



Narrative Assessment: A new approach to evaluation of advocacy for development

Evaluation
2018, Vol. 24(4) 400–418
© The Author(s) 2018



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1356389018796021
journals.sagepub.com/home/evi



Margit van Wessel

Wageningen University & Research, The Netherlands

Abstract

This article proposes an approach to monitoring and evaluating advocacy that integrates theory of change and storytelling. This approach, called Narrative Assessment, addresses the feasibility of objectivity and evidence in the complex context of advocacy and proposes an evaluation methodology rooted in alternative conceptualizations of rigour and of evaluator roles. The approach centres on practical judgment and the construction and examination of stories through interaction between advocates and monitoring and evaluation specialists. The article discusses how Narrative Assessment can be useful in the evaluation of advocacy in terms of (1) monitoring and evaluation specialists' orientation toward programmes; (2) the interpretation of outcomes; (3) the assessment of outcome relevance; (4) reflection and learning and (5) the communication of programme results. The approach builds on lessons drawn from evaluation of eight advocacy programmes in international development.

Keywords

advocacy evaluation, international development, narratives, storytelling, Theory of Change

Introduction

Development organizations have increasingly turned to advocacy as a means to achieve change. A common argument for this is that seeking change through projects may lead to results, but these results often do not touch upon the fundamental conditions that shape lives, such as sustainable ecosystems management, land rights, cultural understandings of sexual and reproductive health and rights, or market relations. Advocacy for development often addresses such structural matters, seeking to transform the legal, political and social conditions that shape development (Barrett et al., 2016). These advocacy activities are often funded

Corresponding author:

Dr Margit van Wessel, Wageningen University & Research, Sub-department Communication, Technology and Philosophy: Centre for Integrative Development, Chair group Strategic Communication (COM), P.O. 8130, 6700 EW, Wageningen, The Netherlands.

Email: margit.vanwessel@wur.nl

by states, private donors and foundations. To learn and adjust, and to be accountable to the institutions, organizations and citizens supporting these activities, advocacy needs to be monitored and evaluated.

However, monitoring and evaluating advocacy is notoriously difficult. Causal relations between actions and results are difficult to establish, achievements tend to be largely invisible and hard to trace, and influencing often takes place behind closed doors. Those influenced by advocacy may not be ready or available to concede being influenced (or not) by specific actors, actions or events. Moreover, intervention effects are likely found among numerous causal strands. The objects of advocacy – policymakers, publics and private sector actors – are moving targets subject to numerous influences. Much of the difficulty of assessing advocacy has therefore been attributed to the complexity of the change processes where advocacy seeks to contribute (Chapman and Wameyo, 2001; Jones, 2011). Key elements of this complexity include the nonlinearity of change and the emergent nature of outcomes (Rogers, 2008; Arensman et al., 2018). Drawing upon Arensman et al. (2015), I define advocacy for development as follows:

[A] wide range of activities conducted to influence decision makers at different levels, with the overall aim of combatting the structural causes of poverty and injustice. This definition follows the widely held belief that NGO advocacy is a tool to fight the causes of poverty or injustice and influence structural change, aiming to change social, political and policy structures and to challenge power structures. This concept of advocacy goes beyond influencing policy and aims for sustainable changes in public and political contexts. This work includes awareness raising, legal actions and public education, as well as building networks, relationships and capacity. (Arensman et al., 2015: 6; see also Morariu and Brennan, 2009: 100)

Advocacy for development can be seen as an extreme case in terms of advocacy evaluation challenges. In addition to the complexity issues facing advocacy evaluation in general, advocacy for development is often multi-level (with differentiated linkages among levels), multi-sited (with differentiated linkages across sites) and multi-actor (with differentiated engagements, understandings and roles in programmes, and with multiple organizational structures, capacities and accountability relations).

This article contributes to an ongoing discussion about how evaluation can be geared to the nature of interventions and the complexity of the relevant contexts and change processes (Arensman and Van Wessel, 2017; Arensman et al., 2018; Bamberger et al., 2016; Rogers, 2008; Stern et al., 2012; Walton, 2016). Building on a large-scale evaluation of advocacy for development, I propose an approach to advocacy monitoring and evaluation that seeks to do justice to the nature of advocacy work in terms of the dynamics between advocates and the environments where they operate. This approach integrates theory of change and storytelling, drawing on narrative analysis and narrative inquiry theory. First, I argue that theory of change, now commonly employed in international development and other contexts (in inception, grantmaking, dialogue and learning), can also play a significant role in advocacy monitoring and evaluation. Here, I build on the possibilities of theory of change processes advanced, envisaged and enacted in many publications, funding structures and everyday organizational practices in recent years. Second, I argue for the use of theory of change in a monitoring and evaluation approach where stories play a central role. In this approach, which I call Narrative Assessment, credible and insightful stories of advocacy strategy, process, achievement and

failure result from co-creation and examination between advocates and monitoring and evaluation (M&E) specialists. The terms ‘narrative’ and ‘assessment’ denote the narrative method (of co-constructing and examining stories) by which advocacy is assessed.

I present five ways in which Narrative Assessment can be useful in monitoring and evaluating advocacy: (1) M&E specialists’ orientation toward programmes; (2) interpreting outcomes; (3) assessing outcome relevance; (4) engaging in reflection and learning and (5) communicating programme results. The analysis builds on experiences from an evaluation of eight advocacy programmes supported by the Netherlands Ministry of Foreign Affairs. I chart how the evaluation team employed theories of change, present the collective lessons learned and present Narrative Assessment as a new approach to advocacy monitoring and evaluation that builds on these lessons and offers solutions to advocacy monitoring and evaluation challenges.

Theoretical framework

Establishing effectiveness

Demonstrating effectiveness is a key challenge for organizations carrying out advocacy programmes. For the international development sector, this challenge is institutionally ingrained. Organizations conducting advocacy activities in the field of development, dependent on public and private donors for funding, are under constant pressure to be effective and provide evidence of this effectiveness. Effectiveness is thus a reality of development work, inspiring a great deal of interaction and investment in the name of accountability and an orientation toward results (Riddell, 2014). Improvement of constituencies’ lives is the ultimate yardstick.

Effectiveness is usually considered in terms of achieving predefined objectives, as required in relations between donors and recipients (OECD, 2005). Approaches to evaluation conceptualized in terms of ‘value for money’ or ‘results-based financing’ are widely adopted by funders (see e.g. DFID, 2011 and World Bank, 2013 in Arensman et al., 2018). Criteria for evaluating development assistance and aid effectiveness focus on measurable results (Riddell, 2014). However, drawing meaningful conclusions about achievements based on available measurements is generally difficult because they often fall short in providing insight into what is achieved and how (Riddell, 2014). Whether development can be captured in measurement is broadly debated in the development field, and such questions are perhaps even more at the forefront for the evaluation of advocacy for development. Firstly, monitoring protocols asking for (often numerical) accounting of outcome indicators do render account, but in a way that conveys little of the actual meaning of the achievements. Typical outcome indicators are abstractions, for example, ‘the number of elected officials who publicly support the campaign’ (see e.g. the Urban Institute, n.d.). Qualitative indicators can also be employed (e.g. strengthened organizational capacity, support base or policy). Such indicators are sometimes used to ‘measure’ the achievements of advocacy programmes (see e.g. Coffman, 2009). However, outcomes like increased credibility of an organization’s positions in the eyes of policymakers, increased public awareness about an issue or increased support for a certain policy position among policymakers are difficult to measure, as resources to carry out research are mostly limited, and information inaccessible. Secondly, although advocacy achievements can be measured in some cases, or to some degree, interpretation is necessary to make sense of them.

For example, the number of elected officials publicly supporting a campaign does not tell us whether the campaign has changed their viewpoints. For many reasons, interpretations of influence can be flawed (Lowery, 2013). Thirdly, and relatedly, for proper interpretation, one would need to assess how meaningful a contribution to change a certain achieved outcome might be. Qualitative approaches to measuring change that take questions of interpretation at heart include 'Most Significant Change' (Dart and Davies, 2005), a participatory assessment of the changes a programme has delivered; the 'Bellwether' approach, based on the idea that highly informed external resource persons can be used to determine the influence of an intervention (Coffman, 2009; Tsui et al., 2014); and Outcome Harvesting (Wilson-Grau and Britt, 2012). However, these approaches are highly intervention-centred, paying little attention to the context in which a particular intervention is intended to make a difference. This easily leads to a decontextualized and thus incomprehensible presentation of a programme and its advocacy results. Relating to context when considering effectiveness is crucial, considering that, in advocacy, one is often faced with complications that undermine one's efforts, including powerful opponents, disinterested targets and shifting political winds.

A fourth issue that needs attention is that existing methods appear to work from the assumption that evidence of achievements and contribution to these will be available and accessible. This is often not the case in advocacy, which typically involves many small steps toward something as substantial as a policy change. Consider, for example, outcomes like enhanced credibility with a target, a target warming to a certain proposal, or input in a stakeholder meeting contributing to a shift in understanding. However, this issue is hardly recognized in current literature on advocacy evaluation methodology (see e.g. Gardner and Brindis, 2017, for an otherwise helpful overview). A fifth issue is that current methods do not consider that advocacy results often lack direct impact or that the true significance of an advocacy outcome can often only be interpreted in light of a theory of change envisioning changes *in the future*.

Theory of change

Complexity has been widely acknowledged as an issue with which both development and development evaluation need to engage. In recent years, theory of change has been hailed as an approach to planning, monitoring and evaluating in the field of international development that can do more justice to complexity. Moving away from the constrictiveness of log frames, theory of change is seen as an approach funders, practitioners and M&E specialists can use to account for the complex nature of the change processes to which development interventions seek to contribute (Prinsen and Nijhof, 2015). Theory of change allows focus on longer-term impacts rather than only short-term results. Because theories of change make assumptions and pathways of change explicit, a theory of change process facilitates reflection, interaction and adjustment in terms of understandings of how change happens, thereby also helping to bring about critical reflection on how an intervention may actually contribute to change. This approach can take into account how changes caused by other influences create new conditions for programmes. Programmes can be expected to gain in effectiveness through the capacity to adjust in such situations instead of holding on to planned actions (Rogers, 2008). Theory of change also facilitates interaction and adjustment among different programme actors, because differences may be confronted and engaged as the theory of change is deliberated (Arensman et al., 2015; Prinsen and Nijhof, 2015). For these reasons, theory of change is hailed as potentially helpful in articulating assumptions and pathways of change and in encouraging dialogue

and adjustment over time, both as programmes develop and in evaluation (Stein and Valters, 2012). Many reports and practical guidelines engage with theory of change in terms of its potential merits regarding the focus on assumptions and pathways of change, seen as fundamental to devising, understanding and adapting interventions. For evaluation purposes, different usages have been advanced, including the identification of meaningful outcome indicators (Reisman et al., 2007) and the evidence-based analysis of interventions' contributions to change (Mayne, 2012). Currently, funders and organizations commonly require that theories of change be included in proposals and project plans (Prinsen and Nijhof, 2015; Stein and Valters, 2012; Vogel, 2012). This includes proposals and projects focusing on advocacy, such as the Dutch government's current (2016–20) one billion-euro support programme for civil society advocacy and advocacy capacity development in the context of international development (Netherlands Ministry of Foreign Affairs, 2014).

Interestingly, the enthusiastic uptake of theory of change in international development has had little relation to theory-based evaluation, in either conceptualization or rationale. Theory-based evaluation, at least in principle, largely focuses on strengthening the articulation of the supposed mechanisms behind change and on testing theory against evidence obtained over the course of programmes, centring on questions of the validity of theories and evidence (Coryn et al., 2011; Rogers, 2007; Weiss, 1997). There are many forms of theory-based evaluation, sharing the key elements of developing the theory behind a programme, establishing whether what the theory stipulates happens and elucidating the causal mechanisms. In short, theory-based evaluation tests *if* and *how* a programme theory works.

Theory-based evaluation has been strongly contested. A prominent discussion point is the feasibility of testing theory, considering the difficulties of developing a theory to serve such a purpose: Obtaining a validated theory against which to assess programmes is a fundamental problem in theory-based evaluation. This certainly applies to advocacy, where there is little to draw on when it comes to evidence-based strategy development. There are many handbooks and manuals on effective advocacy that appear to work from experience-based knowledge. These commonly chart the importance of capacities like relation-building, fostering credibility, monitoring contexts and acting in a timely fashion. However, there is little research establishing causal relations between strategy and outcome.

Another problem is the feasibility of testing causal explanations (Coryn et al., 2011; Rogers, 2007). Even if we assume a theory is valid, it is difficult to isolate the different causal mechanisms that may be behind a change or to assess their absolute or relative roles (Weiss, 1997). This is true even in relatively simple change processes, let alone in more complex ones. The discussion about theory of change in international development often stresses the approach's usefulness for articulating the mechanisms behind change. However, there is little discussion of testing these mechanisms or problems involved with this testing, as seen in the literature on theory-based evaluation. The few publications focusing on theory of change for advocacy evaluation do not incorporate research on the strengths and weaknesses of theory-based evaluation. Rather, this work takes on the limited and pragmatic ambition of systematically gathering information about achievements instead of mechanisms (Jones, 2011; Reisman et al., 2007). This literature often acknowledges the complexities of advocacy and its evaluation, in particular with regard to gathering information about causal mechanisms. For evaluating advocacy for development, such problems are often important. In many cases, it may be possible to find a sequence of events postulated in a theory of change, but evidence on the mechanisms may be very difficult to establish. For example, a public campaign mobilizing public

opinion on land rights in Africa implicating a multinational company may be followed, a year later, by that company announcing that it will change its behaviour. There may be strong signs that the campaign contributed to this change: A company spokesperson may have responded to the campaign in a press conference, and the changes carried out by the company may be in line with demands made in the campaign. Indications of advocacy achievement do not come much stronger than this, but such evidence does not expose the causal mechanism. During the campaign, many other things likely happened that may have played a role: for example, actions by other organizations, declines in sales, an incriminating news item, behind-the-scenes interactions with the government, or internal corporate deliberations. It is not possible to subject all these things to analysis, leading to conclusive results regarding causation (see Weiss, 1997). Similarly, when such a campaign seems to fail in its objectives (e.g. a targeted company did not change its behaviour), it will often be impossible to establish decisively to what extent this failure is because of the theory being faulty, improper execution of a strategy or an unexpected shift in a company's internal strategizing resulting from other causes, to name some possible factors. In fact, it may even be difficult to call the campaign a failure. While the objectives may not have been reached because the company's behaviour did not change, the legitimacy of the company's actions may have been undermined, although this is unmeasured, contributing to future change by the company, its consumers or governmental actors. Evidence of such achievements is often difficult to obtain. Although evidence can be found for contribution to change in some cases, this kind of evidence will often be patchy and questionable.

All this suggests that the usefulness of theory of change for evaluation may best be understood as constricted because of the difficulties of objective assessment confronting evaluation. It may also suggest that theory of change is most useful in facilitating reflection, dialogue and adjustment. Still, there are reasons to seriously consider theory of change for monitoring and evaluation, beyond what has been presented so far in literature and practice. In my view, letting go of questions about mechanisms is unwarranted. I argue that we can do more with theory of change in advocacy monitoring and evaluation, in particular for complex and complicated fields like advocacy for development, and for investigating mechanisms. Moreover, we need to do this to advance monitoring and evaluation quality. Failing to explore these areas discards important possibilities of theory of change. It is exactly *because of* the difficulties involved with objective assessment that we need to do more with theory of change. I propose we do this from an alternative angle, corresponding to the nature of advocacy for development and other complicated and complex interventions, and their results.

Stories

My proposal for advocacy monitoring and evaluation integrates storytelling into a theory of change approach to evaluation, linking with the rise of theory of change in a well-fitting marriage of ideas. A story is an account of events over time, brought together into a coherent whole. It conveys certain meanings from the standpoint of the narrator. A story is built up from different elements. It contains an element of transformation and presents this transformation as a movement over time. A story contains the actions through which this transformation happens, carried out by characters in a specific setting. These different elements are brought together in a plot, which often involves crises and turning points. Finally, this plot communicates a point to take away from the story (Lawler, 2002).

Narrative Assessment involves the systematic co-construction of stories between advocates and Monitoring and evaluation (M&E) specialists, ensuring that all the key elements are included in stories recounting the past unfolding of programmes. By asking questions, encouraging the construction of elements and connections between these elements, and drawing on past observation and fresh reflection, advocates and M&E specialists build the stories through which the nature and meaning of programme achievements, challenges and disappointments are constituted. By consistently relating what is told to the programme's theory of change, the meaning of the story is systematically established in terms of relevance for wider audiences (colleagues, partners, communities and donors). To clarify, stories are symbolic interpretations of human action in practical, concrete situations. They simultaneously function as explanation, justification and instruction. A plausible story creates order in a shifting, unpredictable and incoherent environment. It makes past actions understandable and creates the conditions for further action. It is often only in light of a theory of change, in some form, that the interconnected sequence of past actions and events can be given meaning and relevance for programmes and their stakeholders, as a journey over time, with twists and turns, setbacks and advances. It is only in storied form that this journey can be envisaged and proposed to those same stakeholders in a way that does justice to its complicated and complex reality. By focusing on sequences of developments and on how actors, actions and contexts influence these developments, stories can shed light on mechanisms as they are observed and interpreted by advocate as actors in the middle of the action, making these explicit and laying bare the unfolding of developments. Whereas causal mechanisms are often impossible to establish with certainty, we can co-construct and critically assess the plausibility of advocates' stories of how what they did or did not make a difference, and why. In particular, such stories can clarify how targets' attitudes, understandings or behaviours can be understood, at least partially, as responses to advocacy. This is accomplished through the description, observation and analysis of plausible links between advocacy and these changes. This kind of analysis can also clarify how failures to contribute to such responses came about. The temporal dimension of stories, doing justice to the unfolding of a programme and targets' responses over time, is important (Clandinin and Connelly, 2000: 19–20). Stories, compared with methods seeking more 'objective' truths centring on factual elements like outcomes and evidence, are therefore better able to capture the *true* dynamics of advocacy. This is not to disparage evidence. Evidence is of great value, but the available evidence may not be the most significant information to draw on when monitoring and evaluating advocacy.

However, stories also create order by resisting contradictory information. They provide meaning and direction by countering the threat of becoming overwhelmed and taken over by disorder and indeterminacy (Wagenaar, 2014: 215–16). A story is an account of what happened that idealizes and cleans things up, highlighting certain elements, minimizing others, and imposing causation and meaning (Wagenaar, 2014: 210) that is likely incomplete, ambiguous and impossible to verify fully. Stories will thus offer stylized accounts of the meaning of actions. This clearly invites a bias that Narrative Assessment must address systematically, as will be elaborated in the section entitled 'questions of quality', below.

Stories are already part of evaluation methodology in some contexts. In development, some have argued that stories are potentially helpful in learning about change in complex contexts (Conlin and Stirrat, 2008). Several existing methods for monitoring and/or evaluation work with stories, such as 'Performance Story Reporting' (see e.g. VanClay, 2015) and 'Most Significant Change' (Dart and Davies, 2005). However, existing methods focus mostly on

capturing programme impact through participants' or beneficiaries' stories (e.g. Costantino and Greene, 2003). For advocacy, this is often problematic, as many achievements are removed from impact. Such methods also do not specifically address the potential value of the stories of those carrying out the projects: the specific insider knowledge on advocacy processes and the nature of achievements that advocates can offer. Finally, no existing method capitalizes on the communicative potential of stories beyond the information they offer on effectiveness.

The MFS II Joint Evaluation of International Lobbying and Advocacy

This article draws on a joint evaluation involving eight advocacy programmes executed by eight NGO alliances working on a range of themes and issues within the domains of sustainable livelihoods and economic justice; sexual and reproductive health and rights; and protection, human security and conflict prevention (Arensman et al., 2015). I was the project leader and an evaluator for this evaluation. The evaluated alliances were funded through the Netherlands government's co-financing development cooperation policy programme, known as 'MFS II'.¹ The evaluation was a requirement of the Dutch government, and its setup was partly negotiated among the parties involved. The evaluation was carried out by the Social Sciences Group of Wageningen University & Research and external consultants from the global South and North. The NGOs involved were Dutch NGOs (including, among many others, Oxfam Novib, Cordaid, International Union for Conservation of Nature and Natural Resources Netherlands (IUCN NL), PAX, Both Ends, Hivos, Friends of the Earth Netherlands (FOEI NL), Transnational Institute (TNI) and Centre for Research on Multinational Corporations (SOMO) and hundreds of partner organizations working with them across the globe through partnerships, coalitions and networks.

The evaluation team was given a primarily accountability-centred framework, largely devised by the Netherlands Ministry of Foreign Affairs. The evaluation questions were as follows: What changes did alliances contribute to? What contribution did alliances make to these changes? How relevant were these changes? How efficient were the efforts made by the alliances? What explains the findings? Three priority result areas were predefined: agenda setting, policy influencing and changes in practice. The terms of reference allowed the team to experiment and innovate in terms of approach, conceptualization and methodology.

In our evaluation, we experimented with theory of change thinking as an approach to advocacy monitoring and evaluation, seeking to do justice to the complexity and complicatedness found in this type of work. Partly, what I chart below is an account of this experimentation and the lessons we learned. We made a start with a narrative approach, although it did not centre explicitly on narration, but we did not take this approach further. In this article, I expand this previous work to build a future vision for advocacy monitoring and evaluation. I discuss five ways in which a narrative approach, interpreting past action in light of theory of change in storied form, can contribute to objectives of monitoring and evaluation. Finally, I go into questions of quality and how these can be addressed.

Orientation

The evaluation team decided to make theory of change a fundamental starting point for the evaluation because of the complexity of the processes of change to which the advocacy programmes

sought to contribute and the complicated nature of the interventions. This complexity lay in the unpredictable and emergent nature of change. Complexity arises in systems, which we can see as constellations of interconnected and interacting elements. Important dynamics in systems include emergence (of system characteristics and behaviours) and feedback processes inhibiting or promoting change. Systems are said to self-organize, with agents adapting and co-evolving during their interaction. Change in systems takes place through nonlinear developments and is often unpredictable. Complexity, in this context, comes in through the dynamism of the situations to which the NGOs related in their advocacy (Arensman et al., 2015). In their advocacy, the evaluated organizations had to engage with these complexities, which made their work only partly plannable and highly adaptive in nature. Advocates need to manoeuvre through complex contexts, acting on practical judgment about where possibilities for change lie and what strategy can effectively influence certain targets. Although there are some shared understandings about how to act effectively as an advocate (e.g. based on an analysis of the political context, acting in a timely fashion, engaging with policymakers' needs and understandings, providing credible evidence and usable ideas, and working in coalitions), none of these understandings is sufficient to ensure success. There is also no agreed view on what works under what conditions. As a team, we needed a way to orient ourselves to how advocates approached and manoeuvred in the situations they faced, strategizing not based on sure-fire knowledge of cause and effect, but rather on practical judgment of what could work in a certain situation. For the team, theory of change was a frame that could guide the evaluation process in a way that provided the flexibility needed to do justice to the complexity of the relevant change processes and the complicated nature of the interventions (Arensman et al., 2015).

Following Vogel's (2012) conceptualization of theory of change, we sought to reconstruct the NGOs' theories of change from an outsider's position. Through interviews with NGO staff members and document analysis, we sought to reconstruct, verbally and visually, the theoretical thinking behind the programmes. How did NGOs conceive of context, including social, political, institutional and environmental conditions, the current state of the issue the project sought to influence, and other actors able to influence the desired long-term change? What was the anticipated process/sequence of change, both overall and for the alliance's contribution to this change? What were the assumptions regarding how change happens? In this reconstruction, we sought to remain open to divergence among partners, change over time, gaps and different forms of theorizing (Arensman et al., 2015). For the team, knowing the alliances' theories of change made it possible to understand how, for the NGOs, interventions, targets and objectives made sense. It also made it possible to maintain this understanding as programmes developed and unfolded in different and partly unplanned directions.

Interpretation of outcomes

The team took also theory of change as fundamental for interpreting programme results. The evaluation centred on outcomes, which we approached as changes – intended or unintended – in the policies, practices, behaviour, relationships, actions, activities or mind-set of an individual, group, community, organization or institution. Outcomes were identified through a combination of document review and interviews with alliance staff and with external resource persons. The existence of these outcomes was then verified by further (external) document review, interviews with external informants, and additional discussion and reflection with alliance staff (Arensman et al., 2015).

We aimed to achieve two things. First, starting with outcomes, we worked ‘backward’ to reconstruct the pathways of change, developing a plausible case that the outcomes could be understood as part of a process of change to which the NGOs’ programmes contributed. These outcomes were described and interpreted with consideration of how they related to the NGOs’ theories of change. We worked with outcome indicators only to identify patterns in different types of achievement. This approach, centring on outcomes, is similar to Outcome Harvesting, an evaluation method that collects evidence of achievements and works backward to determine whether and how the project or intervention contributed to the change (Wilson-Grau and Britt, 2012). Outcome Harvesting is flexible, accommodating the complexity of change and adaptation in interventions. However, Outcome Harvesting centres on identifying outcomes and on evidence, involves limited attention to theory and does not address challenges involved in gathering evidence for advocacy monitoring and evaluation. In our evaluation, evidence was sought, but putting sense-making at the centre was an innovative step. This sense-making involved theorizing, narratively constructing the ways in which outcomes could be understood and described from the advocates’ theories of change. Advocacy is often primarily based on practical judgment, in complex situations with many unknowns. It cannot be otherwise, and the nature and quality of that judgment, incorporating an analysis of the chances of success and balancing opportunity, cost and risk (Wagenaar, 2014: 218), can only be conveyed through stories. Stories help to provide a view of the reality of advocacy work and its results in a way that outcome descriptions and ‘evidence’ cannot.

For example, the evaluation of a programme advocating for sexual and reproductive health and rights (SRHR) presented as an outcome the increased mobilization of progressive civil society organizations, including youth organizations, around the International Conference on Population and Development (ICPD) process from 2011 to 2014. As our evaluation report narrates, the alliance’s efforts helped to bring together increasing numbers of civil society organizations in an international network. This contributed to their increased participation in the ICPD process. These efforts helped to build trust and collaboration among civil society organizations. The theory of change presented as a main objective the renewal of the SRHR agenda at the United Nations (UN) level after 2014 through the ICPD and post-2015 processes. The envisaged change process involved multiple governance levels and civil society action addressing and connecting these different levels. The outcome was interpreted in light of this theory, bringing out its meaning vis-à-vis the objectives and the processes of change through which these were to be achieved:

In the complex ICPD process, it is important for likeminded civil society organisations to be organised and linked by a strong network to bring about desired changes. It is also important to have a broader movement with a diverse group of people, including youth, from different countries, so that they can influence their own governments, who play a major role in the UN. The increasingly large presence of many SRHR advocates at the CPD 44, 45 and 47 [the 44th, 45th and 47th Sessions of the Commission on Population and Development] sent a message to delegations worldwide that SRHR is widely supported. Reaching a critical mass of active individuals beginning with CPD 44 also allowed SRHR advocates to divide tasks and accomplish much. Civil society in the different countries was able to work with their governments to prepare them for the CPD process, and then to use the CPD resolution to advocate towards their own governments. (Arensman et al., 2015: outcomes annex 108)

We also worked with a narrative approach because outcomes often do not come in isolation. Rather, they are connected and often occur in clusters. Different outcomes contribute to change

over time, with outcomes building on previous outcomes. Only stories can bring out these connections, which can be made plausible through reasoned accounts, largely drawing on an insider's perspective and knowledge, rather than being objectively assessed. In the MFS II evaluation, reports that we received from organizations did not always clarify how outcomes fit into a theory of change. We analysed and described outcomes in interaction and negotiation with alliance staff, working to articulate how outcomes fit specific pathways of change, while allowing for flexibility in how programmes unfolded and also establishing where the programme theory did not work out as anticipated. In this way, we geared our use of theory of change to the complex and complicated nature of the evaluated change processes.

Interpreting outcomes in relation to theory of change made it possible to specify and clarify the theories of change in a way that was directly relevant for the outcomes at hand. Programmes' reconstructed theories of change were generic enough to be applicable to different sub-programmes and projects within these. These projects could be interpreted as fitting these generic theories but also worked with more specific, contextual, lower-level theories. Starting from outcomes, it is possible for M&E specialists to engage with these and surface their contribution to change within the broader theory of change without taking on the impossible task of surfacing all possible lower-level theories over the evaluated period. In our evaluation, outcomes were interpreted in this way, and lower-level theories contextualizing the outcomes were reconstructed, taking the outcomes as a starting point, as shown in the above example on SRHR.

In practice, possibilities for finding and tracing evidence are often limited; identifying and testing potential competing explanations is resource-intensive, and, again, evidence will often not be available or accessible. However, those carrying out programmes can show how their analysis is based on the best possible understanding of how their actions may have contributed to a certain change, making the account plausible and building stories that show strategic thinking, contextual engagement, execution of action and observation and analysis of targets' responses. Assessment is then not primarily about accumulating evidence, but rather about appraising practical judgment in terms of strategic intent, the informed analysis of options within a context, the deployment of capacities in line with this analysis, and testimony from observation, all in light of a theory of change. Other than existing methods like Contribution Analysis (Mayne, 2012) or process tracing (Collier 2011) that focus on causal inference, Narrative Assessment hereby seeks to bring out advocates' navigation of complex processes, and the results of this.

With stories imposing meaning, evaluation demands the examination of that meaning – testing not the 'truth' of a theory, but rather the plausibility of the story. Stories should show a firm grasp of the institutional contexts and dynamics, as well as an informed sense of the programme's relative contribution to addressing an issue and influencing key audiences. Stories should be concrete in conveying the nature of contributions to change. What difference did a programme plausibly make? On the basis of what practical judgment are statements about this made? What did advocates observe or otherwise learn about what happened that can support a story? How do claims hold up in the face of contradictory information or alternative judgments?

To illustrate: programme staff may claim to have contributed to a shift in the agenda of certain national- or European Union-level policymakers who have embraced a position that the programme spent years advancing – for example, that land-based biofuels threaten food security and the land rights of poor people in developing countries and that the European Union should therefore impose a maximum allowable cap on land-based biofuels (see

Arensman et al., 2015: 133–88). An evaluator's examination of that account, in interaction with programme staff, could address several types of questions, such as: What makes an advocate think the policymakers' position was different before a campaign? What makes her/him think that change in a target that she perceived was because of the programme? What was the relative role of other contributing factors, as she/he observed? Facts matter here, as does the organization's or advocate's interest in ensuring outcomes are interpreted as favourably as possible (Arensman and Van Wessel, 2017; Arensman et al., 2015: 518). Advocates and M&E specialists need to look closely for signs that confirm that the programme made a contribution – their task involves building as well as examining the plausibility of the account. Some outcomes, such as sustained access to a roundtable or other policymaking arena or the evident use of a report written by advocates in a government's policy plan, may be clearly attributable. However, claims of achievements that do not come with such clear-cut evidence also need to be taken seriously. Their plausibility should be assessed based on an exploration of the practical judgment of advocates who have experience manoeuvring in the political arena and who have witnessed relevant events, swings in attention for and engagement with their cause, and the actions of other organizations.

Assessing the relevance of advocacy outcomes

An important part of interpreting an outcome is establishing whether and how it contributes to a process of change envisaged by the advocates involved – its relevance to this process. The assessment of relevance may be a theoretical exercise. To clarify: as found in the MFS II evaluation, the results of advocacy for development often tend to be removed from actual improvements in people's lives or in environmental conditions (i.e. impact). In our evaluation, outcomes were predominantly in the result area of agenda setting, which included a range of achievements that often preceded policy influencing, such as achieving a shared position as an organization/network, gaining access to a policymaking arena, or having an advocated position taken up by a policymaker in public debate. To a much lesser extent, further-reaching changes were achieved, including policy influencing (demonstrable changes in policy including the adoption of a position or prevention of unwanted policy change, or demonstrable changes in accountability structures for government actors or other authorities) and changes in practice (concrete changes in processes of policy formulation or policy implementation).

In advocacy monitoring and evaluation, it is commonly accepted that outcomes are often steps in the direction of desired change. It is likely that this is especially true in advocacy for development. For an important part, the programmes in the MFS II evaluation were oriented toward normative and legal frameworks, often at multiple levels, concerning, for example, gender and reproductive health and rights, land rights and human security. The timeframes in the programme theory were typically many years longer than the duration of the evaluated programmes. With a focus on organizing, articulating, changing minds and agendas through awareness raising, and shifting aspects of long-running policy processes, there was, in many cases, relatively little orientation toward policy development closer to practice or implementation.

Following the approach we took, assessment of outcomes can be centred on theory of change, taking the stance that a relevant step forward as conceived in a theory of change is a relevant step forward in reality. However, relevance, as articulated in theories of change, may

be contested in light of counterarguments. Establishing advocacy outcome relevance should therefore involve a critical and widely inclusive assessment of the theory of change vis-à-vis constituency needs. For example, how would an adjustment to a certain European Union law fought for in an advocacy programme benefit constituencies? What difference would the desired changes really make for constituencies, and on what judgment, following which criteria, is this assessment based? Answering these questions involves critically examining outcome stories, demanding that M&E specialists grasp the political context and have the capacity to ask questions about the stories to the claim-makers to fill in gaps and address contradictions and doubts. For example, in the MFS II evaluation, the evaluation team questioned the relevance of the achievements of an alliance working on palm oil. The organization contributed to changes in the policies of the Roundtable on Sustainable Palm Oil (RSPO), following a theory of change claiming that such changes would contribute to the objective of making the palm oil sector more sustainable. The relevance of these changes was questioned by the team because of a range of issues including, for example, the weak implementation record of the RSPO (Arensman et al., 2015: 244). Although such questioning may not lead to evidence that upholds or rejects relevance claims, Narrative Assessment can establish the plausibility of relevance claims and may lead to a reconsideration of such claims.

In addition, Narrative Assessment would involve addressing questions of impact even when objectives and strategies maintain that impact is to be reached in the distant future, following gains on many fronts through many steps – thereby addressing theories of change *themselves*. Theories of change should cover constituency needs explicitly to help develop plausible arguments about outcome relevance, but we found that this key element is not always articulated clearly. Programmes and actors thereby run the risk of operating in a bubble, focusing on activities and assumptions about impact rather than on substantiated and shared understandings of how a strategy may lead to impact. In our evaluation, advocates sometimes found themselves in doubt as to how even their own participation in a programme could possibly lead to impact ‘on the ground’. For example, an African advocate participating in a campaign for mainstreaming human security at the UN headquarters in New York remarked that he had no clear sense of how his participation could make a difference for people living in villages in his own country whose lives he sought to improve through his advocacy. This issue was not addressed during the development or execution of the campaign. Rather, such impact appeared to be simply assumed as a possible or likely result, without clearly establishing the pathways of change. I propose that we confront the impact question directly in monitoring and evaluation of advocacy, including the examination of theories of change to uncover how understandings of potential impact on constituencies are conceived, articulated and shared among the different actors involved.

With advocacy outcomes remaining largely removed from impact, relevance for constituencies can and should be *theorized*, banking on future changes that, in the end, are to contribute to impact. Outcomes can be accepted as relevant when they can be plausibly constructed as steps in the direction of desired change – individual steps that can be assessed in the context of many other steps, *including future ones*. In Narrative Assessment, outcome stories are thus to be set in a wider, theory-centred argument linking stories on outcomes to future action. The quality of such an assessment is crucial. M&E specialists need to invest in their understanding of issues and contexts, seek out relevant interlocutors (with insider information and the capacity to bring forth constituency perspectives) and learn how to question claims in a respectful yet thorough manner.

Reflection and learning

Through the creation and assessment of stories, Narrative Assessment can facilitate reflection and learning. Advocacy often involves strategic manoeuvring in dynamic situations. Narrative Assessment can make this manoeuvring, as well as the strategic considerations and the underlying assumptions, explicit. In addition, Narrative Assessment can help put claims about the effectiveness of strategy in such dynamic situations to the test. Narrative Assessment involves critical inquiry into the assumptions underlying theories of change, the capacity of those involved to analyse and act when faced with challenging contexts and developments, and the identification of critical issues demanding reflection and learning. This involves addressing a variety of questions: Why did a certain strategy make sense in a particular situation? Based on what analysis? How can this be assessed in hindsight? In this sense, Narrative Assessment is a form of action research. Narrative Assessment thus focuses not only on outcomes, but also on identifying and explaining disappointments and challenges. Different internal and external conditions and developments may contribute to disappointing results. Identifying relevant internal factors can contribute to the development of internal capacities, and identifying relevant external factors can contribute to a reconsideration of strategy and to enhanced understanding and resolve in the face of disappointing progress. In addition, by documenting strategic manoeuvring through stories, organizations – and the field more broadly – can build knowledge about the workings and outcomes of strategy, and reflect and build on these.

Communicating advocacy results

Narrative Assessment also facilitates the communication of results, relating to the nature of advocacy work. Advocacy is often a struggle. Therefore, advocacy achievements lend themselves well to storytelling. There may be little political space in which to act. There may be little or no interest in one's issues among the powers that be. There may be powerful adversaries acting against a programme's objectives. Advocacy often seeks to advance change in situations that, from the advocate's perspective, are an affront to moral principles (e.g. defending rights to clean water, land, justice or protection from violence, as well as basic human rights). Advocates appeal to a shared moral standpoint when they call for support or make propositions about the worthiness of their efforts or the significance of their success. It is in situations of moral affront demanding action that stories emerge (Wagenaar, 2014: 215), as ways to make sense and take position. Narrative Assessment can thereby help to bring forth the significance of advocacy achievements, as well as setbacks or gridlock, in a way true to this set of conditions.

Advocacy results can easily remain meaningless for publics beyond a very small set of insiders because of a lack of direct impact. This lack of impact on constituencies can be mistaken for a lack of significance if the significance of the results is not articulated. Achieving an adjustment to a policy document may result from great effort made over many years. However, clarifying how that change is or may become helpful to constituencies or society at large needs attention. This is not simply because policy processes are technical. Again, the outcomes of advocacy for development are often interim in nature, banking on *future* policy processes to attain the ultimate legitimacy of positive impacts such as increased access to clean water, food security or a living wage. This raises the fundamental question of effectiveness. Narrative Assessment can deal with this question in communicating the meaning of

outcomes to both internal and external audiences. Stories allow you to take others along on your journey. Good stories engage and may thereby contribute to legitimacy and public support for a programme's efforts. A key benefit of Narrative Assessment may thus lie not only in answering *how something happened*, exposing the mechanism behind change, but also in answering *why it matters*. The answer to this latter question will often be theoretical in advocacy monitoring and evaluation, as impact can often only be addressed in the form of theory of change-based propositions of achievements as steps toward longer-term change. These necessarily relate to the future as much as to the past, addressing the need to account for actions while offering a rationale for supporting further actions. Stories offer an interpretation that shows how a change is a step in the right direction, but also how it may help set the stage for further desired change. Stories can also legitimate further investment, which is important considering the public debate regarding the effectiveness of development interventions.

To achieve these objectives, stories must convey at least some promise of a successful way forward, made plausible by the quality of advocates' grasp of the situation and the organization's chances of carrying out significant action in a given situation. Theorizing further about the qualities a story should have to contribute effectively to programme legitimacy, different aspects are potentially important. The added value of an outcome and of the programme thought to contribute to it can be shown more or less compellingly. Likewise, the explanation of how an outcome is to be understood as a step in a pathway of change with potential relevance for constituencies, building on past developments and banking on future ones, can be more or less compelling.

Questions of quality

The proposed approach is grounded in the assumption that M&E specialists and advocates can co-construct and test stories for plausibility through a process of critical reflection and examination. However, narrative analysis research suggests that stories have a stabilizing function. They underpin assumptions to allow decision making in the face of high uncertainty, complexity and polarization. Stories resist contradictory information that would destabilize a key function of the story – offering a basis for decision making and everyday action (Polletta et al., 2011; Roe, 1994). Because stories can have great value, offering a basis for orientation and action, challenging them comes with a loss. We need to learn to what extent space for reflection and examination exists or can be created, and how critical reflection and examination can take place so that that confrontation with contradictory information, gaps and other challenges to stories becomes a key part of the co-construction of stories for advocacy monitoring and evaluation.

Importantly, Narrative Assessment does not mean abandoning rigour as a dimension of good monitoring and evaluation. Focusing on the plausibility of theory, achievements and their relevance, Narrative Assessment offers a new direction for conceptualizing rigour, drawing on the narrative inquiry research tradition (see e.g. Loh, 2013). Plausibility is to be determined by the lifelikeness and believability of stories – their 'verisimilitude'. Stories should be co-constructed between advocates and M&E specialists, with advocates in the lead role and M&E specialists serving as examiners seeking and assessing detail, consistency, embeddedness in the context, clarification, intersubjectivity in interpretation, and evidence to co-develop and test the believability of stories against alternative interpretations, undermining omissions and empty claims. This essential requirement of systematic examination of stories, key to

making the method rise above collection of stories pure and simple, places demands on M&E specialists, whose professional role will be grounded in their capacity to simultaneously build and examine, balancing the need for trusting relations with advocates and their organizations, on the one hand, and for critical distance, on the other. From these starting points, Narrative Assessment can offer rigour in a practicable and logically coherent methodological outlook for advocacy monitoring and evaluation. One issue to be addressed is whether and how plausibility could or should be determined in ways that meet standards for plausibility held by audiences (Polletta et al., 2011).

However, a further question is whether and how funders might be convinced to accept Narrative Assessment in institutional contexts where evaluation is expected to centre on measurement, evidence and ‘objectivity’. It is unknown to what extent, and on what basis, the logically coherent methodology that we foresee will satisfy such audiences. A counterargument here is that also current methods run into problems on this front, in the sense that evidence provided is sometimes very limited or subjective without this being addressed in any systematic way, as we propose here.

A few other issues arise. One is that is that stories may demand too much prior knowledge and acceptance of assumptions to be accessible or to convey meaning as intended. Thus, stories should be constructed differently for different audiences. Finally, the sensitive nature of advocacy may result in a challenge regarding the lifelike quality of stories. Building lifelike, believable stories brings events, developments and actors to life. Audiences come close to advocates, targets and others who may be implicated in some way, suggesting how things happened, what people thought and how they responded. Notably, advocacy often takes place behind closed doors for good reasons. The exact nature of an activity, why it is seen as effective and the effect itself may all be too sensitive to publicize, and shedding too much light on strategies and people may jeopardize possibilities for further change (see Patton, 2008). Stories may therefore need to be adjusted for different internal and external usages.

Conclusion: From measurement to meaning

We need to accept advocacy for what it is: a craft largely rooted in the practical judgment of situations and strategic options within those situations. There is little to draw on when it comes to fool-proof knowledge of ‘what works’ (or even ‘what works where’) or evidence to back up claims of contribution to outcomes. Advocacy is also characterized by a great deal of unpredictability, powerful opponents and results that are often intangible. Narrative Assessment can deal with these unavoidable problems with measurement and evidence facing advocacy monitoring and evaluation. This approach accepts that certainty about causality will often be out of reach and that believability will often be as close to ‘truth’ as advocacy monitoring and evaluation can get, in particular when it comes to insight into how and why certain advocacy activities did or did not make a difference. Believable stories are thus seen not as an inferior stand-in for objective evidence, but rather as a rich and meaningful source of knowledge in their own right.

Presently, Narrative Assessment is being translated into practical guidelines, incorporating the theoretical starting points and questions of quality discussed above, by Margit van Wessel at Wageningen University & Research and Wenny Ho at Hivos.² We hope to launch Narrative Assessment as a full-fledged and tested methodology by the end of 2018.

Acknowledgements

While Narrative Assessment is my own idea, I would not have been able to come up with it without important lessons learned through the team cooperation in the MfS II International Lobbying and Advocacy Evaluation. Thanks therefore to Bodille Arensman, Jennifer B. Barrett, Dorothea Hilhorst, Wolfgang Richert, Cornélie van Waegeningh, Annemarie Wagemakers, Dieuwke Klaver, Elisabet Rasch and Arend Jan van Bodegom. I also thank Bodille Arensman and Wenny Ho for our interesting discussions on translating the theoretical ideas into practice. These have led me to adjust parts of this article. I finally thank two anonymous reviewers for their helpful feedback on an earlier draft of this article.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. MFS II stands for ‘medefinancieringsstelsel II’, a term describing the second, 2010–15 round of the Netherlands’ policy of co-funding international development activity by civil society organizations.
2. Hivos is a civil society organisation working in international development.

References

- Arensman B and van Wessel M (2018) Negotiating effectiveness in transnational advocacy evaluation. *Evaluation* 24(1): 51–68.
- Arensman B, Barrett JB, Van Bodegom AJ, et al. (2015) *MFS II Joint Evaluation of International Lobbying and Advocacy: Endline Report*. Wageningen: Wageningen University.
- Arensman B, van Waegeningh C and van Wessel M (2018) Twinning ‘practices of change’ with ‘theory of change’: Room for emergence in advocacy evaluation. *American Journal of Evaluation* 39(2): 221–36.
- Arensman B, van Wessel M and Hilhorst D (2017) Does local ownership bring about effectiveness? The case of a transnational advocacy network. *Third World Quarterly* 38(6): 1310–26.
- Bamberger M, Vaessen J and Raimondo E (eds) (2016) *Dealing with Complexity in Development Evaluation: A Practical Approach*. London: SAGE.
- Barrett JB, van Wessel M and Hilhorst D (2016) *Advocacy for Development: Effectiveness, Monitoring and Evaluation*. ISSUU. Available at: https://issuu.com/margitvanwessel/docs/advocacy_for_development?e=3921622/33875150 (accessed 14 June 2016).
- Chapman J and Wameyo A (2001) *Monitoring and Evaluating Advocacy: A Scoping Study*. ActionAid. Available at: <http://www.eldis.org/vfile/upload/1/document/0708/DOC21800.pdf> (accessed 10 August 2016).
- Clandinin DJ and Connelly FM (2000) *Narrative Inquiry: Experience and Story in Qualitative Research*. San Francisco, CA: Jossey-Bass.
- Coffman J (2009) Overview of current advocacy evaluation practice. Available at: <http://www.evaluationinnovation.org/publications/overview-current-advocacy-evaluation-practice%20> (accessed 10 October 2016).
- Collier D (2011) Understanding process tracing. *Political Science & Politics* 44(4): 823–30.
- Conlin S and Stirrat RL (2008) Current challenges in development evaluation. *Evaluation* 14(2): 193–208.

- Coryn CL, Noakes LA, Westine CD, et al. (2011) A systematic review of theory-driven evaluation practice from 1990 to 2009. *American Journal of Evaluation* 32(2): 199–226.
- Costantino TE and Greene JC (2003) Reflections on the use of narrative in evaluation. *American Journal of Evaluation* 24(1): 35–49.
- Davies R and Dart J (2005) The Most Significant Change (MSC) Technique. *A Guide to its Use*. Available at: <https://www.kepa.fi/tiedostot/most-significant-change-guide.pdf> (accessed 12 September 2018).
- Gardner A and Brindis C (2017) *Advocacy and Policy Change Evaluation: Theory and Practice*. Stanford, CA: Stanford University Press.
- Jones H (2011) *A Guide to Monitoring and Evaluating Policy Influence*. London: Overseas Development Institute.
- Lawler S (2002) Narrative in social research. In: May T (ed) *Qualitative Research in Action*. London: SAGE, pp. 242–258.
- Loh J (2013) Inquiry into issues of trustworthiness and quality in narrative studies: A perspective. *The Qualitative Report* 18(33): 1–15.
- Lowery D (2013) Lobbying influence: Meaning, measurement and missing. *Interest Groups & Advocacy* 2(1): 1–26.
- Mayne J (2012) Contribution analysis: Coming of age? *Evaluation* 18(3): 270–280.
- Morariu J and Brennan K (2009) Effective advocacy evaluation: The role of funders. *The Foundation Review* 1(3): 100–8.
- Netherlands Ministry of Foreign Affairs (2014) Policy framework dialogue and dissent: Strategic partnerships for lobby and advocacy. Available at: <https://www.government.nl/topics/grant-programmes/documents/regulations/2014/05/13/policy-framework-dialogue-and-dissent> (accessed 28 June 2017).
- Organisation for Economic Co-operation and Development (2005) DAC criteria for evaluating development assistance. Available at: <http://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm> (accessed 8 August 2016).
- Patton MQ (2008) Advocacy impact evaluation. *Journal of MultiDisciplinary Evaluation* 5(9): 1–10.
- Polletta F, Chen PCB, Gardner BG, et al. (2011) The sociology of storytelling. *Annual Review of Sociology* 37: 109–30.
- Prinsen G and Nijhof S (2015) Between logframes and Theory of Change: Reviewing debates and a practical experience. *Development in Practice* 25(2): 234–46.
- Reisman J, Gienapp A and Stachowiak S (2007) *A Guide to Measuring Advocacy and Policy*. Baltimore, MD: Annie E. Casey Foundation.
- Riddell RC (2014) *Does Foreign Aid Really Work? An Updated Assessment*. Development Policy Centre Discussion Paper 33.
- Roe E (1994) *Narrative Policy Analysis: Theory and Practice*. Durham and London: Duke University Press.
- Rogers PJ (2008) Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation* 14(1): 29–48.
- Rogers PJ and Weiss CH (2007) Theory-based evaluation: Reflections ten years on: Theory-based evaluation: Past, present, and future. *New Directions for Evaluation* 114: 63–81.
- Stein D and Valters C (2012) *Understanding Theory of Change in International Development*. London: LSE.
- Stern E, Stame N, Mayne J, et al. (2012) *Broadening the Range of Designs and Methods for Impact Evaluations: Report of a Study Commissioned by the Department for International Development*. Department for International Development.
- Tsui J, Hearn S and Young J (2014) *Monitoring and Evaluation of Policy Influence and Advocacy*. Working Paper 395. London: Overseas Development Institute.

- The Urban Institute (n.d.) *Candidate Outcome Indicators: Advocacy Program*. Available at: http://betterevaluation.org/en/resources/tool/advocacy_program_indicators (accessed 7 October 2016).
- Vanclay F (2015) The potential application of qualitative evaluation methods in European regional development: Reflections on the use of performance story reporting in Australian natural resource management. *Regional Studies* 49(8): 1326–39.
- Vogel I (2012) *Review of the Use of 'Theory of Change' in International Development: Review Report*. London: Department for International Development.
- Wagenaar H (2014) *Meaning in Action: Interpretation and Dialogue in Policy Analysis*. London: Routledge.
- Walton M (2016) Expert views on applying complexity theory in evaluation: Opportunities and barriers. *Evaluation* 22(4): 410–23.
- Weiss CH (1997) Theory-based evaluation: Past, present, and future. *New Directions for Evaluation* 76: 41–55.
- Wilson-Grau R and Britt H (2012) *Outcome Harvesting*. Cairo: Ford Foundation Middle East and North Africa Office.

Margit van Wessel is an assistant professor in the Strategic Communication Chair Group at Wageningen University & Research. She explores the dynamics of communication and governance. Recent projects address citizens' understandings of politics, advocacy for development and advocacy evaluation.