Commentary

# Large-scale randomized control trials of incentive-based conservation: What have we learned?

Nigel Asquith

*Cuencas Sustentables, Calle Rio Totaitu 15, Santa Cruz, Bolivia*

ARTICLE INFO

ABSTRACT

Landscape-scale conservation programs are challenging to implement, and even more difficult to evaluate. Fundación Natura Bolivia and associated researchers have spent the last decade undertaking a series of randomized control trials (RCTs) of an incentive-based conservation program in Bolivia. Large RCTs are complex, perhaps more so in conservation, as they require measurement of multiple kinds of outcomes operating on different timescales. We have learned that successful RCTs of conservation interventions require that program implementers demonstrate seven characteristics, namely that they are able and willing to: replicate a proven intervention at scale, define and measure outcomes, risk their reputation, have patience, access world-class technical research support, inculcate a tight researcher/practitioner collaboration and adapt the intervention based on evaluation results. Importantly, we have shown that large-scale robust RCT-based evaluations are possible in conservation. Learning how to use such evaluation tools is critical if conservation practitioners are to demonstrate attributable impact of their interventions.

© 2019 Elsevier Ltd. All rights reserved.

Landscape-scale conservation solutions are needed to mitigate the global threats of climate change and biodiversity loss. However, at landscape scales, clean causal inference and theories of change are often muddied. Effective evaluation is thus critical in order to understand if and how an intervention can work. The experimental tools for impact evaluation pioneered by Professors Banerjee, Duflo and Kremer have revolutionized the approach to alleviating global poverty: their methods now entirely dominate development economics. Nevertheless, this experimental approach to evaluation has been slow to spread to conservation (Ma et al., 2017).

Randomized control trials (RCTs) – the signature tool of the experimental evaluation paradigm – are complex to design and implement. This is especially so in conservation, as large RCTs often require measurement of multiple outcomes over different temporal and spatial scales (Pynegar et al., 2019). There have thus been few RCTs of landscape-scale conservation programs (c.f. Chaves et al., 2017): one each in Uganda (Jayachandran et al., 2017) and Sierra Leone (Wilebore et al., 2019) and a series of four in Bolivia (Pynegar et al., 2018; Wiik et al., 2019, 2020). In this paper I reflect on lessons learned from the Bolivian RCTs and identify the evaluation challenges faced by conservation practitioners.

*Watershared* is an incentive-based conservation program currently replicating across the Andes. The original *Watershared* pro-

ject (Acuerdos Recíprocos por Agua (ARA) in Spanish) conserved 465 ha of forest in the Bolivian village of Los Negros in 2003 (Asquith et al., 2008). By 2019, 58 Bolivian municipalities had appropriated the model, with 8,000 upstream farmers conserving and reducing cattle grazing in 350,000 ha of forests. This conservation was in exchange for $500,000 worth of development projects annually contributed by 250,000 water users. Similar *Watershared* programs have been initiated in Colombia, Ecuador, Peru and Mexico (Asquith, 2016).

By 2010, *Watershared* appeared to hold great potential for replicable, scalable in-kind transfers in exchange for conservation. Pilot research had described the basic model (Asquith et al., 2008), identified famers' and community leaders' motivations to participate (D'Adda, 2011; Grillos, 2016; Jack & Recalde, 2014) and identified the program's focus on building institutions for sustainable commons management and behavior change (Asquith, 2016). Given *Watershared*'s potential for replicability, Fundación Natura Bolivia (Natura) decided to see if the model could work at scale, and if so, how it could be improved.

In a series of four RCTs evaluating *Watershared*, Natura and associated researchers quantified changes in forest cover and water quality across a landscape, changes in the socio-economic status of participants, and their perceptions about natural resource management (Pynegar et al., 2018; Wiik et al., 2019; Wiik et al., 2020). We learned that large-scale RCTs are frustratingly challenging. Our overall experience suggests that RCTs of conservation

interventions will likely only be successful if project implementers are willing and able to:

- Replicate a proven intervention at scale
- Identify and measure outcomes
- Risk their reputation
- Have patience
- Access world-class technical research support
- Promote and inculcate a tight researcher/implementer collaboration
- Adapt the intervention based on evaluation results

I discuss these issues below and describe their importance for our evaluations. I suggest that the main reason there are so few RCTs of conservation interventions is because few conservation NGOs can comply with these criteria.

*1. Replicate a proven intervention at scale.* In deciding whether to invest in an RCT, there is fine balance between ensuring that an intervention has been sufficiently proven to be worthwhile, but not studied enough to know if it can work at scale. Prior to our research, *Watershared* pilots seemed to have functioned successfully (Asquith et al., 2008). Critically important for our evaluation was that *Watershared* programs are small, locally managed and easily copied. Our RCTs evaluated if modular replication of these successes could demonstrate impact at the landscape scale. Unfortunately, the conservation sector has rarely been able to identify scalable solutions, with most interventions comprising one-off projects adapted to specific local circumstances. Unless conservation practitioners can develop replicable solutions – and rigorously evaluate if they work beyond one location – they will simply be unable to solve the climate crisis.

*2. Define and measure outcomes.* Many conservation organizations monitor outputs (e.g., park guards trained or equipped) but few evaluate outcomes (e.g., additional hectares effectively conserved). This is partly because the complexities of landscape-scale conservation and development make it difficult to attribute the impact of a particular intervention. Our RCTs included the area of forest in conservation as an output and measured the intervention's impact on water quality (Pynegar et al., 2018), deforestation rates (Wiik et al., 2019), and participant behavior change (Bottazzi et al., 2018; Grillos et al., 2019; Wiik et al., 2020). Without randomized controls it would have been more difficult for us to truly tease out and attribute impact. It is perplexing that – unlike in the development sector – few conservation donors require proof of impact. Unless the conservation sector and its funders get better at defining and measuring impact, it will be impossible for them to claim success or improve.

*3. Risk their reputation.* Few conservation practitioners are accustomed to the bracing cold shower of peer review. With no robust evaluation required by funders or partners, it rarely occurs to conservation practitioners that interventions might not be achieving their expected outcomes. In contrast, an RCT mind-set encourages practitioners to take risks, and to make the intellectual leap of recognizing that if a project isn't working as planned, it can be treated as a lesson. Indeed, our first RCT suggested that five years of *Watershared* had not resulted in measurable landscape-scale impacts on water quality and deforestation. Rather than hiding this inconvenient truth, the research team published it widely (Pynegar et al., 2018; Wiik et al., 2019) and Natura redesigned the next program iteration accordingly.

*4. Have patience.* Many conservation outcomes – including those measured in our RCTs – take years to be manifest. For example, even after cattle are removed from riverine forests, coliforms in their feces can lie dormant in stream sediment for years, to be released only when sediment is disturbed (Pynegar et al., 2018). Conservation project and research funding cycles are invariably shorter than impact timelines. This creates a challenge: how to retain investigators and funders for long enough to design the research, collect baseline data, implement the intervention, collect end line data and analyze results? In our case it took almost a decade from conceptualization to first peer-reviewed journal article publication. Our initial student collaborators are now associate professors, two doctorates have been completed, collaborators have published papers and left the team, and a new generation of researchers and graduate students have joined. To maintain momentum, we needed to continually recycle a dozen donors, each of which risked financing an individual investigative step, recognizing they could not fund the entire research program.

*5. Access world-class technical research support.* Field data can be collected in developing countries inexpensively. Because the data requirements of an RCT (Bottazzi et al. 2017) are only slightly more complicated than those for effective before-and-after project monitoring, we naively believed that RCTs could also be inexpensive. We did not understand how expensive data analysis and post-doctoral and graduate student time would be, and how difficult it would be for a developing-world NGO to justify academic support at developed-world prices. We were thus required to blend funding sources: Natura applied for project implementation grants, while investigators applied in parallel for research funds. Only once these complementary funding streams were synchronized were the RCTs possible.

*6. Promote and inculcate a tight researcher/implementer collaboration.* Our researchers were authorized to publish their results independently with Natura having no veto power, but in reality, the research and implementation teams were co-dependent. The pioneering nature of conservation RCTs, the complexities of large geographic scale and long timelines, the unfamiliar reputational risk for Natura and the need for synchronized implementation and research fundraising, meant that researchers could not completely control how the experiments would unfold in the field nor undertake entirely independent evaluations. This co-dependency between researchers and implementers is not what evaluation purists would prefer, but we found that tight collaboration was critical as we jointly wrestled with designing an effective conservation RCT. Resolving unexpected challenges required full trust and compromises from all team members, which would have been impossible with a traditional "arm's length" evaluation.

*7. Adapt the intervention based on evaluation results.* There is little point undertaking an evaluation if its results are not used to change the program. In the first experimental iteration of *Watershared*, individual landowners could decide which land parcels to enroll, and the value of the compensations was moderate, based solely on parcel size. As a result, uptake was low, and so the intervention did not have a significant landscape-scale impact on water quality or deforestation (Pynegar et al., 2018; Wiik et al., 2019). However, the results were quickly analyzed, allowing Natura to immediately adapt the program. In the next iteration, compensation levels were increased, based on the number of hectares conserved and population size, and were made to communities, not individuals. Moreover, participants were required to conserve at least 50% of their watersheds. Coordination and trust between researchers and implementers thus enabled programmatic adaptations long before the results were published.

Our experiences over the last 10 years suggest that these seven characteristics of implementing organizations are essential in order to undertake robust impact evaluations of conservation interventions. Conservation, however, is perhaps not fundamentally different from development. These seven criteria are probably important not just for RCTs in the conservation sector, but for all large-scale experimental evaluations. For example, quantifying impact will always expose practitioners to reputational risk, and the costs of expert evaluation are high for any sophisticated data

analysis. However, what makes these characteristics uniquely important for the conservation sector, is that few environmental organizations yet embrace them.

Although RCTs are challenging, Natura's *Watershared* evaluations show that rigorous analysis of long-term, large-scale conservation interventions is possible. Learning how to use such evaluation tools is critical if conservation practitioners are to demonstrate measurable and attributable impact. Just as Professors Banerjee, Duflo and Kremer revolutionized evaluation of the development sector, conservationists need to revolutionize evaluation of their sector. It is not an exaggeration to suggest that the future of life on earth depends on it.

## Acknowledgements

## References

Asquith, N. M., Vargas Rios, M. T., & Wunder, S. (2008). Bundling environmental services: Decentralized in-kind payments for bird habitat and watershed protection in Los Negros. *Ecological Economics, 65*, 675–684.

Asquith, N.M. (2016) Watershared: Adaptation, Mitigation, Watershed Protection and Economic Development in Latin America. Climate & Development Knowledge Network, London, UK.

Bottazzi, P., Wiik, E., Crespo, D., & Jones, J. P. G. (2018). Payment for environmental 'self-service': Exploring the links between farmers' motivation and additionality in a conservation incentive programme in the Bolivian Andes. *Ecological Economics, 150*, 11–23.

Bottazzi, P., et al. (2017) Baseline and endline socio-economic data from a Randomised Control Trial of the Watershared intervention in the Bolivian Andes. (Data Collection). Colchester, Essex: UK Data Archive. 10.5255/UKDA-SN-852623.

Chaves, W. A., Valle, D. R., Monroe, M. C., Wilkie, D. S., Sieving, K. E., & Sadowsky, B. (2017). Changing Wild Meat Consumption: An Experiment in the Central Amazon, Brazil. *Conservation Letters*. https://doi.org/10.1111/conl.12391.

D'Adda, G. (2011). Motivation crowding in environmental protection: Evidence from an artefactual field experiment. *Ecological Economics, 70*(11), 2083–2097.

Grillos, T. (2016). Economic vs. Non-Material Incentives for Participation in an In-Kind Payment for Ecosystem Services Program in Bolivia. *Ecological Economics, 131*, 178–190.

Grillos, T., Bottazzi, P., Crespo, D., Asquith, N., & Jones, J. P. G. (2019). In-kind conservation payments crowd in environmental values and increase support for government intervention: A randomized trial in Bolivia. *Ecological Economics*, E166. https://doi.org/10.1016/j.ecolecon.2019.106404.

Jack, K., & Recalde, M. P. (2014). Leadership and the voluntary provision of public goods: Field evidence from Bolivia. *Journal of Public Economics, 122*(2015), 80–93.

Jayachandran, S., De Laat, J., Lambin, E. F., Stanton, C. Y., Audy, R., & Thomas, N. E. (2017). Cash for carbon: A randomized trial of payments for ecosystem services to reduce deforestation. *Science, 357*, 267–273.

Ma, Z., Bauchet, J., Steele, D., Godoy, R., Radel, C., & Zanotti, L. (2017). Comparison of Direct Transfers for Human Capital Development and Environmental Conservation. *World Development, 99*, 498–517.

Pynegar, E. L., Gibbons, J. M., Asquith, N. M., & Jones, J. P. G. (2019). What role should randomized control trials play in providing the evidence base for conservation?. *Oryx*, 1–10.

Pynegar, E. L., Jones, J. P. G., Gibbons, J. M., & Asquith, N. M. (2018). The effectiveness of payments for ecosystem services at delivering improvements in water quality: Lessons for experiments at the landscape scale. *PeerJ, 6*. https://doi.org/10.7717/peerj.5753 e5753.

Wiik, E., D'Annunzio, R., Pynegar, E. L., Crespo, D., Asquith, N. M., & Jones, J. P. G. (2019). Experimental evaluation of the impact of a payment for environmental services program on deforestation. *Conservation Science and Practice* e8.

Wiik, E., Jones, J. P. G., Asquith, N. M., Bottazzi, P., Gibbons, J. M., Kontoleon, A., & Pynegar, E. L. (2020). Exploring mechanisms and impacts of an incentive-based conservation program with evidence from a Randomized Control Trial. *Conservation Biology*. registered report accepted at Stage 1.

Wilebore, B., Voors, M., Bulte, E. H., Coomes, D., & Kontoleon, A. (2019). Unconditional transfers and tropical forest conservation: Evidence from a randomized control trial in Sierra Leone. *Amer. J. Agr. Econ., 101*(3), 894–918. https://doi.org/10.1093/ajae/aay105.