield 2005; Rees 2014a; Ticktin 2011), it is not so much a concern with filling-up state's incapacities *per se* that grounds the visions and daily work of those organizations promoting evidence-based development. Instead, this collection of non-governmental organizations worried about the consequences of *all* anti-poverty initiatives, whether state-based, non-governmental, or public-private partnerships, insofar as they are not evidence-based, namely not grounded in evidence of causal impact collected through randomized evaluations (see Chapter 1).

"So would you say that most scale-ups [of "successful" interventions] are actually done through government as opposed to another [entity]?", I had asked Robert after he insisted on the importance of granting technical support to government.

"No", interjects Robert, "scale-ups are done through all different types of organizations, but – this is also really important – we do recognize that the *highest* level of scale and reach will be achieved by governments, they just have the ability to reach *way more* people than *any* NGO, perhaps with the exception of BRAC [the largest NGO in the world in terms of number of employees, based in Bangladesh]. But I mean BRAC just reaches millions and millions of people. But you know, if you take for example of [the non-profit] Seva Mandir who is the partner for the incentives for immunization program and they scaled it up to everyone of their beneficiaries then you'll have a number X. But if you convinced the government of India to scale it up across the entire country, it'll just be *exponentially* larger."[16] Robert was suggesting that collaborating with governments is a necessary ingredient in achieving J-PAL's interest of scaling up 'successful' interventions to millions of people. This was an outlook that I was to encounter again when shadowing the work of Evidence Action.

In an interview in May 2016, Iris, then Director of Global Communications and Advocacy at Evidence Action, told me that the NGO's mandate was to implement programs that each reached at least 1 million people. Iris claimed that "to provide those kinds of services" at that level of scale, "that's the purview of government".

"If we reach 220 million people with a staff of 300 people, I say this is a cost-effective way of doing anti-poverty alleviation work. That's a good multiplier".

Iris talked fast, assertively and exploratively all at once.

---

[16] For instance, the Evidence-Action led Dispensers for Safe Water Program (DSW) in Malawi works through government health centers and their staff to deliver key program resources, while the DSW Program in Kenya relies instead exclusively on a vast network of office and employees of Evidence Action for reasons that have to do with their capacity of forging a closer collaboration with the government in one place versus the other. That does not mean however that the DSW does away entirely with the state insofar as it needs official permissions regulated by the Kenyan government to operate. For a longer discussion of the DSW, see Chapter 2.

"For our programs, she added, "I would consider the Kenyan government and the Indian government a partner, without which the programs would not reach the scale they do."

She said that she doesn't think there is a single NGO that could claim to have touched that number of people in a year without operationalizing its programs with governmental means.

"Well", she conceded, "there are organizations that reach millions and millions of people", and she gives the example of Pratham, an India-based education-focused NGO, "but they have 60,000 staff. Some NGOs who are quasi-governmental, because of their size and reach, but they're so big."

Iris' mention of Pratham reminded me of Robert, back at J-PAL headquarters, who had talked of BRAC. They were both gesturing to NGOs that have the capacity to reach *scales similar to governments* – to touch populations as big as those that the government can reach. Iris calls these kinds of NGOs "quasi-governmental": "they provide services in lieu of governments, but on behalf of governments". These NGOs have extensive financial and human resources, for instance BRAC is said to have more than 100,000 employees. Both Iris and Robert contended that if you are not that kind of NGO, you need the government to scale your program to millions of people, such as to achieve that 6.1 million figure revealed at the Release of Year 3 Results event I attended, which referred to the number of children reportedly covered by the Kenya NSBDP in 2014-2015[17].

In a recent article, anthropologist Manjari Mahajan (2017) describes a similar trajectory for the Gates Foundation in India, from initially "circumventing" the government in its work

---

[17] Another interlocutor, Edward Miguel, founder of the Center for Effective Global Action, an evidence-based research and policy institute at the University of California-Berkeley, also referred to this image of the government as instrument of scale during a presentation when he talked of "public sector solutions" that are currently being evaluated by randomized evaluations (by contrast to NGO solutions). Edward said: "A colleague, Karthik, is leading a project in Andhra Pradesh with governments who introduce biometric smartcards to transfer funds to citizens [to prevent corruption]. What is great about Indian states is that each state is 50-60 million people, so this huge scale…"

on HIV/AIDS in the early 2000s to increasingly seeking partnership with the government in the 2010s as it launched a new Ananya program that targeted diverse outcomes related to family health. Mahajan's analysis suggests that the scale and complexity of the problem their new Ananya program was focused on was an important reason why the Gates Foundation had changed its stance regarding its relationship with the government. Mahajan reports on an interview conducted in 2015 with Girindre Beeharry, the Director of the Gates Foundation: "It is not about several thousand sex workers" [as it was for HIV/AIDs]. Rather we are contending with the health of several millions of children and women. There is no question of going it alone [without the government]" (2017: 8).

Both mine and Mahajan's interlocutors were emphasizing that working with the government becomes a requirement when one aims to reach *millions* of people with a program because governmental means afford to achieve what Robert had called "the highest level of scale and reach". In sum, these descriptions of collaborative work sketch, from the perspective of non-governmental organizations, what I call the *government as instrument of scale*[18]. Of course, this term must be taken somewhat playfully because the intention is not to affirm that national governments, not even in the view of my interlocutors, can be reduced to passive implementation tools of extra-governmental visions. However, I do contend that delving into this idea of the government as instrument of scale might help appreciate how the "new arts of government" – these large projects carried out by NGOs in developing countries – do destabilize state sovereignty, but not in the ways we might expect it. Transnational forms of government might not necessarily erase state sovereignty, but rather strengthen its reach because the operational capacity and the goals of both the NGO and the state are enmeshed. At the same time, transnational forms of government might decenter the state's authority,

---

[18] My colleague Julianne Yip suggested that the metaphor of the government as instrument of scale is reminiscent of other mechanical metaphors, such as pulleys or levers, which enable to exert greater force with less effort.

focused essentially on fostering the lives of its citizens. To elaborate these claims requires knowing what both Robert and Iris meant by *government*, and by *"its ability to reach way more people than any NGO"*. In the next section, I provide answers through the ethnographic material I gathered in Kenya about the NSBDP.

**Government as instrument of scale**

Most likely, this ability to reach "way more people" and "treat at least five million Kenyan children each year" – as the 2014-2015 Year 3 Programme Report outlined – is not grounded in the state's financial capacity given that the Kenya NSBDP did not hold "a budget line" in the Kenyan government's budget. On the contrary, the program was entirely funded by external, non-governmental donors. In addition, as I mentioned earlier, fiscal management of the program was handed to Evidence Action in 2012 following rumors about the misappropriation of funds but also worries about the government's incapacity to efficiently allocate program funds due to the intricate Kenyan public procurement service. If you recall, Siya, program manager in the Deworm the World team at Evidence Action in Kenya, had told me that the government instead provides *in-kind* contributions. She referred to man-hours of all the government staff who in different ways are part of the program but also mentioned the schools where the deworming takes place.

These are schools such as the Kalani school in Kisumu where I observed the Deworming Day proceedings in February 2016, as headmasters and teachers distributed deworming drugs to children and recorded the distribution on monitoring forms, while taking notes of those children who should be exempted from receiving the drugs because their parents have asked so. The program also requires the participation of local officials from the ministries of health and education to coordinate the Deworming Day and the adequate storage of drugs, like that Makueni county director of health I alluded to in the opening vignette, who arrived

hours late to a logistics meeting. There is also Elizabeth, quality inspector at the Ministry of Education in Nairobi whom I met during a three-day long meeting with representatives from ten counties of Kenya to be trained as "master trainers" before they return to their counties to train teachers and health government employees to carry out the NSPDP. At the end of the training, as we wait for the training certificates to be delivered (someone in the Ministry of Health had to still sign them), the participants go over whether the expectations set up earlier in the week were met, if the fears were overcome, and if ground rules respected. One of the fears is that the program would be tainted by corruption.

One month later, I meet with Elizabeth in Jogoo House, which hosts the Ministry of Education in Nairobi's Central Business District. I am asking Elizabeth whether in her view the NSBDP is different from other programs funded by external donors she has worked on because it is based on evidence from randomized evaluation. Elizabeth instead explains to me how the NSBDP is similar to other programs, comparing it to the UNICEF Child Friendly School Program, because all "these consultants have to work with us because they're going to our schools. Basically we have the data and everything about schools. So you find our directors have nominated several of us [in the Ministry of Education] to work with different [programs]". Elizabeth detailed how for the UNICEF program, representatives from the Ministry of Education at the national level would also travel to counties to train education officers and teachers so they can return to their schools and implement what they have been trained to do. "Most programs [external sponsors] are like that", Elizabeth concludes, "because for them to reach *down there*, they have to begin at the national level *where we are* and then they cascade down".

Working with the government therefore secures access to a certain set of public infrastructures (e.g. schools and health facilities) and the people both staffing and operating it (teachers, headmasters, county directors of education, students, or health professionals). When,

therefore, Robert and Iris talked of the *government* as enabling the "highest level of scale", they were referring to that physical embodiment of the state in the form of public infrastructures designed to deliver services to the population: health and education in this case. Indeed, this government infrastructure is widely present across an entire country (even if not always effective or successful), making it possible to deliver an intervention (such as deworming) that the government might not provide were it not for the work of NGOs and their external funders, but in places where the government can (and does) reach thanks to its national infrastructural systems. In my initial description of the NSBDP above, I had outlined that the Year 3 Programme Report stated that it was "the readily available, extensive, and sustained educational infrastructure" that made it possible to successfully provide deworming tablets to children in the country. The report also thanked Evidence Action "for their technical assistance and fiscal management of the programme". From the point of view of the Kenyan state, insofar as it sees itself as delivering a national public health programme, Evidence Action is a technical assistant. Flipping things around however and looking at them from the perspective of the mandate of Evidence Action, the government is an instrument of scale that allows the NGO to reach millions of children *through* this governmental "educational infrastructure".

I refer to this public infrastructure as a sort of a *skeleton of the state* because skeleton conjures simultaneously sturdiness and weakness, the quality of being necessary and yet insufficient, of being both reliable and incomplete. The skeleton is an icon of essentiality: even after death, the skeleton remains a durable witness of past lives. At the same time, the skeleton is also precisely that: it evokes emaciation, a structure that is not complete, an indication of a lack and of potentially being even broken in certain places. In the same way, the NSBDP enrolls an existing governmental infrastructure, both physical and human, that nonetheless must be constantly mobilized.

When I met with him in his office close to the Kenyatta University Hospital in Nairobi, Adrian, program officer in the School Health section of the Kenya Ministry of Health, emphasized that the coordinating role of Evidence Action was central to the functioning of the program.

"The coordination is very critical in such a big programme", he says "and there's a lot of, how do I call it, there's a lot of effort being put to make sure that [it runs well] because this is a Programme being ran by two Ministries, and by the way, it's usually very difficult to coordinate two government departments to act together. Our partners [Evidence Action] they've really played a really key role in trying to make sure that, both the governments, the leadership, as well the mechanisms are really there, that we're able to deliver a programme".

Jane, Global Director of the Deworm the World Initiative, and Iris, Director of Communications at Evidence Action, had both concurred that their organization had to do most of the "heavy-lifting" because "there's no one person in Kenya whose full-time job is school-based deworming". In other words, education and health officials were there but they had to be trained to be capable of carrying the NSBDP, including organizing and carrying out national deworming days, overseeing drug transport and storage, accounting for drug usage, and reporting back through monitoring forms and expenditure claims. Schools had to be set up, adding temporary structures such as white tents and chairs, to welcome the NSBDP proceedings like the Release of Year 3 Results event with which I opened this chapter.

"In some cases, if there's no good coordination or understanding between the two ministries", Adrian adds, "sometimes there are programmes that don't run very smoothly and sometimes you will find that the program might even not take off because of the misunderstanding between the various coordinating government departments or ministries".

So this skeleton of the state, in the form of educational and health infrastructure, made-up of both buildings and people, has to be constantly trained, reminded, at times even nagged–

especially "if civil servants are disenchanted" as I was once told – to make it effective. Personal relations figure centrally here, because different individuals, such as Elizabeth from the Ministry of Education or Adrian from the Ministry of Health, might be differently committed to the program. Similarly, different governments and their ministry cabinets might attribute changing importance to the NSBDP across time. The state therefore emerges as a lasting and cumbersome infrastructure, one whose possibilities are used inasmuch as its limits are recognized.

But if I call the public infrastructure through which the NSBDP is routed as a skeleton of the state, it is also to bring attention to an instrumental relation to the apparatus of the state in evidence-based development. Implementing the NSBDP requires the state's operational capacity, but it is not premised on a perfect overlap between the aims and objects of government of the NGO and those of the state.

"We don't partner with other [than governments] for the Deworm the World Initiative", Iris had underscored in her role as Global Communications and Advocacy at Evidence Action. "Scale is only attainable if we work *through schools*, if we work with government as partners, in our view. So all the other organizations which work on deworming are not suitable, unless these organizations themselves already work with governments".

In Ethiopia for instance, Evidence Action has partnered with the Schistosomiasis Control Initiative (SCI), the latter being the primary entity who works with the government on questions regarding parasitic worm infections. Evidence Action provides assistance to SCI to implement the Ethiopian government's Schistosomiasis and STH Control Program. In fact, Evidence Action's Deworm the World Initiative extends beyond both the circumscribed Evidence Action's Kenyan team and the Kenya NSBDP. Deworm the World operates in multiple countries, through in-country offices that vary in size and scope, including in India, Kenya, and Vietnam. In all these countries, Evidence Action supports national programs by

working with governments, albeit through different levels of partnerships[19]. On their website, Evidence Action introduces the Deworm the World Initiative in the following terms: "It envisions a world where all at-risk children have improved health, increased access to education, and better livelihoods potential as a result of being free of intestinal worms."[20]

To meet its global goal of reaching millions of at-risk children worldwide, Evidence Action mobilizes the spatial reach of the state with its health and educational infrastructures made of people, buildings, knowledges as well as communication and transport routes. More specifically, the project to care for a national population (*the* people in the traditional sense) and the project of caring for global populations (millions *of* people around the world) are here operated through similar technologies of reach. Anthropologists James Ferguson and Akhil Gupta have discussed how the routine operations of state institutions (e.g. organizational charts, inspections, statistics collection) produce spatial and scalar hierarchies, in which the state is imagined as "reaching down into communities, intervening in a 'top down' manner, to manipulate or plan society" (2002: 983). Through these ordering functions, that also involve the routinized violence of monitoring and calculation, the state obtains a quality of transcendence to govern people and things in its territory (Das and Poole 2004b: 6). In the same way, the NSBDP program relies on a very hierarchical organization, what Elizabeth had called "cascading down", which runs from the national level to the local level. It is representatives from the Ministries of Health and Education at *the national level* who organize the programs through logistical meetings with local officials who they train toward implementing Deworming Day throughout the country. In turn, teachers, headmasters, county

---

[19] In India, unlike in Kenya, Evidence Action has nowadays a very light involvement and the organization intended to "phase out" completely in two Indian states at the time of my fieldwork. The Indian government manages all implementation costs (even if there is still light funding from external donors) for their national School-Based Deworming Programme (started in 2015) now that its operational guidelines are in place, which include how states can request a budget from the program at the federal level. Evidence Action states on its website that the 2017 national deworming day in India "treated more than 260 million children".

[20] https://www.evidenceaction.org/the-deworm-the-world-initiative, accessed April 28, 2019.

directors of health and education then report *back* to the national level through data collected, monitoring forms, and expenditure claims.

And when in 2013 Kenya moved toward a "devolved" system of governance, by creating 47 county governments in an effort to reduce inequalities of power and access to resources previously controlled by the central government, the NSBDP re-shaped its mundane activities (location and schedule of trainings and meetings, but also the information imparted at these meetings as well as reminder emails and SMS messages sent to government officials, budget allocation, transport of printed materials and drugs) to fit the redrawn national divisions, both administrative and physical. In sum, when NGOs like as Evidence Action work with governments to reach scale, they mobilize the very claims of overarching superiority and tools of vertical encompassment that serve to legitimize the authority of states as natural centers of power (Ferguson and Gupta 2002: 995).

I refer to these state-based technologies of reach as the skeleton of the state to foreground their importance and persistence even in visions and plans whose targets exceed a national population. There is therefore a certain necessity to work through the state, even if one's project has global ambitions. At the same time however, this is a skeleton of the state because it keeps alive an operational convergence while suggesting that the state's aims are not directly at stake, but are rather somewhat sidestepped and absent. These public institutions can be taken up to intervene upon a population in ways that can fit several aims, some of which might not completely align with those of the state. Rather, from the perspective of Evidence Action, partnering with the government to access health and educational infrastructures emerges as an instrument to reach millions of people, rather than the representative and executant of a full-fledged project of political sovereignty. Governmental means appear as *instruments* here to underline that they can be enrolled in visions that extend beyond and run alongside the state's project. In this portrayal, the operational capacity of the state and that of

Evidence Action are enmeshed. Governmental infrastructures are conduits to implement the NSBDP program, but on the other hand, they require the technical assistance of Evidence Action to carry out all the steps of the programs. In meetings and trainings about the program, and in schools during deworming day, we find both governmental staff and staff from Evidence Action who carry out similar tasks[21].

To say that in my examination of the NSBDP, the government emerges as an instrument of scale does not mean that the government takes up this role, nor that my interlocutors would agree with how I qualify their partnership with the Kenya Ministries of Education and of Health. This caveat must be made explicit to keep my argument empirically accurate. At stake is not to erase the significance of Kenya's nation-state sovereignty, but on the contrary, to caution against accounts in which "NGO strength is a function of state weakness" (Mann 2015: 2; but see also Escobar 1994; Ferguson 1990; Jackson 2005; Li 2007). Even as the government of Kenya was repeatedly tagged as corrupt and more concerned with doing profitable investments than with delivering welfare services to its population, it was also the preferred implementing partner for my interlocutors who recognized that the government allowed the highest level of reach and scale. In this analysis, the government appears as the bearer of crucial, if insufficient, resources, rather than emerging from a failed state.

In an effort to counteract the "taken-for-granted assumption" that the "neoliberal retreating state" is the explanatory background for the proliferation of non-governmental governance, recent literature has drawn attention to how NGO programs often reinforce the state's capacities or people's abilities to claim, expect, and use these capacities from the state (Brown 2015: 351; as well as Gupta 2012; Mahajan 2017). In her ethnographic examination

---

[21] In her article about a partnership between the Kenya Ministry of Health and US-funded NGOs to provide HIV/AIDS healthcare services, Hannah Brown (2015: 357) underscores how the NGOs often recruited staff who had previously worked for the Ministry of Health, insisting that this exchange of staff further blurred the distinctions between the "government side" and the NGO. In my fieldwork, I encountered a few people who had also started their career in different ministries, but it was not something that was heavily emphasized in daily work.

of a partnership between the Kenyan Ministry of Health and a US-funded non-governmental organization to support the delivery of HIV care, Hannah Brown concludes that this collaboration "in many respects expands the capacity of the state" by "working through state infrastructure to strengthen national health systems and deliver improved health services" (2015: 342 & 352). Similarly, when the NSBDP mobilizes governmental structures, it trains and transforms governmental resources into conduits for more effectively tackling the issue of parasitic worms[22]. If programs like the NSBDP solidify the reach of national health and education systems, or in historian Gregory Mann's words, "reinforce African governments rather than work around them" (2015: 207), can they be described as "neocolonial", imperialist or neoliberal – in short, as eroding nation-state political sovereignty? The remainder of this chapter attempts to make visible how difficult it is to answer this question with certainty. This reveals the need to conceptualize, as have done other authors before me, transnational forces not necessarily as jeopardizing state sovereignty, especially in post-colonial settings, but as integral elements of state formation (Gupta 2012; Mann 2015). This gestures to a complex conception of political sovereignty, at the same that it qualifies the real power and impact of global non-state actors on the designs of the state.

**Enmeshed plans, (con)fused aims of government**

Early on in my fieldwork in Kenya back in February 2016 during that logistical meeting in Makueni, the audience composed of local officials from the Kenyan Ministries of Health and Education had been asked: "What are the benefits of deworming?" After the long silence that followed the question, two people finally expressed what I was thinking in my head and what had been presented earlier by the training team, who made reference to evidence drawn

---

[22] I do not delve here into the debate around whether mass deworming as a national policy is ethically questionable insofar as deworming has become a priority in Kenya's national plan for the control of neglected tropical diseases (see the next section in this chapter).

from randomized evaluations: *to increase school presence and improve children's health*. Shortly after, another man confidently added: "to have a healthy nation". As could be expected, the government staff I encountered in fieldwork presented the NSBDP as a traditionally (biopolitical) project, concerned with fostering the lives of Kenyan citizens.

Adrian, the program officer at the Ministry of Health introduced above, directly portrayed mass deworming as a responsibility of the Kenyan state to its population.

"We want the counties [in Kenya] to have a buy-in, to understand what are the negative effects of worms on children, health and education; and children are the future of the country. When you talk about Vision 2030 [Kenya's national development program], or the pillars of economic development, you can't have them when you have a sick nation, and children is a foundation of that. So the counties have to understand that, so they can allocate some budget to be able to do this particular program [once funding from external partners ran out]."

As you might remember, the Kenya NSBDP Year 3 Programme Report outlined that "the goal of the programme is to remove parasitic worms as a public health problem for children in Kenya".

If the state's responsibility and the health of a national population what was at stake in official documents and for government representatives, a more intricate picture of the NSBDP emerged in terms of responsibility, ownership, and aims of government when looked at from the perspective of Evidence Action's mandate. I stated above that Evidence Action introduces the Deworm the World Initiative in the following way: "It envisions a world where all at-risk children have improved health, increased access to education, and better livelihoods potential as a result of being free of intestinal worms." The next sentence on the webpage reads: "We work in close partnership with governments to enable elimination of intestinal worms as a

public health problem." In this description, there are at least two distinct visions, or plans, that overlap[23].

There is a world plan: a world where children are no longer infected by intestinal worms and can lead "better lives", with improved education and higher income over their lifetime[24]. Here deworming is a global cause, one which mobilizes non-governmental as well as governmental resources. There is a national plan that also entails both governmental and non-governmental means: intestinal worms are a public health problem for nation-states, whether in Kenya or India. The world plan is therefore enacted through nationally contained programs, like the NSBDP or the Indian National Deworming Day, which are framed in terms of what are traditionally a state's preoccupation (public health or healthy nation). At the same time, the NGO looks out to implement new deworming programs in other countries by determining "contexts" where programs could be successful[25].

I suggest that this description muddies the distinction between the aims of government of Evidence Action and those of the state. It seems like Evidence Action is invested, materially but imaginatively as well, in both the world plan (reduce worm prevalence in children around the world) and in the national plan (address parasitic worms as public health problem for each nation-state). In other words, Evidence Action's mandate presents their global visions and the state's designs as *enmeshed* into one another, whereby it is not only that Evidence Action enrolls the state's technologies of reach to deliver deworming treatment to millions of children, but that at the heart of its global aims lie national concerns. As an analytical tool, enmeshment complicates vertical "topographies of power" in which the exercise of power is understood essentially in scalar terms, such that actors like NGOs or international organizations are

---

[23] Anthropologist Tobias Rees (2014a: 458) has underscored how "planning oscillates between the feasible and the fantastic". I thus use this term "plan" to indicate that Evidence Action's description of its work, and how it enrolls the government and imagines the role of the state therein as both real and aspirational.
[24] See Chapter 1 for the details on the concept of "better life".
[25] Chapter 4 of this dissertation addresses my interlocutors' understanding of context and generalizability as expressed through their attempts to scale-up and adapt a similar program across the world.

conceptualized to hover above the nation-state, just as nation-states are taken to encompass society (in the forms of regions, towns, and villages) to suggest instead that the programs of Evidence Action and that of the state are coterminous and operating at the same levels (Ferguson and Gupta 2002: 990; and Ferguson 2006b). As I have already sketched with the example of the Kenya NSBDP, the programs whose implementation I followed are often ran through and with national capacities, mobilizing ministries and their resources. The NGO relies on the state's operational capacity and spatial encompassment to achieve its aim of reaching millions of children with its Kenya Deworm the World Program.

In the rest of the chapter, I wish to illuminate how this convergence of technologies of government in turn results in *both* agreements and conflicts of objectives, both between the government and the NGO, but also within the government[26]. One might be willing to fix a certain analytical frame on government-NGO partnerships, but something is to be missed if one conceives of the Kenya NSBDP as solely a state project or, conversely, a global project exclusively. Below, through an examination of advocacy practices, the framing of program goals, and debates about sustainability around the NSBDP, I demonstrate how this enmeshment of goals and responsibilities at the heart of evidence-based development sheds confusion around who governs and what is being governed, forcing us to ask what today is government (Mann 2015). Along with the anthropologist Hannah Brown, I contend that a juxtaposition of these different visions strewing NGO-government partnerships "raise[s] fundamental questions of sovereignty and governance", like "who has the right to act on behalf of Kenyan citizens? Who should be in a position to oversee interventions that responded to

---

[26] I am grateful to the following blog post by economist Jean Drèze for helping me articulate my curiosity about the diversity of objectives present even at the core of one single program, which significantly complicates the task of advising governments as researchers. Drèze (2018) gives the example of "including eggs in midday meals [which] may be good for the government in terms of its stated objective of improving child nutrition, but bad in terms of re-election prospects (for example, because of the power of vegetarian lobbies)."
http://www.ideasforindia.in/topics/miscellany/evidence-policy-and-politics.html, accessed April 30, 2019.

public entitlement?" (2015: 437). In this light, determining what is the state's jurisdiction and what is *not* emerges as an analytic task riddled with ambiguity.

*Evidence-based advocacy*

The Kenya National Deworming Program and the Indian ones did not exist before the work of advocacy carried out by a tight yet diverse group of people[27]. I can provide a condensed timeline of events as they were told directly to me, but also as I read them in reports and blog posts.

As mentioned in Chapter 1, from 1998 to 2001 economists Michael Kremer and Edward Miguel led the randomized evaluation study of a deworming program run through a small Dutch non-government organization then called International Children Fund[28]. In 2004, the study results appeared in the peer-reviewed *Econometrica* journal. Late 2005, the Jameel Poverty Action Lab (J-PAL) authored *Fighting Poverty: What Works?*, a policy brief "which compared the cost-effectiveness of various education interventions and showed that deworming was the most inexpensive way to increase school attendance" (Ashraf, et al. 2010: 6). In 2007, Michael Kremer, along with Esther Duflo (co-founder and director of J-PAL), attended the World Economic Forum in Davos and suggested to the Young Global Leaders to make deworming a central component of its education platform. In that wake, Michael Kremer and Rachel Glennerster, long-time Executive Director of J-PAL, launched Deworm the World, a "more formal organization" to get deworming programs implemented around the world[29]. At that point, Deworm the World was hosted under Innovations for Poverty Action, a research

---

[27] In a careful exploration of global health practices around immunization in Africa, Ehrenstein & Neyland (2018) also detail how evidential practices are closely associated with the enrollment of political support through advocacy techniques. I borrow the term "evidence-based advocacy" from their article, which they also draw from Storeng and Béhague (2014), to serve as the title of this section.
[28] For details about why this randomized evaluation was started and how it was carried out, see Chapter 1.
[29] In January 1, 2018, Rachel Glennerster took an extended leave of absence from J-PAL to join the UK Department for International Development as its chief economist.

(and policy) institute (with offices in at least 15 different countries) that carries out randomized evaluations and disseminate their findings[30]. In parallel, the Kenyan Medical Research Institute had been running a pilot deworming project, in collaboration with the Japanese Development Agency, in the central region of Mwea since the early 2000s.

In Kenya, efforts to establish a National School-Based Deworming Programme occurred throughout 2008 and 2009. It rallied many people: Lesley Drake, then Executive Director of Deworm the World; Valery Levy, then Kenya Country Director of Innovations for Poverty Action; Margaret Ndanyi, Head of the School Health and Nutrition Programme at the Ministry of Education in Kenya; Don Bundy, Lead Specialist for School Health HIV/AIDS & Education at the World Bank, who had himself conducted research on the links between deworming outcomes and education and supported earlier and smaller scale deworming initiatives in Kenya; and Simon Brooker, a tropical diseases epidemiologist who "began a worm mapping project in Kenya" (Ashraf, et al. 2010: 9).

The advocacy efforts "were formalized" in April 2009 with the roll-out of the first year of the National Program, supported by drug donation programs and the Kenya Education Support Programme, a pool fund from different donors such as the World Bank and the UK Department for International Development. The Memorandum of Understanding delineated the role of each Ministry in the program. Larry, one of the earliest members on the Kenya Deworm the World team at Evidence Action, had told me that "the advocacy had placed the national school-based deworming program under the Ministry of Education for the implementation because it was school-based, while the Ministry of Health was to provide technical guidance". In January 2009, Raila Odinga, then Prime Minister of the Republic of

---

[30] It was only in 2013 that Evidence Action was created as an NGO focusing specifically on scale-ups of "successful" interventions (as assessed by randomized evaluations) into large programs in multiple countries. For more information, see also the Introduction of the dissertation, section Field sites.

Kenya, announced that Kenya had committed approximately $1 million to target school children most at risk for parasitic worm infection", with the following speech:

> Every child born in our society irrespective of his or her social or economic background should have an equal opportunity to realize their potential…This cannot happen if children…are...infected with worms and cannot…attend classes…This is…why the government…initiated this program. We want to have a healthy nation and we are very happy with this assistance that we are receiving in our campaign against worms (Ashraf, et al. 2010: 1).

In the years to follow, national deworming was to be entrenched in several official policy documents, such as the National School Health Policy Guidelines, pertaining to the Ministries of Health and Education and the National Multi-Year Strategic Plan for the Control of Neglected Tropical Diseases pertaining to the Ministry of Health.

As for India, "in 2010, J-PAL talked to the State Director of Health in Bihar and that's how we got started in Bihar, through those conversations", Jane, Global Director of the Deworm the World Initiative, told me. Searching online, I found information to specify what Jane meant:

> In 2010, J-PAL hosted a conference with the Government of Bihar, a state in India, to share evidence on school-based deworming and other promising programs. Following these discussions, the Government of Bihar, with support from Deworm the World, launched a state-wide deworming campaign in 2011, which reached 17 million children. Other large-scale deworming programs in India were implemented in Andhra Pradesh, Delhi, and Rajasthan. In February 2015, the Government of India launched a national deworming program and in October 2015, India announced it had treated 89.9 million children in 11 states since the launch of the campaign.[31]

In sum, extra-governmental energies and structures lobbied and worked to put in place national programs to tackle an issue such as worm infections. Without these (extra-governmental efforts) the issue might well have been a general preoccupation, but perhaps it would not have come to form the state priorities, nor would it have become a target where the state allocated attention or resources. This story, I want to suggest, sketches non-governmental

---

[31] https://www.povertyactionlab.org/case-study/deworming-schools-improves-attendance-and-benefits-communities-over-long-term, accessed April 30, 2019

action and global visions as the engine of national plans. Given this, it is important to ask whether the NSBDP should be read through the light neocolonial interference into the state's affairs on the part of NGOs that certainly have offices in these countries but are headquartered in the USA.

But for Adrian at the Ministry of Health, the problematic interference of Evidence Action into what belongs to the state was situated somewhere else than in this non-governmental shaping of the national policy agenda. When I had asked for his overall assessment of the NSBDP, Adrian had replied with a few areas that needed improvements, including this one: "When you have a partner implementing a program, sometimes they forget that this program is a government program. They tend to think it's their program, so sometimes they may even want to dictate some things to the government. Especially, we're doing our fourth year of the program. People forget where we've started, the Memorandum of Understandings, what is the role of the government, the partner, so sometimes we find 'you may overstep with your role'. But the good thing is that we have a good communication to tell each other politely that whatever it is it's not ok. But that's human, sometimes it's allowed to act [like that]."

According to Adrian, the NSPDB is definitely a state program and his words cast doubts over what, if anything, belonged to Evidence Action. By indicating that he sees it as problematic that the NGO partner would want to dictate things to the government, Adrian is placing Evidence Action back into its role as the technical assistant of a program materially and imaginatively in the hands of the national government. In so doing, Adrian sketches the NSBDP as first and foremost a state project, one that falls under the jurisdiction of the state and embedded in the state's design of a healthy and educated population. "You can't have national economic development", you might remember that Adrian had said, "when you have a sick nation". If non-governmental actors had lobbied the government of Kenya to put in place

the NSBDP, it was now Adrian who was trying to "lobby or advocate the counties to take this program as [we] are trying to develop sustainable mechanisms once the partner exists".

Put differently, Adrian hinted at an intrusion but one that seems of a very different nature than accusing non-governmental forces of shaping national priorities and governing instead of the government. By describing Evidence Action's interference as overstepping their role, Adrian indirectly banalizes that interference simply as a to-be-expected occurrence in all forms of collaboration because "it's human to act like that" and it is something that can be addressed through "good communication". I know that I am treading on slippery terrain with this analysis, but I attempt it because I believe it has important repercussions for how to understand political sovereignty. Taking Adrian's view at face value, one might be ill advised to conclude that the work of Evidence Action erases the state's sovereignty, even if the NSBDP is the direct result of advocacy efforts carried out by organizations of evidence-based development that explicitly wanted to influence the policy agenda. There are times indeed where it is the state's goals that sidesteps that of Evidence Action and other organizations engaged in evidence-based development.

*Pragmatism and the state's concerns*

"In the case of India, school attendance might not be the most important priority for government, but health instead", Iris, Global Director of Advocacy and Communication at Evidence Action, said to me in an interview.

Nevertheless, the causal impact of deworming on improved educational outcomes, as reported by randomized evaluation results, were an essential grounding to the creation of the Deworm the World initiative at Evidence Action (see also Chapter 1).

"Who are we to quibble with the Indian government's reason for implementing the deworming program?", had asked Iris, rhetorically. "As long as a high-quality program can be implemented, we're going to be pragmatic about it, not dogmatic".

Iris detailed two other programs that Evidence Action is trying to bring at scale, one in Kenya and one in Bangladesh, to show how the reasons for supporting the program might diverge between the NGO and the government[32]. The G-United program in Kenya is based on evidence gathered from randomized evaluations of remedial education across India and several African countries that briefly trained volunteers to help with learning in classrooms significantly improve educational outcomes for students. Launched in 2014, the program is implemented by the Kenya Ministry of Education with governmental funds and through public procurement channels, with the technical assistance of Evidence Action. On the official website of the G-United Program[33], no mention at all is made of this evidence base. Instead, the website explains how the program centers on college graduates providing remedial education to pupils who have been identified as requiring extra support across schools in Kenya.

"The Kenyan government has an interest in finding meaningful occupation for college graduates who are unemployed and to foster national unity. To increase coherence between tribes. That's how the government is promoting the program within its own government structure. And we're ok with that so long as the educational outcomes are achieved. If that's politically relevant…" Iris underscores what she means by being pragmatic rather than dogmatic. "We [Evidence Action] are going to stay true to our interest, that is [improved] educational outcomes [for children who receives the remedial education], but if the

---

[32] This is a different focus than the difference between evidence and implementation, or how programs are being "tweaked" or "tinkered" with as they are being transformed into large-scale programs across different "contexts", an issue tackled in Chapter 4. I also do not delve here into whether the programs, once scaled, have exhibited positive, mixed, or no results – a question relevant for all programs, beyond evidence-based development. Here I am interested in how a program's aims of government might be differentially formulated.
[33] http://g-united.or.ke/, accessed May 1, 2019.

government wants to communicate to the population that this is a meaningful experience for college education and it increases national cohesion…"

Iris described cases in which the end goals of the NGO and those of the state might diverge (education versus health; education versus employment and national unity) but one finds common grounds in programs capable of filling both objectives at once. We return again to this idea of one program (whether deworming or remedial education) that is "caught" between two plans: one national and one extra-national. This means however that in quite explicit ways global ambitions must inhabit national aims and frame their own work through the state priorities, even when the NGO does not directly have those national aims at stake.

Iris added as a last example: "in Bangladesh, we are currently testing our No Lean Season program at scale. This is a grant-based program to encourage poor rural laborers to migrate to a nearby city in order to find a job and increase their income in the period between planting and harvesting. The 2008 randomized evaluation study was very focused on seasonal hunger and famine."

I met the Bangladesh-based staff of Evidence Action and of RDRS, the very big Bangladesh-based NGO, who is their implementing partner for the No Lean Season program when they visited Kenya in July 2016 to share experience and knowledge with their Kenyan counterparts. During a presentation, the RDRS representatives explained that the small size of farms in many regions of Bangladesh resulted in seasonal famine that could be mitigated by going to larger towns during the low agricultural season to gain income and access goods in town markets.

"The government [of Bangladesh] doesn't like to talk about famine" Iris told me, "so now we and our partners are framing the program as income vulnerability…with the recognition that there are now government food aid programs which tackled food security and are making a good strive in reducing famine. So if we were today talking about [the No Leason program]

the way we used to in the past, that is as famine, we wouldn't go very far. This is what I mean by being pragmatic. The 2008 emphasis on caloric intake may not be true anymore. We would be clueless to talk about this as an anti-hunger program. The same way the government of Kenya won't be talking about G-United as focused on education alone."

Following Iris' description, what is within and beyond the goals of Evidence Action becomes more ambiguous. An initial focus on seasonal hunger is reframed as an emphasis on income vulnerability to take into account Bangladesh governmental programs about food aid and to align with the state's shifting priorities. The National Deworming program in India might be presented as primarily a health program by the government of India even if Evidence Action promotes mass deworming not only for health reasons, but also to enhance educational outcomes and longer-term income effects. Iris concluded: "These things are not static; and we do live and upgrade in political contexts. And it can come with scale and evidence. We can't only do political contexts."

This last sentence encapsulates the ambivalence around the NSBDP's work and Evidence Action's work more broadly. When Iris says, "we can't only do political contexts", she seems to be suggesting two different though related ideas. First, Evidence Action's work cannot focus only on the national level because their ambitions extends across multiple countries to cover the world's poor. Second, Evidence Action's objectives cannot be (entirely) dictated by what politicians, and governments, want. It must and does keep its twin central concerns with *evidence* and *scale* at the forefront of determining what programs to take of as part of its portfolio. Evidence Action takes on programs backed by evidence derived from randomized evaluations and that have the potential to be scaled to millions of people[34]. At the same time, Iris suggested that they *must also* take into account the political context – or country

---

[34] See Chapter 2 for more details on this "potential to reach millions of people" criterion, called scalability in Evidence Action's strategy.

– within which they are implementing a program, especially because several of their programs require state capacities to exist at scale. In Iris' terms, it would be "clueless" to talk about the No Lean Season as an anti-famine program or to refuse G-United's focus on national cohesion, suggesting that to erase or hover above state priorities would only lead to failed partnerships, atrophied programs, or the exclusion of Evidence Action as a relevant actor in large-scale, government-based development programs.

While this might reveal another instrumentalization of the state – to enroll the state's political goals in an NGO-based program not invested in national unity, for instance, but in order to reach millions of people with the G-United program – it is not really clear who is being instrumentalized here, the government or Evidence Action. Political sovereignty is not dispensed with as advocacy efforts and a pragmatic outlook leave, quite explicitly, space for the government to claim these programs as its own, or at least harbor them in the language of what matters to the state (e.g. healthy population, national unity). In this way, the government might affix the state's goals to programs delivering services that do not entail governmental funding, but yet that publicly appear to foster the state's concerns. As I stated in the early pages of the paper, this shows how the project of evidence-based development and the state's project implicate each other, in ways that cannot be easily captured and even less easily explained away by either one being subsumed by the other.

While I was in the field, two organizations closely linked to Evidence Action, Innovations for Poverty Action (IPA) and J-PAL (both mentioned above) were crafting a new organizational strategy, one in which the priorities of the government would inform the work of evidence-based development in a different, more upstream way.

In an interview with Jilal, Kenya Country Director of IPA, he said: "the approach has changed: we don't wait until the end of a project to go to them [the government] and present results, but now right from the beginning, we involve them – we build a policy plan from

beginning. We minimize pilot projects, proof of concepts projects - now we favor piece of work which links to policy."

In other words, IPA but also J-PAL, through its recent Innovation for Government Initiative, *design* programs to be evaluated in collaboration with governments to therefore directly assess programs or policies relevant to governments. Jilal gives me an example:

"In education – we give a blank check to governments; we ask them to identify gaps that can be filled by doing research, e.g. curriculum; [the government of Kenya] had a curriculum which doesn't meet demands."

Jilal is making reference to a *curriculum reform* that the Kenyan government was currently carrying out as part of Vision 2030, the country's development plan. As I learned from Prof. Simuru, then Director General of the Vision 2030, the curriculum reform was a multidimensional plan to reduce inequalities in educational outcomes in the country. Here, partnering with the government is not only about securing access to a certain set of public infrastructures to scale one's pre-existing program. Instead, the government is a partner in the creation of programs pertinent to specific "political contexts" as both IPA and J-PAL "work with governments on policy priorities they have identified"[35]. J-PAL's Innovation in Government Initiative focuses on the sectors of education, health, and social assistance for which there is, it says, a "high government demand" while explicitly stating that they are open to extend to other sectors where there is (government) demand. If the foregrounded goal of these organizations engaged in evidence-based development, just like the mandate of Evidence Action, remains "to improve the lives of millions of people living in poverty in developing countries", this is done in this case by harnessing the state's political goals in more direct ways. It suggests that global ambitions cannot be separated from the state's project if evidence-based development is to enact lasting and large-scale change.

---

[35] https://www.povertyactionlab.org/innovation-government-initiative, accessed May 2, 2019

This outlook complexifies the picture I drew above of *the government as instrument of scale* whereby the NGO needs to enroll the state's public infrastructures to reach millions of people and enact a world plan. Above, I proposed to conceptualize this public infrastructure as a *skeleton of the state* to emphasize that these public institutions can be taken up to intervene upon a population in ways that can fit several aims, some of which might not completely align with those of the state. Examining the state as a "political context" whose priorities must be reckoned with in the framing of existing projects (education vs. national unity), or in the design of programs to be evaluated, adds flesh to this skeleton. At this point, the state is more than merely an infrastructural conduit for programs packaged elsewhere. The movement for evidence-based development demonstrates a certain trust in the state as a designer for "better lives" insofar as governmental priorities themselves are potential pathways toward global poverty alleviation. This sketches a recognition of the nation-state as the dominant political formation, both in terms of pragmatism and legitimacy, because implemented programs are presented around the state's traditional preoccupations and programs to be evaluated are imagined through governmental concerns.

In his last book *Red Tape* about bureaucracy and poverty in India, anthropologist Akhil Gupta (2012: 241) shows how the Indian state, whether in its moment of "high sovereignty" under Prime Minister Nehru or today under "neoliberal globalization", "has always been inflected by transnational processes and ideologies". Whereas centralized, socialist planning ideologies influenced the state's development agenda in the 1950-60s, now it is being shaped by "global neoliberal ideas" about the reliance on markets and self-help. In his history about the entrance of NGOs in the West African Sahel, historian Gregory Mann (2015: 244) makes a similar point, suggesting that in the wake of independence, "once imperial relationships simply became transnational ones". Mann furthermore underscores that "in the 1970s, international NGOs had not muscled their ways into newly independent African states

uninvited", stating that "at least some of them were sought out" (2015: 243). In his book, Mann describes how organizations such as the humanitarian agency CARE shifted its work from an initial focus on relief during famine emergency toward an interest in development through various programs about education and health. Both Gupta and Mann propose to shift from a picture of NGOs and their influence as erasing political sovereignty to conceptualizing the state as a political formation defined by transnational forces.

I contend that non-governmental actors and the state *mutually inflect* each other in the work of evidence-based global poverty alleviation I studied, even if not necessarily in equal nor equitable ways. J-PAL and IPA enroll the government to *determine* and *design* relevant policies to evaluate. In turn, organizations committed to evidence-based development *lobby* the government to take up programs found to be successful. Might this sketch a picture of a less bounded, less impermeable sovereignty? Might we be able to sidestep territorial boundaries as markers of legitimacy, in such a way that local determination from within might not be necessarily the preferred imaginary? In the last section of this chapter, I zoomed back to the Kenya NSBDP to explore how discussions about the sustainability of the program enmesh Evidence Action's future and the priorities of the nation-state, showing state sovereignty under yet other guises.

*Sustainability*

When I attended the Launch of the Second Kenya National Strategic Plan for Control of Neglected Tropical Diseases in June 2016, an event mentioned at the beginning of this chapter, government officials made frequent references to the NSBDP as a featured and successful program toward the eradication of neglected tropical diseases (NTDs) in the country, part of which include parasitic worm infections. In the 136-page long Second National Plan document which outlines the national multi-year strategy for the elimination of NTDs

over the years 2016 to 2020, Evidence Action and the Kenya NSBDP are discussed in the first few pages to underscore the number of children reached by the program across the years: 5.9 million in 2012/2013, 6.4 million in 2013/2014 and 6.1 million in 2014/2015 as I had learned during the Release of Year 3 Results. As I alluded to earlier, the Kenya NSBDP is entrenched in two national policy documents under the umbrella of both the Ministry of Health and the Ministry of Education. Despite this, when I was doing fieldwork in Kenya in 2016, the future of the NSBDP seemed rather quite uncertain.

The program was completing its fourth year of implementation and donors' funding was to run out at the end of the fifth year. Given this, the financial sustainability of the program beyond 2017 was a recurring topic of conversation during my fieldwork, as much in formal meetings as over lunch breaks. At the Release of Year 3 Results, Sheila, director of Kenya's Deworm the World at Evidence Action, had expressed that her team's main concern about the program is "what will happen when we are not there?" She said she was explicitly calling on the government to take the lead. Like others, she shared the idea – and the hope – that in the long run, there would be a government's "budget line" for deworming, meaning that some of the government budget would be allocated toward the NSBDP on an annual basis. At the beginning of my fieldwork, Eric in the monitoring department at Evidence Action had told me that the goal of his NGO was to make itself "obsolete" with regards to the NSBDP and that the program would be fully *institutionalized* when the funding by external funders ran out in 2017[36]. The Year 3 Programme Report, issued by the Ministries of Health and Education, also indicated that beyond the current five-year implementation phase, "sustainability can be achieved through the institutionalization of programme components within existing

---

[36] Of course, there is the widespread idea that NGOs, like Evidence Action, have incentives *not* to institutionalize programs and become obsolete, because this means losing operations and jobs in a country. What I am trying to make visible however is something different that are the different views about the program's future and how these are tainted by uncertainties, both for the NGO and the government officials.

governmental structures - organized, resourced, and coordinated at the national level and implemented locally by the counties" (Kenya Ministry of Education and Ministry of Health 2016: 9). To become part of the government's budget is often taken as a symbol of the state's commitment to a program and a marker of success for having created a relevant program.

It was not clear however whether Sheila's call to the government was being heard, or by whom. There was a general expectation among staff at Evidence Action, accompanied sometimes by both worry and resignation, that it was very "unlikely" that the government would take up the program in its own hands in 2017, if ever. And these feelings and concerns were echoed within the government itself, at least by the officials I interviewed in the Ministry of Health and of Education. Elizabeth, the quality inspector at the Ministry of Education, wondered what priority would be attributed to the NSBDP if its funding and management were handed over to the government.

"No, you see the problem we have is that we have very many programs going on at the same time," Elizabeth explained, "so my worry is: if it is left to the ministry, it will have to be looked at in line of priority areas of the ministry so I'm worried, I don't know which [priority] number it will take… It won't be the first…I'm not saying it won't be, but I'm worried about the position it will take in the priority list of the ministry because if it doesn't get the top priority then…what will happen".

Elizabeth underscored that education was a very expensive undertaking and the government did not have enough funds to do everything.

"Now, as we're talking, we have some children that are out of school; schools are very far, they cannot reach schools: that could be a priority. We have the issue of girl child: being out of school because of poverty, so it keeps [these girls] away from school; so those areas…have to be prioritized. So I don't know, at this moment, I can't say where the deworming would be put [in the list of priorities]".

On the one hand, Elizabeth's words suggest a typical story of partnership in which a cash-constrained government is not be able to sustain a development program once external funding is not available anymore. Several scholars have underscored this "temporal imaginary of global health" in Africa, wherein many health interventions come in the form of projects, externally funded but only for short time frames of a few years (Whyte, et al. 2013). According to anthropologist Uli Biesel (2014), these projects often take the form of "vertical interventions" focused on a single disease, such as parasitic worm infections or malaria, rather than more holistic projects geared toward health system strengthening. According to Biesel, these "spatio-temporal logics of global health" result in forms of care that are patchy, spatially and temporally bounded, and that exclude many places, illnesses, and people. "Under the label of evidence-based medicine", Biesel (2014) writes, "new intervention strategies and technologies are trialed, scaled-up and then all too often handed-over to Ministries of Health that lack the financial and operational means to sustain the interventions (note, most Ministries of Health do not lack the expertise!)". In this story, a weak state has received external, prepackaged, and possibly an uninvited assistance for its health or development projects, but that only covered a limited number of people because it will be stop sooner rather than later. The consequences of this short temporal horizon of a *projectified* provision of welfare services is crucial to underline and speaks importantly to the uncertainties around how, and if, the Kenya NSBDP was to survive beyond 2017.

But there is also another way to tell the story that foregrounds the government of Kenya less as *unable*, but rather as more as *unwilling* to carry on the NSBDP into the future, as a way to counteract social science writings that focus on how global non-governmental actors erode state sovereignty by distorting governmental priorities. Indeed, there is something specific about the deworming project supported by Evidence Action in Kenya because it is also *at the same time* the *National* School-Based Deworming Program. On the one hand, the Kenya

NSBDP is what Biesel calls a "vertical disease project" – focused only on parasitic worm infections – and not on reforming national health or educational infrastructures. And yet, on the other hand, unlike the programs Biesel describes as examples (such as the Gates Foundation's work on malaria), the Kenya NSBDP does *not* circumvent "the messy realities…of local infrastructures" (Biesel 2014), even if it requires an external team to regularly mobilize these. The Kenya NSBDP is furthermore entrenched in two state policy documents as you might remember, the National School Health Policy Guidelines and the National Multi-Year Strategic Plan for the Control of Neglected Tropical Diseases.

Elizabeth's words therefore can also be read through the lens of the state's designs. Funding still figures centrally in the story because, as Elizabeth said, the government has a limited budget and so ministries must decide which (welfare) programs to allocate funds. In this narrative, the Cabinet of the Ministry of Education is an actor who, as Elizabeth portrayed it, "sits, discusses, and prioritizes wat policy matter" to the country. The government also adapts and updates its priorities as the country's situation, with regards to health or education for instance, changes across time and places. Jane, the Global Director of the Deworm the World at Evidence Action, had told me that a very positive point for the sustainability of the Indian national deworming program was that the government had established deworming as a *national priority.* This had indeed facilitated moving the program into the hands of the government, both for operational guidelines and for how Indian local governments can require budget from the program at a federal level. This way of telling the story sketches a Kenyan (and Indian) state which, like other nation-states around the world, must govern with limited resources, prioritizing this or that, based on a multitude of factors that have to do as much with empirical knowledge as with a party's political program. Here, sustainability emerges clearly more than about money only; it is also a direct function of *political will.*

In her discussion about the Gates Foundation's work on HIV/AIDS in India, anthropologist Manjari Mahajan (2017) has also underlined the mismatch between the public recognition of Gates' supported Avahan program and the government's questioning of its sustainability. Mahajan explains how the Indian government could not inject the type of resources that had been injected into Avahan, but also *did not want to* because there were many other demands on government funds and because "its own prevention programmes, operating on far more stringent budgets, had also been successful in slowing down HIV transmission rates" (2017: 9). Mahajan aims to paint a picture that shows the Gates Foundation's "qualified sense of impact in India": it is still helpful for addressing some health issues, but it does not exercise hegemony over the government by shifting its priorities. Mahajan concludes that "donor programmes come and go" but that government "programmes, however flawed, have a staying power that far exceeds that of most of global philanthropic actors" (2017: 10).

While I take cues from Mahajan's portrayal (2017: 9) of the government more as a "player in its own right, and less as a weak victim, within global health exchanges", I want to emphasize that the distinction she assumes between donor programs and government programs might not be so obvious. Wherein funding comes from external donors but the operational capacity relies heavily on public infrastructures both physical and human, the Kenya NSBDP complexifies a clear-cut division between what belongs to the NGO and what belongs to the state. Given this, which entities—if any—hold the responsibility for the future of the NSBDP?

Among my interlocutors, some explored alternative paths of sustainability for the Kenya NSBDP beyond its complete institutionalization by the government, spelling out the many reasons why the government of Kenya might be unable, or unwilling, to take on the Kenya NSBPD. It might have to do with financial means. But some hypothesize it might have more to with the overly bureaucratic procurement service mentioned earlier in the chapter, which might make the program undoable not because it is such an expensive program but

rather because adding a "budget line" in the government budget is an incredibly complicated process. Others—including some working in government—suggested that even if the funds were attributed to the program, this would not prevent the government from "playing politics", referring to fears of corruption, of politicians keeping all the money rather than sending it to counties to run the NSBDP.

Valery, Senior Director of Innovation at Evidence Action, rejected the "implied moralistic argument that if governments cared about it, they would take it up". In her view, this argument was premised on a sidestepping of key structural problems at the core of the Kenyan administration, like the public procurement system. As Valery explained to me, Treasury funds are disbursed separately to the Ministry of Education and the Ministry of Health and cannot cut across ministries without always going back through the Treasury first. For an intervention like the NSBDP, whose implementation depends on these two ministries, this might entail very long delays to access and use funds. As a thought experiment, Valery proposed that it might be more expensive to fix the procurement system in Kenya than to fund the NSBDP through external donors for 20 or 30 years, after which there might not be a need for mass deworming in Kenya anymore as prevalence and incidence rates would have gone down, especially as Kenya might have by then become a middle-income country. In this imagined scenario, donors acquire a longer *staying power,* not of a few years as in Biesel's and Mahajan's discussion of donor programs, but of a few decades.

Valery in fact was holding donors responsible for continuing to financially support the Kenya NSBDP because when donors leave after short periods of time, it is as if they suppose governments "don't make trade-offs" and can suddenly just fund the programs. Echoing Uli Biesel's observation that development interventions are handed to ministries without consideration for their feasibility through a state's means, Valery asked me rhetorically, "Where should the government take the funds for the deworming programs? From its current

malaria program? From more taxes? Or should it be less corrupt, as many people say?"

Valery's words place the future of the Kenya NSBDP within an awareness of complicated state political processes that include bureaucratic configurations, policy priorities, acceptable taxation schemes, and also the country's famous history with misappropriation of public funds through the government's tender and procurement system (Wrong 2009). Precisely because the state's project is infinitely bigger than the circumscribed NSBDP, Valery contended that hoping for the government to take the deworming program in its own hands was the wrong goal. Instead, at stake was to develop a cost-effective program that could reasonably be supported by several cycles of donation across many years.

Building upon Valery's view, it is almost as if there are multiple bearers of longevity, given that Evidence Action and its donor partners task themselves (rather than the state) to care for the circumscribed issue of deworming in the longue durée, hopefully until its elimination. Certainly, the jurisdiction of the state in terms of the responsibility to care for its population is incommensurably other and broader. Yet Valery's thought experiment suggests that a recognition of a state's overwhelming priorities might demand to allocate this or that responsibility beyond the government. In this analysis, the state is not so much erased as it is complemented, but not necessarily, or rather not only, because it is weak or failed, but because resources, financial and otherwise, are always limited, whether one talks of the Kenyan state or any other state. As I write these words, there has been a new partnership signed between Evidence Action and the Government Kenya over the NSBDP through 2022, with funding from external donors (one being the same as in the earlier round, and the other being a pool fund of individual donors) but estimates indicate that the funding secured so far is only enough to sustain the program until 2020[37]. Here, the nation-state might neither be the bearer of all

---

[37] https://www.evidenceaction.org/blog-full/5-years-kenya-deworming &
https://www.givewell.org/charities/deworm-world-initiative#footnote115_fdlhnkk, accessed May 3, 2019

futures within its national territories, nor a failed system to be circumvented[38]. Drawing on the formulation used by historian Gregory Mann to describe the expansion of foreign NGOs into the Sahel, I propose that evidence-based development offers a way to understand non-governmental actors as governing *alongside* states, "not over them, but sometimes in their place" (2015: 244). Here to govern implicates multiple realms of responsibility within the same national territory, some of which are shared but not so clearly, and are therefore enmeshed, between national governments and their non-governmental partners.

**Conclusion: of and not of the state?**

This chapter has argued that it is essential to give heed to both interpretations for the state's absence (remember the joke about the governor's delay and Elijah's words depicted in the introduction) when studying partnerships between national governments and non-governmental actors to provide welfare services to the population. If one relies exclusively on theorizations based on African states' fundamental incapacity or shrinking in the wake of neoliberal globalization, one might fail to recognize the extent to which these partnerships and their perceived successes are premised on already existing governmental initiatives and resources. Conversely, if one depends solely on the image of the busy yet reliable government official, one might miss that certain imaginations about African sovereignty are grounded in unrealistic and ahistorical expectations that nation-states everywhere should be a political form able to "serve the needs of their respective populations" (Redfield 2012: 157). Instead, through a detailed ethnographic analysis of the Kenya NSBDP, I have sought to show how the work of evidence-based development, in partnership with governments, brings political sovereignty in different guises.

---

[38] Similarly, in a recent article, Peter Redfield (2016: 174) eloquently wrote that "'states' and 'governments' (along with 'communities') are as much objects of desire and disappointment as conventions of political order".

For my interlocutors at the NGOs, working with government is as limiting as it is sometimes enabling. The government, that physical embodiment of the state, its skeleton in some ways, was ever present in my fieldwork as an instrument of scale: that which allows to reach millions and millions of people by channeling efforts to infrastructures that have already produced reach across vast expanses of lands and people. While governments are sometimes poor or unreliable, there is nevertheless a necessity to work through the state, even for projects that have global ambitions. The argument of this chapter however is not only concerned with a convergence of operational capacity between Evidence Action and the government of Kenya, as if states were the mere implementers of foreign ideas and prepackaged programs. Instead, I have included perspectives of governmental officials to show how Evidence Action and the Kenyan ministries are enmeshed in much more profound ways around the NSBDP, which sheds ambiguity over what the aims, terrains, and who the actors of government are in a movement like evidence-based global poverty alleviation. If the last decades have "pried open the gap between state and government" in African countries (Mann 2015: 3), among others because of crushing debts, structural adjustments, and privatization, this chapter was an effort to qualify transnational forms of government by showing that even within one development program like the NSBDP there are multiple ways to understand this gap.

In a very different setting, and in a very different way, anthropologist Stephen Collier has written about the "intransigence of things", how neoliberal reform in post-Soviet Russia sought to "work within the stubborn material framework of heat systems", an infrastructure of social modernity whose modes of provisioning cannot be easily changed (2011: 207). In that sense, neoliberal thinking about fiscal equity and the microeconomic devices through it proposed to achieve these (e.g. through progressive subsidies based on income rather than a notion of equity based on equality where everyone pays the same) were limited by what Collier calls the "material and normative framework of heating systems" and their sprawling,

centralized networks of pipes (2011: 234). However, Collier insists, this "is not a story about how neoliberal reform was blocked by the stubborn stuff of post-Soviet Russia", in the sense that there is a pure and ideal neoliberalism "in the minds of misguided ideologues, economists, and reformers" that has "in its contact with "reality, become something else" (2011: 242). On the contrary, Collier wants to argue, neoliberalism is just that: a form of problem making, a style of reasoning, and a mode of governing that takes shape precisely to address situations in which "normal" mechanisms are not workable, where these are older style socialist provisioning or regular mechanisms of competitive markets.

Without pushing the analogy too far, to a point where it becomes flaky rather than generative, may I propose to think about the government, in its network of Ministries that run from national level to village levels, from Ministers to school teachers, as a reluctant, intransigent infrastructure of modernity? It might not be a feature of all global projects to hover above the state, to jeopardize national plans, or to circumvent public resources, but on the contrary some must work through national capacities, account for state priorities, and risk being derailed by the very fact (and lasting power) of the state's political sovereignty. This is not only about friction but also about enmeshment, about the impossibility to tell apart sometimes between what is of the state and what is not governmental.

# Chapter 4 - Lentils don't generalize: on generalizability and contextualization in evidence-based development

**One study and one setting**

"The question I receive the most while I work with governments is: 'ok, but this is one study in one setting, right?'". This is what Carla, then Senior Manager at the Jameel Poverty Action Lab (J-PAL), told the audience at a 2016 conference in San Francisco when prompted to share her experience about how to most effectively present evidence to policy-makers. Over the years, Carla's organization developed several kinds of documentation, including policy-briefs, meant to detail in accessible language the main lessons of a specific piece of randomized evaluation research, such as for instance the effects of a cash transfer program on children's educational attainment in Niger. Yet when they receive these documents, Carla said, governments repeatedly asked: "ok, but this is one study in one setting?", indirectly probing about the relevance of that intervention for the country or region they headed. Carla also told the audience that J-PAL had developed a framework on how to generalize evidence "to teach governments about how they can use the evidence and how they can translate it into their contexts, taking into account not only evidence from randomized evaluations, but also other evidence as well, to be very specific about the causal mechanisms at stake and the key lessons we can take away from the intervention we've tested".

Generalizability is the one limitation my interlocutors always mentioned about their work on evidence-based development, even amid their excitement about evidence and its transformation into programs at scale, because in the end, willingly or not, they always had to confront it. Perhaps because the Kenyan Minister of Education interrogates them on how evidence of positive impacts of a cash transfer program in Niger on education is going to translate in the Kenyan context. Or because the cash transfer program that ran so smoothly in Niger fails to generate improvements or institutional support when exported to Vietnam. Or simply because some colleagues during the weekly economics department seminar series ask the presenter about how useful his research findings are beyond the specific study setting. This issue was present even when I could sense, on the part of the economists I met in the United

States or the project managers I shadowed in Kenya, the hope, the willingness even, that a specific program (say of cash transfers) would turn into a magic bullet across the planet. Indeed, I simultaneously came across their awareness that this is not how the world, with its infinitude of differences, is taken to work. My interlocutors generally referred to the question of how applicable the findings of a study are beyond the setting of that study as the problem of *external validity*, a technical word used in research methods to talk about generalizability.

In this chapter, I zoom in on fieldwork moments where my interlocutors attempt to qualify external validity and to characterize "contexts", and thus to determine across which geographical variations positive program impacts identified as such through randomized evaluations will hold, but also, just how exactly to make them hold. While it is true that those who ran randomized evaluations were always eager to broach with me the topic of external validity partly because as an anthropologist studying development, I was explicitly categorized as an expert of the local context, my interlocutors' concerns for contextualization far exceeded their professional perception of me. The contextualizing practices I describe in this chapter, in other words the act of contrasting contexts as being similar or different, took place across team meetings, conferences, informative websites, but also in interviews, ranging from thought experiments and more formalized frameworks about how to generalize such as the one mentioned by Carla above, to the ways in which programs are being "tweaked" or "tinkered" with when implemented across different contexts.

I seek to underscore that these struggles with generalizability demonstrate an acceptance of "context" as key empirical reality on the part of the development practitioners I worked among. By foregrounding my interlocutors' engagements with the concept of context, I wish to complicate an observation widespread within critical social science scholarship that "quantitative methods are not very context-sensitive" (de Sardan and Piccoli 2018: 18). Instead, I attend to how my interlocutors' contextualizing practices, which differ from more common

substantive understandings of contexts, insofar as they do not primarily invoke political, cultural, religious or historical elements, nevertheless produce a plausible way of ordering the world. On the one hand a self-evident reality, for economists and anthropologists alike, the demands of context appear as that which one cannot ignore when trying to intervene in the world. On the other hand, however, the ethnographic vignettes presented below underscore that it is not easy to characterize contexts, nor to know where one context ends and the other begins. Overall, the aim of this chapter is to raise attention to the multiple, and perhaps equally defensible, ways in which similarity and difference can be produced.

Who weaves together space and time into one context or disjoints them as separate, incommensurable places? The politician? The social scientist? The activist? The passenger next to us on the bus? There are stories about the weaving together of place, men, and things into a context, for multiple purposes: knowledge, truth, representation, but also colonialism, nationalism, exclusion, or forced assimilation. What does it do to the world that we take it as made-up of numerous contexts that are to be identified, contrasted, and compared? What kinds of presuppositions about knowledge, history, and politics are contained therein? Taken together, these are the overarching questions guiding the reflections presented in this chapter.

The chapter is organized along four main parts. The next section looks at the criticism of external validity and the limits of generalizability aimed at randomized evaluations to lay the problem as it is understood within economics and international development themselves. The second section presents a brief overview of the concept of context as it has been addressed within recent scholarship in anthropology, while introducing the need to grapple with context not only as an analytical and hermeneutical tool, but also to pay careful attention to the contextualizing practices of the actors we study for what they might illuminate about disciplinary and conceptual givens used to produce knowledge about the world. The third section discusses how the nation-state emerges as a strong unit of context for my interlocutors

as well to show how this level of analysis, assumed in most comparative approaches, both lay and scholarly, demands to be explained rather than assumed. The last part of the chapter is divided into three case studies – *matching contexts, context scans,* and *variable (of) contexts* – that tackle the issue of context definition, context comparison, and generalization through different ethnographic material weaved across multi-sited fieldwork among those who practice evidence-based development. By interrogating the concept of context, this chapter also aims to unsettle fundamental understandings of what difference and similarity are.

**Criticism of external validity**

The lack of external validity is one of the major criticisms targeted at randomized evaluations and their concurrent enrollment in the project of evidence-based development. A recent paper about a randomized evaluation might report exciting results about the potential of a cash transfer to increase both educational attainment and income in a region of Niger where the study was conducted, but this leaves open the question of how much these findings are replicable elsewhere. The central problem with external validity, critics points out, is that "the effects of an experiment in one place may be different from those effects somewhere else"[1]. The debate about the external validity of randomized evaluations is one that importantly takes places *within* economics and international development. Although himself an economist focusing on poverty and economic development, the 2015 Nobel Prize Winner Angus Deaton is one of the most vocal critics of randomized evaluations. Angus Deaton's name appeared in almost every instance I talked with economists about the good and the bad of randomized evaluations. For instance, Joseph, a well-known economist at the University of California-

---

[1] https://beta-blogs.worldbank.org/impactevaluations/external-validity-musings, accessed June 20, 2019
Of course, the issue of generalizability riddles any form of knowledge production or claim-making about the world: "Does the knowledge I just enunciated about the world hold beyond the setting in which I have gathered it?"

Berkeley, considered that Deaton's criticism is somewhat flawed "because he focuses mainly on external validity".

"But Abhijit's answer", Joseph continues to explain, "is to state that external validity is an issue with all studies, not only with randomized evaluations" .

By "Abhijit", Joseph means Abhijit Banerjee who is one of the founders of J-PAL, the biggest research and policy institution promoting the use of randomized evaluations for worldwide poverty reduction, mentioned in the opening vignette of this chapter.

"Look", Joseph adds to make his argument clearer, "in anthropology, you study a village which you use to draw conclusions at a more general level. Therefore, Abhijit says that if you only have one study, it's better to have a rigorous study, such as a randomized evaluation. And then, you can qualify external validity every time: characterize the context, describe the institutions for every study, so that this informs questions about generalization".

Joseph adds that there is an organization called Innovations for Poverty Action (IPA) whose mandate is specifically "replication" to explore "external validity". Central to IPA's work, Joseph maintains, is to implement randomized evaluations of similar interventions in many places to assess whether and where impacts hold up across geographical settings. While the need to replicate a similar randomized evaluation in multiple countries (in order to ascertain across which contexts results replicate) is often discussed in conferences, meetings, blog posts, or workshops, actual replications – to measure the impacts of a similar intervention in various places – remain very limited[2]. There are a few organizations, such as the International Initiative

---

[2] A well-known example is the evaluation of the Graduation approach targeting the ultra-poor, designed by the NGO BRAC across six countries (Ethiopia, Ghana, Honduras, India, Pakistan, and Peru), which outlined significant positive impacts on several "key" variables, such as consumption or assets (Banerjee, et al. 2015a). There is also an article (Banerjee, et al. 2015b) that reported on the evaluation of microcredit loan programs across six countries, but each program was designed and implemented independently rather than as part of the same coordinated comparative project. The Development Impact Evaluation unit (DIME) at the World Bank recently partnered with the World Food Program to run a series of similar experiments across multiple countries. These experiments share common design and common measurement tools, and attempt to understand where, for whom, and why these interventions work (see https://beta-blogs.worldbank.org/impactevaluations/external-validity-musings, accessed July 1, 2019).

for Impact Evaluation (3Ie), which fund replication studies, or at least systematic reviews to synthesize the available evidence on a specific topic across many studies. Yet, several of my interlocutors discussed about how their journal submission or grant proposal was rejected on the basis that the intervention that they proposed to evaluate had already been assessed "elsewhere". In practice, assessments of external validity remain more speculative in nature, as I will show through ethnographic material in this chapter.

Indeed, within evidence-based development, the prime consideration is often around the question of *internal validity* – the question of whether a specific research design can adequately infer cause-effect relationships. As a research method that works by dividing randomly the sample population into two groups (one that receives the intervention and one that does not) and comparing their outcomes, the randomized evaluation has been elevated in the last two decades as the "gold standard" of economic research to estimate a statistically unbiased impact (Abdelghafour 2017; Donovan 2018; Jatteau 2013)[3]. As geographers Sophie Webber and Carolyn Prouse write, from the World Bank's World Development Report 2015 to TED talks, and in publication and conference requirements, "the RCT is touted as the most reliable, rigorous tool of development economics" (2018: 167).

The more recent emphasis on the critic of external validity must be read as an effort to counteract what several economists and development practitioners take as an inattentive infatuation with randomized evaluations on the parts of academics, donors, development agencies, and policy-makers. In a co-authored article with the philosopher of science Nancy Cartwright, economist Angus Deaton, mentioned above, aims to dispel the "magical thinking" associated with RCTs (2018: 3), which contributes to the perception that RCTs always provide the strongest evidence for causality and effectiveness, potentially transforming circumscribed

---

[3] Chapter 1 of this dissertation tackles the issue of internal validity and of measuring causal relations for randomized evaluations through a focus on the counterfactual.

knowledge into universal truths[4]. Economist Lant Pritchett, another forceful commentator on the limits of randomized evaluations, argued that the significant resources mobilized toward conducting randomized experiments are only justifiable if the knowledge generated through RCTs have external validity.

"Unless I can tell you the range of contexts over which these results apply…I have no idea what to do with it", said Pritchett to his audience at New York University in February 2018. Pritchett stated that experiments, without substantive economic theories behind them, will never satisfy the requirements of external validity. A significant number of RCTs assessing the same intervention in different places cannot guarantee external validity, Pritchett contended, because "context is not place, context is place and time, or more precisely, context is some list of parameters that affects the performance of the response surface."[5] In his view, the same region in India, say Rajasthan, will not be the same context today as in a few weeks from now if, for instance, the state implements some changes in policy that makes it much more difficult to start a business for those who receive microcredit loans. Without a complete delineation of these parameters of context, it is impossible to know whether a result has external validity, according to Pritchett[6].

Whether or not one agrees with these criticisms of randomized evaluations is besides the point. What I wish to underscore here is that the attention given by those who promote evidence-based development to devise concrete tools to, borrowing Joseph's worlds, *characterize contexts* and address the practical issue of generalizability also emerged as a

---

[4] While discussing an issue closer to medicine than to economics, a recent blogpost has drawn attention to the danger of relying on a single study to inform policy through the case of a medical study which was used as an important basis for a Dengue vaccine program in the Philippines. For more information, see the blogpost on the 3ie website.
[5] http://www.nyuafricahouse.org/?risen_event=dr-lant-pritchett-talk-the-debate-about-rcts-in-development-is-over-we-won-they-lost, accessed June 20, 2019.
[6] For RCT-related research that foregrounds "time" as an important element of context, either in the form of offering the program at the right time, or in the form of accounting for "weather shocks" in external validity models (in addition to thinking about external validity "breaking down" because interventions have different impacts in different places and on different individuals), see respectively the 2003-2005 study Nudging Farmers to Use Fertilizers and the 2019 reflections in External Validity Musings.

response to these criticisms about the external validity of their work. If, as Pritchett said, context is "everything", then if you move one foot left or wait one minute, your research results can become obsolete, as you will have changed the study setting. In his analysis of the concept of context across the social sciences, historian Peter Burke (2002: 173) argued that "strong or ambitious contextualism privileges difference over similarity", shedding some confusions as to whether the targets of contextualists are only unjustified generalizations or the possibility of generalization *tout court*. For my interlocutors, to examine context therefore meant concurrently to attend to both the possibility of generalizable knowledge and the necessary conditions for intervening successfully in the world.

Because they look for global solutions to poverty problems, those who are committed to randomized evaluations hope, perceive, and disseminate their results and their poverty anti-interventions as having (potentially) universal relevance, or at least relevance in multiple contexts. The endeavor to interpret context at length is not inherent in their study approach, but rather an after-the-fact engagement to know where to scale or who to advise to take up this or that program. This does not mean however that they do not take "the burden of contextual understanding" earnestly (Huen 2009), but that their doings ask us to apprehend the multiplicity of contexts, and hence difference, as a product that must be examined rather than be considered a given to be assumed at the start of analysis.

**Investigating context**

In taking context as the direct object of my investigation, I follow other scholars who have emphasized how the anthropological project, at least since the 1980s, has become to "understand human life in context", and what little attention has been paid to interrogating context as a central analytical category in contrast, for instance, to the sustained analytic attention given to culture (Burke 2002; Dilley 1999; Dilley 2002; Englund 2004; Fabian 1995;

Huen 2009; Seaver 2015; Strathern 1987, 1995a, 2002). Most examinations of the concept of context have taken place within the literary fields or confined essentially to linguistic anthropology (Duranti and Goodwin 1992)[7]. In his essay titled "Context in Context", historian Peter Burke (2002) traced the history of the concept of context within diverse kinds of scholarships and disciplines.

Burke (2002: 153) starts with the etymology of context, derived from Latin, as a "weaving together". Burke argues that it is around the sixteenth century that the term 'context' was revived to be mobilized as an analytical tool for the interpretation of texts, especially the Bible. Context was conceived as the textual ensemble, or the set of events, that surrounded a passage and illuminated its meaning. Burke (2002: 156) further contends that the Renaissance, and importantly its movement to return to Antiquity, raised awareness about the difference between the two periods. This in turn spurred questions around whether principles and problems of Antiquity were still relevant to Renaissance Europe, and more generally about the question of the exemplarity and general applicability of ancient rules and norms to contemporary periods. Interestingly, Burke (2002: 157) qualifies the Scientific Revolution and the Enlightenment as "anticontextual" enterprises insofar as they were concerned with formulating laws of nature and society, namely generalizations, that would be valid whatever the circumstances of time, place, or persons. Burke (2002: 162) also discusses how today "anthropology is sometimes perceived, not without reason, as the contextual discipline par excellence". Anthropological research "is often justified by the need to 'put things into context, especially where that context is understood as society or culture'" (Strathern 2002: 303).

Bronislaw Malinowski (1923) is often credited with having pioneered one of the early usages of context in social anthropology to disparage the disregard for context and singularity

---

[7] In the last two decades, this attention has increasingly extended to computer science and machine learning insofar as a computer operates through language structures (e.g. Akman 2000; Dourish 2004).

present in earlier evolutionary anthropological projects, in particular in the work of Sir James Frazer. Evolutionary anthropologists were seen to "scramble contexts", bringing ethnographic elements from across the globe only to illustrate different levels of savagery and civilizations of humanity as a whole (Strathern 1987: 63). In contrast, Malinowski objected that social institutions and material culture could not be abstracted from the matrix of social relations and face-to-face interactions within which they took place (Dilley 2002: 438).

As Marily Strathern wrote in her piece titled "The nice thing about culture is that everyone has it", according to Clifford Geertz's definition, "culture *is* context, the frame within which, as he says, social life can be intelligibly described" (1995a: 159)[8]. Within the social sciences, to invoke context has therefore become a "major critical practice" for ascertaining epistemic truth. To contextualize is used as a favored methodological move to repair accounts and actions perceived to be "out of context" (Huen 2009: 151) and to counteract potential "rash or unqualified generalizations" (Burke 2002: 173) with real implications for people's livelihoods. Targets of this line of criticism are now very diverse, especially in science and technology studies[9].

In a recent book about anti-poverty policy titled *Cash Transfers in Context: An Anthropological Perspective*, Emmanuelle Piccoli et Olivier de Sardan (2018) identify randomized evaluations as one such technology which sidesteps an attention to contexts when looking for anti-poverty solutions to privilege instead the analysis of a simple relationship between a basic intervention (say cash transfer programs) and a limited set of outcomes of interest (education, income, etc.). In their view, such a form of knowledge production supports

---

[8] Interestingly, while Clifford Geertz's work is often taken to be among intellectuals most advancing the significance of cultural context, insofar as he equated culture and context, some of Geertz's earlier writings such as *Peddlers and Princes* (1963) can be read into the continuation of a comparative project to modernize faraway, colonial places.

[9] While I do believe that context is used as a central concept in critical studies of science and technology, I should make clear that this portrayal of the anthropological project is far from being exhaustive and certainly does not encompass all anthropology. Like Joel Robbins wrote about his text on the anthropology of the good, "my remarks aim at an exploration of what I think is an…important stream of the discipline over…time" (2013: 459-460).

"traveling models" –such as cash transfer programs or microcredit loans– applied by development experts somewhat indiscriminately around the globe to reduce poverty. Once these standard and general traveling models are implemented across many continents and countries, the difference between model and local realities is made visible as program implementations routinely encounter various problems (e.g. resistance by the government of the country; unintended stigmatization of the beneficiary population). De Sardan and Piccoli qualify this phenomenon as "the revenge of contexts": "being underestimated or ignored, the contexts come back and erupt into the game in an unexpected way, widening the implementation gaps and distorting the objectives of the programs" (2018: 17)[10].

In these critical moves to invoke context, it is as if context is a reality out there, already assembled and waiting to be summoned, and as Marilyn Strathern wrote (2002), often understood as society or culture. Despite key disciplinary debates to problematize the concept of culture[11], anthropological invocations of contexts as critical gesture often refer predominantly to social relations, national or regional culture, small scales, ethnic groups and their histories, advanced as "the unquantified remainder that haunts mathematical models, making numbers [and realities] that appear to be identical actually different from each other" (Seaver 2015: 1101; but see also Adams 2016; boyd and Crawford 2012; Colvin 2016; Cruz 2017; de Sardan, et al. 2017; Merry 2016; Morvant-Roux, et al. 2014; Webber 2015; Webber and Prouse 2018). In that approach to context, *contextualizing* to avoid faulty generalization means therefore appealing to the *right* context, and therefore to have sufficient, good, and accurate knowledge about a certain time or place.

---

[10] This line of analysis, common among qualitative scholars, seems indebted to Karl Polanyi's seminal argument that presented economics as a process of abstraction. In this view, economics is a form of presenting the world which relies on dis-embedding the economic activities under study from the "social and other relationships in which they had occurred" (Carrier 1998; cited in Strathern 2002: 304).

[11] See for instance Clifford (1997) who argued for a shift from a fixed notion of *culture* to that of *travel* as characterizing the locus of human experience, cited in Dilley (2002).

In an article about contextualization in big data science, anthropologist Nick Seaver has argued that "to argue *for* context without recognizing that we are arguing *about* context is to beg the question…Context is not the unique domain of thick describers – everyone has it, or does it, and they do it differently" (2015: 1106). Seaver therefore advocates for attending, as ethnographers, to the contextualizing practices of the actors we study because "our task [however] is not only to relocate mathematical models in the social world" (ibid). This is not an argument to do away with context, and thus leave the stage to either unbounded universalism or extreme individualism. It is rather an invitation to pay attention to context even more, but differently. Overall, this chapter aims, following historian Burke, "both to criticize and to employ the concept of context, to develop or adapt it in response to challenges" (2002: 176). To do so, one must reckon with context not only as anthropological analytical instrument, but also as an actor's category. As Ann Kelly (2008: 99-100) has said about her study on medical clinical trials, the goal is therefore to regard randomized evaluations and ethnography as "similar inductive strategies" instead of merely using anthropological tools to "contextualize" evidence-based development. This means paying attention to how context and contextualization are understood and to what aims they are deployed, even where one least expects them, such as in projects with global ambitions, whether science or global poverty alleviation.

By looking at how my interlocutors engage with the demand of contexts, I seek to raise two points. First, I interrogate the widespread notions of contexts, also taken for granted by my interlocutors, such as equating the nation-state as a strong unit of a context, to challenge geopolitical and conceptual givens that riddle both scholarly and lay understandings of the world. The conceptual move I suggest here echoes Akhil Gupta and James Ferguson's call to renounce the concept of *cultural difference* and its "assumed isomorphism of space, place, and culture" as a ready-made interpretive tool used to bridge between naturalized conceptions of

spatialized cultures (1992: 7). Instead, they suggest a move toward understanding how cultural difference is produced through shared historical processes, which include colonialism but also anthropological accounts[12]. In conversation with other authors (Burke 2002; Dilley 2002; Strathern 1987), this means recognizing that contextualization is not a self-evident endeavor, that context is a problematic rather than a redemptory term "as much in need of elucidation" as the events or claims which are being contextualized (Culler 1988; cited in Fabian 1995: 47).

Second, I wish to take seriously alternative conceptualizations of contexts and the world they present us with. By sketching the possibility that contexts can be constructed in various genuine ways, how does this analysis force us to rethink the modes and instruments we use to delineate between the similar and the dissimilar? To advocate for such an outlook is also to foreground the "issue of power which underlines the process of context definition" by emphasizing the ways in which things are framed, delimited, demarcated, and connected or disconnected (Dilley 2002: 453). Discussions around contextualization and generalization, difference and similarity, matter because they are far from being confined to scholarly debate but permeate all domains of lives. Every day, everywhere, people constantly interpret and draw conclusions about the world. To mention just one example, this struggle over generalizability and the significance of context is contained in the popular warning against engaging in clichés or stereotyping. Given this, how do certain understandings of similarity or difference come to dominate, and turn into naturalized givens? Who decides what is in and what is "out of context"? The ways in which both context and the limits to generalization are understood bear on how the world is lived, on the actions people take, but also on the kinds of policies designed to manage people and things across the globe, for global poverty alleviation and beyond. It is important to acknowledge, I believe, that our own anthropological exercises of speaking about

---

[12] Also see Amselle (1990) for a similar analysis emerging from francophone anthropology around the same years as the seminal 1992 Cultural Anthropology issue on "Space, Identity, and the Politics of Difference" put together by Akhil Gupta and James Ferguson.

and for contexts can fail as much as my interlocutors' practices to characterize context. The remainder of this chapter brings together ethnographic material spanning across multiple field sites to examine how context, but also similarity and difference, are transient yet powerful realities, differentially assembled by enrolling specific practices and different parameters.

## Unit of context: the nation-state

Replicating and expanding a successful evaluated pilot to "similar contexts" is called *scale-up* by my interlocutors[13]. Scaling-up, I came to realize, is an ambiguous expression. My interlocutors can use it to mean transposing the pilot program from the study setting to another context, often abroad. Often however, the term "scaling-up" is used to refer to initiatives to make the program "larger", with the assumption that the program is brought "at scale" in the same context. Here the emphasis is to operationalize programs in ways that they can reach *millions* of people – an ambition which is the focus of Chapter 2 of this dissertation. That does not mean that the program "at scale" is bound to succeed, for different reasons might disrupt its success.

For example, there are programs that work on a small scale but do not hold when brought at a national level. One can think of a job readiness program that helps participants to secure employment. But if the goal is to make the whole population employed, and there are simply not enough jobs because employers do not recruit much, that job readiness program will not solve the issue. Another example is when the quality of monitoring might decrease when the intervention is transformed from a project ran by a small team of researchers and field staff into a governmental program, with the consequence being that the program is not implemented as well, and thus positive impacts plunder. Nevertheless, there is the notion that one program can be made smaller or bigger within *one* context.

---

[13] See also https://www.povertyactionlab.org/evidence-to-policy, accessed February 27, 2019

Traveling across Kenya and India in vans, jeeps, motorbikes, or cars to visit villages and schools targeted by anti-poverty interventions included in randomized evaluations, I kept asking myself whether I was crossing contexts. Where and when did one context stop and the other begin? I shadowed a non-governmental organization Evidence Action as they installed chlorine dispensers for water sanitation purposes. The program ran in three countries: Kenya, Malawi, and Uganda. Were these three different contexts? For this specific intervention, was Kenya a unitary context? But the region of Mwea in central Kenya, with its low rice fields and well-tended roads, offered a vastly different scenery, both visual and political, to the western region of Kakamega which is hilly, with luxuriant and humid vegetation, and which requires constant zigzagging to the left and the right to avoid entire stretches of unpaved roads marked by deep potholes. In both regions, however, I remember the color of the dirt as a bright red and regularly seeing low school buildings with children wearing colorful uniforms. Was Kenya, then, made of multiple contexts that needed to be characterized, to know whether the sanitation intervention would be as successful in different parts of the countries, for various reasons, whether behavioral, institutional, operational, or any other?

Expanding a program in the same context is an endeavor captured in numerical form by the development practitioners among which I conducted fieldwork. There were talks of bringing a pilot program of remedial education in Kenya at a scale that is 50 times bigger, by increasing a project size from 200 villages to tens of thousands of villages, and moving from reaching 1000 people to reaching several millions of them. During the 2017 Ely Lecture addressed at the American Economics Annual Meeting, Esther Duflo (2017: 30), one of the pioneers of the RCT-for-development movement, contended that when economists work directly with governments evaluating a policy and determining how to best design it, "the stakes are high enough that it is perhaps less important to know whether the findings will generalize beyond that particular work" (2017: 30). Esther Duflo claimed:

To take just one extreme example, the Affordable Care Act [a health insurance scheme in the USA] is a unique institution, and perhaps a less-than-ideal result of a long and complicated political process. It is very unlikely that anybody designing a health insurance system from scratch would adopt anything resembling it, so in a sense there is very little external validity to any work about the Affordable Care Act. Yet, when Handel, Koldstad and others write papers that are concerned with detailed issues of implementation of a system like the [Affordable Care Act], it does not really occur to us to wonder whether this is perhaps a little too specific (ibid).

In fact, Duflo insisted, such policies affect millions of people, and so the lessons derived from evaluating these policies are often acted upon, and together these two elements lower the burden for the critique of external validity which requires comment on how the results obtained for this one study setting will generalize beyond it. In other words, when the study setting itself is very large, it might be less important for my interlocutors to qualify the external validity of the findings. Yet, when the initial randomized evaluation is conducted on a small sample, the need to scale up but also to prove the generalizability of the findings beyond the study setting are all the more important if one wishes to transform an intervention found successful through a small pilot study into a bigger intervention implemented at a larger scale.

Scaling *across* different contexts, by contrast, is an aspiration expressed in qualitative terms, as in Joseph's injunction above to *qualify external validity* and *characterize the context*, or in the recognition that to inform change at scale across the world requires a "deep understanding of context and systems"[14]. But this form of scaling is also expressed in actual attempts already alluded to above to develop frameworks that determine where else the evidence from one context can apply. It also appears in less explicit ways, such as during a meeting between the Kenyan team of Evidence Action and its Bangladesh counterpart to learn about a new program, still in its trial phase, called No Lean Season.

"We've found amazing impacts with this small loan program in certain parts of Bangladesh to mitigate the risks of destitution between agricultural seasons, in encouraging

farmers to migrate and look for temporary work in urban centers", says Pierre, the program manager chairing the meeting.

"But this was a pilot", he underscores. "We're now working to bring it at scale within other parts of Bangladesh and want to include different regions: some places where people anyway have the propensity to migrate, others where they wouldn't normally, to see if the loan is successful in helping them seek work elsewhere despite their initial reluctance. But then the situation in Bangladesh is very different from that of Kenya: people don't have big farms there like here, there have been episodes of famine, there's need for cheap labor, like rickshaw drivers."

As this little vignette stresses, my interlocutors operate with a widespread and ready-made conception of context, whereby scaling one intervention across contexts means crossing national boundaries, or at least administrative lines. Here, one context equals one country (i.e. Bangladesh is different from Kenya), or follows official subdivisions (i.e. the Bangladeshi region of Dhaka versus the Bangladeshi region of Rangpur). When the Senior Policy Manager for the South Asia office of J-PAL talked about scaling up a promising education program across India, he took each Indian state as an indicator of local context. "A state will agree to pilot the program to 300, 1000 thousand schools…tweak the program…adjust the model to their localization…and then scale-up phase by phase to cover the state; so different states are at different phases of that scale-up". Here, context is a spatial territory for which one specific government or administration is responsible. In other words, my interlocutors operate within what anthropologist Liisa Malkki (1992: 28) has termed a "national geographic" map of the world into distinct, naturalized nation-states, in which the nation-state emerges as a strong unit for defining a context.

Several authors have discussed how the "nation-state constitutes a preeminent category of and for comparative analysis" (Mongia 2007: 384). Indeed, comparative analysis, whether

emerging from scholarship (say comparative literature or international relations) or lay discussions (Who has the best cuisine, Italy or Japan? Which country in the world has the best healthcare system?), often relies on the nation-state as the level of analysis (Dilley 2002; Fourcade 2009; Malkki 1992). This is because nation-states are geopolitical formations produced as simultaneously "similar *yet* distinct and separable, and thus available for comparison" (Mongia 2007: 386). Each nation-state is its separable unit, marked by its own state sovereignty and demarcated by national borders. At the same time, all nation-states are part of the same *geopolitical formation*; each nation-state is therefore a case of a "larger principle" (Mongia 2007: 385). Therefore, each nation-state is one specific *context* for the purpose of comparison, for establishing the grounds of similarity and difference, and for drawing possible generalizations.

In the last decades, authors have insisted on the need to problematize this conventional analytical approach that takes nation-states as given, natural, and autochthonous entities (Geschiere 2009; Gupta and Ferguson 1992; Mann 2015; Mongia 2007; Torpey 2000). Instead, they have called for examining how such demarcations are produced through a complex array of historical events. The national geographic map of the world (Malkki 1992) and the isomorphism of place, space, and culture (Gupta and Ferguson 1992), which underpin anthropology's foundational attention to and expectation of difference, predominantly as cultural difference, should themselves be explained before they can serve as explanatory terms in attempts to make sense of the world, including by my interlocutors.

Sociologist Radhika Mongia (2007) underscores how both colonialism (and issuing post-colonial struggles) were central to expanding and universalizing the norm of state sovereignty and triggering the eventual partitioning of the entire world into nation-states[15].

---

[15] I am grateful to Nassima Abdelghafour for pushing me to articulate further the equivalence between ready-made contours of context as nation-state and the production of nation-states from heterogeneous realities through various socio-technical techniques.

Mongia (2007: 396) discusses how colonial administration linked sovereignty and territory to "delegitimize the claims of nomadic peoples", enrolling civilizing discourses, military means, and the use of force. As a telling example, Mongia provides the 1884 Berlin Conference whereby African geopolitical borders were arbitrarily drawn by the European colonial powers without "a single African representative present" (ibid)[16]. Similarly, anthropologist Jean-Loup Amselle (1990) has discussed how colonization in West Africa, by fixing people onto specific territories, has crystallized multiple, flexible ethnic belongings into immutable identities, thereby giving way to the development of numerous conflicts in the region based on ethnicity.

Put differently, the nation-state – and the related taken-for-granted contours of differences, such as context and ethnicity – are produced through technologies of control and further consolidated through management techniques, such as the existence of a national government that governs the national economy, an object constituted at least partially through statistics and the ebbs and flow of the GDP (Appel 2017; Mitchell 2005). Borders and the regulation of migratory flows also contribute to entrenching the idea of the nation-state as a well-bounded unitary context. As historian Gregory Mann wrote, citing John Torpey's (2000) historical work on travel documents, "the passport worked to establish, rather than [simply] bound, political formations that have become naturalized" (2015: 91).

To sum up, my interlocutors have an ambivalent positioning vis-à-vis this widespread equivalence between *context* and the *nation-state*. On the one hand, they take this national geographic vision of the world as a fact. Niger and Burkina Faso are therefore two different contexts, and when scaling one cash transfer program from Niger to Burkina Faso, one must therefore account for external validity and the differences, or similarities, between these two places. But at the same time, precisely because my interlocutors try to generalize "successful"

---

[16] Some scholars of Africa insist that the partitioning of Africa happened also before and after the Conference through bilateral agreements between colonial powers, in an effort not to overestimate the role of the Berlin Conference (see for instance Craven 2015; Katzenellenbogen 1996).

programs across contexts, devising strategies to bring forth such global ambitions despite working within this national geographic vision, their efforts complexify and indirectly put into question this rigid understanding of *context* as a geopolitical, cultural formation. The following sections turn to ethnographic moments wherein my interlocutors are constructing contexts in various ways, weaving diverse elements together or disjoining them. In recounting these, the point is to trouble contextualization as a self-evident and unproblematic process, always reliant on the social and cultural qualities of a situation, event, or place[17].

**Matching contexts?**

In mid-November 2015, I attended a staff meeting at the Center for Effective Global Action (CEGA) at the University of California-Berkeley. The discussions at this meeting spotlighted once again the preoccupation with external validity and a concrete proposition to confront the demand of context. Chiara, Program Associate for Financial Inclusion, is presenting the policy lessons derived from the many research projects conducted under the Agriculture Technology Adoption Initiative (ATAI), a collaboration between CEGA and J-PAL at MIT. A research fund launched in 2009 and supported in important part by the Bill & Melinda Gates Foundation, ATAI's mandate is to generate and disseminate "rigorous evidence about what helps and hinders farmers to adopt technologies and access markets, in a manner that increases their profitability and improves their welfare"[18]. To this aim, ATAI funds *randomized evaluations* to assess the impacts of multiple agriculture development strategies on farmers' livelihoods. As stated in Chapter 1, as of now, 51 such research studies have been carried out across Africa and South Asia.

---

[17] In his book *The Perils of Belonging,* anthropologist Peter Geschiere (2009) conducts a similar analysis about the paradox that surrounds the concept of autochthony between apparent certainty and yet a practice, or reality, of great uncertainty put into question by historical records.
[18] See https://www.atai-research.org/, accessed February 27, 2019.

I already know a good chunk of the key points shared by Chiara because I have been following the wrap-up of these results in the last few weeks: sitting in a call with the ATAI board and helping Chiara edit the ATAI Policy Insight Compilation document. I am thus not surprised when I hear her repeating multiple times throughout the presentation that the ATAI team "is very careful in the phrasing of the policy lessons and what they are comfortable stating based on the bodies of evidence". This uneasiness had already been broached at length by the ATAI board during their last conference call when discussing how to draw general conclusions from a smaller set of studies.

Today, Chiara insists on the difficulty for academics to "distill findings" considering that this is not a task traditionally part of their professional repertoire. To literally illustrate her point, Chiara brings up a slide that shows a page from the ATAI Policy Insight Compilation document filled with comments by multiple people, edits in "track changes", alternative phrasings, and a work-in-progress bibliography. Chiara pulled-up this unfinished document, simultaneously shared among and edited by at least five people, to illustrate the difficult labor enrolled in deriving general claims and policy recommendations from a circumscribed set of studies. "We have to be careful [about what we claim],", Chiara concludes, because we don't want someone to scale an [agricultural intervention] unless we are sure". Paula, CEGA's Director of Partnership and Innovation, intervenes with a question.

"Are you able to assess the external validity of these findings, are you attempting to comment on these, or are you rather more focusing on local context?"

Paula is asking whether the ATAI's team is making any conclusions about how generalizable the research findings are to contexts that are not the "local context" of the study setting. Chiara replies that this issue is still under discussion. A fellow colleague, Helen, wonders whether the ATAI board members are willing to serve as advisors to governments: "Is there going to be a consulting side?" Chiara adds that this is a question also for CEGA as a

whole: how to make findings accessible in a real sense, not just online, so that they are taken up and put into practice to inform large programs ran by governments or other actors with the capacity to scale. At this point, Jake, yet another colleague, suggests almost as a thought experiment: what if in order to increase the relevance of these policy lessons, "we used census data to see if contexts match", implying that where "contexts match", the projects could be recommended for scale-ups and even implemented?

Was Jake not proposing a concrete tool to practice what Joseph above called "characterizing the context" when assessing external validity and the capacity to replicate findings elsewhere? To play along with Jake's idea, I started wondering just what kind of census data one would want to collect to "see if contexts match". What kind of data could be deemed suitable to compare contexts? Joseph had mentioned information about institutions as an example. Could we talk about "matching contexts" if two "contexts" exhibit similar institutional realities, for instance similar governmental systems and services? Would we need data on the availability of infrastructure, such as roads and hospitals? Should we contrast culture, norms, and worldviews, across two contexts? And what main parameters would we use to define similar institutions (e.g. social security, free school system, immigration policies) or similar culture (e.g. language, religion, food, arts)? In other words, through Jake's proposition, a question was taking shape: what *made* a context? What is a context defined by, and by whom?

In thinking about these questions, I was reminded about a conversation I had in 2016 with Leonard, an epidemiologist from the Kenya Medical Research Institute. Leonard had explained to me that both snails and mosquitoes have ecological preferences, linked to water and soil chemistry, that directly determine their geographical distribution, which in turn affects the patterns and prevalence of snail- and mosquito-borne diseases across space, such as parasitic worm infection or malaria. Could we, or should we, include disease prevalence as a

marker of context? Somewhat more fundamentally, should we consider animals (and the relative presence or absence of certain species) as features of a context? If "what makes a context" is the subject of nested questions, rather than the outcome of a preset list of elements, how does this illuminate the critical move to invoke *context* to repair or ameliorate accounts that ignore context and over-generalize?

Jake's proposition challenges what Johannes Fabian (1995) has called the "positivity of context": the assumption that contexts are self-evident aspects of realities that are out there, rather than in our heads. Within anthropology, according to Fabian, this has taken the form of assuming that contextualization should always be grounded on the same relevant sources of knowledge such as cultural or linguistic competence, orality, or the native's point of view. This black-boxing, I argue, lends context both tremendous analytical power and important explanatory limitations. Context has tremendous analytical power because it can always be invoked, to use de Sardan and Piccoli's formulation (2018), as a *revenge* over mistaken forms of knowledge and intervention, with reference to its epistemological and political superiority to know the world. Yet, because appealing to context is deployed as a critical stance repeatedly and over vastly different domains, the concept of context seems simultaneously to encompass everything and anything, with the risk of losing its interpretative edge. For even if context is assumed to include culture, history, politics, and the native's point of view, these are broad fields and circumstances that can always be differentially assembled. Evoking the Kenyan context to challenge the purview of global indicators or a specific intervention to address health raises the issue of what it means to take Kenya, an entire country with its manifold dimensions, as a single context?

In his reflection on ethnographic misunderstanding, Fabian (1995: 43) discusses how accurate explanation of events must sometimes be gained through a *break* with this usual and assumed anthropological conception of *context*. Fabian provides an example from his own

experience around the interpretation of a song recorded while doing fieldwork in Zaire, known today as the Democratic Republic of Congo. Because he was having difficulty with the translation of the song, Fabian checked his transcription "with the help of that provider of context par excellence – a native speaker of Shaba Swahili". The native speaker confirmed Fabian's interpretation and supplied additional information to enrich the chosen explanation. It turns out, Fabian writes, they had both misunderstood what they heard: "the song in question was not in Swahili but in Kizela (a language spoken east of Lake Moero). And it was not a fighting song but a marching song" (1995: 42). Fabian was able to obtain that accurate interpretation specifically because, still facing issues of translation with other recorded songs, he decided to suspend trust in the "'natural' workings of context" which privileges interpretation by native speakers, granted with cultural and language competence (1995: 43). Fabian's reflection demonstrates that if we base our analysis on a preconceived notion of context, it is not certain whether we will succeed or fail to understand the empirical worlds we both study and inhabit. In light of this, it might seem less appropriate to talk of *the revenge of contexts* when describing the corrective move to bring qualitative methods to correct conclusions or policies derived from quantitative analysis; rather, we can warn against, using Fabian's formulation (1995: 42), the *perils of context* if we take context as a taken-for-granted reality with undeniable epistemic edge.

By juxtaposing, on the one hand, field-based interrogations generated through Jake's suggestion to use census data to see where contexts match, and on the other hand Fabian's argument, I want to underscore that context cannot be called upon as a ready interpretive instrument, neither for the anthropologist nor for the development practitioner. This is because contexts do not exist independent of the practices and people that are there to assemble them. I insist on attending to context therefore not as an end-product, but rather as a process, one that involves crafting relevant features together. And drawing playfully on the etymology of the

word, this leads to a specific type of craft of *contextere*, the Latin verb meaning to weave together (Burke 2002). Talking of contextualizing as a craft of weaving together also brings back the visual image displayed by Chiara at the staff meeting of a document in track changes, subjected to multiple rounds of revisions, in iterative cycles of editing by cautious academics somewhat uncertain about which general conclusions to draw from their randomized evaluation evidence. Here, context appears as a reality that is never solidified, because it is constantly assembled, disassembled, and reassembled as research findings are repeatedly reinterpreted or new ones are published so as to modify previous conclusions.

Asides from raising problems about context definition, Jake's thought experiment directly addressed the issue of *comparison* between contexts. What kind of data is to be used to contrast two or more contexts to determine whether they are sufficiently alike to warrant implementing the same intervention in both of them? Do two contexts have to match on all features to make generalizability relevant, or do certain characteristics of context take on increased relevance according to the results or the intervention that one seeks to generalize? Overall, I use these questions to examine how similarity and dissimilarity are constituted within my field sites, through a set of ethnographic material that mostly deals with what my interlocutors called "context scans", and looking also at how they concretely adapt programs to fit new contexts.

**Context scans**

Throughout my fieldwork in Kenya, I was shadowing programs led by Evidence Action, created in 2013 to transform interventions identified as effective via randomized evaluations into impactful programs at scale across the world. In 2016, Evidence Action's portfolio was composed of two main programs: the water sanitation program mentioned earlier implemented across Kenya, Uganda, and Malawi, and "Deworm the World" –an initiative to provide mass deworming treatment to reduce parasitic worms in children to increase health and

educational outcomes[19]. At the time of my fieldwork, Evidence Action worked with both the government of Kenya and of India on national mass deworming programs and was looking to expand into Ethiopia. During my first meeting with Eshan, then Director of Evidence Action for the African Region, I asked him how his organization established partnerships to scale up programs across countries with donors and implementing agencies, whether governmental or otherwise.

In his response, Eshan said that the first step was to find contexts where programs could be successful. With this in mind, he detailed, the organization conducts desk researches to assess the geographic conditions. This involves a mapping of countries to see for instance where the prevalence of these kinds of problems, by which Eshan meant water sanitation and parasitic worm infections, are significant. Eshan cited five countries that had been researched, among which were Pakistan, the Philippines, Nigeria, and Ethiopia. Based on the desk research to inquire about prevalence and implementation logistics, Ethiopia was deemed the easiest context for the next scale-up. Once this has been established, Eshan further explained, the next step is to reach out to an implementing organization in that country and, when possible, to partner with them. The last step is then to look for donors. In short, concluded Eshan, the main concerns when thinking about new scale-ups are context suitability and who will implement the program, which implies whether or not Evidence Action will have to carry out the heavy burden of operational logistics. Reading over the website of Evidence Action one day, I realized that examining context suitability has been institutionalized as a core strategy by the organization under the heading of "context scans"[20].

When identifying which interventions to take up next as part of its portfolio, Evidence Action combines *evidence* scans – in other words literature reviews on a theme, say education

---

[19] For more details about the sanitation program and the Deworm the World initiative, see Chapter 2 and 3 respectively.

[20] https://www.evidenceaction.org/how-beta-works, accessed February 28, 2019

or financial inclusion – with *context* scans. The latter are "geared towards helping [us] understand the socio-political and economic dynamics of a prospective implementation area." One learns that in addition to desk research, context scans also entail "interviews with development researchers and practitioners who are located in or familiar with the area in question." Context scans and the need to assess context suitability are increasingly integrated in the work of scaling up programs to reach millions of people across the world, carried out by organizations promoting anti-poverty policy based on evidence derived from randomized evaluations.

To provide a recent example, Teaching at the Right Level (TaRL), an educational approach jointly led by J-PAL based at MIT and the Indian NGO Pratham, is one of the five initiatives that have been chosen as part of a US$80 million grant provided in January 2019 by Co-Impact, a new global collaborative focusing on what they call "system change", and who include as their donors the Bill & Melinda Gates Foundation and the Rockefeller Foundation[21]. J-PAL and Pratham received that grant to scale up the TaRL approach in multiple countries across Africa. As an intervention, the TaRL was initially designed and evaluated in India through six randomized evaluations[22]. One way to describe the TaRL is to say that it consists mainly in re-organizing classrooms and educational settings by learning *needs* rather than formal age or grade, with the recognition that inputs alone (such as textbooks) do not significantly enhance educational outcomes. The TaRL approach proposes different "implementation models" with adjustable elements ranging from whether learning sessions organized by learning needs take place in the classrooms or in parallel settings, ran during or

---

[21] The TaRL approach is also discussed in Chapter 1 in reference to theories of change and poverty.
[22] Randomized evaluations have also been conducted in Ghana and Kenya to assess similar versions of the TaRL program. The Zambian government also launched a pilot program and the J-PAL's website informs us that "with support from Pratham and J-PAL Africa, Governments in Côte d'Ivoire and Nigeria have launched TaRL pilots to adapt the approach to their local contexts" (https://www.povertyactionlab.org/case-study/teaching-right-level-improve-learning, accessed March 1 2019).

after school hours, by trained volunteers[23]. At the same time however, the approach also involves four necessary components that must be implemented together: building up capacity and trainings teachers and instructors, assessing the levels of children at the start of the intervention, adopting a new classroom methodology around learning levels, and setting-up clear and regular monitoring and assessment cycles[24]. At the bottom of the homepage of the TaRL website, there is a circular icon with a hand-raised black figurine, hailing us with a written question on the right: "Is TaRL right for my context?"[25]
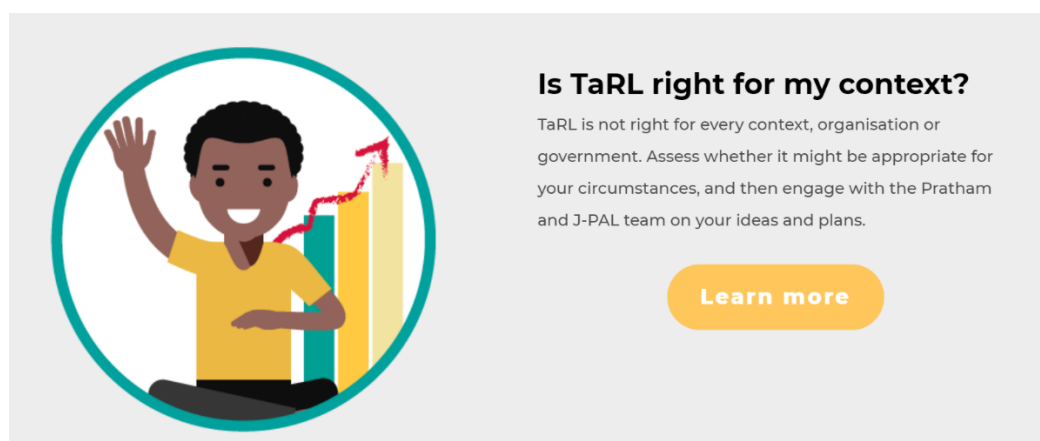


**Figure 13: Is TaRL right for my context? (Screenshot taken from https://www.teachingattherightlevel.org)**

If one clicks on "Learn more", one is led to another web page with a set of questions that should be used to determine whether one's context fits the conditions under which TaRL has been shown to work best. The questions are diverse in nature, some having to do with the school environment, while others focus on the broader education system and on governmental priorities. "Are children lacking foundational reading and mathematics skills?" "Are classes fairly heterogeneous?", "Does the government prioritise learning and basic skills?" "Do current education systems mainly serve top performing learners?" "Can the necessary systems of

---

[23] https://www.teachingattherightlevel.org/, accessed July 1, 2019
[24] https://www.teachingattherightlevel.org/the-tarl-approach/, accessed July 1, 2019
[25] https://www.teachingattherightlevel.org/, accessed March 1, 2019

support be activated or built?" There are also targeted questions depending on whether "you are considering a government-led programme" or whether, as an NGO, "you are considering a direct implementation". "Does your organisation have time to dedicate to TaRL training and field practice? Is your organisation flexible enough to adjust the approach, methodology, materials and classroom practices?"

This list of questions, one could conclude, allows to *characterize one's context* and to assess, as Jake had talked about, whether one's context *matches* the context in which the TaRL has been demonstrated to have impactful outcomes. The TaRL context scan might not be using *census data,* like Jake had proposed, to see whether two contexts match, but it is still offering a tangible proposal to compare contexts and determine where and when TaRL might be replicated while maintaining positive results on children's educational outcomes. In light of the TaRL program, a context becomes defined by such items as children's average educational level, structure of classrooms, interest of government in basic skill acquisition, as well as the nature and organization of the implementing body. Indirectly, political and economic aspects are given some attention in the form of governmental priorities and whether the necessary systems of support can be built to sustain the program. Yet, regional history and broader cultural and social realities – usual markers of context deployed as anthropological analytic – are absent from consideration[26].

I therefore want to raise the possibility that my interlocutors are not depriving the world of substantial variations, especially insofar as they hold onto a national geographic vision of the world in which each country is a unit of context with its cultural, historical, political, economic, and ecological elements. Yet, they have concluded that the latter are not the pertinent dimensions for discerning the generalizability of the TaRL intervention. Authors have talked

---

[26] The TaRL context framework raises another issue around spatial scale: it is unclear whether by "your context", the framework means to refer necessarily to the country/nation-state level, or whether it allows for comparison between contexts at different scales, let us say contrasting Ghana to one specific region of Uganda.

about the "dilemma of context" (Scharfstein 1989) or "unboundedness of context" (Culler 1983) to refer to the fact that there is no limit to the number of contexts that can be invoked once *context* has been enrolled as an analytical lens (Dilley 2002; Englund 2004; Huen 2009). For instance, the recent militarized confrontations between Pakistan and India in February 2019 can be explained through reference to India's current ruling party, or to Pakistan's government, or seen through the prism of longstanding religious clashes, by bringing in the history of partition, or the conflict over Kashmir, and so on and so forth, ad infinitum. Anthropologist Chi W. Huen (2009: 151) underscores therefore that the very same critique of being "out of context" can be potentially invoked against the critic herself, in an infinite exchange of critical moves.

As historian Peter Burke wrote, "context is often regarded as local, and yet we talk of 'global context'", suggesting that it might be essential to ask "what is not context?" (2002: 17). Put differently, how and where might we find world elements amenable to the possibility of generalizability, because their existence exceeds contextual demands? One can be tempted to argue that this outlook toward *context* reveals, borrowing loosely from de Sardan and Piccoli's words again (2018: 17), context "insensitivity" –a lack of both knowledge and concern for distinct places and times that allows for transforming a complex and intricate social context requiring monographs and deep analyses into a list of distinctly identifiable parameters. One could portray such a contextualizing practice only as a show devised by pragmatists to eschew received criticism of overgeneralization when they in fact remain committed to applying their results indiscriminately everywhere in the world. It is true that some moments of my fieldwork raised this question in complicated ways.

When I was in Kenya in 2016, Evidence Action was supporting the implementation of G-United, a Kenya governmental programmed based on the TaRL approach. Explained briefly in Chapter 1, the program revolves around pairing recent university graduates with schools to

provide remedial education to primary school students. According to the TaRL website, the G-United program aimed to reach 40,000 learners across Kenya in 2019[27]. When I met with Sara, who managed the G-United program on behalf of Evidence Action, she presented it in the following terms: "These last 15 years of research have shown that lightly trained volunteers increase performance and decrease drop-out in schools. All the research projects are different, but the basic evidence", Sara says, "is that when pairing lightly trained volunteers with children who are struggling, the children improve on various educational outcomes. There is evidence but it's not scaled up a lot [yet]".

When Sara talks of *basic evidence*, one wonders whether she intends to mean evidence that is so fundamental that it holds true everywhere in the world. Here, this basic evidence derived from randomized evaluations – that minimally trained volunteers increase educational outcomes – could be taken to have become, to borrow de Sardan and Piccoli's expression, a "traveling model" applied across the world by development practitioners. There is a tension however because just after, Sara added that "evidence is a third of the program. There is more that needs to be worked on [to have a successful program]. For instance, we must account that we work with governments, they must release funds. (…) The evidence is properly packaged, but it's another thing to make it work in the field, it involves a lot of iterative learning."

By this, Sara means that there "are different ways in which the evidence can be interpreted to reach the same results". Sara tells me that people at Pratham, the NGO that first developed the TaRL approach, "change their programs according to regions, that they're willing to adapt". For instance, in one state of India, the remedial education might happen in the classrooms, while in another region, they might set up special weekly "camps" to dispense remedial education. She says that Pratham, in collaboration with J-PAL, offers different models of "teaching at the right level", which in her view, included reorganizing classrooms (by levels

---

[27] https://www.teachingattherightlevel.org/tarl-in-action/kenya/, accessed June 27, 2019

instead of age) and even included performance-based teachers. "But in certain contexts, like Kenya, this would not work – very strong teachers' unions would oppose it [reorganizing the classrooms and teacher-based performance]." So instead of these two alternatives, the G-United program focuses on remedial education, brought about by volunteers outside of class time.

Overall, this process of adapting the evidence to a given context is called "tweaking" or "tinkering" by my interlocutors. "The RCT gives you a black box from input to impact, and our job is to unpack that black box", Sara explained to me. Sara insists that the process of tweaking helps understand better the theory of change: what are the exact (and context-specific) causal mechanisms that lead from remedial education to better educational attainment? But, I ask Sara: "how do you make sure that the 'tweaks' you put in place do not affect the causal impacts?" Sara says that they look at what is "the *core part* [of the evidence] we should not change. For instance, in the case of remedial education, studies have shown it's not different if you do it before or after class, or in (summer) camps. But there is a difference whether you keep children in class or take them out of class". And of course, Sara says, you monitor and evaluate your program as it changes to assess whether outcomes are still obtained.

Sara's description of her work exhibits an interesting tension between generalization and specificity. On the one hand, her mention of *basic evidence*, of the *core part* of evidence, signals a wish to both believe in and find generally applicable laws – e.g. pairing minimally trained volunteers with children who are struggling will always and everywhere lead children to improve on various educational outcomes – at the core of traditional economics and its (universal) models of human behaviors. On the other hand, Sara emphasized, that evidence is only *a third of the program*, explaining clearly how varying contexts and their demands must be considered if one wishes to run a successful program. In Kenya, teachers' unions emerged

as a relevant parameter of a context, one which importantly dictated what exact model of remedial education to put in place in Kenya.

In this analysis, my interlocutors do harbor universally applicable evidence and traveling models, but to travel well, these models *must* be adapted to specific contexts. While Sara did not state it explicitly, this leaves open the possibility that contexts, at least some, put into question the general applicability of some piece of evidence. Iterative learning, as Sarah called it, in other words anticipates that contexts *might derail* program implementation. This is precisely why organizations such as Evidence Action exist as incubators: to test programs at scale and see when, where, and why they might succeed, or fail. For instance, in June 2019, Evidence Action decided to terminate its No Lean Season program in Bangladesh, a program of small loans to encourage seasonal work-related migration, for reasons that ranged from evaluations that showed a lack of impact at scale to others that had to do with operational issues, including financial irregularities and accidents[28]. In this perspective, contexts are not so much taking revenge, but instead they are considered to modulate aspirations to have found anti-poverty interventions that could be effective at a global scale.

The interesting point is not whether the TaRL contextualization is any more true or false than other conceptions of context that would go against it. The significant question is whether my interlocutors' engagements with contexts can be granted the possibility of being as often right and as often mistaken as qualitative social-science scholarship and its enrollment of context as a key analytical tool, and whether one can avoid discarding my interlocutors' outlook from the get-go as the product of technical and standardizing practices that are disjointed from local systems of experience. The last section focuses on another attempt at "unpacking the black box of the RCT evidence" by contrasting Sierra Leone and India around the replication

---

[28] Fore more information, see https://www.evidenceaction.org/blog-full/why-we-are-shutting-down-no-lean-season, accessed June 20, 2019

of a program to increase immunization rates. The intention is to show that the contextualizing practices of my interlocutors raise important doubts about the capacity to discriminate easily between what constitutes similarity and what is dissimilarity (or difference).

**Variable (of) contexts**

I am listening to Rachel Glennerster, then Executive Director of J-PAL through a computer screen as she gives a talk at the 2016 Effective Altruism Global Conference in California. Already mentioned in Chapter 1, effective altruism is a movement that links together people – academics, NGOs, and donors mainly – who seek to use "rigorous" evidence and analysis to find answers to the overarching question of "how can we use resources to help others the most"? It is a movement grounded in the recognition that many people want to do something about suffering, poverty, or injustice, but that figuring out what that "something which helps is" turns out to be a very challenging problem[29].

It is now the Questions and Answers session and Rachel is sitting on a metal chair along with two other members of the J-PAL policy team, one being Carla, working in Latin America, with whom I opened this chapter, and the other the Senior Policy Manager for the South Asia office. The panel chair, relaying a question from the audience, is asking Rachel: "could you tell more about the framework you're helping develop about how you can determine whether evidence from one context can be applied to another?" As I described in the opening vignette of this chapter, Carla had referred to this framework "on how to generalize evidence" a few moments earlier, mentioning it as a necessary tool because the question she receives the most often from governments when presenting evidence that could inform their policy decisions is: "ok, but this is one study in one setting, right?".

---

[29] https://www.effectivealtruism.org/articles/introduction-to-effective-altruism/, accessed March 1, 2019

Rachel prefaces her answer by mentioning that she will provide only a summary of a longer framework, which she has detailed elsewhere. To do so, she takes the example of a program called "incentive for immunization" in India, well-known to anyone who knows about randomized evaluations for development because it is routinely cited as a success case by J-PAL[30]. The study by Banerjee and colleagues (2010) reported that giving small incentives for parents to come to immunize their child led to an important increase in immunization rates. Now, Rachel tells the audience, she is also working with governments in West Africa, especially Sierra Leone, that worry about low levels of immunization. This means she has to think about it: "is this intervention going to work with a government in West Africa?" The program in India was done through an NGO in India, Rachel points out, but now a partnership with a government is at stake in West Africa. "Different continent, different implementer. So you think: how on earth would that evidence generalize?"

Rachel continues laying out her framework, with a pace that signals she has faced that question before. "One part is: do humans respond to small nudges? And that is something that has been found to generalize a lot. It's about fundamental human behavior. If you take another bit of the question: if you just take the program as a package – do I think the lentils generalize to West Africa? No, they don't eat lentils in West Africa, it's not a good incentive. So when we say does research generalize, we have to be a bit careful about which questions we're asking. Lentils don't generalize. Underlying general human behaviors, well, actually turn out to be very similar across different contexts." The third part of the framework, according to Rachel, is whether the implementer in the new context can also run the program efficiently, especially without leakage, and again, she says "this is going to be very context-specific". Finally, she adds: "when you're thinking about whether research generalizes, you need to ask

[30] See https://www.povertyactionlab.org/sites/default/files/publications/incentives-for-immunization_0.pdf, accessed July 1, 2019

yourself: is the need the same in the old and new context? In this case the issue in India [in the initial randomized evaluation] was high level for first immunization (first shot), but it did not persist [beyond the first shot, after which immunization rates dropped]."

Responding to nudges and incentives is therefore, as Sara would put it, the *basic* evidence, *the core* of the evidence, which is reckoned to have potential generalizability around the world. Interestingly, in a recent article about RCTs for development, geographers Webber and Prouse argued that "behavioral economics….in particular, is one of the new forms of expertise that many of our interviewees draw on to help solve the external validity issue, unpacking what happens within the black box of the RCT" (2018: 181). By referring to nudges and underlying human behaviors, Rachel is enrolling both a recent field of economics – behavioral economics – and the longer lasting assumptions, both within economics and across other social sciences, that there are universal behaviors. In Rachel's framework, nudges are an essential part of explaining why people go get their children immunized.

In her recent work, anthropologist Doerte Bemme has discussed this process of disaggregating RCT evidence as that of identifying "active ingredients" (2018: 182). The willingness on the part of our interlocutors such as Rachel or Sara to make visible such "active ingredients", Bemme claims, emerges precisely from the recognition that "evidence is never wholly reproducible; evidence does not travel well". By discussing active ingredients, Bemme wishes to underscore that scaling necessitates these ingredients that "become mobile and transferable, not because they are uniform and unchanging, but because they are generic, broad and hyper-changeable, as they are expected to be realized, fleshed out, and made effective within a new complex situation upon arrival" (ibid). If you recall, Rachel had said that "when we say does research generalize, we have to be a bit careful about which questions we're asking." Rachel had insisted that nudges, acting on fundamental human behaviors, generalize, but that "lentils don't generalize" to West Africa, and that therefore it would be a mistake to

take the intervention as a package without changing the content of the incentives. In other words, to make the intervention successful in Sierra Leone, it would be necessary to adapt the exact contours of the active ingredient: the "nudges".

Through Rachel's presentation, we are left with the idea of fundamental human behaviors, lentils, implementers, and needs as relevant parameters for characterizing a context. Rachel concludes by saying that "I think it's very important as you think about bringing evidence to different contexts to think through those different components [human behaviors, content of incentives, implementing organization, and the prevalence of the issue] and what evidence you have. And you will use RCT evidence potentially for human behaviors; and you may need descriptive evidence for assessing what the need is; and process monitoring for [gaging] the logistics questions." Overall, Rachel insists that addressing all these questions requires more than evidence from randomized evaluations, including qualitative evidence to identify the needs of the context, for instance a place where there are high immunization levels for the first vaccine shot but families do not return for follow-up shots.

Interestingly, on the one hand, if one takes the two dimensions of "fundamental human behaviors" and "needs" (responsiveness to nudges and low levels of immunization respectively), India and Sierra Leone do emerge as similar contexts. In other words, if you take "prevalence of the immunization issue" and "human behaviors" as a way to characterize context, Sierra Leone and India would fall in the same "context" category. On the other hand, if you pay attention to two others parameters, which are the content of the nudges (i.e. would people respond well to being offered lentils?) and the implementing agency for the immunization program, then Sierra Leone and India are placed further away from each other, given that people do not care for lentils in West Africa and that the Indian NGO is nowhere to be found in Sierra Leone, which relies on different institutions to carry out development projects. Are India and Sierra Leone therefore similar or different contexts? Does the new

context of Sierra Leone suit or match that of India in terms of vaccine-related health challenges? Is exporting the "incentive for immunization" program from India to Sierra Leone likely to be a successful endeavor? Rachel's exposition leaves these questions somewhat unanswered.

At stake for me here is not to be able to provide definitive answers. Of relevance to this chapter is that Rachel's attempt to characterize Sierra Leone and India shows how the same contexts can appear as different or similar from one another depending on the lens one uses to look at them. Put differently, I use this ethnographic vignette to make visible an idea perhaps already known, but I think well-worth reiterating: the outcome of any comparison varies depending on the categories used to draw it in the first place. Sociologist Marion Fourcade captures this tellingly when she writes that "the question of *how different things are* is not a simple, hard empirical "fact": it is an intellectual construct that has to be produced through particular methodological, analytical, and narrative strategies" (2009: 12).

This suggests that just like context, difference and similarity are not obvious categorizations, already out there and simply waiting to be grasped and detailed by the analyst. The very elements one enrolls to establish comparison impinge on the issue of what similarity and difference mean at a foundational level. It also raises the question, both epistemological and political in content, of "why do we perceive difference at all?" For instance, why have we come to see India, this expanse of land and this temporal reality, as one bounded context? And this circles us back to an earlier section of this chapter that was concerned with the historical production of the nation-state as the paramount unit used for comparison. By drawing Sierra Leone and India as similar contexts, where the same (or a very similar intervention) would be effective, Rachel's exposition also unsettles the reality of a demarcated nation-state as a naturalized unit.

I argue that in academic disciplines and beyond there have been crystallized coordinates through which to understand the attribution of difference or similarity. Within anthropological practice since the 1980s, cultural and societal circumstances mark a world as though made of pockets of often incommensurable contexts. Within economics or psychology in their extreme forms, but also in evolutionary anthropology, fundamental human nature can serve to explain people while leaping over national borders. Paradoxically, the anthropological axiom of human difference is itself already "enunciated in the modality of the universal": as a fact about the world worthy of general application (Balibar 2016: 46). This disciplinary move implicitly turns cultural difference into an analytical category reckoned to have constant epistemic edge over knowledge enterprises that standardize or generalize. The intention here is not to challenge difference from the viewpoint of universalism or globality, nor to foreground unfettered individualism. It is likely that one modality or the other carries its own load of dangers.

The aim instead is to raise attention to the multiple ways in which similarity and difference can be produced. If one problematizes the idea of context without offering universal truth as a an alternative, what modes of knowing the world can one propose instead? For those, like my fieldwork interlocutors who want to intervene and "make the world a better place", what modes of intervening could this train of thought sketch for them? By delving into Rachel's comparison between India and Sierra Leone, I sketch a perspective on generalizability (its value and its limits) that is not premised on unsolvable musings between universal validity and (cultural) specificity. Drawing on historian Peter Burke's suggestion (2002: 174) that the term context "is best used to refer to phenomena which are not in focus at a given moment", I wish to propose approaching contextualization as a dynamic process of *focusing*.

In photography for instance, focusing is a skill that involves many moves and modalities: to alternatively zoom in and out, move left or right, or show more details here and less there. Presenting contextualizing as the work of focusing emphasizes that each object of

study will bring to life a particular way of characterizing context, even if one holds onto an idea of the world as demarcated by national borders. What kind of context Kenya is, and how it compares to Bangladesh, depends on whether one asks this question to generalize the "Teaching at the Right Level Approach" or the "immunization program" described by Rachel, or whether one wonders about the history of colonialism or about the status of women. It does not mean one should stop pondering about whether both the goals and the contextualizing practices are valid or ethical, if this is what is at stake. When questioning them, one nevertheless should not forget to question one's own contextualizing moves, inherited from disciplinary training but also other kinds of lived experience.

**Conclusion: and then the social context became part of epistemology…**

This chapter has dealt primarily with how similarity and difference between contexts are being understood by my interlocutors. It has not examined at length how programs are being derailed when implemented across contexts. This is mainly because the ways in which contexts derail program implementation is not only a descriptive reality, but a performative one, which needs to be explained. Contexts derail precisely in the way contexts are expected to derail, whether one thinks about national sovereignty or ethnic groups. The relationship with the state is likely to be an important factor of whether a program succeeds or fails, not necessarily because the state is a provider of essential resources, but because of the legitimacy it is granted as the administrator of a context in a national geographic world in which nation-states are the prime geopolitical formation. Similarly, resistance to traveling programs will be conceptualized in terms of cultural or ethnic resistances, even though it remains unclear whether such ethnic differentiations, such as that among the Luo and the Kikuyu of Kenya, are natural markers of a context, or a product of how context and difference are both understood and sanctioned in today's world (Wrong 2009).

To problematize these critical moves, this chapter investigated contextualizing practices within economics and global poverty alleviation to discuss how context cannot be taken for granted by attaching to it "cultural" or "social" qualifiers. Instead, context must be repeatedly qualified to confront the kinds of questions one is trying to answer. The intention has not been to multiply indefinitely the number of possible contexts, operating through different levels and dimensions, nor to do away fully with essential knowledge categories such as culture, difference, border, or the nation-state. Nor do I mean to suggest that my empirical field sites replace a culture-place understanding of context with a truer one. Yet, because my interlocutors' engagements with context scramble together universalism and specificity, rendering blurry and difficult the adjudication of similarity and difference, I believe they can be taken as an invitation to produce careful and nuanced explanations about the world and generate practices of knowledge that are not only about representing the world and others but also "confronting reality", as messy and unruly as it is (Fabian 1995: 48).

At an event hosted by the Berkeley Initiative for Transparency in the Social Sciences, which brought mostly economists, psychologists, political scientists, and data scientists together, someone said something that stuck with me. It was not a particularly new idea. But I think it was the fact that it was uttered, not during the course on the history or sociology of science, but rather by Paul Romer, who later became chief economist at the World Bank, reflecting on the practice of the social sciences today. Romer, well actually his voice on a superbly recorded podcast, so well recorded in fact that after writing some notes in my notebook I raised my head to look for him in the room, forgetting he was not there, talked about epistemology. It underscored that "at the beginning the idea was that it was straightforward to know Nature, and then the social context became part of epistemology with the idea there is no such thing as truth, and now the current outlook is a combination of both...".

At the end of it, insofar as we produce statements about the world, we will be confronted with having to choose what to take into account, how to compare, and whether to generalize. The hope is to emphasize that both the world and our knowledge of it – anthropological, economic, and otherwise – are organized around conceptual divisions that, perhaps once were novelty (such as granting the "other" its own contextual space), become routinized, taken-for-granted, and perhaps as much tools of inquiry as instruments of convenience.

# Conclusion - What does evidence do?

**The quotes**

"We're looking for quotes that support the notion that governments think our work is essential", said Will, CEO of Evidence Action, at the start of our second Skype meeting in March 2016. He was letting me know that GiveWell, an organization that ranks global development charities to guide donation decisions, had recently asked him to provide evidence that the work carried out by Evidence Action is influencing governments.

As mentioned in Chapter 1, GiveWell has consistently ranked Evidence Action's Deworm the World Initiative among its top charity programs. This signals to potential donors that the national deworming programs ran by Evidence Action in partnership with governments in Kenya and India are, in the words of one of GiveWell's team members, among the best-giving opportunities in terms of lives saved or improved by dollar invested. "We believe there is relatively strong evidence for the positive impact of deworming", the GiveWell website writes in its summary page for the Deworm the World Initiative, referring both to evidence derived from randomized evaluations and to data obtained through ongoing monitoring of the national programs in Kenya and India[1].

But now, GiveWell wanted to know something else. It wanted to know whether the work of Evidence Action is influencing governments. The request from GiveWell suggests that it is one thing to provide convincing evidence about the causal impact of mass deworming on worm prevalence and school attendance, and yet another for that evidence to impact policy-making. As Will told me: "GiveWell is looking to understand the counterfactual: what the governments would have done without us, Evidence Action." Would Kenya, for instance, still

---

[1] https://www.givewell.org/charities/deworm-world-initiative/january-2017-version, accessed July 29, 2019. The pairing of the term "believe" and "strong evidence" might struck as odd, but it refers to the fact that charity recommendations are based around different criteria, established by GiveWell, to identify reliable evidence.

have put in place a mass school-based deworming program? Or were the evidence base and the lobbying efforts harbored by Evidence Action indispensable to the existence of national programs of school-based mass deworming?

"This is why", Will further explained, "the managers at Evidence Action are wanting to garner quotes from our partners which state something along the lines of: *thank you, we could not have done it without your work*".

"Makes sense", I was thinking while listening to Will. What caught me unprepared was what he said next. From his office in Washington, Will asked me whether I could share such quotes if I have garnered any during my field research in Kenya, only, of course, if it would not compromise my research integrity. I mumbled a few words to avoid an answer. I was still pondering his request when I realized, with some relief, that Will had switched topics.

I never encountered such quotes directly, or if I did, it was at events that were also attended by some of my interlocutors, who could relay them back to Will. I open the conclusion to this dissertation with the dialogue between Will and me because it makes explicit how difficult it is to grasp to what extent, and exactly in which ways, evidence-based development shapes policy-making, whether carried out by governments or other non-governmental institutions. In important ways, Will's asking for quotes and demonstrations that evidence worked in shaping policy-making indirectly rendered the power of evidence less obvious and more uncertain (than even these Chapters might have led to believe so far).

Several organizations committed to evidence-based development keep meticulous records in data-bases, reports, or policy briefs, of all the scale-ups and collaborations carried out with governments and other bodies. For instance, the Center for Effective Global Action and the Jameel Poverty Action Lab have internal monthly newsletters where they detail new partnerships. But what Will's words were suggesting is that there remains a nagging uncertainty over whether, for instance, it really was the advocacy work carried out by the

Deworm the World Initiative and the technical assistance they provide, which *caused* governments in India and Kenya to implement country-wide programs of mass deworming. What would have been the counterfactual otherwise, Will wanted to know, because GiveWell had asked for this information. Would the programs have been conducted anyway?

In Chapter 1, I discussed how randomized evaluations that assess development programs aim to capture the counterfactual: what would have happened in the absence of the intervention? The control group, I documented, serves as the incarnated proxy for an impossible counterfactual to be contrasted with the treatment group. If the treatment group exhibits improved outcomes compared to the control group, this will be attributed to the intervention. In other words, the difference between the treatment and control groups serves to measure the causal impact of interventions. Now, Will and GiveWell were willing to capture another counterfactual: what would have happened without the work of Evidence Action? Would policy-making, and hence people's livelihoods, be any different? Would priorities of national governments be different? In other words, my interlocutors at Evidence Action and GiveWell were eager to know the causal outcomes of Evidence Action for managing poverty worldwide. What difference, if any, did evidence-based development make in the world?

**The power of evidence?**

In a recent article titled "Facts, power, and global evidence: a new empire of truth" anthropologist Ann H. Kelly and sociologist Lindsey McGoey talk about "the great 'failure' of RCTs: the way that even their clearest 'triumphs', such as the ability to furnish proof that microfinance gains are negligible, does not necessarily lead to policy change" (2018: 12). Kelly and McGoey argue that "RCTs adherents tend to presume that the revelation of policy 'failure' is sufficient to lead to policy reversal, but often this is not the case" (ibid). They further insist that RCT evidence alone is not sufficient to discourage government actors or private donors

from pursuing a policy that might not meet a stated objective (e.g. microfinance should lift people out of poverty) but that perhaps serves their interests and goals in other ways (e.g. for instance just because of the fact that supporting microfinance schemes enhances government popularity in opinion polls). While policy stakeholders might "find creative ways to strategically ignore inconvenient evidence", Kelly and McGoey further contend that "importantly the RCT method is incapable of providing a useful *theory of inaction*, because it is incapable of investigating the motives of policy actors who have unstated reasons for avoiding uncomfortable facts" (ibid).

Yet, I would argue that this is far from what I encountered during fieldwork when following the daily work of people and projects that work to transform RCT results into programs "at scale". While the RCT as a research design indeed might only make visible a purified version of how an intervention affects livelihoods, practitioners of evidence-based development are painfully aware of how difficult it is, for instance, to collaborate with governments and influence policy, and how often this fails (see Chapter 3). In the last decade, organizations committed to evidence-based development have launched the *Journal of Development Effectiveness* which has featured at least a couple of publications discussing the impact of evidence and its dissemination on policy-making around the world. There is for instance the article "Sound expectations: from impact evaluations to policy change" which opens with the following questions:

> To what extent can evidence from evaluations really affect the policy-making process? What are the purposes and potential roles of impact evaluations (IEs) in the design and implementation of new policies? How can those who commission and produce IEs improve their potential use and value? Which are the factors and forces that determine IE's uptake in policy-making processes? We believe that answering these questions can help in increasing existing knowledge on the real benefits of producing IEs (Díaz Langou and Weyrauch 2013: 269).

The authors, Langou and Weyrauch, do not focus solely on randomized evaluations, but on impact evaluations more generally, to devise "an analytical conceptual framework" for

understanding two related issues: *whether* an impact evaluation had impact on policy or not, and *why* that influence was achieved, namely, identifying the factors and forces that condition those policy shaping efforts.

Another article titled "What is the impact of a policy brief?", published in the same journal reports on an experiment conducted to estimate the effectiveness of "policy briefs" on changing people's beliefs (Masset, et al. 2013). As you might know, the policy brief is a document of a few pages summarizing key findings and policy implications of one or more pieces of research, often making use of visual data such as images or graphs. The authors argue that while policy briefs are a tool very commonly employed by development agencies or NGOs to disseminate findings, "the policy impact of policy briefs is unknown" (Masset, et al. 2013: 50). A 2015 World Bank blog post which also addressed the question "Do policy briefs change beliefs?" reported on a survey conducted with about 500 policy stakeholders in Latin America and South Asia which ranked the policy brief as *the least* useful form of exchange and pathway to support policy, in contrast for instance to in person events or online statistical databanks[2].

While I do not dwell on these examples, they make apparent that the impact –on funding decisions or policy priorities– of the vast body of evidence derived from impact evaluations of development programs has become a significant preoccupation for those who carry out evidence-based development, including my direct interlocutors during fieldwork[3]. For my interlocutors, it seemed that evidence about the impact of evidence on policy-making was elusive and hard to grasp.

At the end of a panel on "Engaging developing countries' governments" presented in 2016 in San Francisco, the moderator asked Rachel Glennerster, then executive director of the

---

[2] See http://blogs.worldbank.org/impactevaluations/do-policy-briefs-change-beliefs, accessed August 7 2019
[3] For further research that looks at how policy-makers, politicians, and leaders respond to research, either by updating their belief or adapting their actions, see also (Hjort, et al. 2019; Vivalt and Coville 2017) and https://www.psychologytoday.com/us/blog/getting-along-and-getting-ahead/201906/why-don-t-politicians-admit-when-more-research-is-needed, accessed August 13, 2019.

Jameel Poverty Action Lab (J-PAL), the question that received the most interest from the audience:

"Lant Pritchett [a development economist] has argued that we lack evidence on whether research *actually* improves policy", the moderator said, "and that was a key part of this presentation [you and your colleagues just gave], so perhaps we can give other examples in addition to the ones shared so far about the role of research in shaping policy. The second part of the question gets at the theory of change about how that happens. For example, do we feel that policy-makers lack information or they're choosing bad policy because that promotes their self-interest? One part of the question that I'd like you to answer is what other factors are there in addition to evidence that shape policy-making?"

Rachel's answer was lengthy, emphasizing that "it is important to recognize that not all actors in the policy space are altruistic and want to improve the lives of the poor for that reason and evidence is not the only thing that is going to come on their calculation. However, I think it is wrong to assume that it's *not* part of their calculation, that it doesn't play *some* role. Especially in democracies, where there is a political cycle, where if you deliver better results on the ground, there's a benefit to the governments; if you improve your programs and people do better, you're more likely to be elected. We're not saying that all policy-makers are only driven by evidence; but it's also very true from our experience that we find a lot of people do care about improving their programs, but they don't know how improve it."

Rachel further underscored that even in cases where "some actors in the program had a clear incentive in keeping the corruption going, we still see massive improvements and some actors within governments are really responding to that evidence", citing a program to provide

subsidized rice to poor households in Indonesia, which made use of identity cards to improve access to targeted social assistance and reduce leakage[4].

"Does research improve policy?", Rachel asked again to herself and the audience as a way to gather all her thoughts. "Something like 300 million lives have been touched through the scale-ups of programs that we have evaluated at J-PAL and have been scaled-up", she says. "We haven't done a randomized controlled trial, and I can't say with absolute certainty: 'that's because of the trial', but the work that we've put in –there are many, many examples where we can show a pretty clear chain where there wasn't evidence, people didn't know the evidence, and now there's evidence, and we can show you quotes from people who said: 'I took this action because of this evidence'. So we can't provide 100% certainty in every case that it was the research that changed people's mind, but if you think about the huge volumes of money, the huge changes that we've seen that appeared to be linked to this evidence, there are now a lot of examples that people are responding to this evidence."

I briefly examine here the significance of doing fieldwork at that specific conjuncture: to study a global poverty alleviation project that on the one hand is grounded in *evidence of impact* about development programs produced through "rigorous" quantitative methods (the randomized evaluation), and on the other hand relies on anecdotal, qualitative evidence (in the form of quotes or stories for instance), the very kind of knowledge it seeks to get away from in the first place, to justify the policy impact of the evidence it produces. With evidence-based development and randomized evaluations, one is presented with a form of evidence that asserts itself to constitute certain knowledge about causality. And yet, the effect of that knowledge – the (policy) impact of that evidence – is variable and ambiguous. Translating evidence into policy action is a process that involves a notable infrastructure made of people, things, steps,

---

[4] https://www.povertyactionlab.org/case-study/using-identification-cards-improve-national-social-assistance-indonesia, accessed July 31, 2019

and negotiations. Evidence Action and J-PAL are two organizations doing that work, as discussed in earlier chapters, through advocacy, lobbying, pitching, finding donors, and providing technical assistance: all complex tasks whose outcomes are far from certain. How does this help to qualify the "power" of evidence?[5]

In the last decade, scholarly literature has drawn attention to how metrics and statistical indicators are today central to guiding decision-making and policy around the world (Adams 2016; Jerven 2013; Merry 2011; Rottenburg, et al. 2015; Storeng and Béhague 2014). These authors have importantly studied how global numbers are produced, often through the elision of complex realities, and therefore are used in ways that might distort what issues are politically relevant (Geissler 2013; Merry 2016). In taking up the investigation of how evidence-based development affects actions and livelihoods around the world (by detailing how its body of knowledge is being discussed at conferences and trainings, turned into programs at scale, or exported and adapted from one context to another new one), this dissertation was importantly informed by this scholarship that has documented how indicators are increasingly used to manage people and lives.

In these concluding pages, I bring together a few ethnographic vignettes that might make us think differently about the "power" of randomized evaluation evidence, both when present and absent. I am not hereby discrediting the value of evidence-based development or any other evidence-based endeavor, nor am I suggesting that we should abandon all pretense to ground our actions in evidence (and how could I myself as a researcher, especially in this era of populistic politics?). Instead, I take cues from those scholars who have underscored the need to study the productive and powerful effects not only of rational and communicable knowledge, but also of ignorance, secrets, uncertainty, ambiguity, and the absence or disguise

---

[5] In the conclusion of her own dissertation, my colleague Doerte Bemme offers a similar line of questioning on evidence, albeit with different empirical claims than the one I suggest here. She asks: "In the light of my findings I began wondering more generally: what if evidence harbors a rather different ethos and politics than commonly ascribed to it?" (2018: 186).

of evidence (Anand 2015; Geissler 2013; High, et al. 2012; Kelly and McGoey 2018; Knaapen 2013; McGoey 2007, 2009, 2017; Taussig 1999). The following reflection is purposefully anecdotal and tentative, and is not aimed to provide any final words on the extent to which evidence-based development influences anti-poverty policy-making around the planet. I conceive it rather as an attempt to illustrate how my interlocutors themselves grapple with this question. I provide the anecdotes in turn, almost like a list, and tie them together at the end in a short commentary about what it might mean for anthropology to follow, as suggested by the title of this dissertation, "evidence in action".

**Scaling down, not up**

Several months before my conversation with Will, I had met with Idris, an economics professor at the University of California-Berkeley, who suggested that it would be interesting to have evidence about whether the evidence produced by impact evaluations persuades policy makers (i.e. affects their planning and budget decisions) or whether they take the evidence as it suits them (i.e. they use the evidence to support programs that they would have put in place anyway).

Louise, who worked at the Center for Effective Global Action (CEGA) at the University of California-Berkeley, just a few doors down from Professor Idris' office, provided a piece of that evidence in the form of her professional experience she shared with me. When I was doing fieldwork at CEGA, the Center was coming up with its logic model, which involved the difficult task of defining, collectively, the ultimate goals of their work. The team had tentatively agreed, albeit not unanimously, to define "building a body of rigorous evidence to guide decision-making on approaches of global poverty" as a key objective. That day, Louise told me that she disagreed with what some colleagues had proposed as an adequate measure of success: number of scale-ups. As discussed in earlier chapters, this refers to transforming

interventions found to be effective via RCTs into programs at large scale in partnerships with NGOs, governments, and donors.

"Indeed", Louise said, "projects often use evaluation to confirm that they work. Evaluations are thus used often to prolong projects already in place, but projects are rarely stopped as a result of a new body of evidence".

To illustrate what she means, she refers to a large project she worked on for over five years, from start to end, which she calls her formative evaluation[6].

"Throughout this evaluation of a healthcare system delivery", Louise continues, "it was very hard to keep the government from scaling up the intervention to control groups, those not supposed to receive the intervention, because they found their program to be great, so they did not see why they shouldn't do so", even though Louise said she tried to explain that assessing *whether* the program is effective was the whole point of the evaluation.

"When the study found positive impacts", Louise added, "the report was framed as *the evaluation supported the scale-up of the healthcare system*, when in fact the government had planned the scale-up all along".

Louise tale's presents a case in which the evaluation concluded that there was evidence of impact (the healthcare system delivery was effective on whatever outcomes were measured), but the impact of that evidence on policy-making seemed slim, even thought it might look like it had an effect, since the program was supported by the government all along. For that reason, Louise contended, the ways to measure success for centers working on evidence-based development should be to provide evidence capable of fostering better decisions on the part of policy makers. It should be evidence, going back to GiveWell's request, that influences governments. It should be evidence that is essential to governments or other stakeholders in a

---

[6] Louise did not provide me with specific details about the program such as a location or a period on purpose to preserve anonymity of others, especially given the content of her argument. She provided the name of the funding agency but for reasons of anonymity, I have also chosen not to disclose this information.

particular way, not simply to support projects already receiving political favor, but also to affect action plans. This can mean scaling up or persevering, if the evidence is considered positive, like in the case of the deworming intervention in Kenya, but also discontinuing a program if the evaluation provides evidence of negative or no impact. It is evidence that enables a counterfactual – an otherwise – capable of disrupting[7].

In recent years, the importance of "scaling down" in light of evidence, and not only of scaling up, has started to be emphasized in public discussions about evidence-based development. For instance, one of the six components in J-PAL's new Evidence to Policy pathway is "Scaling back an evaluated program" premised on the notion that "incorporating evidence into decision-making is not only about scaling up effective programs. Results that show a program doesn't work can be just as critical."[8] The webpage also provides examples of programs that have been discontinued as a result of the program being evaluated (via RCTs) and found ineffective, including an initiative to increase doctor's attendance at primary health care centers through biometric monitoring technology in the state of Karnataka in India. Another example given by the website is a policy from the French government that would require firms to make hiring decisions based on anonymized resumes, and whose impact evaluation showed that this actually harmed minority applicant's employment chances because removing identifying information posed an issue for firms who wished to contextualize résumés and discriminate positively.

---

[7] Among development practitioners, the continuation of Progresa, a government social assistance program in Mexico, across distinct political administrations has been cited as a sign of the ability of independent impact evaluations to guide political decision-making. In fact, Progresa was started in 1997 under the socialist-inclined government of President Zedillo and (as mentioned in the Introduction) evaluated around those years by a team from the International Food Policy Research Institute. The results of the first set of impact evaluations were made public after the presidential elections (in July 2000) but before the new administration took over in December 2000. The results are said to have influenced the incumbent right-wing government of President Fox to keep in place a program of social assistance more traditionally associated with social party programs, while rebranding it as Oportunidades in the early 2000s (see for instance Levy 2006: 12).

[8] https://www.povertyactionlab.org/evidence-to-policy/scaling-back-evaluated-program, accessed July 31, 2019

**Absent evidence**

The absence of evidence as well might have power to reach desired political ends. It allows the status-quo to go unquestioned. Idris, the economics professor mentioned earlier, argued that in some cases there might be zero interest in running a randomized evaluation, or even an interest in *not* doing so. He provided the example of toilet construction being the "pet policy" of the current Prime Minister of India. He was referring to current Narendra Modi's "Clean Mission", which, since 2014, aims to build millions of (household) latrines in a pledge to ensure universal sanitation by the end of 2019[9]. Idris told me: "there should be RCTs conducted about that [initiative], because the evidence we have says there is limited evidence about the efficacy of that initiative, but no official wants to sign on this because it's the pet project of the Prime Minister, because it would not be good if a negative impact is found".

In a recent article, Lindsey McGoey (2017: 276) discusses the power of "absent evidence" in shaping current policy agendas around wealth inequality through the example of economist Thomas Piketty's now world-famous treatise *Capital in the Twenty-First Century*. In the book, Piketty calls for policy mechanisms that would increase taxation on the world's highest earners to reduce global wealth disparities. But what McGoey – along with other critics of Piketty – point out is that Piketty "does not consider the possibility that [their] wealth was ill earned in the first place", for example through abusive appropriation of land or through patents that restrict the right to a factor of production to a single company or individual (2017: 269). This omission, McGoey further underscores, is not the result of Piketty ignoring countervailing and inconvenient evidence, but it is because the datasets that disaggregate between "unproductive" and "productive" wealth are simply not there (2017: 276). And,

---

[9] See for instance https://www.independent.co.uk/life-style/design/india-biggest-toilet-building-spree-narendra-modi-a8512026.html, accessed August 1, 2019

McGoey continues, "one of the reasons why they don't exist empirically is because generations of neoclassical economists did not see the need to generate them" (ibid).

In fact, according to McGoey (2017: 264), Piketty's work – but also that of mainstream economics– is indebted to the theory of the marginal productivity of income distribution, developed by economist John Bates Clark in the nineteenth century and which stipulates that "under ideal conditions, laborers receive an economic contribution that reflects their "natural" contribution to the production process". This theory, despite having been disputed for decades by mainstream and heterodox scholars, remains highly influential today, sustaining long-entrenched cultural maxims that are pervasive in the West: "incomes reflect talent, top managers deserve top pay, and hard work pays off" (2017: 276). This reliance on Clark's theory therefore sidestepped the need to trace the *origins* of wealth and to measure wealth only *after* it reaches the hands of private individuals or corporations, therefore leading Piketty and most policy makers to focus solely on income redistribution rather than wealth acquisition (2017: 272).

Similarly, Idris was raising the possibility that perhaps toilet construction was after all perhaps not such a magical pathway to reducing issues of poverty in India. Idris meant that there should be more systematic evidence collected about its causal impact, whether on health, gender gap, or other outcomes. While the project has granted the government important popularity, with news releases and official statements reporting that the Clean Mission initiative increased women empowerment or helped generate employment[10], such systematic evidence was lacking. What Idris was underscoring is that this absence was perhaps intentional, entrenched by those who had a political interest in not knowing,  asides from news stories and

---

[10] For example: https://asia.nikkei.com/Spotlight/Cover-Story/World-s-biggest-toilet-building-project-empowers-India-s-women, accessed August 1, 2019

individual or village-based stories, the (statistical) impact of that program on people's livelihoods.

Importantly, the *absence of evidence* is not the same as *evidence for the absence of impact*, such as the null result findings from a randomized evaluation: e.g. deworming does not improve test scores or building toilets across India does not reduce open defecation and its associated negative health outcomes[11]. As alluded to with the example of Louise and scaling down mentioned above, most of my interlocutors agreed that evidence about the lack of impact of an intervention should have an influence on policy decisions – to discontinue a project, to seek for alternatives, or in the least foster public debate – just in the same way that evidence of positive or negative impacts might guide governments or NGOs in their budgeting and priorities. What comes into view when one thinks about the absence of evidence, instead of the evidence itself, as powerful?

**The ambiguous meanings of evidence**

On a July morning, leaving behind the Kisumu office of Evidence Action, I head to the airport in the organization's van along with its permanent occupant, the organization's driver, Francis. We pass through Kisumu's industrial zone lined along the shore of the lake in which two days ago, I saw beautiful kingfishers flying relentlessly over fishermen's catch of the day. As detailed in Chapter 2, I had just spent a week in western Kenya working along field staff who install, maintain, record, and track chlorine dispensers installed near small rivers or

---

[11] Differentiating between the absence of evidence and evidence for the absence of impact might, for some readers, summon an (in)famous quote by Donald Rumsfeld, United States Secretary of Defense under George W. Bush and during the 2003 invasion of Iraq. Answering a question about terrorism and weapons of mass destruction at a 2002 press conference at the North Atlantic Treaty Organization, Rumsfeld said: "The absence of evidence is not evidence of absence. It is basically saying the same thing in a different way. Simply because you do not have evidence that something exists does not mean that you have evidence that it doesn't exist. And yet almost always, when we make our threat assessments, when we look at the world, we end up basing it on the first two pieces of that puzzle, rather than all three." In light of the ensuing war waged in Iraq by the USA, this line of thought –and its invocation of the absence of evidence as not being equivalent to the evidence absence of evidence– seemed to have had tremendous effects. See https://www.nato.int/docu/speech/2002/s020606g.htm, accessed August 14 2019.

boreholes in an effort to secure "safe" water for households throughout the region. As our van entered the Kisumu airport, the policeman who opens the gate says "*Evidence Action*" out loud. He is reading the NGO's logo displayed in pink on the side of the van facing him, right by my window. Harboring an inscrutable smile, the policeman asks looking at me in the eyes: "Where is the evidence? What kind of evidence?" Uneasy, I chuckle, then look at Francis in the driver's seat who shows no sign of bringing the vehicle to a halt, and finally I utter the only answer that comes to mind: "I don't know. I'm just a passenger travelling."

Our van continues its journey towards the airport, leaving behind the policeman with my excuse in lieu of an answer. Francis was likely, politely, waiting for me to reply on behalf of Evidence Action, me being the researcher and him the organization's driver. I could have provided the policeman with a real rather than an evasive reply. I could have told him that Evidence Action is a non-governmental organization whose mandate is to scale development programs that are supported by "rigorous evidence", for instance survey data and water quality tests that support the causal claim that installing chlorine dispensers at village water sources does indeed improve sanitation around Kisumu. But I did not. Time was running out before my flight's departure. And considering that I was not a staff member of Evidence Action but an anthropologist gnawed by the persistent worry of being a burdensome outsider, I did not want to speak for the NGO that day. Nor any other day, really.

The evidence at stake here is thus made, among other things, of paper questionnaires, water samples and their accompanying water quality measurements, field staff who carry this material back to the main offices in Kisumu and Nairobi to be entered into data entry software, and then analyzed, transformed, published as academic papers and policy briefs to be discussed, debated, disseminated and used as the basis to forge government or private initiatives in Kenya and elsewhere. This morning, none of these people and things are in the van.

As we can still discern the policeman in the rear windshield, Francis lets me know that this is not the first time a policeman asks about the logo *Evidence Action* when encountering the van. "These policemen", Francis tells me, "worry that we carry evidence of corruption, and especially about their doings". I wished Francis could have told me more, but we have reached the airport terminal and I must rush to catch my flight. Landing in Nairobi, I therefore seize the opportunity to tell Bernard about the encounter, the taxi-driver who became a friend over my months of fieldwork there. Before I had the time to mention Francis' interpretation, Bernard interjects: "Of course the policeman asked because he was worried you had some evidence of corruption. Policemen are often guilty of something, otherwise they would not react like this."

In some ways, through the policeman's questions, evidence about the effectiveness of anti-poverty projects (collected through surveys, measurements, and interviews by field staff on behalf of development organizations and academics) acquired a new life: transformed, in however a fickle and transient manner, into evidence about corruption (the domain of journalists and special taskforces).

Did the policeman really fear that Francis and I were part of an initiative to investigate public sector corruption? Was he really concerned that the van might be the bearer of compromising evidence about acts of corruption committed by the police? Were his questions triggered by fear or simple curiosity at the sight of a pink logo stating Evidence Action in capital letters? To Francis and Bernard, the policeman's enquiry was most clearly a sign of nervousness on his part.

On the day of my encounter with him, the policeman ignored everything about the nature of the evidence that Evidence Action is working to gather. He does not even know whether, really, the van transports any evidence at all. He merely read *evidence* on the van, paired in that Evidence Action logo with the term *action*, which does not say much, except that some undertaking is hoped to be derived from that evidence, serving perhaps to enhance an

outsider's anxiety. This triggered him to ask questions ("Where is the evidence? What kind of evidence") taken by Francis and Bernard as a sign of worry from a policeman wondering whether our van was carrying compromising information. Even when he directed his questions at me, the evidence in its materiality (documents, recording, graphs, water samples, etc.) was not shown to him. It is precisely this ambiguity that causes the confusion: it enables the policeman to imagine that "we", as Evidence Action, are working to fight corruption in Kenya. The sight of evidence, in the form of a word, sufficed to produce an effect. But it is not what my interlocutors aim for when they talk of evidence having an impact in the world. While this dissertation has been concerned with a specific kind of evidence, what evidence becomes when it travels beyond the experimental site is not so obvious because it can take on several valences and meanings, becoming as much about statistical estimates derived from RCTs as about legal proofs, or something that is as much produced by scientific practices as by courts of justice.

## What does evidence do?

I presented this last anecdote with the policeman because it is a playful way to make visible that the power of evidence cannot be known in advance because its effects are multiple, unexpected, unknowable, and uncertain. As the vignettes presented above gesture to, evidence from randomized evaluations can be taken up to convince a government to put in place a new program, or on the contrary to discontinue a policy if it is shown to be ineffective. But evidence can also be used to justify a program, *as if* the government program was based on evidence, when in fact government actors had already committed to it for other reasons such as perhaps electoral prospects, unrelated to the evidence. At yet other times, evidence-based development might lack traction, or power, against political forces that have incentives not to produce statistical evidence at all about certain popular policies, such as in the example of the Clean Mission in India.

I chose to conclude with a final reflection on the impact of evidence produced by randomized evaluations because this dissertation has extensively documented the technologies and aspirations mobilized to scale, expand, and generalize evidence-based findings and programs to reach millions of people across the globe. While I showed how my interlocutors are centrally preoccupied with reaching the highest level of scale, not measured in thousands of lives touched but in millions, my first exchange with Will and other moments during my fieldwork raised the significant question of how much evidence-based development stirs up "politics as we know it", and how much statistical indicators really do guide political efforts worldwide rather than other factors, such as beliefs, opinions, or group interests.

In recent years, sciences and studies scholarship has come around the important critical question of "what counts as Evidence?" Like Loes Knaapen (2013: 683) in her examination of evidence-based medicine, I put here a capital E to Evidence to denote that anthropologists have traced what has come to count as valid and legitimate scientific Evidence in global health, development, or other policy efforts around the world. Concerned with the rising importance and ubiquity of numbers, scholars have payed particular attention to the epistemic exclusions and inclusions at stake in these practices, pointing out to widespread Evidence hierarchies, which encourages the RCT while disqualifying other methods of knowledge. In the process, these global indicators are portrayed as "translat[ing] the buzzing confusion of social life into neat categories that can be tabulated" and thus as potentially distorting complex social phenomena and policy objectives (Merry 2016: 3).

My interlocutors' musings on the impact of evidence on policy making seemed to raise another parallel line of inquiry. Throughout the few anecdotes I have weaved together, those who practice evidence-based development ask: "How does our evidence shape policy?" "Does evidence do enough, does it disrupt policy?" "Do we have enough evidence, or do we need more of it in places which are giving way to autocratic governments?" From my interlocutors'

perspective, these diverse queries seem to encapsulate a daunting interrogation: "do our imaginations succeed or fail, and how can we know it?" To the vast research agenda of "What counts as Evidence?", the ethnographic material presented in this conclusion seems to juxtapose an equally complex question that one could summarize as "What does evidence do?" or "What does evidence *manage* to do?" In starting to examine this problem, it seems to matter less whether we are talking only about evidence from randomized evaluations or about the potential of research in general to shape politics today.

My intention is not to entrench the image of failed states in the global South, persisting in being led by corruption schemes rather than scientific evidence. Instead, I wonder about the "power of evidence" today everywhere, even in so-called Western democracies, where the exploitation of fears, racism, prejudices, and verbal aggression seem to have become common features of political strategies. Following what evidence does might request us to qualify the power of evidence and importantly document what other kinds of forces– and with what intricate consequences– evolve, contest, and struggle alongside the purview of evidence-based development to shape policy around the world today.

# References

Abdelghafour, Nassima
    2017   Randomized Controlled Experiments to End Poverty?. A Sociotechnical Analysis. *Anthropologie & développement* (46-47):235-262.

——

    n.d.   The micropolitics of development: Experimental approaches to anti-poverty policy-making in developing countries.

Abdul Latif Jameel Poverty Action Lab
    2011   *The Price is Wrong*. J-PAL Bulletin.

Adams, Vincanne
    2013   Evidence-Based Global Public Health: Subjects, Profits, Erasures. *In* When people come first : critical studies in global health. J.o.G. Biehl and A. Petryna, eds. Pp. pages cm: Princeton University Press.

——

    2016   *Metrics: What counts in global health*: Duke University Press.

Adams, Vincanne, and Tobias Rees
    2014   GATESKEEPING: History, Markets and Metrics in Global Health. *In* Workshop Proposal Draft.

Agamben, Giorgio
    1998   *Homo sacer: Sovereign power and bare life*: Stanford University Press.

Ahuja, Amrita, Michael Kremer, and Alix Peterson Zwane
    2010   Providing safe water: Evidence from randomized evaluations. *Annu. Rev. Resour. Econ.* 2(1):237-256.

Aiken, Alexander M, et al.
    2015   Re-analysis of health and educational impacts of a school-based deworming programme in western Kenya: a pure replication. *International journal of epidemiology* 44(5):1572-1580.

Akman, Varol
    2000   Rethinking context as a social construct. *Journal of Pragmatics* 32(6):743-759.

Alkire, Sabina, and James Foster.
    2011.   Counting and multidimensional poverty measurement. *Journal of public economics* 95, no. 7-8 (2011): 476-487.

Amselle, Jean-Loup
    1990   *Logiques métisses: anthropologie de l'identité en Afrique et ailleurs*. Paris: Payot.

Anand, Nikhil
    2011   Pressure: The PoliTechnics of Water Supply in Mumbai. *Cultural Anthropology* 26(4):542-564.

—

 2015 Leaky States: Water Audits, Ignorance, and the Politics of Infrastructure. *Public Culture* 27(2):305-330.

—

 2016 Hydraulic Publics. *In* limn. S.J. Collier, J.C. Mizes, and A. von Schnitzler, eds, Vol. Issue 7: Public Infrastructures/Infrastructural Publics.

Appel, Hannah
 2017 Toward an ethnography of the National Economy. *Cultural Anthropology* 32(2):294-322.

Arndt, H. W.
 1981 Economic Development: a Semantic History. *Economic Development and Cultural Change* 29(3):457-466.

—

 1987 *Economic Development : the History of an Idea*. Chicago: University of Chicago Press.

Ashraf, Nava, Neil Shah, and Rachel Gordon
 2010 Deworming Kenya: Translating Research into Action (A). *Harvard Business School NOM Unit Case* (910-001).

Baird, Sarah, et al.
 2016 Worms at work: Long-run impacts of a child health investment. *The quarterly journal of economics* 131(4):1637-1680.

Balibar, Étienne
 2016 *Des Universels: Essais et conférences*: Éditions Galilée.

Banerjee, Abhijit, et al.
 2015a A multifaceted program causes lasting progress for the very poor: Evidence from six countries. *Science* 348(6236):1260799.

Banerjee, Abhijit, Dean Karlan, and Jonathan Zinman
 2015b Six randomized evaluations of microcredit: Introduction and further steps. *American Economic Journal: Applied Economics* 7(1):1-21.

Banerjee, Abhijit V, et al.
 2007 Remedying education: Evidence from two randomized experiments in India. *The Quarterly Journal of Economics* 122(3):1235-1264.

Banerjee, Abhijit V., and Esther Duflo
 2011 *Poor Economics : a Radical Rethinking of the Way to Fight Global Poverty*. New York: PublicAffairs.

Banerjee, Abhijit Vinayak, et al.
 2010 Improving immunisation coverage in rural India: clustered randomised controlled evaluation of immunisation campaigns with and without incentives. *Bmj* 340:c2220.

Barbu, Alain, and Howard White
    2006  *Impact Evaluation - The Experience of the Independent Evaluation Group*. Independent Evaluation Group, World Bank.

Bashford, Alison
    2014a *Global population: history, geopolitics, and life on earth*: Columbia University Press.

—

    2014b Introduction: Life and Earth. *In* Global Population: History, Geopolitics, and Life on Earth. Columbia University Press.

Beaudoin, Steven M.
    2007  Poverty and poor relief after 1945. *In* Poverty in World History. Abingdon, Oxon ; New York: Routledge.

Bemme, Dörte
    2018  Contingent universals, aggregated truth:  an ethnography of knowledge in Global Mental Health. Department of Anthropology, McGill University. PhD.

—

    2019  Finding "What Works": Theory of Change, Contingent Universals, and Virtuous Failure in Global Mental Health. *Culture, Medicine, and Psychiatry*:1-22.

Bentham, Jeremy
    1970 [1789]  *An Introduction to the Principles of Morals and Legislation*. University of London: Athlone Press.

Berndt, Christian, and Marc Boeckler
    2016  Behave, global south! Economics, experiments, evidence. *Geoforum* 70:22-24.

Best, Jacqueline
    2013  Redefining Poverty as Risk and Vulnerability: Shifting Strategies of Liberal Economic Governance. *Third World Quarterly* 34(1):109-129.

—

    2017a  Governing Development Failure. *In* limn. S.J. Collier, J. Cross, P. Redfield, and A. Street, eds, Vol. Issue 9: Little Development Devices / Humanitarian Goods.

—

    2017b The rise of measurement-driven governance: The case of international development. *Global Governance* 23(2):163.

Bhatt, Jigar
    2011  Causality and the Experimental Turn in Development Economics. *New School Economic Review* 5(1):50-67.

—

    2017  Learning to learn again: Restoring relevance to development experiments through a whole systems approach. *In* The Routledge Companion to Planning in the Global South. Pp. 368-379: Routledge.

—

    2018  Politics by Other Means: Economic Expertise, Power, and Global Development Finance Reform. Columbia University.

Biehl, João, and Adriana Petryna, eds.
    2013    When People Come First: Critical Studies in Global Health. Princeton and
    Oxford: Princeton University Press.

Biesel, Uli
    2014    On gloves, rubber and the spatio-temporal logics of global health.
    Somatosphere: Ebola Fieldnotes. http://somatosphere.net/2014/rubber-gloves-global-
    health.html/96, August 10, 2019

Bornstein, Erica, and Peter Redfield
    2011    *Forces of compassion: humanitarianism between ethics and politics*: School for
    Advanced Research Press.

boyd, danah, and Kate Crawford
    2012    Critical questions for Big Data: Provocations for a cultural, technological, and
    scholarly phenomenon. *Information, communication & society* 15(5):662-679.

Brown, Hannah
    2015    Global health partnerships, governance, and sovereign responsibility in western
    Kenya. *American Ethnologist* 42(2):340-355.

Brown, Hannah, and Maia Green
    2017    Demonstrating development: meetings as management in Kenya's health sector.
    *Journal of the Royal Anthropological Institute* 23(S1):45-62.

Brown, Peter
    2002    *Poverty and Leadership in the Later Roman Empire*: UPNE.

Brown, Theodore M, Marcos Cueto, and Elizabeth Fee
    2006    The World Health Organization and the Transition from" International" to"
    Global" Public Health. *American Journal of Public Health* 96(1).

Burke, Peter
    2002    Context in Context. *Common knowledge* 8(1):152-177.

Çalışkan, Koray, and Michel Callon
    2010    Economization, part 2: a research programme for the study of markets. *Economy
    and Society* 39(1):1-32.

Callon, Michel
    1998    The laws of the markets. *Sociological review monograph*.
——
    2007    What does it mean to say that economics is performative? *Do economists make
    markets? On the performativity of economics*:311-357.

Campbell, Donald T
    1969    Reforms as Experiments. *American Psychologist* 24(4):409.

Carr, E. Summerson, and Brooke Fisher

2016    Interscaling Awe, De-escalating Disaster. *In* Scale: Discourse and Dimensions of Social Life. E.S. Carr and M. Lempert, eds: University of California Press.

Carr, E. Summerson, and Michael Lempert, eds.
2016    Scale: Discourse and Dimensions of Social Life: University of California Press.

Carrier, James G
1998    Introduction to D. Miller and J. Carrier. *In* Virtualism: A New Political Economy. J.G. Carrier and D. Miller, eds: Berg publishers.

Cherrier, Beatrice
2015    History of policy evaluation: a few questions. The Undercover Historian. https://beatricecherrier.wordpress.com/2015/02/03/history-of-policy-evaluation-a-few-questions/56, August 9, 2019

—

2017    From physicists to engineers to meds to plumbers: Esther Duflo rediscovering the lost art of economics @ASSA2017. The Undercover Historian. https://beatricecherrier.wordpress.com/2017/01/07/from-physicists-to-engineers-to-meds-to-plumbers-esther-duflo-rediscovering-the-lost-art-of-economics-assa2017/76, August 9, 2019

Clifford, James
1997    *Routes: Travel and Translation in the Late Twentieth Century*. Cambridge, MA: Harvard University Press.

Cochrane, Archibald Leman
1972    *Effectiveness and efficiency: random reflections on health services*. Volume 900574178: Nuffield Provincial Hospitals Trust London.

Cohen, Jessica, and Pascaline Dupas
2010    Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment. *Quarterly Journal of Economics* 125(1):1.

Collier, Paul
2007    *The Bottom Billion : Why the Poorest Countries Are Failing and What Can Be Done About It*. Oxford ; New York: Oxford University Press.

Collier, Stephen J.
2011    *Post-Soviet Social : Neoliberalism, Social Modernity, Biopolitics*. Princeton N.J.: Princeton University Press.

—

2013    Two Forms of Vulnerability. *In* Presented to the CRESC Annual Conference London, September 2013.

—

2016    Email message to author on 'state/government'.

Collier, Stephen J., et al.
2017    Preface: Little Development Devices / Humanitarian Goods. *In* limn. S.J. Collier, J. Cross, P. Redfield, and A. Street, eds.

Colvin, Christopher J.

    2016  New Political Ecologies of Knowledge?: The Integration of Qualitative Evidence in Global Health Guideline Development and Policymaking. *In* Workshop on Evidentiary Practices: Testing, Measuring and Accounting in Global Infrastructures. Berlin

Cornia, Giovanni Andrea, Richard Jolly, and Frances Stewart

    1987  *Adjustment with a Human face: Protecting the Vulnerable and Promoting Growth: A study by UNICEF*: Oxford University Press.

Crane, Johanna T

    2010  Unequal 'Partners'. AIDS, Academia, and the Rise of Global Health. *BEHEMOTH-A Journal on Civilisation* 3(3):78-97.

Craven, Matthew

    2015  Between law and history: the Berlin Conference of 1884-1885 and the logic of free trade. *London Review of International Law* 3(1):31-59.

Crow, Ben, and Edmond Odaba

    2010  Access to water in a Nairobi slum: women's work and institutional learning. *Water International* 35(6):733-747.

Cruz, Katie

    2017  Sally Engle Merry: The Seductions of Quantification: Measuring Human Rights, Gender Violence, and Sex Trafficking. *Journal of Law and Society* 44(4):726-731.

Cull, Robert, and Jonathan Morduch

    2017  Microfinance and economic development: The World Bank.

Culler, Jonathan D.

    1983  *On Deconstruction: Theory and Criticism after Structuralism*. London: Routledge & Kegan Paul.

——

    1988  *Framing the Sign: Criticism and its Institutions*. Norman, OK: University of Oklahoma Press

Curtis, Bruce

    2002  Foucault on governmentality and population: The impossible discovery. *Canadian journal of sociology* 27(4):505-535.

Das, Veena, and Deborah Poole

    2004a *Anthropology in the Margins of the State*: Oxford University Press.

——

    2004b State and its Margins: Comparative Ethnographies. *In* Anthropology in the Margins of the State. V. Das and D. Poole, eds. Pp. 3-33: Oxford University Press.

Davey, Calum, et al.

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(5):1581-1592.

de Laet, Marianne, and Annemarie Mol
2000    The Zimbabwe Bush Pump: Mechanics of a Fluid Technology. *Social studies of science* 30(2):225-263.

de Sardan, Jean-Pierre Olivier, Aissa Diarra, and Mahaman Moha
2017    Travelling models and the challenge of pragmatic contexts and practical norms: the case of maternal health. *Health research policy and systems* 15(1):60.

de Sardan, Jean-Pierre Olivier, and Emmanuelle Piccoli
2018    Cash Transfers and the Revenge of Contexts: An Introduction. *In* Cash Transfers in Context: An Anthropological Perspective. J.-P.O. de Sardan and E. Piccoli, eds: Berghahn Books.

de Souza Leão, Luciana, and Gil Eyal
2019    The rise of randomized controlled trials (RCTs) in international development in historical perspective. *Theory and Society* 48(3):383-418.

Deaton, Angus, and Nancy Cartwright
2018    Understanding and misunderstanding randomized controlled trials. *Social Science & Medicine* 210:2-21.

Deaton, Angus S
2009    *Instruments of development: Randomization in the tropics, and the search for the elusive keys to economic development*. National Bureau of Economic Research.

Desrosières, Alain
2002    *The politics of large numbers: A history of statistical reasoning*: Harvard University Press.

Díaz Langou, Gala, and Vanesa Weyrauch
2013    Sound expectations: from impact evaluations to policy change. *Journal of Development Effectiveness* 5(3):269-304.

Dilley, Roy M, ed.
1999    Perspectives from Social Anthropology and Elsewhere: Berghahn Books.
—
2002    The problem of context in social and cultural anthropology. *Language & Communication* 22(4):437-456.

Donovan, Kevin P.
2018    The rise of the randomistas: on the experimental turn in international aid. *Economy and Society* 47(1):27-58.

Donzelot, Jacques
1984    *L'invention du social: essai sur le déclin des passions politiques*: Fayard.

Douglas, Mary
    1966    *Purity and Danger: an Analysis of the Concepts of Pollution and Taboo*.
    London: Ark Paperbacks.

Dourish, Paul
    2004    What we talk about when we talk about context. *Personal and ubiquitous
    computing* 8(1):19-30.

Drèze, Jean
    2018    Evidence, policy, and politics. Ideas for India for more evidence-based policy.
    https://www.ideasforindia.in/topics/miscellany/evidence-policy-and-politics.html88,
    August 9, 2019.

Duflo, Esther.
    2010    Social    Experiments    to    Fight    Poverty.    TED    Talk    at
    https://www.ted.com/talks/esther_duflo_social_experiments_to_fight_poverty?langua
    ge=en, November 27, 2019.

Duflo, Esther
    2017    Richard T. Ely Lecture: The Economist as Plumber. *American Economic
    Review* 107(5):1-26.

Duflo, Esther, Rachel Glennerster, and Michael Kremer
    2007    Using randomization in development economics research: A toolkit. *Handbook
    of development economics* 4:3895-3962.

Duflo, Esther, Michael Kremer, and Jonathan Robinson
    2011    Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from
    Kenya. *American Economic Review* 101(6):2350-90.

Duranti, Alessandro, and Charles Goodwin, eds.
    1992    Rethinking Context: Language as an Interactive Phenomenon: Cambridge
    University Press.

Edwards, Paul N.
    2017    The Mechanics of Invisibility: On Habit and Routine as Elements of
    Infrastructure. *In* Infrastructure Space. I. Ruby and A. Ruby, eds: Ruby Press.

Ehrenstein, Véra, and Fabian Muniesa
    2013    The conditional sink: counterfactual display in the valuation of a carbon
    offsetting reforestation project. *Valuation Studies* 1(2):161-188.

Ehrenstein, Véra, and Daniel Neyland
    2018    On scale work: evidential practices and global health interventions. *Economy
    and Society* 47(1):59-82.

Elyachar, Julia
    2005    *Markets of dispossession : NGOs, economic development, and the state in
    Cairo*. Durham: Duke University Press.

---

2010    Phatic labor, infrastructure, and the question of empowerment in Cairo. *American Ethnologist* 37(3):452-464.

Englund, Harri
2004    Conflicts in context. *In* Violence and Belonging: The Quest for Identity in Post-Colonial Africa. V. Broch-Due, ed. Pp. 54-67. London and New York: Routledge.

Escobar, Arturo
1994    *Encountering Development: The Making and Unmaking of the Third World*: Princeton University Press.

Ewald, François
2016    François Ewald on Foucault & Neoliberalism. 13 Years at the Collège / 13 Seminars                            at                            Columbia. http://blogs.law.columbia.edu/foucault1313/2016/01/24/ewaldneoliberalism/#sthash.E 380Q3i1.uxfs&st_refDomain=t.co&st_refQuery=/cTH5jps9se166, August 12, 2019

Fabian, Johannes
1995    Etnographic Misunderstanding and the Perils of Context. *American Anthropologist* 97(1):41-50.

Fadaak, Raad
2019    Prevent, Detect, Respond: An Ethnography of Global Health Security. Department of Anthropology, McGill University. PhD.

Fassin, Didier
2007    Humanitarianism as a Politics of Life. *Public culture* 19(3):499-520.
—

2012    *Humanitarian Reason: A Moral History of the Present*. Berkeley and Los Angeles: University of California Press.

Feldman, Ilana, and Miriam Ticktin, eds.
2010    In the Name of Humanity: The Government of Threat and Care. Durham & London: Duke University Press.

Fennell, Catherine
2016    Are We All Flint? *In* limn. S.J. Collier, J.C. Mizes, and A. von Schnitzler, eds, Vol. Issue 7: Public Infrastructures/Infrastructural Publics.

Ferguson, James
1990    *The Anti-Politics Machine : "Development," depoliticization, and bureaucratic power in Lesotho*: University of Minnesota Press.
—

1997    Anthropology and its Evil Twin. *International development and the social sciences: Essays on the history and politics of knowledge*:150-75.
—

2006a    *Global Shadows: Africa in the Neoliberal World Order*: Duke University Press.
—

2006b  Transnational Topographies of Power: Beyond "the State" and "Civil Society" in the Study of African Politics. *In* Global Shadows: Africa in the Neoliberal World Order. Pp. 88-112: Duke University Press.

——

2010    The Uses of Neoliberalism. *Antipode* 41(S1):166-184.

——

2013    Declarations of dependence: labour, personhood, and welfare in southern Africa. *Journal of the Royal Anthropological Institute* 19(2):223-242.

——

2015    *Give a Man a Fish: Reflections on the New Politics of Redistribution*: Duke University Press.

Ferguson, James, and Akhil Gupta
2002    Spatializing states: toward an ethnography of neoliberal governmentality. *American ethnologist* 29(4):981-1002.

Ferguson, Yale H., and Richard W. Mansbach
2007    Post-internationalism and IR Theory. *Millennium* 35(3):529-549.

Finnemore, Martha
1997    Redefining Development at the World Bank. *International development and the social sciences: Essays on the history and politics of knowledge*:203-27.

Fisher, Ronald Aylmer
1935    *The Design of Experiments*. Edinburgh: London Oliver & Boyd.

Fishlow, Albert
1972    Brazilian Size Distribution of Income. *The American Economic Review* 62(1/2):391-402.

Foucault, Michel
1977    *Discipline and punish: The birth of the prison, trans. Alan Sheridan*. New York: Pantheon Books.

——

1982    The Subject and Power. *Critical inquiry* 8(4):777-795.

——

2004    *The Birth of Biopolitics: Lectures at the Collège de France, 1978-1979*. New York: Palgrave Macmillan.

Fourcade, Marion
2009    *Economists and societies: Discipline and profession in the United States, Britain, and France, 1890s to 1990s*: Princeton University Press.

Gal, Susan
2016    Scale-Making: Comparison and Perspective as Ideological Projects. *In* Scale: Discourse and Dimensions of Social Life. E.S. Carr and M. Lempert, eds: University of California Press.

Garland, David

2014    What is a "history of the present"? On Foucault's genealogies and their critical preconditions. *Punishment & Society* 16(4):365-384.

Geertz, Clifford
1963    *Peddlers and Princes: Social Development and Economic Change in Two Indonesian Towns*. Chicago: University of Chicago Press.

Geissler, P Wenzel
2013    Public secrets in public health: Knowing not to know while making scientific knowledge. *American Ethnologist* 40(1):13-34.

Geissler, P Wenzel, and Catherine Molyneux
2017    *Evidence, ethos and experiment: the anthropology and history of medical research in Africa*: Berghahn Books.

Geremek, Bronisław
1994    *Poverty : a history*. Oxford ; Cambridge, Mass.: Blackwell.

Gertler, Paul
2004    Do conditional cash transfers improve child health? Evidence from PROGRESA's control randomized experiment. *American economic review* 94(2):336-341.

Gertler, Paul J, et al.
2011    *Impact evaluation in practice*: World Bank Publications.

Geschiere, Peter
2009    *The Perils of Belonging: Autochthony, Citizenship, and Exclusion in Africa and Europe*. Chicago: University of Chicago Press.

Green, Maia
2006    Representing Poverty and Attacking Representations: Perspectives on Poverty from Social Anthropology. *The Journal of Development Studies* 42(7):1108-1129.

Gupta, Akhil
1995    Blurred boundaries: the discourse of corruption, the culture of politics, and the imagined state. *American ethnologist* 22(2):375-402.
—
2012    *Red Tape: Bureaucracy, Structural Violence, and Poverty in India*: Duke University Press.
—
2015    An Anthropology of Electricity from the Global South. *Cultural Anthropology* 30(4):555-568.

Gupta, Akhil, and James Ferguson
1992    Beyond "Culture": Space, Identity, and the Politics of Difference. *Cultural Anthropology* 7(1: ):6-23.

Gutton, Jean Pierre

1971    *La Société et les pauvres (Lyon)*. Volume 26: Société d'édition" Les Belles Lettres".

Hanlon, Joseph
    1991    *Mozambique: Who Calls the Shots?*: James Currey.

Harvey, Penny, and Hannah Knox
    2016    Infrastructural Incursions. *In* limn. S.J. Collier, J.C. Mizes, and A. von Schnitzler, eds, Vol. Issue 7: Public Infrastructures/Infrastructural Publics.

Hawthorn, Geoffrey
    1993    *Plausible worlds: Possibility and understanding in history and the social sciences*: Cambridge University Press.

Hearn, Julie
    1998    The 'NGO-isation'of Kenyan society: USAID & the restructuring of health care. *Review of African Political Economy* 25(75):89-100.
—
    2007    African NGOs: The New Compradors? *Development and Change* 38(6):1095-1110.

Hetherington, Kregg, and Jeremy M Campbell
    2014    Nature, Infrastructure, and the State: Rethinking Development in Latin America.

High, Casey, Ann H Kelly, and Jonathan Mair, eds.
    2012    The Anthropology of Ignorance: An Ethnographic Approach. New York: Palgrave McMillan.

Himmelfarb, Gertrude
    1984a  The Idea of Poverty. *History Today* 34(4).
—
    1984b  *The idea of poverty : England in the early Industrial Age*. New York: Knopf.

Hirschman, Albert O.
    1981    The Rise and Decline of Development economics. *In* Essays in Trespassing: Economics to Politics and Beyond. Pp. 1-24. Cambridge: Cambridge University Press.

Hjort, Jonas, et al.
    2019    *How Research Affects Policy: Experimental Evidence from 2,150 Brazilian Municipalities*. 0898-2937. National Bureau of Economic Research.

Howe, Cymene, et al.
    2016    Paradoxical Infrastructures: Ruins, Retrofit, and Risk. *Science, Technology, & Human Values* 41(3):547-565.

Huen, Chi W.
    2009    What is context? An ethnophilosophical account. *Anthropological Theory* 9(2):149-169.

Jackson, Stephen
    2005    The state didn't even exist": Non-governmentality in Kivu, Eastern DR Congo. *Between a Rock and a Hard Place: African NGOs, Donors and the State*:165-196.

Jackson, Steven J.
    2014    Rethinking Repair. *Media Technologies: Essays on Communication, Materiality, and Society*:221-39.

Jackson, Steven J., Alex Pompe, and Gabriel Krieshok
    2012    Repair Worlds: Maintenance, Repair, and ICT for Development in Rural Namibia. Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work, 2012, pp. 107-116. ACM.

Jain, S. Lochlann
    2010a   The mortality effect: Counting the dead in the cancer trial. *Public Culture* 22(1):89-117.

——
    2010b   The Mortality Effect: Counting the Dead in the Cancer Trial. *In* In the Name of Humanity: The Government of Threat and Care. I. Feldman and M. Ticktin, eds. Pp. 218-237. Durham & London: Duke University Press.

——
    2013    *Malignant: How cancer becomes us*: University of California Press.

Jatteau, Arthur
    2013    *Les Expérimentations aléatoires en économie*. Paris: La Découverte.

——
    2017    Faire preuve par le chiffre ? Le cas des expérimentations aléatoires en économie. Sciences de l'homme et de la société, École Normale Supérieure Paris-Saclay. Ph.D.

Jensen, Casper Bruun
    2017    Pipe Dreams: Sewage Infrastructure and Activity Trails in Phnom Penh. *Ethnos* 82(4):627-647.

Jerven, Morten
    2013    *Poor numbers: how we are misled by African development statistics and what to do about it*: Cornell University Press.

Kabeer, Naila
    2005    Is microfinance a'magic bullet'for women's empowerment? Analysis of findings from South Asia. *Economic and Political weekly*:4709-4718.

Kaplan, Martha
    2011    Lonely Drinking Fountains and Comforting Coolers: Paradoxes of Water Value and Ironies of Water Use. *Cultural Anthropology* 26(4):514-541.

Kapur, Devesh, John P. Lewis, and Richard Charles Webb
    1997    *The World Bank : its First Half Century*. 2 vols. Washington, D.C.: Brookings Institution.

Karim, Lamia

    2011   *Microfinance and its discontents: Women in debt in Bangladesh*: U of Minnesota Press.

Karlan, Dean, and Jacob Appel
    2011   *More Than Good Intentions: Improving the Ways the World's Poor Borrow, Save, Farm, Learn, and Stay Healthy*: Penguin.

Katzenellenbogen, Simon
    1996   It didn't happen at Berlin: politics, economics and ignorance in the setting of Africa's colonial boundaries. *African boundaries: Barriers, conduits and opportunities*:21-34.

Kelly, Ann H
    2008   Pragmatic Evidence and the Politics of Everyday Practice *In* How Do We Know? Evidence, Ethnography, and the Making of Anthropological Knowledge L. Chua, C. High, and T. Lau, eds. Pp. 97-117. Newcastle, UK: Cambridge Scholars Publishing.

Kelly, Ann H, and Linsey McGoey
    2018   Facts, power and global evidence: a new empire of truth. *Economy and Society* 47(1):1-26.

Kelty, Christopher
    2017   The Participatory Development Toolkit. *In* limn. S.J. Collier, J. Cross, P. Redfield, and A. Street, eds, Vol. Issue 9: Little Development Devices / Humanitarian Goods.

Kenya Ministry of Education, Science, and Technology,, and Ministry of Health
    2016   *Kenya National School-Based Deworming Programme Year 3 (2014 / 2015) Results*.

Knaapen, Loes
    2013   Being 'evidence-based'in the absence of evidence: The management of non-evidence in guideline development. *Social Studies of Science* 43(5):681-706.

Koolwal, Gayatri, and Dominique Van de Walle
    2013   Access to water, Women's work, and Child outcomes. *Economic Development and Cultural Change* 61(2):369-405.

Larkin, Brian
    2013   The Politics and Poetics of Infrastructure. *Annual Review of Anthropology* 42:327-343.

Leigh, Andrew
    2018   *Randomistas: how radical researchers changed our world*: Black Inc.

Levy, Santiago
    2006   *Progress Against Poverty: Sustaining Mexico's Progresa-Oportunidades Program*. Washington D.C.: Brookings Institution Press.

Li, Tania Murray
    2007    *The Will to Improve: Governmentality, Development, and the Practice of Politics*: Duke University Press.

MacKenzie, Donald A, Fabian Muniesa, and Lucia Siu, eds.
    2007    Do Economists Make Markets? On the Performativity of Economics: Princeton University Press.

Mahajan, Manjari
    2017    Philanthropy and the nation-state in global health: The Gates Foundation in India. *Global public health* 13(10):1357-1368.

Malinowski, Bronislaw
    1923    The Problem of Meaning in Primitive Languages. *In* The Meaning of Meaning. C.K. Ogden and I.A. Richards, eds. Pp. 1-10. New York: A Harvest Book.

Malkki, Liisa H.
    1992    National Geographic: The Rooting of Peoples and the Territorialization of National Identity among Scholars and Refugees. *Cultural Anthropology* 7(1):24-44.

——

    2002    News from nowhere: Mass displacement and globalized 'problems of organization'. *Ethnography* 3(3):351-360.

Mann, Gregory
    2015    *From Empires to NGOs in the West African Sahel*: Cambridge University Press.

Masset, Edoardo, et al.
    2013    What is the impact of a policy brief? Results of an experiment in research dissemination. *Journal of Development Effectiveness* 5(1):50-63.

McGoey, Linsey
    2007    On the will to ignorance in bureaucracy. *Economy and Society* 36(2):212-235.

——

    2009    Pharmaceutical Controversies and the Performative Value of Uncertainty. *Science as Culture* 18(2):151-164.

——

    2017    The Elusive Rentier Rich: Piketty's Data Battles and the Power of Absent Evidence. *Science, Technology, & Human Values* 42(2):257-279.

McKay, Ramah
    2012    Afterlives: Humanitarian Histories and Critical Subjects in Mozambique. *Cultural Anthropology* 27(2):286-309.

McNamara, Robert S.
    1973a  Address to the Board of Governers. *In* World Bank. Nairobi, Kenya.

——

    1973b  *One Hundred Countries, Two Billion People: the Dimensions of Development*. New York: Praeger Publishers.

Merry, Sally Engle

2011    Measuring the World: Indicators, Human rights, and Global Governance. *Current Anthropology* 52(S3):S83-S95.

——

2016    *The seductions of quantification: Measuring human rights, gender violence, and sex trafficking*: University of Chicago Press.

Miguel, Edward, and Michael Kremer
2004    Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities. *Econometrica* 72(1):159-217.

Mill, John Stuart
2009 [1879]    Utilitarianism: The Floating Press.

Mirowski, Philip
1992    Do economists suffer from physics envy? *Finnish Economic Papers* 5(1):61-68.

Mitchell, Timothy
2002    *Rule of Experts : Egypt, Techno-Politics, Modernity*. Berkeley: University of California Press.

——

2005    The Work of Economics: How a Discipline Makes its World. *European Journal of Sociology / Archives Européennes de Sociologie* 46(02):297-320.

——

2008    Rethinking economy. *Geoforum* 39(3):1116-1121.

Mongia, Radhika V
2007    Historicizing State Sovereignty: Inequality and the Form of Equivalence. *Comparative Studies in Society and History* 49(2):384-411.

Moore, Adam
2008    Rethinking scale as a geographical category: from analysis to practice. *Progress in human geography* 32(2):203-225.

Morduch, Jonathan
2017    Microfinance as a credit card. *Limn* (9).

Morvant-Roux, Solène, et al.
2014    Adding Value to Randomization with Qualitative Analysis: The Case of Microcredit in Rural Morocco. *World Development* 56:302-312.

Murphy, Michelle
2017    *The economization of life*: Duke University Press.

Nguyen, Vinh-Kim
2010    *The Republic of Therapy: Triage and Sovereignty in West Africa's Time of AIDS*: Duke University Press.

Null, Clair, et al.

2018    Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Kenya: a cluster-randomised controlled trial. *The Lancet Global Health* 6(3):e316-e329.

Nyanchaga, Ezekiel Nyangeri, and Kenneth S. Ombongi
2007    History of Water Supply and Sanitation in Kenya, 1895-2002. *In* Environmental History of Water. P.S. Juuti, T.S. Katko, and H.S. Vuorinen, eds. Pp. 271-320. London: IWA Publishing.

O'Connor, Alice
2009    *Poverty knowledge: Social science, social policy, and the poor in twentieth-century US history*: Princeton University Press.

Ogden, Timothy N
2016    *Experimental conversations: Perspectives on randomized trials in development economics*: MIT Press.

Orr, Julian E
1999    *Talking about Machines: An Ethnography of a Modern Job*. Ithaca and London: Cornell University Press.

Pawson, Ray
2006    *Evidence-based policy: A realist perspective*: Sage.

Pearl, Judea, and Dana Mackenzie
2018    *The book of why: the new science of cause and effect*: Basic Books.

Pearson, Lester
1970    *The Pearson Report: A New Strategy for Global Development*. Paris: UNESCO, United Nations.

Petryna, Adriana
2007    Experimentality: On the global mobility and regulation of human subjects research. *PoLAR: Political and Legal Anthropology Review* 30(2):288-304.
—
2009    *When Experiments Travel : Clinical Trials and the Global Search for Human Subjects*. Princeton: Princeton University Press.
—
2010    The Politics of Experimentality. *In* In the Name of Humanity: The Government of Threat and Care. I. Feldman and M. Ticktin, eds. Pp. 256-289. Durham & London: Duke University Press.

Pfeiffer, James
2013    The Struggle for a Public Sector: PEPFAR in Mozambique. *In* When People Come First: Critical Studies in Global Health. J. Biehl and A. Petryna, eds. Pp. 166-181. Princeton and Oxford: Princeton University Press.
—
2014    How the aid and development industry helped cause Africa's Ebola outbreak. Humanosphere.        G.        Spitzer        and        T.        Polson.

http://www.humanosphere.org/podcasts/2014/09/how-the-aid-and-development-industry-helped-cause-africas-ebola-outbreak/93, August 10, 2019.

Philips, Susan U.

    2016    Balancing the Scales of Justice in Tonga. *In* Scale: Discourse and Dimensions of Social Life. E.S. Carr and M. Lempert, eds: University of California Press.

Rabinow, Paul

    2008    *Marking time: On the anthropology of the contemporary*. Princeton and Oxford: Princeton University Press.

Rankin, William J

    2009    Infrastructure and the international governance of economic development, 1950–1965. *Internationalization of infrastructures* 61.

Ravallion, Martin

    2009    Should the randomistas rule? *The Economists' Voice* 6(2).

———

    2013    *The Idea of Antipoverty Policy*. National Bureau of Economic Research.

Rayzberg, Margarita S

    2019    Fairness in the field: The ethics of resource allocation in randomized controlled field experiments. *Science, Technology, & Human Values* 44(3):371-398.

Redfield, Peter

    2005    Doctors, Borders, and Life in Crisis. *Cultural Anthropology* 20(3):328-361.

———

    2012    Bioexpectations: Life Technologies as Humanitarian Goods. *Public Culture* 24(1 66):157-184.

———

    2013    *Life in Crisis: The Ethical Journey of Doctors Without Borders*. Berkeley: University of California Press.

———

    2016    Fluid technologies: The Bush Pump, the LifeStraw® and microworlds of humanitarian design. *Social studies of science* 46(2):159-183.

Rees, Tobias

    2014a  Humanity/Plan; Or, On the "Stateless" Today (Also Being an Anthropology of Global Health) *Cultural Anthropology* 29(3):457-478.

———

    2014b  What is Global Knowledge. *In* American Anthropological Association Annual Meeting. Washington D.C.

Reid, Julian

    2010    The Biopoliticization of Humanitarianism: From Saving Bare Life to Securing the Biohuman in Post-Interventionary Societies. *Journal of Intervention and Statebuilding* 4(4):391-411.

Rigillo, Nicole

2014    Doing Well by Doing Good: Logics of Corporate Social Responsibility in Bangalore, India. Department of Anthropology, McGill University. PhD.

Rist, Gilbert
2008    *The History of Development: From Western Origins to Global Faith*. Londong & New York: Zed Books.

Rittel, Horst WJ, and Melvin M Webber
1973    Dilemmas in a general theory of planning. *Policy sciences* 4(2):155-169.

Robbins, Joel
2013    Beyond the suffering subject: toward an anthropology of the good. *Journal of the Royal Anthropological Institute* 19(3):447-462.

Rock, Melanie
2000    Discounted lives? Weighing disability when measuring health and ruling on "compassionate" murder. *Social science & medicine* 51(3):407-417.

Rodgers, Daniel T.
2011a   The Rediscovery of the Market. *In* Age of Fracture. Pp. 41-76: Harvard University Press.
—
2011b   *Age of Fracture*. Cambridge: The Belknap Press of Harvard University Press.

Rottenburg, Richard
2009a   *Far-fetched facts: A parable of development aid*: MIT Press.
—
2009b   Social and public experiments and new figurations of science and politics in postcolonial Africa. *Postcolonial Studies* 12(4):423-440.

Rottenburg, Richard, et al.
2015    *The world of indicators: The making of governmental knowledge through quantification*: Cambridge University Press.

Roy, Ananya
2010    *Poverty Capital : Microfinance and the Making of Development*. New York: Routledge.
—
2012    Subjects of risk: Technologies of gender in the making of millennial modernity. *Public Culture* 24(1 (66)):131-155.

Roy, Ananya, et al.
2016    *Encountering poverty: Thinking and acting in an unequal world*. Volume 2: University of California Press.

Savedoff, William D, Ruth Levine, and Nancy Birdsall
2006    When Will We Ever Learn? Improving Lives Through Impact Evaluation. Report of the Evaluation Gap Working Group.

Scharfstein, Ben-Ami

      1989    *The Dilemma of Context*: NYU Press.

Seaver, Nick
      2015    The nice thing about context is that everyone has it. *Media, Culture & Society* 37(7):1101-1109.

Seers, Dudley
      1969    The Meaning of Development. *Institute for Development Studies Communication* 44.

Sen, Amartya
      1999    *Development as Freedom*. New York: Anchor Books.

Sharma, Aradhana, and Akhil Gupta
      2006    *The Anthropology of the State: a Reader*: Blackwell Publishing.

Simone, AbdouMaliq
      2004    People as Infrastructure: Intersecting Fragments in Johannesburg. *Public culture* 16(3):407-429.

Singh, Bhrigupati
      2015    *Poverty and the Quest for Life: Spiritual and Material Striving in Rural India*: University of Chicago Press.

Sneath, David, Martin Holbraad, and Morten Axel Pedersen
      2009    Technologies of the imagination: An introduction. *Ethnos* 74(1):5-30.

Solesbury, William
      2001    Evidence based policy: Whence it came and where it's going: ESRC UK Centre for Evidence Based Policy and Practice London.

Staples, Amy LS
      2006    *The birth of development: how the World Bank, Food and Agriculture Organization, and World Health Organization changed the world, 1945-1965*. Volume 1: Kent State University Press.

Star, Susan Leigh
      1999    The Ethnography of Infrastructure. *American Behavioral Scientist* 43(3):377-391.

Stewart, Kathleen
      2008    Weak Theory in an Unfinished World. *Journal of Folklore Research* 45(1):71-82.

Storeng, Katerini T, and Dominique P Béhague
      2014    "Playing the Numbers Game": Evidence-Based Advocacy and the Technocratic Narrowing of the Safe Motherhood Initiative. *Medical Anthropology Quarterly* 28(2):260-279.

Strathern, Marilyn

1987    Out of Context: The Persuasive Fictions of Anthropology *Current anthropology* 28(3):251-270.

——

1995a  The nice thing about culture is that everyone has it. *In* Shifting contexts. M. Strathern, ed. Pp. 153-176. London Routledge.

——

1995b *The Relation: Issues in Complexity and Scale*. Cambridge, UK: Prickly Pear Press.

——

2002    Abstraction and Decontextualization: An Anthropological Comment. *In* Virtual Society? Technology, Cyberbole, Reality. S. Woolgar, ed: Oxford University Press.

——

2004    *Partial Connections*: Rowman Altamira.

Street, Alice
2014a Rethinking Infrastructures for Global Health. Our Changing World. https://www.youtube.com/watch?v=HehA7XmUzHM, August 10, 2019

——

2014b Rethinking Infrastructures for Global Health: A View from West Africa and Papua New Guinea. Somatosphere: Ebola fieldnotes. http://somatosphere.net/2014/rethinking-infrastructures.html/104, August 10, 2019

Taussig, Michael T
1999    *Defacement: Public Secrecy and th Labor of the Negative*. United States of America: Stanford University Press.

Taylor, Janelle S
2018    What the word 'partnership'conjoins, and what it does. *Medical Anthropology Theory* 5:1-6.

Teele, Dawn Langan
2014    *Field Experiments and Their Critics: Essays on the Uses and Abuses of Experimentation in the Social Sciences*. New Haven: Yale University Press.

Terry, Fiona
2002    *Condemned to repeat?: The paradox of humanitarian action*. Ithaca and London: Cornell University Press.

Ticktin, Miriam
2011    *Casualties of Care: immigration and the Politics of Humanitarianism in France*: Univ of California Press.

Torpey, John C.
2000    *The Invention of the Passport: Surveillance, Citizenship, and the State*. Cambridge UK: Cambridge University Press.

Toye, J. F. J.
1987    *Dilemmas of Development : Reflections on the Counter-Revolution in Development theory and Policy*. Oxford, Oxfordshire , UK ; New York, NY, USA: Blackwell.

Trouillot, Michel-Rolph
    2001    The Anthropology of the State in the Age of gGobalization: Close Encounters of the Deceptive Kind. *Current anthropology* 42(1):125-138.

Tsing, Anna Lowenhaupt
    2005    *Friction: An Ethnography of Global Connection*. Princeton and Oxford: Princeton University Press.

—
    2012    On Nonscalability: The Living World Is Not Amenable to Precision-Nested Scales. *Common Knowledge* 18(3):505-524.

—
    2015    *The Mushroom at the End of the World: On the Possibility of Life in Capitalist Ruins*. Princeton and Oxford: Princeton University Press.

UN Women
    2018    *UN Women Annual Report 2017-2018*. United Nations, Oisika Chakrabarti.

United Nations
    2015    *The World's Women 2015. Trends and Statistics*. United Nations.

Van de Walle, Dominique, Kimberly Nead, and World Bank.
    1995    *Public spending and the poor : theory and evidence*. Baltimore: published for The World Bank by the Johns Hopkins University Press.

Venkatesan, Soumhya, and Thomas Yarrow, eds.
    2014    Differentiating Development: Beyond an Anthropology of Critique: Berghahn Books.

Vivalt, Eva, and Aidan Coville
    2017    How Do Policymakers Update? *Unpublished manuscript, Berkeley, CA: University of California, Berkeley*.

von Schnitzler, Antina
    2013    Traveling technologies: Infrastructure, ethical regimes, and the materiality of politics in South Africa. *Cultural Anthropology* 28(4):670-693.

Webber, Sophie
    2015    Randomising Development: Geography, Economics and the Search for Scientific Rigour. *Tijdschrift voor Economische en Sociale Geografie* 106(1):36-52.

Webber, Sophie, and Carolyn Prouse
    2018    The new gold standard: The rise of randomized control trials and experimental development. *Economic Geography* 94(2):166-187.

Whyte, Susan Reynolds, et al.
    2013    Therapeutic Clientship: Belonging in Uganda's Projectified Landscape of AIDS Care. *In* When People Come First. J. Biehl and A. Petryna, eds. Pp. 140-165. Princeton and Oxford: Princeton University Press.

Wooldridge, Jeffrey M.

2006    *Introductory econometrics : a modern approach*. Mason, OH: Thomson/South-Western.

World Bank
2015    *World Development Report: Mind, Society, and Behavior*. Washington, DC: World Bank.

World Health Organization
2015    *Kenya: Sanitation, drinking-water and hygiene status overview*. UN-Water Global Analysis and Assessment of Sanitation and Drinking-Water (GLAAS).

Wrong, Michela
2009    *It's Our Turn to Eat: The Story of a Kenyan Whistle-Blower*: Harper Perennial.

Yarrow, Thomas
2006    Friction: An Ethnography of Global Connection *PoLAR: Political and Legal Anthropology Review* 29(2):291–296.