Operationalizing integrated landscape approaches in the tropics

Editors
James Reed
Mirjam Ros-Tonen
Terry Sunderland

5

Theories of change and monitoring and evaluation types for landscape approaches*

Colas Chervier, Marie-Gabrielle Piketty and James Reed

*Elements of this chapter have previously been published in the article 'A tentative theory of change to evaluate jurisdictional approaches to reduced deforestation.' (Chervier et al. 2020).

Introduction

As highlighted in previous chapters, landscapes and sub-national jurisdictions such as states, provinces, municipalities or districts are increasingly recognized as strategic levels of governance for climate and sustainable development action. Such action, characterized as landscape approaches supposedly lead to improved effectiveness in finding locally acceptable compromises in terms of environmental, social and economic outcomes through broad stakeholder engagement and removing misalignment and inconsistencies between interests regulations and incentives (DeFries and Rosenzweig 2010; Reed et al. 2016). Efforts to develop and implement landscape approaches have been underway for more than three decades (Noss 1983) and recent reviews have identified over 500 examples worldwide (Estrada-Carmona et al. 2014; García-Martín et al. 2016; Milder et al. 2014; Zanzanaini et al. 2017). However despite increasing popularity amongst scientists and practitioners, there remains a lack of robust empirical data documenting their effectiveness in delivering environmental, social and economic outcomes (Boyd et al. 2018; Reed et al. 2017; Sayer et al. 2017). We postulate that this is partly due to the lack of appropriate or standardized evaluation approaches which therefore inhibits robust assessment of interventions over time or comparison of projects across different ecological and/or socio-economic contexts.

This, and the following chapter, are aimed at providing guidance for landscape approach proponents to make informed decisions about the evaluation approach and methods they will use. To do so, this chapter provides an overview of the main types of evaluation that are likely to be needed in a diversity of cases. Therefore, this guide is not meant to be prescriptive but rather to be used as a decision-making tool for practitioners. The following chapter then documents a selection of methods that broadly correspond to the evaluation approaches. Neither chapter is meant to equip landscape approach promoters with all the necessary knowledge to implement each method but rather to provide enough information about pros and cons and above all the domain of applicability of each approach or tool so that they are able to make decisions about which one to use. Implementing these approaches and tools can be then undertaken in collaboration with an evaluation expert.

5.1 Main types of landscape approach evaluation: comparison and complementarity

Before designing any data collection plan for monitoring and evaluation purposes, landscape approach managers and/or promoters need to make decisions regarding the evaluation question they target and subsequently the type of evaluation they want to use. In this section, we identify three suitable types of evaluation for landscape approaches and show that each of them emphasizes a particular evaluation question.

The first type of evaluation corresponds to performance monitoring tools, which consider the evolution of key variables and compares the level reached by each of them relative to a standard/target value. Thus, the evaluation question deals with the best indicators and corresponding targets to be defined in order to reflect the sustainability of a landscape/jurisdiction. Most of the available tools focus on quite long-term outcomes (e.g. IDH-the sustainable trade initiative's Verified Sourcing Areas[1] and European Forest Institute's Terpercaya initiative[2]) but some of them also deal with intermediary outcomes and conditions/determinants for sustainability (e.g. the landscapes rating tool developed by the Certification of Capability in Business Analysis™ - CBBA®)[3]. These tools provide relevant information to consumers, traders and investors willing to purchase or invest in sustainably produced commodities or to make green investments. This kind of information can also be used as criteria to target incentives, such as fiscal transfers, aimed at accelerating transitions towards sustainability. A clear advantage of this type of evaluation is that it is relatively straightforward in terms of design, once the spatial boundaries of the landscape have been defined. However, these tools are not helpful to answer questions such as how or why the value of these indicators change.

1 https://www.idhsustainabletrade.com/landscapes/verified-sourcing-areas/.

2 http://www.euredd.efi.int/documents/15552/460846/The+Terpercaya+Initiative+Briefing+6_WEB.pdf/.

3 https://www.iiba.org/certification/iiba-certifications/ccba/.

The two other types of evaluation understand effectiveness as the extent to which a given intervention or a group of interventions achieve(s) its/their intended outcomes. These approaches place significant emphasis on causality between purposive actions and outcomes, and thus apply to situations where changes in outcomes are intentional, i.e. the result of interventions. These approaches are particularly useful for decision-makers who are interested in improving current models of landscape approaches, and possibly replicate and scale them up (as is often the case with contemporary landscape approaches).

The second type of evaluation is usually called *impact evaluation*. It aims to assess the additional impact of landscape approaches or conservation policies targeting a landscape on generally rather long-term outcomes (e.g. deforestation, well-being, revenues, etc.). In other words, it is helpful to understand if the intervention did make a difference over the medium/long term. This type of approach generally relies on counterfactual methods, which compare the outcomes of a policy in intervention sites with outcomes in "control" sites with similar observable characteristics. Control sites are selected to provide credible information on what would have happened to intervention sites if they had not received the intervention (which is the counterfactual outcome). Impact evaluation has been applied to landscape-scale analysis, particularly focused on assessing deforestation, following different methodologies such as the matching and synthetic control method (Cisneros et al. 2015; Sills et al. 2015; see also Table 5.1). Such counterfactual approaches are considered to be statistically robust, although their capacity to provide evidence for the underlying mechanism and impact pathways is often presented as limited. Using heterogeneity analysis (differential impact according to initial conditions) or assessing the impact of the intervention on intermediary outcomes have proved particularly informative (Chervier et al. 2017; Cisneros et al. 2015; Hanauer and Canavire-Bacarreza, 2015). It is worth noting, however, that finding appropriate counterfactuals for such evaluation objects implemented at large (landscape) scale remains difficult, particularly because of the small pool of potential controls (Roopsind et al. 2019; Sayer et al. 2017).

The third type of evaluation puts emphasis on why an intervention works and in what context. An example of such *evaluation of processes* is the evaluation of underlying processes of jurisdictional REDD+ with the aim to identify challenges, pitfalls and in turn improve outcomes (Ravikumar et al. 2015; Rodriguez-Ward et al. 2018; Sanders et al. 2017). These studies provide evidence for the influence of various governance interventions on landscape approach outcomes. Evidence for causal mechanisms is based on the triangulation of various data sources, including in particular semi-structured interviews with a large number of stakeholders and relevant secondary data. However, empirical knowledge about how landscape approaches work in practice and under what conditions remain scarce.

Figure 5.1 illustrates the differences between these three types of evaluation, but also demonstrates the potential for complementarity by combining them. It shows that each of them tends to focus more intensively on a specific aspect of the evaluation (in bold red lines). While monitoring tools would focus on the definition of indicators and monitoring their change, impact evaluation would focus more specifically on causality

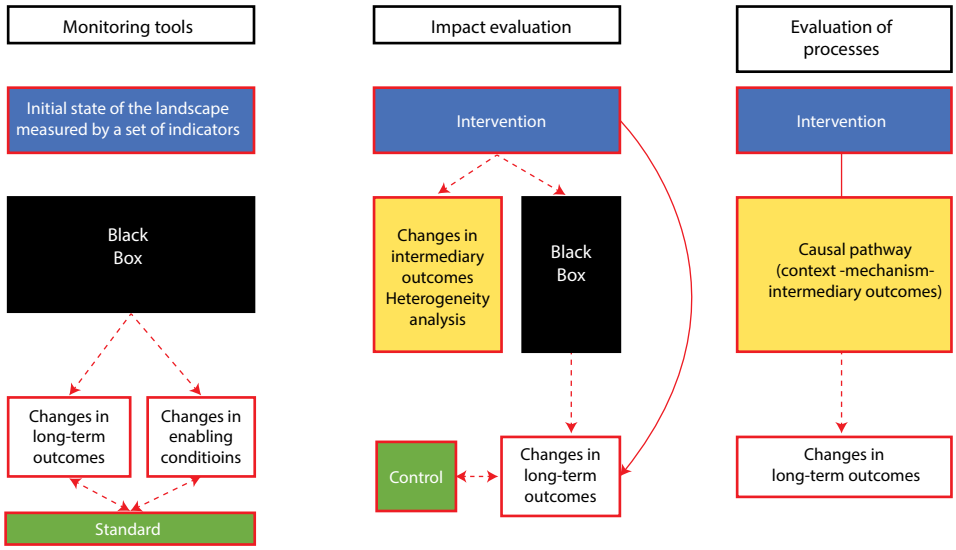

Figure 5.1 Comparison of the three types of evaluation.

between an intervention and long-term outcomes and evaluation of governance processes on disentangling causal pathways. The suggestion is thus to combine these types of evaluation in order to have a comprehensive and robust evaluation of landscape approaches.

5.2 Building a theory of change for the evaluation of landscape approaches

There is a wide consensus in the evaluation literature regarding the importance of using theories of change to evaluate interventions (Astbury and Leeuw 2010; Leeuw and Donaldson 2015), particularly for complex interventions (Rey et al. 2012). A theory of change in evaluation generally articulates causal linkages between interventions and their desired effects. The use of theories of change improves all three types of evaluation (Margoluis et al. 2013; Miteva et al. 2012). For example, the use of theories in impact evaluation would improve the definition of the counterfactual and would help generalize the results. In this section, we present four key steps that are necessary to build and then use a theory of change for evaluation purpose.

5.2.1 Design a causal chain

A theory of change describes the causal sequences of effects associated with the implementation of landscape approaches. Several studies and reports attempt to define generic causal chains for landscape or jurisdictional approaches based on extended empirical knowledge gained from studying and implementing them (Hovani et al. 2018; Sayer et al. 2017). In particular, The Nature Conservancy produced a generic Theory of Change for jurisdictional programs that clearly articulates intermediary outcomes

showing "how improved cross-sector collaboration can translate into improved landscape performance" and define enabling and hindering conditions for a number of causal links (Hovani et al. 2018; Leeuw and Donaldson 2015).

There is a need to extend this work and we suggest that doing so would require combining two types of theories (Leeuw et al. 2015). Participatory methods are generally encouraged to build a causal chain because they enable implementing agents to tailor the design to the specificities of local conditions (Barret et al. 2018; Qiu et al. 2018). The resulting "bottom-up theories" allow capturing stakeholders' and practitioners' perception of how things are supposed to unfold as a result of the implementation of landscape approaches. Evaluation experts argue that it is useful to combine bottom-up theories with knowledge from formal research and theorizing in order to draft meaningful theories of change. It allows putting into perspective the subjective insights provided by the professionals involved in these interventions and incorporating a broader knowledge base.

Figure 5.2 shows an example of a generic causal chain for a landscape approach that is based on a review of the empirical and grey literature on environmental governance theories and landscape and jurisdictional approaches.

5.2.2 Identify the evaluation object, the intervention

A theory of change should clearly identify the intervention(s) that are supposed to bring about change. In the context of landscape approaches, the intervention dimension is broadly defined as: a range of multi-stakeholder interventions, initiatives, programs applied at sub-national/landscape levels in order to achieve lasting, landscape-wide improvements to natural resource management by catalyzing collaborative action of a group of stakeholders working with local government to institutionalize improved land-use governance and practice (Hovani et al. 2018; Paoli et al. 2016). Landscape approach interventions typically focus on strengthening local capacity and governance by building actor networks, improving accountability, investing in knowledge development, etc. (see Chapter 1).

We suggest that, in any landscape engaged in a landscape approach, such coherent systems of interventions can be identified. To be coherent, these actions and changes must have been designed in relation to one another, i.e. as contributions to the achievement of a common goal or a set of interrelated sustainability goals. This implies that these goals have been formally conceptualized and can thus be identified.

Identifying the intervention dimension of landscape approaches is particularly challenging as landscape and actor complexity (i.e. multiple and often conflicting actors and actions across systems, sectors and scales) make it difficult to define intervention capacities, limits and dimensions. Indeed, landscape approaches may encompass multiple interventions that spread across multiple sectors and institutional scales. Furthermore, they span over long-time periods and are characterized by the emergence of new goals and interventions during the implementation phase (Sayer et al. 2017). A causal chain, even if built retrospectively, is a good way to describe such a system of interventions and justify the coherence between its components (Craig et al. 2008).

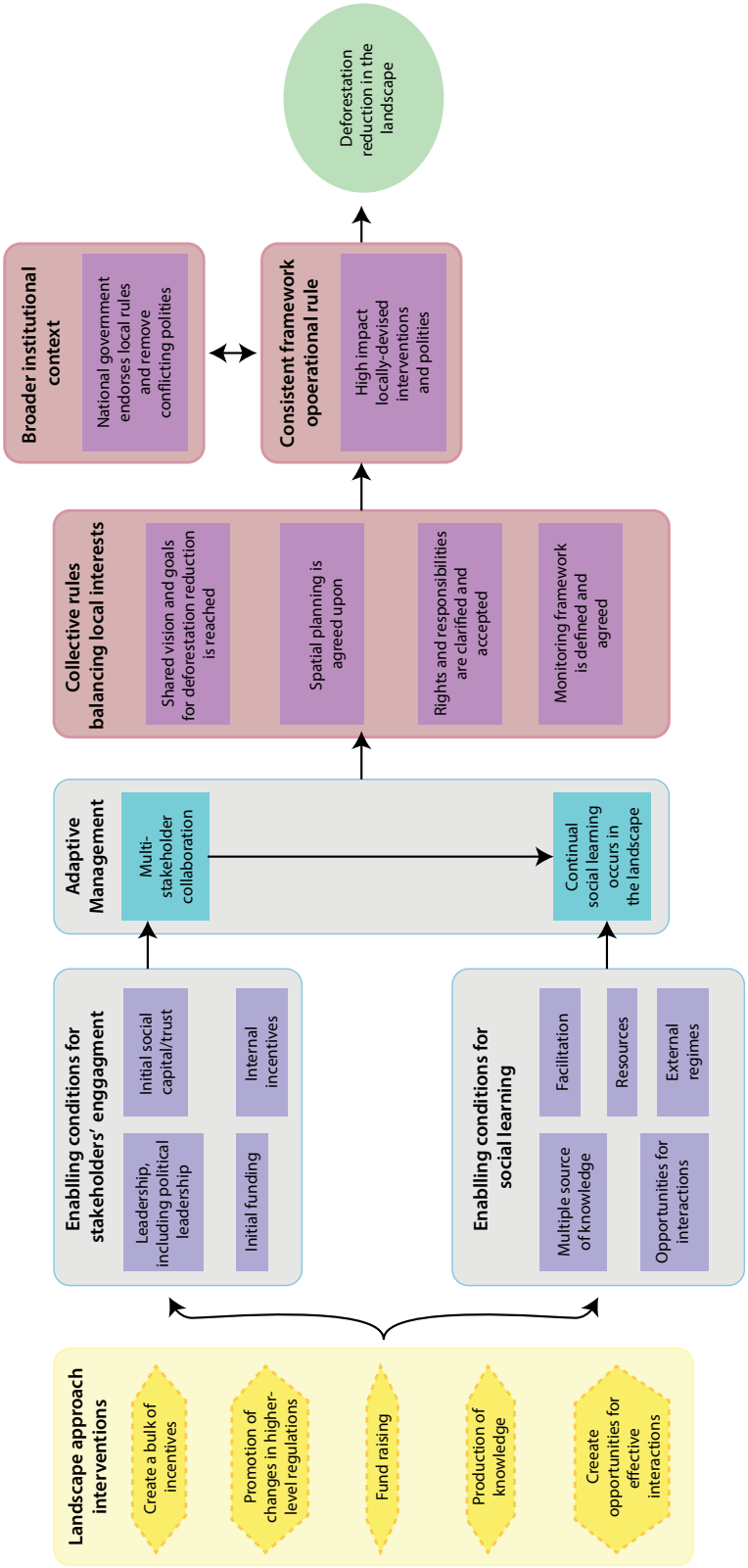

Figure 5.2 Example of a generic causal chain for a landscape approach.
Source: Chervier et al. 2020

5.2.3 Scope the amount of data needed

Data collection requirements can make evaluation relatively expensive, especially in contexts where reliable data is not readily available. It is thus necessary to find compromises between what is desirable in terms of development of knowledge and information, what existing data is already available and accessible, and what additional data collection is feasible provided the limited resources generally allocated to evaluation.

Program theories are generally composed of a significant number of links and boxes and would thus require a gigantic amount of data to be informed exhaustively. Landscape approach leaders or promoters will generally need to prioritize what indicators and what causal links are the most important to be informed. Doing so will narrow down the evaluation (causal links to be informed) and so the monitoring (variables to be monitored) needs. However, careful consideration needs to be given to the selection process such that the indicators selected can adequately capture the full range (or as close to the full range) of potential intervention impacts (Agol et al. 2014).

Besides, the type of causal inference method chosen (see Section 3.4) might necessitate collecting data outside the target landscape. This would be the case if landscape approach promoters want to use counterfactual methods to demonstrate the impact of landscape approaches. Such collection of data in 'control' sites can be viewed as a waste of resources. However, with-without comparison remains the least biased way to measure impact. Adjusting the design of outcome variables and covariates so that they can easily be extracted from available data sources such as remote sensing and census data can help pull the costs down.

5.2.4 Choose the appropriate method(s) for causal inference

Theories of change help making hypotheses about causal linkages associated with the effects of landscape approaches. In turn, landscape approach managers and promoters might be interested in providing evidence for these causal links. They would then face decisions regarding which method to use.

Quasi-experimental design is often viewed as the best possible option for providing evidence for causal links, when experimental design is not possible, which will typically be the case for landscape approaches. There are plenty of quasi-experimental methods available, including matching, difference in differences, instrumental variable, discontinuity analysis, and the synthetic control method (EC Evalset Sourcebook[4]; Sills et al. 2015; see Table 5.1 for a selection). Selecting the most appropriate method will depend on the evaluation setting and more specifically on the availability of appropriate data and the number of treated units (landscapes where the landscape approach is implemented) and control units (landscapes not engaged in a landscape approach). Table 5.1 summarizes three counterfactual methods that are mostly used to evaluate landscape or jurisdictional approaches.

4 https://ec.europa.eu/regional_policy/sources/docgener/evaluation/guide/evaluation_sourcebook.doc

Table 5.1 The most used counterfactual methods.

Matching	Matching is based on the possibility of observing all the characteristics of both treated and control units that influence the probability for a unit to be selected for the intervention and to display higher or lower levels for the outcomes of interest (e.g. deforestation). It is also based on the possibility to find a group of control units that "look alike" the group of treated units along these characteristics, using metrics such as the propensity score (this is the matching). For this reason, matching generally requires quite a high number of treated and control units in order to reach a satisfactory balance. Once the matching is performed, the effect of the intervention is identified by the difference in outcomes between treated and control units, under the assumption that matching has also eliminated differences between treated and control units that may explain differences in outcomes not attributable to the intervention. This assumption cannot be tested: it becomes more credible as more and more characteristics related to the selection process are observable.
Difference-in-Difference	Difference-in-difference or double differencing is based on the precondition that outcome data are available for treated and control units, both before and after the intervention. Effects are obtained by subtracting the pre-intervention difference in outcomes between beneficiaries and non-beneficiaries from the post-intervention difference. The core assumption is that the differences between the treatment and comparison groups that affect either selection into the intervention or the outcome of interest are constant in time. The result of the double difference can be interpreted as a causal effect only if the pre-post trend for control units is a good approximation for the (counterfactual) trend among treated units. The plausibility of this assumption can be tested if more periods of pre-intervention data are available.
Synthetic control method	The Synthetic Control Method (SCM) defines 'similarity' between treated and control units based on both observed characteristics and historical outcomes. SCM creates a counterfactual group obtained as a weighted combination of control units. Weights are assigned to control units so that their combination is as close as possible to the treated unit's and minimize differences in pre-treatment outcomes. SCM assumes that the best fitting weighting of units in terms of pre-treatment outcomes would follow a time trend similar to the treated unit without the intervention. Thus, impact is obtained by comparing a weighted average of control units' outcomes to the outcomes over time in the treated unit. Although this method is useful when the number of treated unit is small, the difficulty in identifying a satisfactory control group arises when the treatment affects large units like regions or countries for which a limited number of untreated units are available.

Source: Drawing from the EC Evalset Sourcebook and Sills et al. 2015

It can also happen that quantitative methods are not applicable (the local context does not create the right empirical setting, poor availability of data and high cost of new data collection). In this case, qualitative methods for causal inference such as process tracing can be used.

> "Process tracing involves the examination of "diagnostic" pieces of evidence within a case that contribute to supporting or overturning alternative explanatory hypotheses. A central concern is with sequences and mechanisms in the unfolding of hypothesized causal processes. The researcher looks for the observable implications of hypothesized explanations, often examining at a finer level of detail or a lower level of analysis than that initially posited in the relevant theory. The goal is to establish whether the events or processes within the case fit those predicted by alternative explanations."
> (Bennett 2010, p. 208)

Conclusion

In this chapter, we highlighted the importance of combining various types of evaluation of landscape approaches. Practitioners should not only consider monitoring tools, but also assess the effectiveness and pathways to effectiveness of landscape approaches, using process analysis or impact evaluation. This will allow practitioners and policymakers to draw important lessons for the replication of landscape approaches in other contexts.

We believe that the use of theories of change would improve the quality and credibility of all types landscape approach evaluation. In order to allow for generalization, replication or adaptation of early landscape approach experience, such theories of change should integrate local knowledge and perspectives with scientific knowledge and theories. Theories of change are also a useful tool to identify the intervention dimension of landscape approaches and to make decisions about the methods and data to be used (see next chapter).

References

Agol D Latawiec AE and Strassburg BB. 2014. Evaluating impacts of development and conservation projects using sustainability indicators: Opportunities and challenges. *Environmental Impact Assessment Review* 48, 1–9.

Astbury B and Leeuw FL. 2010. Unpacking black boxes: Mechanisms and theory building in evaluation. *American Journal of Evaluation* 31, 363–381. https://doi.org/10.1177/1098214010371972

Barret D, Blundo Canto G, Dabat M-H, Devaux-Spatarakis A, Faure G, Hainzelin E, Mathé S, Temple L, Toillier A and Triomphe B. 2018. ImpresS methodological guide. Methodological guide to ex post evaluation of agricultural research in developing countries. CIRAD.

Bennett A. 2010. Process tracing and causal inference. *In* Brady HE and Collier D, eds. Rethinking Social Inquiry: Diverse Tools, Shared Standards. Lanham MD: Rowman & Littlefield Publishers.

Boyd W, Stickler C, Duchelle AE, Seymour F, Nepstad D, Bahar NHA and Rodriguez-Ward D. 2018. Jurisdictional approaches to REDD+ and low emissions development: Progress and prospects. Working Paper. Washington DC: World Resources Institute.

Chervier C, Le Velly G and Ezzine-de-Blas D, 2017. When the implementation of payments for biodiversity conservation leads to motivation crowding-out: A case study from the cardamoms forests, Cambodia. *Ecological Economics* 156: 499-510.

Chervier C, Picketty M-G and Reed J. 2020. A tentative theory of change to assess the effectiveness of jurisdictional approaches to reduced deforestation. *Frontiers in Forests and Global Change*. https://doi.org/10.3389/ffgc.2020.498151.

Cisneros E, Zhou SL and Börner J. 2015. Naming and shaming for conservation: Evidence from the Brazilian Amazon. PLOS ONE 10, e0136402. https://doi.org/10.1371/journal.pone.0136402.

Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I and Petticrew M. 2008. Developing and evaluating complex interventions: The new Medical Research Council guidance. *BMJ* 337, a1655.

DeFries R and Rosenzweig C. 2010. Toward a whole-landscape approach for sustainable land use in the tropics. *Proceedings of the National Academy of Sciences* 107, 19627–19632. https://doi.org/10.1073/pnas.1011163107.

Estrada-Carmona N, Hart AK, DeClerck FA, Harvey CA and Milder JC. 2014. Integrated landscape management for agriculture, rural livelihoods, and ecosystem conservation: An assessment of experience from Latin America and the Caribbean. *Landscape and Urban Planning* 129, 1–11.

García-Martín M, Bieling C, Hart A and Plieninger T. 2016. Integrated landscape initiatives in Europe: Multi-sector collaboration in multi-functional landscapes. *Land Use Policy* 58, 43–53.

Hanauer MM, Canavire-Bacarreza G. 2015. Implications of heterogeneous impacts of protected areas on deforestation and poverty. *Philosophical Transactions of the Royal Society B: Biological Sciences* 370(1681), 20140272. https://doi.org/10.1098/rstb.2014.0272

Hovani L, Cortez R, Hartanto H, Thompson I, Fishbein G, Madeira EM and Adams J. 2018. The Role of Jurisdictional Programs in Catalyzing Sustainability Transitions in Tropical Forest Landscapes. Arlington VA: The Nature Conservancy.

Leeuw FL and Donaldson SI. 2015. Theory in evaluation: Reducing confusion and encouraging debate. *Evaluation* 21: 467–480. https://doi.org/10.1177/1356389015607712.

Margoluis R, Stem C, Swaminathan V, Brown M, Johnson A, Placci G, Salafsky N and Tilders I. 2013. Results chains: A tool for conservation action design, management, and evaluation. *Ecology and Society* 18 (3). https://doi.org/10.5751/ES-05610-180322

Milder JC, Hart AK, Dobie P, Minai J and Zaleski C. 2014. Integrated landscape initiatives for African agriculture, development, and conservation: A region-wide assessment. *World Development* 54, 68–80.

Miteva DA, Pattanayak SK and Ferraro J. 2012. Evaluation of biodiversity policy instruments: what works and what doesn't? *Oxford Review of Economic Policy* 28: 69–92. https://doi.org/10.1093/oxrep/grs009.

Noss RF. 1983. A regional landscape approach to maintain diversity. *BioScience* 33: 700–706.

Paoli GD, Palmer B, Schweithelm J, Limberg G and Green L. 2016. Jurisdictional Approaches to Reducing Palm Oil Driven Deforestation in Indonesia. Scoping Study of Design Considerations and Geographic Priorities. Bogor, Indonesia: Daemeter.

Qiu J, Game ET, Tallis H, Olander LP, Glew L, Kagan JS, Kalies EL, Michanowicz D, Phelan J, Polasky S, Reed J, Sills EO, Urban D and Weaver SK. 2018. Evidence-based causal chains for linking health, development, and conservation actions. *BioScience* 68: 182–193. https://doi.org/10.1093/biosci/bix167.

Ravikumar A, Larson A, Duchelle A, Myers R and Tovar JG. 2015. Multilevel governance challenges in transitioning towards a national approach for REDD+: Evidence from 23 subnational REDD+ initiatives. *International Journal of the Commons* 9(2).

Reed J, van Vianen J, Barlow J and Sunderland T. 2017. Have integrated landscape approaches reconciled societal and environmental issues in the tropics? *Land Use Policy* 63: 481–492. https://doi.org/10.1016/j.landusepol.2017.02.021.

Reed J, Van Vianen J, Deakin EL, Barlow J and Sunderland T. 2016. Integrated landscape approaches to managing social and environmental issues in the tropics: learning from the past to guide the future. *Global Change Biology* 22: 2540–2554. https://doi.org/10.1111/gcb.13284.

Rey L, Brousselle A and Dedobbeleer N. 2012. Logic analysis: Testing program theory to better evaluate complex interventions. *Canadian Journal of Program Evaluation* 26(3): 61–89.

Rodriguez-Ward D, Larson AM and Ruesta HG. 2018. Top-down, bottom-up and sideways: The multilayered complexities of multi-level actors shaping forest governance and REDD+ arrangements in Madre de Dios, Peru. *Environmental Management*. 62(1): 98–116. https://doi.org/10.1007/s00267-017-0982-5.

Roopsind A, Sohngen B and Brandt J. 2019. Evidence that a national REDD+ program reduces tree cover loss and carbon emissions in a high forest cover, low deforestation country. *Proceedings of the National Academy of Sciences* 116: 24492–24499. https://doi.org/10.1073/pnas.1904027116

Sanders AJP, Hyldmo HDS, Prasti H RD, Ford RM, Larson AM and Keenan RJ. 2017. Guinea pig or pioneer: Translating global environmental objectives through to local actions in Central Kalimantan, Indonesia's REDD+ pilot province. *Global Environmental Change* 42: 68–81. https://doi.org/10.1016/j.gloenvcha.2016.12.003.

Sayer JA, Margules C, Boedhihartono AK, Sunderland T, Langston JD, Reed J, Riggs R, Buck LE, Campbell BM, Kusters K, et al. 2017. Measuring the effectiveness of landscape approaches to conservation and development. *Sustainability Science* 12: 465–476. https://doi.org/10.1007/s11625-016-0415-z.

Sills EO, Herrera D, Kirkpatrick AJ, Brandão A, Dickson R, Hall S, Pattanayak S, Shoch D, Vedoveto M, Young L and Pfaff A. 2015. Estimating the impacts of local policy innovation: The synthetic control method applied to tropical deforestation. *PLoS ONE* 10, e0132590. https://doi.org/10.1371/journal.pone.0132590.

Zanzanaini C, Trần BT, Singh C, Har, A, Milder J and DeClerck F. 2017. Integrated landscape initiatives for agriculture, livelihoods and ecosystem conservation: An assessment of experiences from South and Southeast Asia. *Landscape and Urban Planning* 165: 11–21. https://doi.org/10.1016/j.landurbplan.2017.03.010.