



Commentary

Good will hunting: Challenges of theory-based impact evaluations for climate investments in a multilateral setting

Jyotsna Puri, Archi Rastogi, Martin Prowse, Solomon Asfaw*

Green Climate Fund, Independent Evaluation Unit, Songdo, Incheon, South Korea



ARTICLE INFO

Article history:

Accepted 1 December 2019

Keywords:

Climate finance
Theory of change
Impact evaluation
RCTs
Evidence
Real-time
Multilateral

ABSTRACT

In 2018, the Green Climate Fund (GCF) initiated a multi-year Learning-Oriented Real-Time Impact Assessment (LORTA) programme to understand what works in climate interventions, for whom, how much and why. LORTA supports quality data collection, learning and causal impact measurement. It requires leadership from project teams alongside an openness by management to build measurement structures and learning into their decision making. We highlight the institutional dynamics and challenges encountered when institutionalizing LORTA within the GCF ecosystem of international and national actors. These challenges may also apply in other multilateral settings.

© 2019 Elsevier Ltd. All rights reserved.

1. Introduction

The Green Climate Fund (GCF) was established by the United Nations Framework Convention on Climate Change (UNFCCC). It aims to contribute to a paradigm shift towards low-emission and climate-resilient development pathways in developing countries (Green Climate Fund, 2011). As of the 14th November 2019, the GCF had secured just under US\$10 billion in pledges to implement its Strategic Plan for 2020–2023. Two important features characterize the Fund's plans and strategies. First, it is assumed that the Fund will use its scarce resources effectively: although it is the largest climate fund, the GCF is relatively small compared to global annual climate finance flows that were \$579 billion in 2017–18 (CPI, 2019) and climate needs of at least US \$1.8 trillion (IPCC, 2018). Second, the Fund is expected to crowd-in climate finance from the private sector. Both these features assume that the GCF will be able to identify, measure and demonstrate its impact credibly.

In an early review, the Independent Evaluation Unit (IEU) of the GCF found that there is scarce evidence on what works, how much, why and for whom in climate adaptation and mitigation (IEU, 2018). More and better evidence on the impact of climate interventions is therefore clearly required. In the past, other disciplines have confronted a similar scarcity of causal evidence and subse-

quently have variously turned towards evidence-based policy (in medicine see Schulz, Chalmers, Hayes, & Altman, 1995; in international development see Jonathan & Sumner, 2014).¹ To help contribute to building a similar evidence base for climate-related interventions while also aiming eventually to measure the causal impact of GCF investments using theory based impact evaluation methods, the IEU initiated a Learning-Oriented Real-Time Impact Assessment (LORTA) programme in 2018. In this short article, we discuss the institutional dynamics and challenges of this work.

2. The GCF and the challenge of credible measurement

The structures and processes within the GCF present four major challenges for credibly measuring attributable change caused by its investments (also called theory-based impact evaluations).

First, the GCF makes investments in adaptation and mitigation in developing countries. The GCF's current portfolio is heterogeneous and projects (or investments) have multiple objectives and causal pathways and are distributed over 100 countries. Indeed, each investment has all the *characteristics of a complex system*. Such complexity poses challenges for effective evaluation and attributable impact measurement. These include a diversity of actors and interactions, networks and feedback loops, openness of systems, non-linearity and emergence (De Coste and Puri, 2019).

* Corresponding author.

E-mail addresses: jpuri@gcfund.org (J. Puri), arastogi@gcfund.org (A. Rastogi), mprowse@gcfund.org (M. Prowse), sasfaw@gcfund.org (S. Asfaw).

¹ Jimenez and Puri (2017) find that climate action are challenging to evaluate because of distal impact, multiple objectives, and their focus on an inherently public good.

Second, all GCF investments are required to demonstrate country ownership and be country-led. In practice, this means that to be considered by the GCF, investment proposals require prior and continuous participation (and sign-off) from a diverse set of in-country and international stakeholders including staff in the country's National Designated Authority (usually a line ministry such as finance or environment), staff of an Accredited Entity (AE) (who receive GCF funds), staff of Executing Entities (who implement the project) and civil society organizations. Since any effort to measure impacts/change rigorously and credibly, requires unbiased and systematic measurement and independent data systems, it means that all key actors must agree on the production, use, ownership, analysis, and release of data and findings. *This means there need to be multi-partner communications and agreements.*

Third, in its review of the GCF's Results Management Framework (RMF) in 2018, the IEU found that *fundamental requirements for measuring and reporting impact credibly had not been incorporated in GCF systems and processes (IEU, 2018)*. The GCF relies on 'self-reports' from AEs for reporting results. The IEU found that recommended indicators were not well-defined. Additionally, there was no requirement for projects to build measurement systems at the beginning and there was also an absence of guidance on definitions of indicators, methods for tracking implementation, measuring results, time frames and frequencies of data collection. There was also varied understanding of who would be held accountable for deficient or absent reporting. Since systems for credible measurement and reporting were not incorporated in the design of proposed investments, assessments of impact need to rely solely on *ex post* and mostly descriptive data.

Finally, a review by the IEU (Fiala, Puri and Mwandri, 2019) found that *the quality at entry of GCF investment proposals left much to be desired*. Of the 93 proposals that had been approved by the GCF Board (equivalent to \$US 4.8 billion), 20% had well-defined theories of change, 15% contained plans for credible measurement, 13% had impact indicators that were judged to be capable of measuring the magnitude of causal change and 10% included plans for collecting data of sufficient quality for a causal evaluation.

In summary, there are many challenges to measuring causal change and impact.

3. LORTA – What does it do and how?

LORTA is a technical assistance programme for approved GCF projects that helps to build high-quality datasets, aids real-time learning, and helps measure causal change/impact. More specifically, LORTA helps projects measure the causal impact of GCF investments using theory-based mixed methods that include randomized control trials and quasi-experimental designs with 2–3 waves of survey data for targeted populations and, where relevant, for comparison populations. All data collection efforts are informed by well-defined protocols which, in turn, are predicated on well-planned formative field work.² Since the programme is led by the IEU, LORTA helps support data systems and analyses that are independent. To overcome institutional layers and navigate the institutional landscape and its associated challenges discussed in Section II, LORTA has set up three structures and processes.

First, the IEU works closely with the GCF Secretariat (who lead interactions with the AEs, executing entities, national designated authorities etc.) to create awareness and capacity on theory-based impact evaluation designs. Over time, the GCF Secretariat has started to provide guidance on designing investment proposals that incorporate data systems and include considerations of these theory-based designs.

Second, LORTA requires that staff of AEs and project teams buy into the overall idea of incorporating causal designs and theory-based impact evaluations. In practice, hesitation and resistance from them is due to costs, timing and an absence of capacity and understanding of theory-based impact evaluations. It is well recognized that theory-based impact evaluations require additional budgets (including data and measurements in counterfactual/comparison areas) and a higher quality of survey and monitoring data than projects usually cater for (Puri and Rathinam, 2019). They also require an understanding amongst project staff of why these are important (LORTA Synthesis report, 2019). In practice, it is also critical that proposed impact evaluation designs conform with planned operations and programme roll-out. Frequently this means that impact evaluators do not have the ability to design and roll-out the preferred theory-based impact evaluation design (that include both identification designs that involve randomization or quasi-experimental designs but also high-frequency data collection systems and qualitative approaches). This in turn has implications for internal validity of results. Recognizing this, LORTA's approach has been to identify champions within AEs and amongst projects, to help demonstrate the possible benefits of theory-based impact evaluations. Indeed, it has taken advantage of the fact that today many international agencies want visibility on impact measurement and have staff that are trained in theory-based impact evaluations. The IEU has worked hard to identify and build relationships with these champions so that they can, in turn, advocate, build and budget for theory-based impact evaluations within investment proposals.

Third, the IEU team has also developed credibility by focusing on areas of measurement and data collection that helps project teams realize some of their shorter-term goals while still contributing to the overall goal of impact evaluation and measuring attributable change. This has taken the form of organizing 'design workshops' with project teams, providing customized training on theories of change and focusing on how to plan and measure better in real-time so project teams avoid delays and close gaps between planning and implementation (Legovini, Di Maro, & Piza, 2015). In effect, LORTA has responded to the needs of a diversity of actors and has developed an 'in-kind incentive system': project teams learn about real-time tracking and adaptive management; AEs learn about setting better indicators, better planning for project teams and more accurate impact reporting; the GCF Secretariat helps to select projects that participate in the LORTA programme; and the IEU meets its mandate on providing evidence on what works, why and how much.

4. Opportunities and learning

There are three broad lessons from the LORTA process. First and foremost, the success of theory-based impact evaluation requires partnerships and mutual trust between implementers and evaluators. Implementers must understand and trust that these evaluations will assist them in improving the programme and that better real-time measurement will enhance programme quality and results. Evaluators, in turn need to inform implementers of their plans, and build their capacity for understanding and grasping key concepts of theory-based impact evaluations. Our experience in the LORTA programme is that once mutual trust is built, implementers are more likely to welcome advice and act upon it, understanding that such evidence provides important opportunities for learning and improving the programme.

A second important lesson is the necessity of adopting a flexible approach that can help minimize the burden of theory-based impact evaluations on programme implementation staff. Open discussions on trade-offs between design options on one hand, and

² Please refer to: <https://ieu.greenclimate.fund/evaluations/lorta>.

threats to internal validity help. AEs and Executing Entities can that are implementing the programme can then make informed decisions about how much they value a sound design and how much they will 'pay' in terms of budgets and roll-out or other implementation plans. Frequently impact evaluation designs can also take advantage of implementation challenges and highlight these for planners (by suggesting phased roll-outs, or lotteries) that also present options for impact evaluations. Discussions on theories of change also alert programme staff to potential bottlenecks and the importance of real-time tracking. Additionally, discussions about sample sizes frequently indicate that many programme staff over-estimate their measurement challenges (and costs). It is important to note that LORTA activities are drawn from IEU budgets, baseline and end-line data collection budgets themselves are supported by projects and AEs.

Last but not least, as use, uptake and learning specialists know, informing policy through evidence is a non-linear process not least as policy-making itself is iterative and interactive. Cooperative relationships between evidence creators, advocates, and policy targets are important. How evidence is generated, and findings are communicated is key and uses multiple channels over time, increasing the chance of positive outcomes. The effective integration of research findings, and even the research process itself, into the policy process, is important in order to ensure that evaluative research can help inform the evidence needs of the GCF and, most importantly, the climate needs of developing countries.

References

- Climate Policy Initiative (CPI) (2019). *Global Landscape of Climate Finance*. London: Climate Policy Initiative.
- DeCoste, S., & Puri, J. (2019). Complexity, climate change and evaluation. IEU Working Paper No. 02, Green Climate Fund. Songdo, South Korea
- Fiala, N., Puri, J., Mwandri, P. (2019). Becoming bigger, better, smarter: A summary of the evaluability of Green Climate Fund proposals. IEU Working Paper No. 1, Green Climate Fund. Songdo, South Korea
- Green Climate Fund (2011). *Governing Instrument for the Green Climate Fund. Conference of the Parties to the United Nations Framework Convention on. Climate Change (UNFCCC)*.
- IEU. (2018). Independent Evaluation of Green Climate Fund Readiness and Preparatory Support Programme. Evaluation Report No. 1, Green Climate Fund, Songdo, South Korea
- Intergovernmental Panel on Climate Change (IPCC) (2018) Global Warming of 1.5° C: An IPCC Special Report on the impacts of global warming of 1.5 °C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty
- Jimenez, E., & Puri, J. (2017). The wicked cases of education and climate change: the promise and challenge of theory-based impact evaluations, in "Evaluation for Agenda 2030: Providing Evidence on Progress and Sustainability", edited by Rob D. van den Berg, Indran Naidoo and Susan D. Tamondong
- Jonathan, G., & Sumner, A. (2014). *The \$138.5 Billion Question: When Does Foreign Aid Work (and When Doesn't It)?* Washington, DC: Center for Global Development.
- Legovini, A., Di Maro, V., & Piza, C. (2015). *Impact evaluation helps deliver development projects. Policy Research working paper; no. WPS 7157; Impact Evaluation series*. Washington, DC: World Bank Group.
- Puri, J., & Rathinam, F. X. (2019). Challenges in Real-World Impact Evaluations: Some Learning on Costs and Timeliness. *IEU Working Paper No., 03*, 2019.
- Schulz, K. F., Chalmers, I., Hayes, R. J., & Altman, D. G. (1995). Empirical evidence of bias: Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *Jama*, 273(5), 408–412.