2020

Connecting grassroots development evaluation to achieving the United Nations sustainable development goals

Maya Sterling
Vassar College

Connecting Grassroots Development Evaluation to Achieving
the United Nations Sustainable Development Goals

A Thesis submitted in partial satisfaction of the requirements for the degree
Bachelor of Arts in International Studies

Maya Sterling

Under the Care and Guidance of:
Katherine Hite and Light Carruyo, *Advisors*

May 1, 2020

Table of Contents

Acknowledgements

This thesis would not exist without some very important people. I would like to thank the many Vassar College International Studies professors whose courses have challenged and inspired me. Specifically, I would like to thank Professors Joe Nevins and Ismail Rashid, who first introduced me to the Sustainable Development Goals. I would like to thank my two advisors, Katherine Hite and Light Carruyo, for providing continuous support, feedback, and guidance as my ideas shifted about how to frame this topic during the last academic year. Thank you to the Inter-American Foundation (IAF), specifically Raquel Gomes, whose words of wisdom I took to heart when writing this thesis. A special thank you to Kevin Healy for providing historical background information about the IAF. Thank you to Tim Koechlin for his flexibility and positivity. I owe a huge thank you to my senior year housemates Shannon Russo, Emma Wiley, and Rachel Strout for their unwavering friendship and making our time at Vassar special. To Emma, the library, first floor, right side, second table on the left, always. Thank you to Gretchen Lieb. Thank you to the Human Development Conference organizers from the University of Notre Dame for graciously giving me the opportunity to present my research in February of 2020. Finally, thank you to June Beittel, Eric Sterling, and David Olivar for being my rocks.

Prologue

This thesis is the culmination of my undergraduate academic journey and professional opportunities over the past four years. As an International Studies major intrigued by sustainable development, I have repeatedly been drawn to trends in global development and power imbalances between the Global North and Global South. This thesis is informed by my experience working for Amigos de las Americas in 2015 and 2017 and for the Inter-American Foundation in 2019. I would like to acknowledge that my experiences often influence my sympathetic critique of the Inter-American Foundation and grassroots development generally. This thesis is also informed by my positionality as a white, US-raised, college-educated woman. Despite the limitations my sympathies, positions, and experiences embody, it remains my belief that supporting grassroots development evaluation as an alternative approach to traditional methods can help improve the development industry. This thesis serves as an important capstone of my bachelors-level study and a catalyst into my career as a development professional.

Chapter One: Introduction

International development evaluation is crucial. Carefully monitored results, clearly identified impacts, and meticulously evaluated learning outcomes matter. Why? Because billions of taxpayer dollars fund these projects. Because politically motivated development programs are a product of government policy, both in donor and recipient countries. Because without international development evaluation, there is no clear way to gauge the extent to which programs succeed. Because all 193 member countries of the United Nations (UN) have adopted the 2030 Agenda for Sustainable Development, committing to achieve the Sustainable Development Goals (SDGs) in 10 years time. Evaluation matters because it is being done inadequately.

Tension within international development evaluation over modernity, expertise, progress, and technology begs the question of whether quantitative data are necessary to achieve and measure sustainable development. Some practitioners argue that quantitative data should be treated with skepticism because it perpetuates a computational way of thinking about development progress. Whether quantitative data perpetuates the comparative, hierarchical, and colonial legacy of development or if it promotes grassroots development is a key debate among post-development thinkers. These questions must be answered before creating an inclusive development evaluation paradigm that empowers grassroots organizations and works to achieve the SDGs.

This thesis explores the challenges faced by the global community to create such a development evaluation paradigm. This introduction discusses how the development paradigm has shifted over the past 70 years due to pressure from post-development thinkers and promoters of grassroots development. Chapter two explores the use of data by development professionals in

evaluation procedures, dives into the complexities of measuring program impact, and explains how development organizations use data evaluation as a source of internal organizational learning. Chapter three focuses on the evaluation of the SDGs, the frameworks used to measure their implementation, and the limitations within this process. This chapter advocates for a method of data gathering and evaluation that affirms grassroots development efforts. Chapter four explores the role of national and local governments in evaluation; the chapter concludes by reaffirming the challenges and recommendations made in chapter three. Chapter five serves as a case study of the grassroots development efforts of the Inter-American Foundation (IAF). Finally, the conclusion weaves together the best practices presented throughout this thesis and calls for an improved industry-wide effort to evaluate progress towards the SDGs.

Section One: A Framework for Discussing Development

International development, as understood today, began 70 years ago. During President Harry Truman's First Inaugural Address in 1949, he laid out a vision of the United States acting to "solve the problems of the 'underdeveloped areas' of the globe" (Escobar 1995, 3). In the 70 years since that declaration, theorists have coined terms such as the "Global North" and the "Global South" to distinguish between countries with power, influence, and expert knowledge.[1]

The traditional development paradigm is characterized by "structural shifts in national economies from the predominantly rural-agricultural to the urban-industrial" and movement away "from self-employment to wage-earning employment" and from "local markets to national markets" (Santos 1997, 12-13). Previously, leaders from the Global North implemented

---

[1] There are many divisive terms. "First World" and "Third World." The West and the rest. Developed, developing, or underdeveloped. Colonizers and colonized. High-income and low-income. The terms "Global North" and "Global South" are currently used in the literature to distinguish wealthier, more "modern," and more internationally influential countries (Global North) from the historically "undeveloped" countries, within which international development efforts are focused (Global South). References to "top-down" approaches refer to the expert knowledge and policy direction recommended and implemented by experts in the Global North.

development initiatives within the Global South and dictated how development occurred in these countries. While there is some recognition that this system is problematic, there has been no widespread reorganization of development initiatives to focus on grassroots organizations.

During the 20th century, development professionals relied on their preconceived notions of Eurocentrism to categorize countries into "developed" and "developing" based on the dependency theory. Dependency theory stated that "attempts by poor nations to be self-reliant and independent in their development efforts were rendered difficult, and almost impossible, given the coexistence of rich and poor nations in an international system dominated by unequal power relationships between the center (developed countries) and the periphery (developing countries)" (Santos 1997, 18). Development experts in the Global North implemented top-down development processes to combat this state of dependency. These top-down approaches relied on the assumption that benefits would trickle down to local areas. Angelo Santos states that this "development from above," constituted the creation of an "administrative state" whose goal was ultimately to "modernize but not to uproot the existing structures and processes, thereby keeping the poor, monoproducer developing societies locked into a cycle of dependence" upon the knowledge and methods of the Global North (internal quotations removed) (Santos 1997, 20, 23). This ideology cemented a power dynamic resembling colonialism that continued to negatively affect the Global South into the 21st century. Top-down development typically required large amounts of external capital and large technological inputs, defining success through a Western idea of modernity. Lending requirements by global financial institutions and multilateral donors forced Global South countries to comply to receive funding.

Although this ideology increased living conditions for millions around the world, programs and policies have often "left local people displaced, disempowered and destitute"

(Oliver-Smith 2010, 2). Development efforts failed to reach the "last mile," meaning the poorest, least accessible, and most in-need population; "most programs simply do not see them" (Ajmera 2016, 80).[2] Studying the SDGs is just one way to explore the complexities of development evaluation.

Section Two: The MDGs to the SDGs

In September 2000, the UN General Assembly adopted the precursor to the SDGs, the Millenium Development Goals (MDGs). The creation of these seven goals was climactic because it marked the first instance of global partnership focused on global development. The UN mandated specific development objectives which controlled development efforts in the Global South over the next 15 years. The MDGs included aspirations such as eradicating extreme poverty, achieving universal primary education, improving maternal and child health, and ensuring environmental sustainability. A systematic flaw of the MDGs was that they functioned as mandates for the Global South, rather than objectives for all countries. Some scholars from the Global South compare the imposition of the MDGs to the colonialism of the 19th and 20th century. These critics condemn the MDGs' furtherance of the paternalistic historical practices of the Global North towards countries in the Global South. Before the 15-year mandate of the MDGs was complete, the UN began to prepare for the development of a second, better iteration of these global objectives.

The Sustainable Development Goals (SDGs), adopted in 2015, used a three-pronged holistic approach to improving lives through "economic development, environmental sustainability, and social inclusion" (Sachs 2012, 2206). These 17 goals function as "an urgent call for action by all countries - developed and developing - in a global partnership" to achieve

---

[2] Debates about development theory and practice are numerous; not all of these debates could be included here. Please refer to the Appendix for a select list of recommended resources.

goals such as ending poverty, achieving gender equality, and combating climate change (United Nations 2020a). The SDGs represent a high level of global partnership that did not exist in the development industry until their adoption.

One such area in which partnership emerged was in "data, monitoring, and accountability" (United Nations 2020b). The methodology promoted by the United Nations for tracking the progress towards these goals recommends "scientific methods" and "support for data and statistics" for "developing countries" (United Nations 2020b). This UN recommendation prioritises data collected by national governments and supranational organizations like the World Bank and OECD. These entities operate at a large scale, often failing to take grassroots development efforts into account. The resulting evaluation of progress towards sustainable development inevitably prioritizes quantitative data and analysis by supranational organizations while ignoring grassroots evaluation, despite grassroots development having existed for 50 years. Section three introduces grassroots development.

Section Three: A Call for Grassroots Development

Grassroots development has emerged as an alternative to traditional development. Sheldon Annis and Peter Hakim define grassroots development as an "approach to economic and social development in which the best way to help those in need is to finance the organizations that they themselves create and control" (Annis and Hakim 1988). Santos notes that grassroots organizations allow "participants to empower themselves and better their condition" and carry on autonomous efforts even when imported programs have ended (Santos 1997, 42). Rather than being designed in distant boardrooms of multinational institutions, grassroots development relies on local leaders to design and implement programs. Local organizations are empowered to grow,

evolve, and deepen their work, even after external funding expires. Their autonomy is central to their financial, ethical, and political appeal.

One organization that highlights the success of grassroots development is Capiibary. The Paraguayan cooperative combats outmigration fueled by agro-industrial soy projects that destroy forests and water resources. Researchers found that Capiibary was able to be a "learning organization with the capacity to generate and share knowledge in the pursuit of long-term sustainability" (León et al 2017, 70). Capiibary's work developing and managing a sustainable small-scale farming strategy "challenges practitioners and academics to question their own often romanticized notions of sustainable development among smallholders" by proving that grassroots organizations can be successful (León et al 2017, 72). Evidence from cooperatives like Capiibary help challenge traditional development thinking.

Many grassroots development efforts stem from communities where citizens are mobilized to collective action. According to Albert O. Hirschman, "collective action is provoked by some common, usually adverse, experience to which a group of people is subjected" (Hirschman 1984, 27). This communal struggle motivates individuals to take action they feel is best suited to solve the problem. It is important to study the capacity of organizations not only to enact their program, but also to measure their program's impact, in whatever way they define it. The learning conducted at the grassroots level is typically reserved for organizations and funders. Limited learning prevents grassroots organizations from sharing their evaluation with the larger development industry. By more widely sharing knowledge, grassroots organizations can inspire supranational organizations and governments to support more grassroots development programs.

Section Four: Underlying Theory — A Post-Development Framework

Grassroots development stems from the post-development understanding of social change that is "rooted in the agency of local communities who identify their own problems and … move away from development as a Western-dominated ideology and a form of cultural imperialism" (Escobar and Harcourt 2018, 6). Post-development pushes back against the many practices of traditional development. Aram Ziai summarizes post-development criticisms of traditional development: (1) it is a western ideology that attempts to prevent the expansion of communism, (2) it wrongly universalizes a "developed" way of life on a global scale, (3) it is Eurocentric and assumes non-Western, non-modern, and non-industrial ways of life are inferior, (4) it focuses on wealth accumulation and perpetuates capitalist logic, (5) and finally, it legitimizes external interventions into the lives of the "less developed" people for the sake of the common good as defined by "experts" (Ziai 2017, 2547-2548). Ziai reflects that the post-development movement is about reclaiming the commons and seeking autonomy. Ziai further argues that one of post-development's most important contributions is breaking "the consensus about 'development' being necessary, self-evident, positive and unquestionable, and thus pave the way for more nuanced analyses" (Ziai 2017, 2550).

Gustavo Esteva, the father of post-development, has shifted his view on the role statistics can play to evaluate development. He posits, it is "right … [to] open a decent door" to the people who are genuinely "concerned with the real problems of the world [and] interested in making a difference" (Esteva and Escobar 2017, 2562). He argues that "people studying development can accompany and support … the many paths people are following around the world beyond development" (Esteva and Escobar 2017, 2562). This expanded role in the post-development framework for quantitative data, research, and statistics is an important shift.

Arturo Escobar joins Gustavo Esteva's call for shifting the power dynamic between development providers and recipients. Escobar and Esteva propose rejecting cooperation as development aid. They feel "the standard form of cooperation, practiced by institutions such as USAID, the World Bank and mainstream NGOs" enforces colonialism and dispossession (Esteva and Escobar 2017, 2564). Escobar and Esteva first advocate for cooperation *as,* or *for, social justice*. Cooperation for social justice embraces human rights, environmental justice, and direct support for grassroots groups. However, Escobar primarily advocates for cooperation *for autonomy*. Escobar describes cooperation for autonomy as "embrac[ing] all sides in the same ... movement for civilizational transitions and inter-autonomy" by the meshing of communities in both the Global North and Global South (Esteva and Escobar 2017, 2564). Escobar acknowledges that there are "no ready-available models for this third kind of solidarity cooperation" (Esteva and Escobar 2017, 2564). One potential way to facilitate the creation of this kind of coalition-building is to bring more attention to grassroots evaluation. By learning from community successes at the local level, the development industry will be more capable of achieving the cooperation for autonomy proposed by the fathers of post-development.

Chapter Two: Using Data for Impact Evaluation and Learning

Two types of data, quantitative and qualitative, exist for the purposes of development. Qualitative data have been used as the main source of knowledge of development professionals for decades. However, development evaluation increasingly incorporates quantitative data. In 2016, the UN Global Pulse Initiative, which focuses on using data and monitoring to improve development programs, commissioned an independent consultant, Michael Bamberger, to report on the role that big data[3] might play in facilitating development evaluation in the future. While this report focuses on the impact of big data, it reveals some important insights about generating data and creating knowledge. The conclusion of the report finds that there is a "need for bridge building between data scientists and evaluators to allow for the development of a common language and to identify promising areas where big data analytics can be applied in development evaluation contexts" (United Nations Global Pulse 2016, 22). The report notes that before big data can be fully embraced, "many agencies are still in the early stages of understanding big data and its potential applications in development" (United Nations Global Pulse 2016, 38). While there are uncertainties about how big data will be used in the future, it is believed that it will support and enhance the work already done with data. Data in evaluation can be used for monitoring, evaluation and learning.

Monitoring is collecting data about what occurs in projects. How many beneficiaries attended the training? How much money was distributed as microloans to women entrepreneurs? What was the crop yield on the disease-resistant coffee trees? Answers to these questions constitute monitoring. Evaluation, however, takes the knowledge from data to another level.

---

[3] Big data is understood as part of the "digital revolution," a term coined in 2013. This revolution consists of "an explosion in the volume of data, the speed with which data are produced, the number of producers of data, the dissemination of data, and the range of things on which there is data, coming from new technologies such as mobile phones" (United Nations Global Pulse 2016, 30).

According to a report by USAID, evaluation goes a step beyond monitoring to ask questions

such as "why and how a project or activity is achieving, or not achieving, its objectives" (U.S.

Agency for International Development 2016, 8). Ultimately, evaluations are different from

monitoring because they make a "judgment about project performance" (U.S. Agency for

International Development 2016, 8). Furthermore, there are two primary types of evaluations:

impact evaluations and performance evaluations. Performance evaluations measure "how" a

program is being implemented, perceived, and valued. However, evaluation in the context of this

thesis refers to impact evaluations. Impact evaluations "measure the change in a development

outcome that is attributable to a defined intervention" (U.S. Agency for International

Development 2016, 9). An intervention only has an impact when it can be proven "how a

program or organization has changed the world around it" (Gugerty and Karlan 2018b, 42). In

the last 30 years, the industry gold standard has shifted to conducting impact evaluations with

rigorous counterfactuals through a process called randomized control trials (RCTs).

RCTs originated in the medical field, but in the 1980s were incorporated into

development evaluation. A RCT is successfully completed when there are both an intervention

group and a control group. Control groups are needed to determine if something changed as a

result of an intervention or as a result of an external factor. In development, proponents of RCTs

argue that this is needed for rigorous evaluation because in complex environments, many

potential variables could be causing changes. After observing an outcome without a RCT,

organizations might jump to the conclusion that their intervention has caused the observed

change. However, if there is no control group, it may not be clear what factor caused the shift in

the experimental group. Potentially, the result is due to better weather or an improved

macroeconomic landscape and not the development intervention at all. Proponents of RCTs argue that without the counterfactual, it is impossible to "prove" if a program has created impact.

RCTs also have their critics. Evaluators who center rights-based approaches argue that there is a "need to listen to multiple voices" in evaluation. They assert, "that there is no one way to identify or assess programme outcomes" and worry that RCTs impose a one-size fits all method (United Nations Global Pulse 2016, 54). These critics point out that in many instances, capturing quantitative data are impossible or inefficient to understand a development intervention. RCTs come from evaluating medical interventions where it is easier, compared to poverty interventions, to determine the results of a program. An example of a medical intervention would be an HIV vaccine or distributing bed nets to families. Here, it is easy to track who received the intervention and who did not. But poverty alleviation interventions are not as tangible. It is much harder to determine the extent to which farmers have adopted the lessons from a workshop focused on planting and harvesting drought-resistant crops. A bed net program seeks one outcome, less mosquito-borne illness. However, an economic empowerment and sustainable agriculture intervention could be striving to achieve 20 different economic, social, and political outcomes. Designing counterfactual scenarios for development interventions takes many financial and human resources. This is where RCTs fall short.

Some development professionals criticize impact evaluations citing that the efforts are unsuccessful and meaningless. According to Mary Kay Gugerty and Dean Karlan, measuring impact is an "out of balance trend" that has led to a "proliferation of poor methods … resulting in organizations wasting huge amounts of money on bad 'impact evaluations'" (Gugerty and Karlan 2018a, viii). Gugerty and Karlan question why conduct impact evaluations at all if they are extremely expensive and turn out badly. Gugerty and Karlan may criticize the current methods

for evaluation, but they support why organizations evaluate: organizational learning and improvement.

Organizations engage in evaluation for two reasons: accountability and learning. Accountability is "transparently disclosing findings to all stakeholders" about the progress an organization has promised (U.S. Agency for International Development 2016, 8). Learning is "systematically generating and sharing knowledge about how best to achieve development outcomes" and "using that knowledge to inform decisions, adapt ongoing projects and improve the design of future projects" (U.S. Agency for International Development 2016, 8). This distinction in audience between accountability and learning is important. The audience for accountability is an organization's funders, board of directors, and individuals who download an organization's annual review off of their website. For learning, the audience tends to be the internal employees of the organization. Unfortunately, many audiences are grossly excluded from receiving evaluation information. This includes the people who are being monitored (the beneficiaries of programs) and the larger overall development industry (the UN, supranational organizations, and national governments). In most impact evaluation, the learning gained from evaluation remains confined to a limited group of people.

How organizations learn is an important component of evaluations. On the base level of an organization's work are the target beneficiaries. Typically, a funding organization will mandate the use of a framework of indicators to track the results of a project. After the organization has decided on a set of indicators, measured results are gathered from the project location. This information is passed up to the professionals employed by the funder. Organizations then proceed with the evaluation based on the data they have collected. Traditionally, the knowledge from the evaluation is reserved for those internal to the funder

organization. If beneficiaries are not given access to the evaluation or the tools to understand it, they are unable to learn for themselves. This further inhibits learning for grassroots leaders.

Grassroots organizations are additionally disadvantaged by the delayed evaluation and learning timeline of development. When organizations do evaluations, they draw conclusions across their portfolio of projects. This allows an organization to observe trends with the hope of improving their programs in a future portfolio, not the current one. This after-the-fact learning is mainly due to the limitations of funding cycles. Funding cycles are very short in development, typically financing one- or two-year projects. The results of short projects might not be evident until many years after the official expiration. Therefore, evaluation and learning is completed after the opportunity to make changes to current projects has passed. Development impacts are felt by those at the grassroots level before the organizations that manage the programs. This leaves beneficiaries caught in a potentially adverse program with no opportunity to change the funder's behavior.

Data clearly contributes to overall development evaluation and learning. Conducting impact evaluation for learning is a vital part of the development process, but one that falls short for grassroots organizations. Even when organizations continue efforts to evaluate, learn, and improve *their* programs, the promise of achieving the SDGs remains elusive. The following chapter will explore how evaluation and learning efforts relate to the SDGs.

Chapter Three: How To Measure Progress Towards The SDGs

Section One: The Foundation of the SDGs

On January 1, 2016, the 17 Sustainable Development Goals (SDGs) came into effect. These goals established the development objectives that all countries would work to achieve in 15 years time. The update to the SDGs ushered in a new focus of using quantitative data, statistics, technology, and evaluation to track global development.

Some authors, such as Wolfgang Sachs, hoped that the SDGs would correct the errors of the preceding MDGs. Sachs calls the MDGs "visionary" but "non-binding" and lacking a "sanction mechanism" (Sachs 2017, 2574). Sachs worried that the SDGs could solidify the old practice of foraging a "disconnect between international rhetoric," which continued to support land grabbing, coal mining and foreign financing, and "national measures" (Sachs 2017, 2574). Sachs praised the SDGs for ending the aspiration of sky-high growth that was traditionally promised by development. The SDGs are "more or less narrowed down to requirements for survival" instead of seeing development as a linear path that Global South countries follow to achieve the modernity of the Global North (Sachs 2017, 2575). The SDGs' focus is improving livelihoods in non-prescriptive ways by giving all countries the freedom to decide what their own development might look like while still striving towards shared goals. Fundamentally, the SDGs allow the industry to engage more fully in grassroots development.

Using the SDGs' social indicators to measure development progress "replaced GDP in determining the performance of a country" (Sachs 2017, 2578). Supporters of using data to measure social indicators believe that "one can only improve what one has previously measured" (Sachs 2017, 2578). Sachs is not an advocate for using data to measure progress. He fears that "the simplification and complex reduction that numbers and quantification offer" leads to a

"homogenizing" effect (Sachs 2017, 2578). Sachs sees this as boiling down global diversity into a scale of numbers. Sachs argues that relying on "quantitative data" furthers a "dictatorship of comparison … that constructs deficits along the time axis between … nations" (Sachs 2017, 2578). Sachs' concern is echoed by Thomas Carroll. Carroll writes that Latin American development practitioners and academics "mistrust" quantitative methods "partly because of the feeling that they fail to take into account unique aspects of individual cases" (Waters 2004, 57).

Some authors embrace quantitative data. Jeffery Sachs asserts that technology will be able to collect data more quickly from rural areas enabling Global South governments to enact data-informed policy. Sachs specifically notes that "mobile phone, wireless broadband, and remote sensing" have an important future role to play in quantitative data collection (Sachs 2012, 2210). Jeffery Sachs, unlike Wolfgang Sachs, advocates for quantitative data as a global priority to achieve sustainable development (Sachs 2012, 2206).

Post-development thinkers take the side of Wolfgang Sachs and Thomas Carroll over Jeffery Sachs on the debate of technology and quantitative data. According to Esteva and Escobar, improved development is not achieved through quantitative data collection, but by completely rejecting the ideology underlying development. Unfortunately, for post-development theorists, quantitative data have taken center stage in evaluating the SDGs.

Section Two: Using Data to Evaluate the SDGs

Many organizations evaluate progress towards the SDGs using quantitative methods. J-PAL, an evaluation research organization founded at MIT in 2003, argues that data should be at the center of development evaluation. J-PAL researches the benefits of "survey instruments, data collection, statistical analysis, and data publishing" ("J-PAL," n.d.). They conclude that quantitative methods like these help to translate "research into action" and promote "a culture of

evidence-informed policymaking around the world" ("J-PAL," n.d.). J-PAL's hope is to increase

the global standard of government policy informed by evidence and data. J-PAL's advocacy for

quantitative data reflects Jeffery Sachs' belief that it produces improved development evaluation.

The SDGs and supranational organizations embrace J-PAL and Sachs' method. The embrace of

quantitative data has established an industry practice of top-down data collection and learning.

Evidence of this effect is seen in data projects like SDG-Tracker.org.

SDG-Tracker.org is a data visualization tool focused on the SDGs run by researchers at

the University of Oxford and the Global Change Data Lab. The Global Change Data Lab is home

to Our World in Data, a resource that compiles data from a variety of "official, high-quality

sources" including the World Bank, World Health Organization, and various UN branches

(Ritchie et al 2018). The creators describe the website as "an interactive hub where users can

explore and track progress … in such a way that everyone can engage" (Ritchie et al 2018).

SDG-Tracker.org presents data within the framework of the SDGs' targets and indicators. As an

example, SDG 1, to "end poverty in all its forms everywhere," is broken into 7 targets and 14

indicators by the UN. To evaluate progress towards an indicator, researchers at SDG-Tracker.org

identify data that can be used to measure it. Following the example of SDG 1, the first target

(abbreviated 1.1) is to "eradicate extreme poverty." This target includes one indicator of the

same name, "eradicate extreme poverty" (abbreviated 1.1.1) (Ritchie et al 2018). To evaluate

indicator 1.1.1, those living in extreme poverty, on less than the international poverty line of

$1.90 international dollars a day, are counted (Ritchie et al 2018). To track this indicator, the

researchers at SDG-Tracker.org pull data from the World Bank Development Research Group.

These data come from household surveys "obtained from government statistical agencies and

World Bank country departments" (Ritchie et al 2018). Researchers turn to the World Bank

because it, along with other supranational organizations, hold a monopoly on available data. This

limits what information is available to conduct evalutions of SDGs progress.

The researchers at SDG-Tracker.org are aware of the data inadequacy. They note that

"timeliness, frequency, quality, and comparability of household surveys need to increase

substantially" for a complete evaluation of indicator 1.1.1 (Ritchie et al 2018). While they use

World Bank and national government data, they know it is insufficient. An additional

impediment to tracking poverty reduction is disagreement on which data best measure SDG 1.

Household *income* is currently used, yet, some researchers believe that measuring household

living standards better predicts poverty because "income can vary over time even if living

standards do not" (Ritchie et al 2018). The researchers worry that the World Bank and national

government data they use fails to account for countries' "different definitions of poverty"

(Ritchie et al 2018). Data collection is further complicated by the fact that:

> "Most low-income countries do not have systematic data collection processes, and
> especially in marginalized communities populations are fluid … Accordingly, it is
> exceedingly difficult … to establish baselines, identify trends, and quantify impact with
> complete surety … The most marginalized often remain outside the reach of systemic
> interventions and the data that tracks their impact" (Ajmera 2016, 80).

The inability of top-down evaluation and data collection methods to access information

from the most marginalized populations is further evidence that data procedures need to improve.

The monopoly of available data and the unresolved question about data sources

illuminates concerns regarding who controls data. SDG-Tracker.org exposes supranational

organizations like the World Bank as development knowledge tyrants. These organizations

centralize the production of knowledge. This would not be as problematic if their data did not

originate from their own sources and methods or if they duly accounted for poverty in the hardest

to reach places. SDG evaluators should be concerned that global knowledge about development progress omits local realities.

Additional concerns arise regarding the evaluation for SDG 1. According to SDG-Tracker.org, many of the 14 indicators that make up SDG 1 lack target reduction guidelines which are needed to inform development policy. As an example, take indicator 1.A.3, "Inflows directly allocated to poverty reduction programmes" (Ritchie et al 2018). This indicator advises practitioners how to direct financing more effectively to target poverty reduction. The indicator reads:

> Ensure significant mobilization of resources from a variety of sources, including through enhanced development cooperation, in order to provide adequate and predictable means for developing countries, in particular least developed countries, to implement programmes and policies to end poverty in all its dimensions (Ritchie et al 2018).

The indicator provides no definition of what constitutes a "significant" mobilization. Nor does it indicate which funding sources should be included in a "variety." No benchmark for enhanced development cooperation is established. There is no definition for what comprises an "adequate" or "predictable" means to implement programs. This indicator leaves countless unanswered questions. The SDGs are intended to offer a clear roadmap to development actors about how they should design their programs to achieve sustainable development. The UN is right to not mandate a specific policy direction that threatens to infringe on state autonomy, but undermining the clarity of the SDG indicators with vague language is a failure of the UN's intention to provide global leadership.

One would hope that in the five years since this indicator was adopted, some organization or government would have decoding this language and improved their allocation of inflows. One would also hope that data to evaluate this indicator would exist. Unfortunately, the researchers at SDG-Tracker.org, using the highest quality sources, "are currently not aware of data for this

indicator" (Ritchie et al 2018). This is astounding given that thousands of organizations work every day to allocate more than "significant" amounts of funding towards poverty reduction. Small government agencies, international NGOs, grassroots groups, and agricultural cooperatives learn more every day about the success and failure of their poverty reduction programs. Yet, their data is not displayed on the SDG-Tracker.org website. One significant barrier to changing that reality is overcoming the supremacy afforded to data gathered and published by the World Bank and national governments.

SDG-Tracker.org reveals that traditional evaluation methods have fallen short. First, development knowledge production and official data collection is controlled by supranational organizations. This generates a disconnect between the information that is widely acknowledged as SDG progress and the nuanced contexts of poverty reduction programs in local communities. Second, the language of indicators meant to clarify how to engage funding for sustainable development is cryptic. Therefore, the development evaluation industry needs to adopt previously unacknowledged data sources and improve how SDG progress is conveyed. If the purpose of development evaluation is learning, SDG-Tracker.org clearly is insufficient. The next section examines whether evaluative frameworks are better for tracking progress towards the SDGs.

Section Three: Using Frameworks to Evaluate the SDGs

Researchers at the Swiss Sustainability Hub (SSH) have developed the GAPFRAME (GF) to evaluate SDG progress. The GF is a "normative framework" that serves the purpose of "translating" the SDGs "into nationally relevant issues and indicators" (GAPFRAME 2017b). Instead of comparing countries' progress to one another like SDG-Tracker.org, this method compares countries to themselves to analyze "where a country is today as compared to where it

should be in the future" (GAPFRAME 2017b). The GF recommends concrete, practical actions that each country should be able to understand and adopt into their national policy.

The GF research team brings together academics, non-profit organizations, government leaders, and company executives, dubbed "sustainability thought leaders," to devise a "pragmatic answer to the question of how we all can make a difference to the problems in this world" (GAPFRAME 2017c). The central component of the GF is the "target Safe Space," a development sweet spot that corresponds to the framework's 68-indicator design (GAPFRAME 2017a). The Safe Space is not directly related to the 17 goals, 169 targets, or 232 indicators established under the SDGs. Instead, the Safe Space exists within the GF's four sustainability dimensions — planet, society, economy, and governance — which incorporates 24 underlying issues, and is tracked by 68 indicators. In the GF, the 232 SDG indicators becomes a more manageable 68. The GF pitches its 24 issue areas as "relevant to all nations" (GAPFRAME 2017b). On each of the 24 issues, a country receives a score out of 10. A score of 10 represents the "ideal value," meaning a country is performing perfectly regarding this issue. This is practically impossible to achieve. Therefore, the GF creators design the Safe Space range of values from 7.5 to 8.8 on any issue as the target. A country is performing "'good enough'" on an issue if their score falls within the Safe Space range of values (GAPFRAME 2017b). If they score below 7.5, they need to improve their policy in that area.

The "economy" dimension and the "society" dimension are used to evaluate poverty alleviation. In the economy dimension, the top four performers are Sweden, Switzerland, France, and Austria. Again, unsurprisingly, the top for countries ranked in the society dimension are Norway, Sweden, Iceland, and Finland.
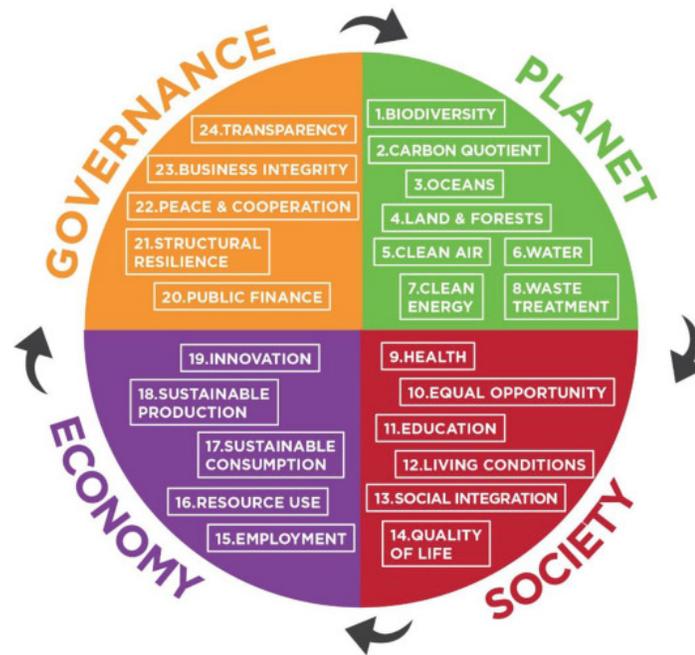
Figure 1 - The GAPFRAME 4 dimensions (planet, society, economy, and governance)
and their 24 underlying issues

The 17 SDGs are not explicitly included in the GF. To investigate the evaluation for a

single SDG, one has to identify which of the 24 issues most represents it. For SDG 1 that is the

"quality of life" issue in the society dimension (See Figure 1). The "quality of life" issue

quantifies how "nations provide the basis for sustainable human and economic development,

lowering poverty and helping people afford a better quality of life" (GAPFRAME 2017d).

Similar to SDG-Tracker.org, limited data exists to evaluate quality of life. The GF is intended to

provide data for each issue of each country, yet, only 36 out of 193 countries receive a Safe

Space score for quality of life. An evaluation of quality of life does not exist for most countries

because the underlying proxies, life satisfaction, quality of support network, work-life balance,

and poverty among the population, all lack data. The data for the first three of these proxies

comes from the Organisation for Economic Co-operation and Development (OECD) and the data

for poverty among the population comes from the World Bank.

The OECD comprises 37 member countries and, along with "key partners," represents 80% of the world trade and investment (OECD 2019).[4] There are only three Global South country members, Mexico, Chile, and Colombia. Countries are accepted as members if they adhere to the OECD "mission and values" and are ready to take on the "demanding task" and "responsibilities and requirements of active membership" (OECD 2019). This standard precludes all but 37 countries. While the OECD boasts that its members spread the globe from "North and South America to Europe and Asia-Pacific," there is a jarring omission of countries from the Global South (OECD 2019). Slovenia gained membership in 2010, the same year as Chile. Slovenia's GDP is just below $50 billion. Brazil, not even listed as a "candidate for accession," has a GDP of $2 trillion, four times the size of Slovenia (OECD 2019).

Brazil is, however, identified as an OECD "key partner." Key partners "participate in the OECD's daily work, bringing useful perspectives and increasing the relevance of policy debates" (OECD 2019). It is shocking that an organization whose mission is to use shared knowledge and global partnership to track global challenges relies on countries not suitable enough to be members to make their policy more relevant. These key partners who provide these "useful perspectives" and "increase the relevance" of policy are none other than five countries from the Global South: Brazil, China, India, Indonesia, and South Africa.

It is concerning that this high ranking and well-regarded institution is predominantly insulated by European groupthink. The OECD and World Bank are not simply membership organizations that collect and publish unbiased data. They are key supranational organizations whose efforts focus on knowledge generation and policy proposals. It is wrong that the same organizations that influence millions of dollars of development funding also generate some of the only globally available development data.

---

[4] Colombia became the 37th member on April 28, 2020

The gaps in the World Bank and OECD data prevent the GF from being effective. For a GF dimension to receive a score, 60% of data must exist. For the society dimension, only 143 countries reach the data threshold to receive a calculation. For an issue to be calculated, 50% of the data must exist. Only 36 countries receive a quality of life calculation because "some indicators have available data only for the OECD countries" (GAPFRAME 2017b). Of the 36 countries with a quality of life calculation, 35 are OECD countries. Vital development evaluation is conducted at times only from OECD data. This overwhelmingly prioritizes the successes of Global North countries while leaving out information about much of the Global South.

To test the usefulness of the GF, take the example of a country like Bolivia and run through a hypothetical scenario. Bolivia is hoping to enact evidence-informed development policy. According to the GF, the largest priority area for Bolivia is society under which the quality of life issue falls. For the 24 issues, quality of life happens to be one of three without sufficient data. Of the four proxies that make up what would be Bolivia's quality of life score, only poverty among the population has data (from the World Bank). The other "universal" proxies receive no score because the data comes from the OECD of which Bolivia is not a member. Therefore, the proxies life satisfaction, quality of support network, and work-life balance do not have data.

According to the GF, Bolivia should improve its society dimension following the recommendation of the GF's "nationally relevant issues and indicators." By default, one of Bolivia's top priorities should be achieving an improved work-life balance among its population. An educated guess would assume that grassroots organizations and the Bolivian government define a better quality of life using a different proxy that is not improving work-life balance. Ultimately, the GF is disappointing for Bolivia's hopes to develop evidence-informed

development policy. This is because very little data exists to evaluate its performance and the nationally relevant benchmarks do not apply to its local context.

The central barrier to achieving improved development evaluations is the lack of diverse data sources. SDG-Tracker.org and the GF reveal a glaring vacancy: data, monitoring, and evaluation done by organizations at the grassroots is absent from the global discussion of SDG progress. One proposal to improve SDG evaluation is to incorporate technological methods. The following section explores the complexities of incorporating technology into measuring SDG progress.

Section Four: The Technological Challenge of Evaluating the SDGs

Emerging technologies are opening up new opportunities for using data to evaluate progress towards the SDGs. According to Michael Bamberger, the SDGs bring a unique complexity to evaluators already trying to measure and evaluate the impact of development interventions. The SDGs encompass "multiple interventions in multiple contexts and with multiple stakeholders involved in large and ambitious programmes" (United Nations Global Pulse 2016, 33). Bamberger writes that achieving the SDGs "requires the use of new, and still evolving complexity–responsive evaluation methods" (United Nations Global Pulse 2016, 33). In 2015 when the goals were adopted, Bamberger reports that the evaluation techniques required to comprehend SDG progress did not fully exist.

Technology for data collection evolves extremely fast. Bamberger writes, "the volume of information is beyond the computing capacity of conventional computers, and the complexities introduced by the need to assess interactions among many components also requires new forms of data analytics" (United Nations Global Pulse 2016, 33). Conventional monitoring and evaluation methods may not be capable of keeping up. Even before the adoption of the SDGs,

development professionals "recognized the limited ability of current evaluation methodologies to gauge complex development programmes" (United Nations Global Pulse 34, 2016). Since current evaluation methods fall short, exploring the potential benefits of evaluation aided by technology needs to be a top priority.

The extent to which data collection and evaluation aided by technology will bring the needed improvement to SDG evaluation is up for debate. Some worry that the use of "new information technologies will be used extractively by governments, large development agencies and corporations" going forward (United Nations Global Pulse 2016, 31). If measures to empower grassroots groups are not implemented, the result would be "poor and vulnerable groups having less, rather than more, information and control over decisions and policies affecting their lives" (United Nations Global Pulse 2016, 31). New technology that continues to ignore data from grassroots organizations furthers the Eurocentrism of traditional development.

Bamberger's report calls for incorporating non-tech-enabled data collected by grassroots organizations *with* the mass-collected quantitative data collected by supranational organizations. Both resources together paint a more complete picture of the realities in disadvantaged areas. Both data sets are necessary because in many instances, data cannot be collected without human interaction. The human dimension to conduct case studies, in-depth interviews, and focus groups requires visiting communities. Only after a level of trust is reached, especially with marginalized members of a community, will those members feel confident to engage in a technological data collection method that can be introduced later (United Nations Global Pulse 2016, 49). A combination of analog surveying with tech-enabled tracking will ensure the best data can be collected and evaluated.

According to Bamberger, big data is not incompatible with development work. He is confident that big data can be combined with the information development agencies are already collecting through their conventional analysis (United Nations Global Pulse 2016, 34). Bamberger hopes that development practitioners will come to more fully embrace big data and quantitative methods. He argues that "big data is often combined with, or validated through mixed method techniques and other qualitative approaches" (United Nations Global Pulse 2016, 37). However, hiring the talent to carry out evaluations using big data is beyond the capacity of most grassroots organizations and funders. Big data analysis requires highly-specific expertise and expensive hardware and software.

If and when big data analytics are incorporated into development evaluation, there are inevitably going to be growing pains. According to Bamberger, "most data scientists operate within a very different research and management paradigm than do most [development] evaluators" (United Nations Global Pulse 2016, 43). Professionals with different expertise and who come from different industries will struggle to adapt their work to goals as large as the SDGs. Overcoming this challenge will require "bridge–building to create a space for development of a common understanding" (United Nations Global Pulse 2016, 43).

Another challenge of incorporating big data analytics into development is that most data analytics is currently carried out for business and commercial interests. If big data is used for development, data will be referring to vulnerable populations, not consumers. Conducting development research using big data will require "a broader focus and the need to dig deeper" than marketing and business evaluations call for (United Nations Global Pulse 2016, 46). The development industry and data analytics industry will need to combine their theory of change.

Achieving improved evaluation aided by big data will require careful management to bridge these divides.

A further step required to successfully shift evaluation practices includes finding better ways to publish evaluation results. Currently, high-level technical skills are needed to not only collect big data and complete the evaluation, but also to publish and interpret the results. This barrier of entry prevents beneficiaries, who tend to be education- and resource-poor, from accessing evaluative information. The educational and technical divide is most clear when evaluation findings are released and only certain stakeholders are prioritized (United Nations Global Pulse 2016, 56-57). Evaluators are not incentivised to produce reports for any additional stakeholders beyond donors and the organizations that conduct the study. According to Bamberger, "often many interested NGOs, civil society and community organizations never receive the findings" (United Nations Global Pulse 2016, 56-57). The difficult task of "translating" evaluation findings into a language that can be widely understood means that many beneficiaries never receive the reports or are not able to offer feedback. The unfortunate result is that "the views of large sectors of the population affected by different interventions are never received" (United Nations Global Pulse 2016, 57). If program beneficiaries and grantees cannot provide feedback, development programs and evaluations will remain flawed. To achieve comprehensive development evaluation with equitable global access, there needs to be improved efforts to make evaluation knowledge accessible to all.

Chapter Four: Shifting Views on Development Evaluation

This chapter explores which organizations take the lead in both enacting development programs as well as conducting evaluations. Grassroots organizations are increasingly funded to carry out development programs but remain absent from evaluation procedures. Luckily, the industry is on the brink of a major shift.

Section One: The Role of Governments in Development and Evaluation

International development projects are carried out on behalf of the United States by USAID. Many writers have criticized the agency, including Lars Schoultz. He believes that US government efforts have always been about "uplifting" people through development. Schoultz describes how USAID turned development funding from something that was ad-hoc into an institutionalized bureaucracy. Many of the millions of dollars spent by USAID to implement their programs flow to US-based contractors instead of local organizations.

Aid funding through governments and agencies like USAID is a common practice, and before the boom of NGOs in the 1990s, was the main mechanism to deliver development aid funding. Federal and local governments were previously thought of as ideal project implementers who ensured grants made their way to local actors.

USAID has faced criticism and pressure from writers, academics, and researchers like Schoultz for their failure to spend money locally. In 2010, they instituted a procurement reform after making a shocking discovery: of the money they spent on development programs worldwide, only 10% of their "benefiting country partners" were listed "in a top-line implementing role" (Dunning 2013, 1). An internal assessment discovered that 65% of "grants and contracts flowed to U.S.-based organizations," not partners abroad where one would expect

the money to go (Dunning 2013, 1). A reform was flagrantly needed; USAID was continuing to operate with an expired ideology.

The role of governments in development extends to data collection and evaluation. While governments collect a significant amount of national statistics, including those on development indicators, much is left to be desired. The concern is that national data collection projects, similar to large-scale development projects, are incapable of reaching those at the last mile. According to Maya Ajmera, people at the last mile are only properly accounted for by community-based groups (Ajmera 2016, 118).

Building an opposition to government involvement in data collection and monitoring is hard given the history of governments monopolizing the practice in the past. According to B.K. Pattanaik, "the level below the state neither have adequate technical manpower" to undertake the exhaustive exercise of data collection nor do they have the "necessary functions, functionaries and funds to carry out this work" (Pattanaik 2016, 77). Pattanaik's pessimism about the lacking capacity of organizations that exist at the "level below the state" should be taken with a grain of salt. If local organizations do not at the moment have the capacity to conduct their own evaluation, that does not mean they lack the capacity to learn it or develop their own method. Many organizations will need support, for example financial resources and training, but they fundamentally have the capacity to conduct rigorous evaluations.

Initiating capacity building efforts for grassroots groups is a large undertaking. While initiatives like this would traditionally fall on governments, writers like Thomas Carroll feel that governments lack the ability to successfully serve as the chief facilitators of capacity building efforts. Grassroots groups rely on a strong partnership with national governments to be successful, but governments repeatedly let them down. This is because they rarely confer a

significant and meaningful amount of autonomy to *campesino* (farmer) federations, and second, because of changing government regimes, are unable to provide reliable and continuous state support over long periods (Carroll 2004, 60). Another critic of governments, Raul Zibechi, argues that "it is virtually impossible for grassroots movements … to overcome their dependence on and subordination to the state" (Zibechi 2009). According to Zibechi, wide-reaching social programs claiming to provide assistance to the poor are forms of social control in disguise. Grassroots groups, while they rely on national governments, are systematically undermined by them.

Despite this concerning outlook, the international development community continues to encourage a greater role played by national governments in enacting and evaluating the success of development projects while not pushing for a similar role for grassroots groups. This ideology can clearly be seen in 2019's Sustainable Development Report (SDR). The SDR is published in conjunction between the private foundation Bertelsmann Stiftung and the Sustainable Development Solutions Network (SDSN), a non-profit directed by Jeffery Sachs, the proponent of the SDGs cited previously. The SDSN was set up in 2012 "under the auspices of the UN Secretary-General" and "works closely with United Nations agencies" to mobilize "global scientific and technological expertise to promote practical solutions for sustainable development" (UN Sustainable Development Solutions Network n.d.). Because of the SDSN's close connection to the UN, it can be assumed that this report is most closely aligned with the framework, thinking, and evaluation of the SDGs expected by the UN.[5]

---

[5] The SDR notes that "the views expressed in this report do not reflect the views of any organizations, agency or programme of the United Nations" (Sachs et al 2019, ii). While this report is not an official publication of the UN, it is a relevant and well regarded evaluative source that presents a framework closely resembling UN objectives and recommendations.

Unlike SDG-Tracker.org and the GF, the SDR presents a unique evaluation method of "transformations." In the SDR, the 17 SDGs fall into six transformations that promote "effective implementation strategies by governments, business, and civil society" (Sachs et al 2019, 2). From these transformations, organizations should be able to "develop a clear-eyed implementation strategy" for how they can work towards SDG progress in their own country (Sachs et al 2019, viii). These transformations have names such as "Education, Gender, and Inequality" or "Sustainable Cities and Communities." Unlike the four dimensions of the GF, these transformations do not divide the 17 goals up into six categories. Instead, each transformation "contributes to several SDGs and is synergistic with others" (Sachs et al 2019, 2). As an example, the transformation "Education, Gender, and Inequality" directly targets SDGs 1, 2, 4, 5, 8, 9, and 10. Yet, SDG 1 does not find a home solely in this transformation. All six of the transformations target SDG 1. In some of the transformations, such as "Sustainable Cities and Communities," virtually all SDGs are targeted (Sachs et al 2019, 3). This is the SDR's "clear-eyed" methodology.

The SDR calls for "modern technologies" to play a significant role in monitoring the SDGs (Sachs et al 2019, viii). In the case of SDG 1, the report argues "more timely data is needed to inform policy intervention[s]" that aim to eradicate extreme poverty (Sachs et al 2019, xi). The report calls for "digital technologies, such as artificial intelligence" to be used to "make major contributions towards virtually all SDGs" (Sachs et al 2019, 3). While data-informed policy decision making is a development industry best practice, relying on technology too heavily can be dangerous.

According to Langdon Winner, "'technologies *in themselves* have political properties'" and widely adopting technologies leads to "relative distribution of power, authority, and

privilege" (Winner 1986, 20, 27). Winner suggests that implementing new technologies produces hierarchical systems which have led to a "centralization of social control in large business corporations, bureaucracies, and the military" (Winner 1986, 47). He argues that advancements in technology have not been guided by legal nor moral limits. Instead, businesses and the military have continued to invent "sociotechnical system[s]... in the blind faith that each will turn out to be politically benign" and therefore unharmful (Winner 1986, 58). Before society is able to articulate moral guidelines on technological pursuits, i.e. some sort of "technical constitution," pursuits of technological accumulation, must be done carefully and under regular monitoring (Winner 1987, 57). However, there is no universally accepted mandate for governments, or anyone else, to monitor technological advancements. That makes it so that embracing technological solutions inherently accepts the knowledge hierarchy that technology perpetuates. It is a failure of the SDR to unconditionally advocate for the use of technology like artificial intelligence. Without careful leadership, retrofitting development data collection with technology will fail to bring those at the last mile into evaluation processes.

The SDR offers a unique call for "localized assessment" (Sachs et al 2019, ix). The SDR authors recognize that the SDGs cannot be reached without "significant involvement of mayors and local policy makers" (Sachs et al 2019, ix). This is an important acknowledgement of the role multiple groups need to play in development interventions. However, while the report calls for local government involvement, there is no mention of organizations that do not hold formal power. The SDR's embrace of "local assessment" does not include the buy-in of grassroots organizations who remain yet to be tapped for their extensive knowledge of local contexts. A genuine local assessment would require evaluating progress in every hamlet, foothill, valley, and

basin. Local governments, which can be anything from a municipality to an administrative division, are not sufficiently capable of conducting evaluations at the last mile.

Current development practice promotes more stakeholder involvement in program design and management. Usually, project "stakeholders" refers to all groups affected by interventions. The SDR at times alludes to the role that local organizations can play by merging them with all stakeholders. For example, when describing how to design more effective strategies, the report notes that "governments and other stakeholders need to determine how to organize interventions … and how to deploy them for the SDGs" (Sachs et al 2019, 1). It is concerning how vaguely the SDR defines "other stakeholders." Stakeholders can include anyone from international corporations to foundations to cooperatives to nature. If "governments and other stakeholders" are seen as the only two groups responsible for organizing policy interventions, it seems that 50% of responsibility falls with the government and 50% to an endless list of stakeholders. The role of grassroots organizations does not seem to merit even a specific mention.

While the SDR pays lip service to an increased role of local governments and stakeholders, overwhelmingly, the report focuses on the role of national governments in development. The report argues that "tracking … government policies and commitments to implement the goals" is a way to gauge whether a country is on "track to achieve the goals by 2030" (Sachs et al 2019, 4). The report proposes three levels of measuring government efforts. These include, high-level public statements in support of sustainable development, strategic use of budget, procurement and data, and, "content of government strategies and policy actions" (Sachs et al 2019, 4). To track this third level, the report analyses evidence in 43 surveyed countries on five points: whether high-ranking officials have made public statements about the SDGs, whether they have submitted a voluntary national review of SDG progress, whether SDGs

are mentioned in the federal budget, whether there exists a central or federal institution to

monitor SDG implementation, and whether there have been methods for "stakeholder

engagement" (Sachs et al 2019, 6).

The four Latin American countries surveyed (Argentina, Brazil, Chile, and Mexico) have

all adopted national monitors for implementing the SDGs. Only Canada (244) surpasses

Argentina (243) and Brazil (237) with the most indicators. This demonstrates that key leading[6]

Latin American countries respond well to taking ownership and contextualizing their

development efforts with indicators. Contextualizing SDG efforts with country-specific

indicators is important because "successful models [to ensure public buy in] will differ across

countries, as they must be mindful of history, customs, and government capacity" (Sachs et al

2019, 7). While national context is important, local context is arguably more so. There are many

instances of local history and customs differing from national practices. Local ways of knowing

should be valid to the same degree as expert and professional knowledge when it comes to

development implementation and evaluation.

The SDR is a forceful proponent of government involvement to achieve the SDGs. The

UN's focus on offering recommendations to national governments is common sense given that

the UN is composed of member countries. However, evaluation frameworks about SDG progress

are limited when they overwhelmingly focus on national government action. The SDR notes that

mobilizing "the machinery of government for the SDGs" has been generally unsuccessful. This

conclusion should insinuate that it is time to rethink development evaluations that prioritize

information transfers between governments and supranational organizations. Grassroots

evaluation remains largely untapped and absent from international discourse.

---

[6] In GDP per capita in Latin America, Brazil and Argentina rank first and third respectively.

The SDR is yet another example of an evaluation framework failing to bring awareness to development enactment and evaluation happening successfully at the grassroots level. The SDR's Western-centric validation of expert knowledge, faith in technology, and hyper-focus on government policy perpetuates traditional development ideology. An alternate method to create the intellectual and political space for grassroots organizations to impact development evaluations remains unexplored.

Section Two: Counter-proposals, Shifts, and Alternate Methods

Luckily, SDG-Tracker.org, the GAPFRAME, and the SDR are not the only ways to evaluate progress towards the SDGs. Thomas Carroll imagines an alternate method to affirm grassroots evaluation where grassroots organizations have the capacity to generate valuable information about their projects. Carroll writes "that to achieve their goals and objectives, local groups must develop the same skills as other kinds of organizations, and they are best evaluated on the basis of similar criteria" (Waters 2004, 56). Carroll argues "clear, objective criteria" be used to "monitor and evaluate how well local organizations are functioning to meet the needs of their members" (Waters 2004, 56). Carroll's proposal calls for increasing the capacity of local groups to develop their evaluation skills and then offering the opportunity for them to share their progress throughout their development sector.

Carroll "tests a methodological approach for measuring different facets of organizational capacity by using objective criteria in the form of previously-determined indicators" (Waters 2004, 56-57). This methodology focuses on bringing to the forefront development indicators that are based on organizational capacity. Organizational capacity includes leadership, participation, organizational culture, resource mobilization and use, sustainability, mediation and negotiation, and relationships and alliances. Attempting to measure an organization's organizational capacity

is difficult given that these outcomes are typically regarded in the development field as "intangible" and therefore extremely hard to measure and evaluate. However, if Carroll's proposal for a methodology that could measure capacity building indicators was widely adopted, this would be a breakthrough for development evaluation worldwide. Unfortunately, determining the "objective criteria" and designing a methodology that is adaptable to local contexts is a difficult undertaking. Carroll recognizes the shortfalls of his methodology noting that it is not "set in stone" (Waters 2004, 58). While his book grapples with the tensions of malleability and rigidity, he does not concretely conclude how his vision might be brought into practice.

Carroll's writing lays an important intellectual foundation for others studying organizational capacity building. He argues that slightly larger non-governmental organizations, which he calls intermediary NGOs, are "key protagonists in rural development" by helping to catapult the knowledge of grassroots efforts to a more powerful audience (Waters 2004, 58). According to Carroll, intermediate organizations are not at the grassroots level, but "represent groups of communities and may belong to national organizations" (Waters 2004, 56). Based on Carroll's research of grassroots development,[7] he found that "campesino federations with the best scores [on capacity building are] the result of patient, long-term support by external agents" (Carroll 2004, 59). Carroll observed that success is evident when grassroots groups manage programs "but with ongoing advisory assistance" which decreases overtime in later phases of projects (Carroll 2004, 59).

For capacity building to succeed, organizational learning is essential. Carroll states that "the external agents," i.e. grantors, have to adopt a "flexible attitude that frequently" allows for "changes in their strategy midstream" (Carroll 2004, 59). Carroll highlights that external agents working with the administration of intermediary NGOs catapults the "intimate knowledge of the

---

[7] Carroll's research was commissioned by the Inter-American Foundation

unique internal characteristics of each" grassroots organizations to larger development organizations (Carroll 2004, 60). Carroll reflects, "few donors can function this way" (Carroll 2004, 60). Carroll's proposal of a possible universal methodology to fortify grassroots capacity for evaluation may feel aspirational. However, there is evidence that the development industry is moving in this direction.

The 2010 USAID reforms demonstrate a highly influential institution implementing a new policy to "directly work with and build the capacity of local governments, civil society, and the private sector" (Dunning 2013, 1). The 2010 procurement reform set a goal of channeling 30% of grants and contracts to local partners by 2015.

Immediately, the biggest feasibility concern about increasing local transfers centered around risk. US policy makers, pushed by for-profit lobbying firms on behalf of the US-based contractors lamenting the loss of their contracts, argued that the corruption in grantee countries was a significant risk (Dunning 2013, 6). While to some extent concerns about financial risk are a factor when considering funding government projects, the hesitation to finance local organizations is problematic. To some extent, the distrust of Global South local partners to manage finances and responsibly advance development projects can be attributed to subtle racism and xenophobia towards citizens of Global South countries. Typically, USAID bases its financial partnerships with governments on "data and metrics that test government institutions and their fiduciary systems before resources are transferred" (Dunning 2013, 7). USAID is accustomed to measuring the trustworthiness of governments. However, it entered this new phase of local spending with unease. USAID appeared skeptical that local organizations would be capable of fiduciary responsibility and rigorous evaluation.

Casey Dunning's Center for American Progress report found that local organizations were capable, especially if USAID pushed reforms even further.[8] For example, the report found that transparency of "tracking the local subgrantees of contracts awarded to international implementers" is still absent in USAID publications despite the 2010 procurement reforms (Dunning 2013, 2). USAID's main model is to give grants to US-based organizations and then those organizations allocate sub grants to their "local partners" to carry out development projects. It can be assumed that these USAID contractors, just like any responsible donor, collect data on their project success and use that for their internal learning. If USAID pushed their US-based contract recipients to make sub-grantee information public (i.e. transparency), then USAID could identify which local organizations have the best success implementing programs. The result could be an increase in resource allocation towards these proven successful and innovative organizations. This would allow USAID to streamline capacity building efforts to enable trustworthy grassroots organizations to carry out their own evaluation if they are not doing so already. Reforming the behavior of multilateral donors like USAID may be an effective way to reform development programs and heighten the prominence of grassroots evaluation data.

This chapter examined alternate methods of attempting to evaluate development progress. Evidence from the SDR demonstrates that focus on national indicators is not sufficient. The methodology proposed by Thomas Carroll echoes the call for improved methods to evaluate development that are responsive to the grassroots. Only when grassroots evaluation practices can be lifted to the attention of international actors can it be said that grassroots organizations, knowledge systems, and contexts are truly centered in the development discourse. USAID procurement reforms may be a step in that direction. The following chapter examines a notable

---

[8] It should be noted that Dunning's report may be offering a sympathetic critique of USAID. The Center for American Progress is a DC-area think tank that is considered left learning but not especially radical.

attempt at grassroots development and evaluation by a different US government agency: the

Inter-American Foundation.

Chapter Five: The Inter-American Foundation Grassroots Development Framework

In 1969, the US Congress founded the Inter-American Foundation (IAF) to fund "grassroots development through community-based organizations" (Fiennes and Escobar 2016, 4). For 50 years, the organization has invested in community-led projects in 5,439 communities in 32 Latin American and Caribbean countries (Inter-American Foundation 2019, 10). Their grants are almost all less than $400,000 and unlike the model used by many development organizations, they require grantees to contribute "counterpart" to complement the IAF's investment. This means, the IAF only makes a grant to the grantee if they mobilize resources, a combination of in-kind, cash, or grant donations, to complement the grant given by the IAF. On average, for every $1 that the IAF contributes, grantees contribute $1.30 in counterpart (Inter-American Foundation 2019, 10).

The IAF uses an evaluative framework called the Grassroots Development Framework (GDF) adopted in 2004 which continues to guide the organization's method to evaluate the success of their grantees. Prior to the GDF, the IAF did not use an integrated system and evaluation was only done on a case by case basis. In adopting the GDF, the IAF felt that they were tackling two of the biggest problems with evaluation methods: that traditional evaluations did not capture intangible changes in society and that traditional methods are rarely adaptable to diverse projects and organizations in various multilingual countries. The GDF found a solution to both.

The GDF framework organizes both the tangible and intangible indicators into three categories of change: at the individual, institutional, and societal levels (Adriance 2004, 47). As a part of this model, the grassroots organizations have the opportunity to define for themselves which of the indicators proposed by the IAF they would like to adopt. A Foundation

Representative (FR) from the IAF visits each organization to work with grantees to decide which

indicators they want to work towards. The FR works with the community organization closely to

set up a plan for implementing their desired program while also achieving the goals they have set

for themselves. This capacity building meeting includes coaching leaders through the GDF

process, selecting the indicators, and clarifying the procedures that will be necessary for

thorough and rigorous data collection (Adriance 2004, 47). The FR then departs, leaving

grassroots leaders with the skills and information they need to carry out and evaluate their own

projects.

Every 6 months, a country national data verifier hired by the IAF, returns to the

community until the completion of the funding. Their role is to "confirm the data compiled by

the project administrators" and verify through "meetings with the grantee" that the organization's

internal records match what members of the community say has occurred (Adriance 2004, 47).

The IAF describes this process as "high touch;" grassroots leaders build their capacity to conduct

evaluation through ongoing support by the IAF's local staff. At the end of the grant period, a

final report is submitted by the data verifier to the FR where the "results and the lessons learned"

are kept in a database so that there is "the possibility of examining data from a wealth of

experiences, allowing for in-depth statistical analysis and comparison" (Adriance 2004, 47).

Then, "the IAF will put these improvements as well as its growing database and institutional and

cluster analyses at the service of grassroots development" (Adriance 2004, 47).

The IAF model incorporates many grassroots evaluation best practices. When the GDF

model was developed in the late 1990s, it was groundbreaking. While some organizations are

currently shifting to updated practices for evaluation knowledge sharing, cross-sectoral learning

has been a part of the IAF mandate since its inception. The IAF mandate states that the

organization has a duty to "coordinate its undertakings with the developmental activities in the Western Hemisphere" led by other organizations including bodies of the US government (Inter-American Foundation 1991, 212). Moreover, their mandate extends their intent to collaborate to include "international organizations, and other entities engaged in promoting social and economic development of Latin America" (Inter-American Foundation 1991, 212). It is uncommon to find an organizations' founding documents requiring a wide breadth of accountability and collaboration. This language positions the IAF in a critical place to advance global conversations about development in Latin America.

While the IAF model is successful, it harbors several imperfections. Some development practitioners express concern over the nature of high-touch methods similar to the IAF's work. Some may see the "hand-holding" required by the IAF FRs and data verfiers as overstepping the role international funders should play. This criticism is not explicitly documented in regards to the IAF. On the other hand, there have been studies applauding the IAF's method. A 2011 and 2014 Grantee Perception Report conducted by the Center of Effective Philanthropy of the IAF and hundreds of other funders, found that the oversight and assistance by the IAF is one of the elements that grantees appreciate the most.

In 2016, Caroline Fiennes and Diego Escobar conducted an external review of the IAF answering the question of what "makes IAF's approach to evaluation and reporting so helpful to its grantees" (Fiennes and Escobar 2016, 4). To conduct their study, Finnes and Escobar used evidence from the Grantee Perception Report. The Grantee Perception Report noted that the IAF received the highest score ever recorded for the question: "'How helpful was participating in the foundation's reporting/evaluation process in strengthening the organization/program funded by the grant?'"(Fiennes and Escobar 2016, 4). On a scale from one to seven, seven being extremely

helpful, in 2014, the IAF scored 6.00. Compared to the results from other funders, the metric achieved by the IAF surpasses other organizations by a significant margin; the funders in the second and third position for that question scored 5.80 and 5.72 (Fiennes and Escobar 2016, 4). These impressive results inspired Fiennes and Escobar to investigate further.

The Fiennes and Escobar report identified key aspects of the IAF framework that led to its success. First, the IAF prioritizes data collection that many grassroots organizations had not previously been collecting. After working with the IAF, grassroots organizations were empowered with an "empirical basis" to continue and improve their programming. Second, the report finds that IAF grantees learn management and analysis skills to "collect, handle, interpret, present and use data" which had not previously been known to them. Third, grantees gain confidence and courage in their ability to collect accurate and complete data which helps them secure future grants from funders. And fourth, the grassroots groups gain credibility with their beneficiaries, the broader community, and other organizations. To achieve these results requires a large investment by the IAF; the IAF spends 8% of its total budget (a high amount) on maintaining the training and guidance component of its reporting process (Fiennes and Escobar 2016, 5).

Across the board, the 2014 results of the Grantee Perception Report for the IAF are positive. The IAF scored in the 90th percentile or above on many of the questions asked. This includes scoring in the 99th percentile for the question; "How helpful has the Foundation been to your organization's ability to assess progress towards your organization's goals?" and in the 98th percentile for the question, "After submission of your report / evaluation, did the Foundation or the evaluator discuss it with you?" where 90% of the grantee respondents said "yes" (Fiennes and Escobar 2016, 11). There are, however, two key questions that the IAF did not score well in.

To the question, "Overall, how fairly did the Foundation treat you?" the IAF scored in the 40th percentile and for the question, "To what extent has the Foundation affected public policy in your field?" it scored in the 14th percentile (Fiennes and Escobar 2016, 11). While generally the results of the Grantee Perception Report are positive, Finnes and Escobar note the considerable range that exists: "IAF's grantees are certainly not uniformly positive about everything" (Fiennes and Escobar 2016, 8). Both of these low scores should push IAF's leaders to tweak and adjust the model based on feedback by grantees; it is especially important that no grantee feel they are treated poorly by funders.

The IAF's low score on affecting public policy is an important point to explore. Unfortunately, like many organizations previously discussed, the IAF engages in an internal learning feedback loop. The IAF should be using "statistical analysis and comparison" of evaluation "from a wealth of experiences" to collaborate with other industry funders. This is, unfortunately, not a reality. The GDF is currently under an internal review partly because it does not function as it is intended to.[9] A weak point within the IAF is connecting the evaluations collected by individual FRs to influence development work beyond the agency. Yes, individual grantees are benefitting and yes, individual FRs learn best practices across the portfolio of the grantees that they manage, but learning that is done internal to the organization does not expand beyond its programs. Just how there is a lack of public transparency from USAID about which subcontractors are achieving the most positive outcomes, the IAF does not have a system to share their learning with other donors, NGOs, and government agencies. The IAF's work to an extent operates in isolation. The organization is small and understandably focused on delivering

---

[9] My knowledge of this comes from my experience as an intern in the summer of 2019. This is my conclusion based on observations and conversations with colleagues.

their evaluation results to the stakeholder that matters most: the US Congress. Yet, they are trapped in a system of internal learning.

A significant disappointment is that the SDGs are nowhere to be found in the IAF narrative. As the IAF mandate delineates, the IAF is supposed to coordinate with "international organizations, and other entities engaged in promoting social and economic development of Latin America" which should include the UN and the SDGs. However, since there is no mention of the SDGs in office conversations or the official organization methodology, grantees and FRs do not strive to connect their work to greater global efforts. This shortfall may explain why the IAF scored so low on the question regarding "affecting public policy" in the broader field. If the IAF participated in a framework to extend its learning to reach audiences outside of its organization and direct stakeholders, maybe it could have a tangible impact on the industries where its grantees operate. If other funders are scoring higher than the IAF on this metric, the assumption is that grantees wish that funders expanded their reach. Therefore, the IAF must continue to be the nimble, collaborative, responsive, and learning organization it was founded to be by responding to the valid concerns from their grantees and to the global need for a broader sharing of best practices.

This chapter explored how even the efforts of a well positioned and highly regarded grassroots development organization can fall short of fulfilling the SDG evaluation gap. Even the IAF, an organization that whole-heartedly supports grassroots evaluation and performs capacity building at a very high level, does not universally receive high marks from grantees and fails to engage their learning outcomes in the broader scope of advancing towards the SDGs. If organizations like the IAF are going to be supported to break out of a cycle of internal learning, a new development industry paradigm for information sharing must come to fruition. The

concluding chapter summarizes the challenges of achieving that paradigm shift while offering

hope that it is possible. Central to that possibility is implementing best practices to capture and

validate grassroots evaluation through a global channel.

Chapter Six: Concluding Best Practices and Hope for a New Paradigm

No international consensus has identified which supranational organization is capable of taking the lead to ensure development evaluation reporting prioritizes grassroots organizations. Many Global South countries maintain permanently negative views of global financial institutions like the World Bank.[10] Therefore, global leadership must originate from a more neutral institution. With reforms, the UN Global Pulse Initiative can adopt a credible international mandate capable of compiling and disseminating global knowledge about progress towards the SDGs.

Some groups (not the Global Pulse Initiative) have attempted to take on the challenge of "global reporting" for gathering evidence towards SDG progress, but they are not appropriate for capturing grassroots evaluation. One example is the Global Reporting Initiative (GRI) which asks organizations to submit a report outlining their economic, environmental, and social impacts. Unfortunately, the GRI reporting guidelines include hundreds of pages of instructions, inevitably excluding most non-profits and grassroots organizations, who already have low staffing capacity, from completing a report. The institution that fulfills the role of gathering global evidence of progress towards the SDGs needs to have the capacity to receive data coming in from small nonprofits, cooperatives, and grassroots organizations. These groups typically only have a capacity large enough to evaluate their own programs, not fill out the hundreds of pages required by the GRI.

A concern of creating a "universal" global reporting scheme is whether or not it is ethical to institute one given the diverse contexts, desires, and methodologies of local organizations. Chandra Mohanty is an author that discusses the tension between the local and the universal. She

---

[10] See the Appendix for recommended resources to learn more about these criticisms.

writes that it is possible, but of course difficult, to do "multilayered, contextual analysis to reveal how the particular is often universally significant—without using the universal to erase the particular or positing an unbridgeable gulf between the two terms" (Mohanty 2003, 501). Mohanty says that there is value in conducting an analysis across local differences; the learning generated from local efforts can positively affect universal understanding. In development, generalized analysis, for example by the World Bank and OECD, has historically existed from the top-down, not the bottom-up. The tool that the development industry lacks is an information channel to capture a shared learning from the grassroots to enable a multilayered, contextualized analysis. Following Mohanty's call for "a framework of solidarity and shared values," the development industry must embrace the shared value of the SDGs while following through on solidarity with the knowledge generated from the grassroots (Mohanty 2003, 502). Retooling an institution like the Global Pulse Initiative to compile the localized knowledge from the grassroots through a flexible system of indicators and disseminate it for all funders, governments, NGOs, and grassroots groups, will enable the most equitable transfer of development information around the globe.

The challenge of adopting a new development evaluation paradigm sheds light once more on the complexity of development interventions. The reforms to development evaluation will not be easy according to Paul Collier. He writes that, "the key obstacle to reforming aid is public opinion. The constituency for aid is suspicious of growth, and the constituency for growth is suspicious of aid" (Collier 2007, 183). Development interventions are built on a trust between those benefiting from development programs and those funding it. That trust has eroded over 70 years of promises broken. However, what is broken about development can be fixed. The future

development evaluation paradigm is not doomed to repeat the failures of the past. To achieve success, grassroots development evaluation must play a role in that future paradigm.

The future evaluation paradigm cannot blindly embrace technology as a means of capturing better data. Technology will not magically fix the gaps of data that currently do not exist. Supranational organizations and governments have a responsibility to validate data that for them may not be easy to access. In the spirit of the SDGs embodying genuine global partnership, all development institutions, especially Global North and supranational organizations, must pave the way for all levels of evaluation to become globally significant. Grassroots organizations, NGOs, and funders have the data the world needs but are trapped operating in their own evaluation silos. Breaking down those silos to integrate the entire development evaluation ecosystem is necessary so that all efforts can be counted in the progress towards the SDGs.

For grassroots evaluation to become center stage in the efforts to achieve sustainable development, first, capacity building for local organizations needs to include training on how to monitor and evaluate development results. Local actors have the capacity to track their own progress and conduct analysis, but need the resources and support to get started. Second, an institutionalized channel needs to exist to enable data collected at the grassroots (both qualitative and quantitative) to reach supranational organizations that publish progress updates. A channel housed in the Global Pulse Initiative may fulfill this requirement. Third, industry leaders need to set in motion an industry wide shift that will encourage and allow organizations to break down their silos of keeping learning internal. Finally, the promise of big data analytics and tech-enabled data collection needs to be embraced with caution.

Organizations like the IAF, which have made huge impacts in small communities using a grassroots method, deserve appreciation for being industry pioneers. Yet, efforts by grassroots

organizations cannot remain a drop in the bucket. The development industry must more fully acknowledge that traditional methods have fallen short. Development practitioners have learned that collaboration, knowledge sharing, and listening to diverse voices enables vital improvements to the work. Now, more than ever, the elements exist to make the next decade one that proudly achieves the SDGs. Knowledge from grassroots evaluation, however, will have to take a center role if there is any chance of achieving them.

Appendix

Recommended Readings

Bhagwati, Jagdish. "In Defense of Globalization: It Has a Human Face." Lecture transcript, Angelo Costa Lecture, Rome, Italy, 2005.

Carroll, Thomas F. *Intermediary NGOs: The Supporting Link in Grassroots Development*. West Hartford, CT: Kumarian Press, 1992.

Chang, Ha-Joon. *Bad Samaritans: The Myth of Free Trade and the Secret History of Capitalism*. N.p.: Bloomsbury Press, 2007.

Conteh, Charles, and Lana Wylie. "Globalization and Africa's Experience: 'Fetching Water from a Broken Cistern.'" In *Globalization and Its Regional Impact*, 244-63.

De La O, Ana Lorena. *Crafting Policies to End Poverty in Latin America: The Quiet Transformation*. New York, NY: Cambridge University Press, 2015.

Escobar, Arturo. *Encountering Development: the Making and Unmaking of the Third World*. Princeton, NJ: Princeton University Press, 1995.

Esteva, Gustavo, Salvatore Babones, and Philipp Babcicky. *The Future of Development*. N.p.: Bristol University Press, 2013. https://doi.org/10.2307/j.ctt1sq5vsw.

Gugerty, Mary Kay, and Dean Karlan. *The Goldilocks Challenge: Right-fit Evidence for the Social Sector*. New York, NY: Oxford University Press, 2018.

Healy, Kevin. *Llamas, Weavings, and Organic Chocolate: Multicultural Grassroots Development in the Andes and Amazon of Bolivia*. Notre Dame, IN: University of Notre Dame Press, 2001.

Held, Daivd, and Anthony McGrew. "The Great Globalization Debate: An Introduction." Introduction to *The Global Transformations Reader: An Introduction to the Globalization Debate*, edited by David Held and Anthony McGrew, 1-50. 2nd ed. N.p.: Polity, 2003.

Hershberg, Eric, and Fred Rosen, eds. *Latin America after Neoliberalism: Turning the Tide in the 21st Century?* New York, NY: New Press, 2006.

Hirschman, Albert O. *Getting Ahead Collectively: Grassroots Experiences in Latin America*. Elmsford, NY: Pergamon Press, 1984.

Isbister, John. *Promises Not Kept: The Betrayal of Social Change in the Third World*. 5th ed. Bloomfield, CT: Kumarian Press, 2001.

Marglin, Stephen A. "Development as Poison: Rethinking the Western Model of Modernity." *Harvard International Review*, Spring 2003, 70-75. PDF.

Moyo, Dambisa. *Dead Aid: Why Aid Is Not Working and How There Is a Better Way for Africa*. New York, NY: Farrar, Straus and Giroux, 2009.

Reuter, Peter, ed. *Draining Development? Controlling Flows of Illicit Funds from Developing Countries*. Washington, D.C., USA: World Bank, 2012.

Sachs, Wolfgang. *The Development Dictionary: A Guide to Knowledge as Power*. Johannesburg, South Africa: Witwatersrand University Press, 1992.

Sen, Amartya. *Development as Freedom*. N.p.: First Anchor Books, 1999.

Sheppard, Eric, and Richa Nagar. "From East–West to North–South." *Antipode* 36, no. 4 (September 2004): 557-63. Accessed April 29, 2020. https://doi.org/10.1111/j.1467-8330.2004.00433.x.

Stiglitz, Joseph. "Globalism's Discontents." *American Prospect*, January 14, 2002.

Tendler, Judith. *Good Government in the Tropics*. Baltimore, MD: Johns Hopkins University Press, 1997.

Uphoff, Norman, Milton J. Esman, and Anirudh Krishna. *Reasons for Success: Learning from Instructive Experiences in Rural Development*. West Hartford, CT: Kumarian Press, 1998.

Bibliography

Adriance, Joel. "IAF's GDF in a Nutshell." *Grassroots Development Journal* 25, no. 1 (2004): 47. https://archive.iaf.gov/home/showdocument%3Fid=189

Ajmera, Maya, and Gregory A. Fields. *Invisible Children: Reimagining International Development at the Grassroots*. N.p.: Palgrave Macmillan, 2016.

Annis, Sheldon, and Peter Hakim, eds. *Direct to the Poor: Grassroots Development in Latin America*. Boulder, CO: Lynne Rienner Publishers, 1988.

Carroll, Thomas. "Some Lessons for Organizational Capacity-Building." *Grassroots Development Journal* 25, no. 1 (2004): 59-63. Accessed February 16, 2020. https://archive.iaf.gov/home/showdocument%3Fid=189

Collier, Paul. *The Bottom Billion: Why the Poorest Countries Are Failing and What Can Be Done About It*. Oxford, UK: Oxford University Press, 2007.

Dunning, Casey. *Is Local Spending Better? The Controversy over USAID Procurement Reform*. N.p.: Center for American Progress, 2013.

Escobar, Arturo. *Encountering Development: the Making and Unmaking of the Third World*. Princeton, NJ: Princeton University Press, 1995.

Escobar, Arturo, and Wendy Harcourt. "Post-development Possibilities: A Conversation." *Development*, November 5, 2018. Accessed September 29, 2019. https://doi.org/10.1057/s41301-018-0184-3.

Esteva, Gustavo, and Arturo Escobar. "Post-Development @ 25: on 'being stuck' and moving forward, sideways, backward and otherwise." *Third World Quarterly* 38, no. 12 (2017): 2559-72. Accessed February 17, 2020. https://doi.org/10.1080/01436597.2017.1334545.

Fiennes, Caroline, and Diego Escobar. *Frankly Speaking: The Inter-American Foundation's Reporting Process: Lessons from A Positive Outlier*. January 2016. Accessed December 2, 2019. https://givingevidence.files.wordpress.com/2016/01/iaf-giving-evidence-case-study-january-2016-1.pdf.

GAPFRAME. "About the GAPFRAME." GAPFRAME. Last modified 2017. Accessed November 12, 2019. http://gapframe.org/about-the-gapframe/.

———. "Methodology." GAPFRAME. Last modified 2017. Accessed November 12, 2019. http://gapframe.org/methodology/.

———. "People behind the GAPFRAME." GAPFRAME. Last modified 2017. Accessed November 12, 2019. http://gapframe.org/people-behind-the-gapframe/.

———. "Quality of Life - Society." GAPFRAME. Last modified 2017. Accessed November 12, 2019. http://gapframe.org/by-issue/quality-of-life/.

Gugerty, Mary Kay, and Dean Karlan. *The Goldilocks Challenge: Right-fit Evidence for the Social Sector*. New York, NY: Oxford University Press, 2018.

———. "Ten Reasons Not to Measure Impact—and What to Do Instead." *Stanford Social Innovation Review* 16, no. 3 (Summer 2018): 40-47. Accessed April 25, 2020. http://libproxy.vassar.edu/login?url=https://search-proquest-com.libproxy.vassar.edu/docview/2043232014?accountid=14824.

Hirschman, Albert O. *Getting Ahead Collectively: Grassroots Experiences in Latin America*. Elmsford, NY: Pergamon Press, 1984.

Inter-American Foundation, SUBCHAPTER XXI—INTER-AMERICAN FOUNDATION, H.R. Doc., at 212 (Oct. 28, 1991). Accessed February 23, 2020. https://www.iaf.gov/wp-content/uploads/2019/04/USCODE-2009-title22-chap7-subchapXXI-sec290f.pdf.

Inter-American Foundation. *2019 in Review*. By Inter-American Foundation. 2019. Accessed February 23, 2020. https://www.iaf.gov/wp-content/uploads/2019/11/IAF-in-Review-digital.pdf.

"J-PAL." J-PAL. Accessed February 19, 2020. https://www.povertyactionlab.org/.

León, Marcela Vásquez, Brian J. Burke, and Timothy J. Finan, eds. *Cooperatives, Grassroots Development, and Social Change: Experiences from Rural Latin America*. Translated by Maisa C. Taha. Tucson, AZ: University of Arizona Press, 2017.

Mohanty, Chandra Talpade. "'Under Western Eyes' Revisited: Feminist Solidarity through Anticapitalist Struggles." *Signs* 28, no. 2 (Winter 2003): 499-535. Accessed April 29, 2020. https://blogs.baruch.cuny.edu/shortnovel/files/2016/01/Chandra-Mohanty.pdf.

OECD. "Where: Global Reach." OECD.org. Last modified 2019. Accessed November 12, 2019. https://www.oecd.org/about/members-and-partners/.

Oliver-Smith, Anthony. *Defying Displacement: Grassroots Resistance and the Critique of Development*. Austin, TX: University of Texas Press, 2010.

Pattanaik, B.K. "Decentralized development planning: an approach towards grassroots development." *Political Economy Journal of India* 25, no. 1 (Spring/Summer 2016): 77-87.

Ritchie, Roser, Mispy, Ortiz-Ospina. "Measuring progress towards the Sustainable Development Goals." *SDG-Tracker.org, website* (2018). https://sdg-tracker.org/.

Sachs, Jeffrey, Guido Schmidt-Traub, Christian Kroll, Guillaume Lafortune, and Grayson Fuller. *Sustainable Development Report 2019*. 2019. Accessed April 29, 2020.

https://s3.amazonaws.com/sustainabledevelopment.report/2019/2019_sustainable_develo pment_report.pdf.

Sachs, Jeffrey D. "From Millennium Development Goals to Sustainable Development Goals." *Lancet* 379 (June 9, 2012): 2206-11.

Sachs, Wolfgang. "The Sustainable Development Goals and Laudato si': varieties of Post-Development?" *Third World Quarterly* 38, no. 12 (2017): 2573-87.

Santos, Angelo Agustin Rivero. "Grassroots and the State: Perspectives from the Neighbors' Movement in Caracas, Venezuela." PhD diss., Syracuse University, 1997.

United Nations. "Sustainable Development Goals." Sustainable Development Goals Knowledge Platform. Accessed February 24, 2020. https://sustainabledevelopment.un.org/sdgs.

———. "Sustainable Development Goal 17." Sustainable Development Goals Knowledge Platform. Accessed February 24, 2020. https://sustainabledevelopment.un.org/sdg17.

United Nations Global Pulse. *Integrating Big Data into the Monitoring and Evaluation of Development Programmes*. By Michael Bamberger. 2016. Accessed February 17, 2020. https://beta.unglobalpulse.org/wp-content/uploads/2016/12/integratingbigdataintomedpwebungp-161213223139.pdf.

UN Sustainable Development Solutions Network. "About Us." Sustainable Development Solutions Network. Accessed February 21, 2020. https://www.unsdsn.org/about-us.

U.S. Agency for International Development. *Strengthening Evidence-Based Development: Five Years of Better Evaluation Practice at USAID 2011–2016*. March 2016. Accessed April 25, 2020. https://pdf.usaid.gov/pdf_docs/PBAAE204.pdf.

Waters, William F. Review of *Construyendo capacidades colectivas: Fortalecimiento organizativo de las federaciones campesinas-indigenas en la Sierra ecuatoriana*, by Thomas F. Carroll. *Grassroots Development Journal* 25, no. 1 (2004): 56-58. Accessed April 29, 2020. https://archive.iaf.gov/home/showdocument%3Fid=189.

Winner, Langdon. *The Whale and the Reactor*. Chicago, IL: University of Chicago Press, 1986.

Ziai, Aram. "Post-development 25 years after The Development Dictionary." *Third World Quarterly* 38, no. 12 (2017): 2547-58. Accessed February 17, 2020. https://doi.org/10.1080/01436597.2017.1383853.

Zibechi, Raúl. "Autonomy or New Forms of Domination? The Complex Relationship Between Governments and Movements." Translated by Monica Wooters. Upside Down World. Last modified February 23, 2009. Accessed February 21, 2020. http://upsidedownworld.org/archives/international/autonomy-or-new-forms-of-domination-the-complex-relationship-between-governments-and-movements/.