

American Journal of Evaluation

<http://aje.sagepub.com>

Six Guiding Principles for Evaluating Mode-2 Strategies for Sustainable Development

Barbara J. Regeer, Anne-Charlotte Hoes, Mariette van Amstel-van Saane, Francisca F. Caron-Flinterman and Joske F. G. Bunders
American Journal of Evaluation 2009; 30; 515
DOI: 10.1177/1098214009344618

The online version of this article can be found at:
<http://aje.sagepub.com/cgi/content/abstract/30/4/515>

Published by:



<http://www.sagepublications.com>

On behalf of:

[American Evaluation Association](#)

Additional services and information for *American Journal of Evaluation* can be found at:

Email Alerts: <http://aje.sagepub.com/cgi/alerts>


Subscriptions: <http://aje.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations <http://aje.sagepub.com/cgi/content/refs/30/4/515>

Six Guiding Principles for Evaluating Mode-2 Strategies for Sustainable Development

American Journal of Evaluation
30(4) 515-537
© The Author(s) 2009
Reprints and permission: <http://www.sagepub.com/journalsPermissions.nav>
DOI: 10.1177/1098214009344618
<http://aje.sagepub.com>


Barbara J. Regeer,¹ Anne-Charlotte Hoes,¹
Mariette van Amstel-van Saane,¹ Francisca F. Caron-Flinterman,² and
Joske F. G. Bunders¹

Abstract

A new generation of public programs emerges, which specifically addresses complex societal problems we witness today. Programs for these types of complex issues—in this article, we consider more closely the challenge of sustainable development—are characterized by emergent design, learning processes between diverse actors, and adaptive management. Managers of these kinds of programs have new demands for evaluation and evaluators. This article describes prevailing evaluation methods for sustainable development (progress assessment, goal-oriented program evaluation, and program theory evaluation) and the challenges they meet when confronted with the complexity of designing and conducting systemic intervention programs for sustainable development. The evaluation framework that we propose offers guiding principles to assist evaluators in evaluating complex programs.

Keywords

systems change, sustainable development, mode-2 evaluation, learning and action, methodology development

Introduction

Most evaluators would agree that their work should contribute to public change of some kind—be it by providing health professionals with quantitative data, by monitoring CO₂ levels, or by measuring effects on student learning as a result of educational reform. As King (2008) recently stated, “We may even regain a piece of paradise through the process and knowledge we create with our evaluations [. . .] in hopes of creating a better world.” A large part of evaluation work (as presented in this and other journals) indeed aims to contribute to the improvement of public programs and thereby foster societal change. Evaluators can choose from, and skill themselves in, a wide range of available

¹ Athena Institute at VU University Amsterdam, Amsterdam, Netherlands

² Animal Sciences Group, Wageningen UR, Wageningen, Netherlands

Corresponding Author:

Barbara J. Regeer, De Boelelaan 1085, Amsterdam, 1081 HV, Netherlands. Email: barbara.regeer@falw.vu.nl

methods and techniques. However, what if the problems to be addressed, and the corresponding programs to be evaluated, are very complex and surrounded with uncertainties? What if problem definitions are highly contested, such as is the case in discussions around health efficiency and quality of life, world food security, the ethics of biotechnology, or sustainable development? What if intervention programs are more like complex and experimental interaction processes between actors from different institutions than linear processes of problem formulation, project design, and implementation? What role can evaluation play in this emerging field of systems change? How then can evaluation contribute to a better world?

In this article, we look at monitoring and evaluation in the context of emerging intervention strategies for systems change, or system innovation. Various scholars from different academic fields have argued that some of the most complex problems society faces today are in need of new intervention strategies (which we will refer to as mode-2 strategies, see below). Complex, unstructured problems cannot be solved by the interventions of single organizations; they require cooperation of actors from different institutional backgrounds. Moreover, complex problems require strategies that entail changes in established patterns of action as well as in the structures in which they take place. Realizing and guiding these systems changes demand new types of monitoring and evaluation. Dutch government adopted the concept of *system innovation* to address complex societal problems. Large-scale innovation programs are initiated to assist system innovation (see SenterNovem, 2007). We were asked to develop and test a monitoring and evaluation approach for system innovation projects in the field of sustainable development, based on years of experience with the interactive learning and action (ILA) approach; a participatory research approach developed by Bunders (1990) and Bunders and Broerse (1991), aimed at supporting more inclusive (agricultural, health and biotechnological) innovation processes. The ILA *monitoring* approach featuring in this article is a further extension of the ILA approach.

In this article, we will describe prevailing methods for monitoring and evaluating sustainable development (progress assessment, goal-oriented program evaluation, and program theory evaluation) and articulate the challenges they meet when confronted with the complexity of designing and conducting mode-2 intervention programs for sustainable development. Throughout the article we will, for the sake of brevity, use the term “evaluation,” even though we take a broad view on “monitoring” and “evaluating” new approaches to sustainable development, incorporating both the regular collection of information as well as the assessment and evaluation of program activities at particular moments in time. The specific approach we developed, however, has been referred to as the ILA *monitoring* approach, which, for the sake of clarity, we will maintain.

To explicate our assumptions with respect to sustainable development, we will first discuss the rationale for new types of intervention strategies in the context of changing perceptions on sustainable development.

Sustainable Development Requires New Intervention Strategies

Since the early 1970s, there is growing consensus that there are limits to the capacity of the (natural) world to adjust to the increasing use of fossil fuels, production of synthetic materials, and use of natural resources (Meadows, Meadows, & Randers, 2004; Meadows, Meadows, Randers, & Behrens, 1972). With the publication of the Brundtland report in 1987, the concept “sustainable development” was introduced, being defined as “meeting the needs of the present without harming the ability of future generations to meet their own needs” (World Commission on Environment and Development [WCED], 1987). Although there is a widespread understanding of and agreement on the abstract idea of sustainable development (being a combination of the conservation of natural resources, quality of life issues, and a commitment to future generations), there is continuous disagreement not only about the extent of the problem but also about the actual problem definition.

By referring to the fundamental changes that have taken place in the way in which we regard and govern the world since the “end of modernity,” the 1997 United Nations Educational, Scientific, and Cultural Organization (UNESCO) policy paper on sustainability states that “Contrary to the conventional modernisation and development theory, the present situation is characterised by the agreement that there is no single or universal path of sustainable development” (Becker, Jahn, Stuess, & Wehling, 1997, p. 20). Or, in a terminology scholars and practitioners of evaluation may be more familiar with, there is no “theory of change” for dealing with the challenges of sustainable development (see also Walker & Kubisch, 2008).

Environmental problems cannot be solved in isolation; they are connected in a complex system of interacting physical, chemical, and biological processes, affecting each other at different scales. Moreover, the biophysical environment can hardly be considered separate from societal development; climate change directly affects the livelihood of millions of people, while industrial disasters (e.g., Chernobyl) destroy the health of generations. The activities that lead to pollution, degradation, and depletion are part of our way of life, embedded deeply in patterns of consumption and production (e.g., Beck, 1986). This understanding of the inherent complexity of the sustainability challenge has resulted in pleas for new strategies (theories of action) that entail changes on many different levels, realized by heterogeneous groups of actors, concerning a diversity of aspects. Correspondingly, new types of evaluation approaches for these new strategies are being developed, notably approaches using systems concepts in evaluation (Williams & Imam, 2007) and developmental evaluation (Patton, 2008). To understand the kinds of demands these new strategies place on evaluation, let us consider the characteristics of these new programs in more detail.

Scholars of science, technology, and society (STS) have argued that seemingly intractable types of problems, such as those associated with sustainable development, require new ways of knowledge development and a new type of relationship, or “contract,” between science and society (notably Funtowicz & Ravetz, 1993; Gibbons et al., 1994; Jasanoff, 2004; Nowotny, Scott, & Gibbons, 2001). They have advocated an inclusive and responsive science; a type of knowledge production that starts from real-life problems and aims to devise solutions in collaborations with multiple stakeholders. Similarly, scholars of policy processes have emphasized that managing for sustainability requires new ways of governance: it is not an exclusive activity of governmental bodies but rather reflects principles of network steering. Traditional intervention modes and policy instruments are replaced or supplemented by notions of reflexive governance (Voß, Bauknecht, & Kemp, 2006), deliberative democracy (Hajer, 2003), and transition management (Loorbach, 2007; Rotmans, Kemp, & Asselt, 2001).

Thus, from different perspectives, we hear urgent calls for new ways of approaching problems. Rather than specifying the differences between these emerging strategies, in this article, we shall refer to them as mode-2 strategies. Like Patton says in his discussion of the many different terms for describing program theory: “that label is best that makes the most sense to primary intended uses—the term they resonate to and has meaning within their context” (Patton, 2008: 339). In our experience, some of the more specific designations (such as transition management, transdisciplinary research, systems change, reflexive governance) do not resonate well for a variety of reasons, which is why we have chosen for the less evocative label “mode-2,” but it can be replaced by more appropriate terminology depending on the context. We loosely refer to the new modes of knowledge production as proposed by Gibbons et al. (1994, see also Nowotny et al., 2001) and have found the distinction between mode-1 and mode-2 helpful as a heuristic tool in our evaluation activities.

Mode-2 strategies share a commitment to addressing complex sustainability problems by involving multiple actors (notably social and natural scientists, entrepreneurs, administrators and governors, activists, citizens), acknowledging the multilevel nature of the problems, and articulating and connecting multiple perspectives. Subsequently, mode-2 *intermediaries* are defined as the change agents that facilitate (and participate in) deliberative processes of learning, knowledge cocreation, and developing

Table 1. Features of Mode-1 and Mode-2 Strategies

	Mode-1	Mode-2
Nature of issue	Structured, well-defined problem and goal	Unstructured, inherently complex
Relevant knowledge	Disciplinary knowledge (particularly from natural sciences)	Knowledge from natural sciences, social sciences, and experiential knowledge
Steering philosophy	Central steering, top-down	Reflexive governance, network steering
Relevant actors and relations	Central role project team from established institution, principal—agent relations	Collaboration between entrepreneurs, scientists, civil servants, societal organizations, citizens
Intervention process	Execution of a linear process of problem formulation, planning, implementation, and evaluation	Emergent design focuses on the stimulation of experimenting and (double loop) learning between different actors
Intervention impact	Direct project outputs	Beyond project level: impact on system level as well as on individual level

problem-solving strategies for sustainable development. Mode-2 intermediaries can be individual actors or institutionalized organizations, but most often we mean project teams or temporary program organizations set up to create an impact on sustainable development. We also use the mode-2 prefix when we refer to emerging approaches to evaluation.

Some characteristics of the mode-2 strategies, in relation to characteristics of mode-1 strategies, are summarized in Table 1. While mode-1 strategies would focus on seeking solutions for relatively concrete problems using well-established methodologies, mode-2 strategies emphasize the creation of conditions for structuring problems and supporting the search for possible directions for solutions (Regeer & Bunders, 2007/2009). Particularly in the case of complex problems, characterized by inherent uncertainty and long-time horizons, effective strategies should be based on adaptive, iterative, and flexible experimentation. Thus, a linear process of problem formulation, project design, and implementation is in mode-2 replaced by an experimental approach with a number of alternative frameworks that induce variation and offer diverse experiences to learn from and evaluate (Voß, Kemp, & Bauknecht, 2006).

Experience shows that implementing mode-2 strategies is not straightforward, and continuous learning (e.g., through monitoring and evaluation) is required. First, mode-2 strategies must be flexible and sensitive to the context to respond to the historically contingent and context dependent nature of sustainability; general guidelines do not suffice. Depending on the particular cultural, political, and ecological starting points, appropriate strategies for sustainable development allow for a variety of transformation paths. Second, sustainable development implies making different choices and trade-offs, often outside the realm of the usual behavior (whether personal or institutional). Connections need to be made between reflecting and acting, which does not imply that reflection can simply lead to change in action. A mode-2 strategy must acknowledge the interrelatedness of action and structure (Giddens, 1984); interactions among agents generate emergent structures that in their turn constrain future behavior of agents (see also writings on complex adaptive systems [CAS] in relation to evaluation, e.g., Eoyang, 2007; Parsons, 2007). Although fundamental changes are required, existing structures cannot be ignored, overridden, or simply bypassed. Hence, these new intervention strategies must act on, or induce change at, multiple levels, ranging from individual actors to incumbent institutional structures, and thus take a broader system perspective (e.g., Churchman, 1970) with the corresponding inherent unknowables and uncertainties that inhibit a predictable course of action.

As a consequence, there is a growing sense of awareness that besides optimizing intervention strategies, the difficult route toward a more sustainable world demands the support of careful monitoring trajectories and evaluation activities. Since the coining of the term sustainable development, the number of monitoring and evaluation projects and programs on sustainable development has grown vastly, as has the number of *approaches* to monitoring and evaluating sustainable development, all varying in terms of scope, unit of analysis, purpose, tools, and role of the evaluator.

Below, we describe three prevailing approaches in more detail and elaborate on their potential contribution to sustainable development as well as on the challenges they meet when confronted with the increasing complexity, multilevelness and actor pluralism that sustainable development entails. We suggest that an additional type of evaluation may further the pursuit of sustainable development by contributing to the learning processes of practitioners in their attempts to develop and implement mode-2 strategies for sustainable development. We explore the outlines of a mode-2 evaluation framework, by building on experiences with developing and experimenting with the ILA monitoring approach.

Three Approaches to Monitoring and Evaluating Sustainable Development, Their (Potential) Contribution and Their Challenges

Taking sustainable development as illustrative domain for evaluating mode-2 intervention strategies, a rough distinction into three types of evaluation can be made. Many evaluation programs were set up to assess national or global trends and facts related to sustainable development (progress assessment). Other evaluation projects and programs have been launched with the purpose to assess the goals and results of interventions aimed at contributing to sustainable development (goal-oriented program evaluation). Finally, there are studies that aim to investigate and assess the (often implicit) theories that underlie those interventions (program theory evaluation). For other domains, such as health, education, and development, similar distinctions between approaches to evaluation can be made (compare Abma's three variations in policy evaluations, 2001; and Guba & Lincoln's three generations preceding their fourth, 1989).

Progress Assessment

Many evaluation programs have been set up in an attempt to formulate an answer to the question: "*If there are limits to the capacity of our (natural) world, where do we stand with respect to them?*" These types of studies evaluate the state-of-the-world with respect to sustainable development by identifying trends and facts, based on longitudinal measurements or complex models (e.g., Intergovernmental Panel on Climate Change, Millennium Ecosystem Assessment, Biodiversity Outlook). They can be deployed in general or in relation to an intervention program. Indicator frameworks have been developed to aid the, primarily quantitative, description of the current status of the environment on global, national, and local level as well as of individual sectors, companies, households, and ecosystems. Indicator-based sustainability monitoring is rather successful in providing feedback. Through reports and public communication, we have become aware of trends in "the state of sustainability," both on global, national, or local scale and of a need to change behavior. Al Gore has been able to draw on many different studies, models, and empirical results to support his "inconvenient truth" (Gore, 2006).

Although progress assessment provides an opportunity to establish a baseline of the existing conditions as well as to identify progress in the natural world (by monitoring at regular intervals), it has proven difficult to accommodate the complex and interwoven nature of various societal (political,

social, economical) and ecological subsystems presupposed by a mode-2 view on sustainable development. Thus, the need for developing innovative indicator frameworks arises (Kemp, Parto, & Gibson, 2005). "If societal processes are to be evaluated in terms of their sustainability, then the different indicator systems cannot be merely reviewed, reformulated or supplemented." (Becker et al., 1997).

We can distinguish at least two development paths for new indicator frameworks required by the complex challenges of sustainable development. First, by stating that change of action is the desired response to the application of sustainability indicators, the issue of *agency* is introduced into the indicator discussion. "Who wants sustainable development indicators? Who owns the process of their creation and who uses them?" (Bell & Morse, 2004). Becker et al. contribute to this discussion by proposing that "it should be further examined how the factor of agency might lead to a more self-reflexive approach to the development of indicators" (1997: 31) by exploring the institutional and political contexts within which sustainability measures take place. Thus, there is a suggested shift of focus from the "what" of measurement to considerations of participative strategies and institutional frameworks.

At the same time, the need is expressed for new sets of indicators themselves, which is the second path; indicators that allow for the assessment of complex relationships between social, economic, political, and ecological systems (Hildén & Rosenström, 2008; Kemp et al., 2005). The challenge to developing innovative indicators thus has a process and a content dimension.

Goal-Oriented Program Evaluation

A second type of evaluation approach aims to answer the question, "If we want to change our position with respect to the limited capacity of our world, **what** interventions can be considered appropriate and successful?" A vast range of (global, EU, and national) policy programs (e.g., Millennium goals, Kyoto protocol), citizen initiatives (e.g., Local Agenda 21, Ecoteams), nongovernmental organization (NGO) collaborations (e.g., Antiglobalization movement), and changes in business management (corporate social responsibility) have been initiated since the publication of the Brundtland report "our common future" (WCED, 1987). Result-oriented management in the context of both businesses and governments has led to a growth of performance measurement and (goal-oriented) program evaluation to provide evidence of impact and effectiveness (De Lancer Julnes, 2006). The dominant focus of this type of approach is on measuring the outcomes of policy programs and/or corporate initiatives, in terms of achievement of predetermined goals and objectives, after the program has been completed (i.e., summative); resulting in, among others, annual sustainability reports of companies based on the global reporting initiative (GRI) indicators or the environmental performance review.

The dominant view on program evaluation presupposes a linear and rather technocratic policy process, which expects evaluation results to be used by policy makers in a direct instrumental manner (Lehtonen, 2005). Some have argued that the potential contribution of these so-called goal-oriented evaluations to subsequent decision making about future action is limited (Patton, 2008). Although program evaluations are known to have an indirect use through an enhanced awareness (a form of Patton's process use, 2008), the direct use of evaluations in policy decision making is an exception rather than a rule (e.g., see Lehtonen, 2005). Decision making is influenced and informed by many different sources of knowledge, of which evaluation is only one: "Programme designers and programme managers interact with a number of different actors: evaluators constitute one group and evaluation a source of knowledge that is to share among many others existing within the organisation" (Marra, 2004, p. 264). Moreover, as interventions are executed, intended and expected as well as nonintended and nonexpected effects become apparent (e.g., Dart & Davies, 2003). If these responses in the environment are closely monitored, they may lead to new insights

into the aspired challenge or objective, with a change of strategy as a result. This has implications for the role of evaluation in the intervention program; more emphasis is placed on incorporating evaluation in the intervention process (see also Friedman, 2001). Program evaluation that focuses on evaluating the program itself with the purpose of assisting in making timely adjustments (formative) rather than focusing on the end result of the program (summative) seems expedient in cases of mode-2 approaches to sustainable development and other complex problems alike.

Another important challenge of program evaluation is to accommodate the *multiple* and *dynamic* understandings of sustainable development. Many evaluators advocate that a rigorous evaluation requires reference to declared goals, even if they are recognized as changing, unstable, and fluid (McCoy & Hargie, 2001). In the case of mode-2 strategies for sustainable development, and other complex issues alike, it is suggested that the actual problem articulation becomes an important objective of the program itself, rather than settling down with an incomplete or inadequate set of goals and measuring the outcomes of the program against them. The challenge of unstructured problems is not to structure them but to make their articulation a matter of discussion and part of a learning process between a variety of (policy) actors (Hisschemöller & Hoppe, 1996). Similarly, Friedman, Rothman, and Withers (2006) advocate to make goal setting a process of systematic and participative inquiry into goals, particularly in cases of controversy or dissent.

Program Theory Evaluation

In contrast to goal-oriented program evaluation, “program theory evaluation” focuses on understanding how programs work, addressing the question: “*why certain interventions are devised?*” Implicit “theories of intervention” are explicated, unveiling the set of assumptions underlying policy goals and objectives. The rationale underpinning a particular intervention program is specified in terms of its inputs, its expected outcomes, the assumed relationships between them, and the underlying mechanisms relating expected program ends and means to each other (Chen, 1990). These mechanisms, explaining how program inputs and activities are intended to create the desired outcomes, are referred to as *logic models*. Logic models can be constructed using various techniques such as “if . . . then” propositions or concept mapping (e.g., Rosas, 2005; Yampolskaya, Nesman, Hernandez, & Koch, 2004). Program theories and logic models can be determined deductively on the basis of academic theory or constructed inductively on the basis of fieldwork (e.g., Patton, 2008, p. 344). User-focused and participative approaches are increasingly used to (collaboratively) construct and articulate the theory of intervention of program teams (e.g., Christie & Alkin, 2003; Patton, 2008).

Program theory studies have many potential benefits; they can inform program evaluation on relevant variables to be included in the analysis, explain events and behavior, assist practitioners in understanding the workings of their programs, and inform future actions. However, three characteristics of mode-2 intervention strategies pose challenges to program theory evaluation: their emergent design, the complexity of the endeavor, and their experimental nature (see Table 1).

First, program theory evaluation presupposes a relatively stable program, of which the activities, goals, and intended effect can be univocally described (even if informed by a wide range of stakeholder perspectives). However, as mode-2 strategies are characterized by emergent design, appropriate interventions can only be determined beforehand to a limited extent, and they change considerably during conduct.

Second, the complexity of mode-2 strategies toward sustainable development (characterized by a plurality of values and perspectives, permanent uncertainties, and pervasive interconnectedness between ecological, social, institutional, political, and economical system) makes it extremely difficult to identify causal links (even hypothetically, as in logic models) between interventions and effects. The number of activities and actors involved as well as the various scales at which

interventions take place add to the complexity of attributing possible effects to the intervention program. As CAS theory states, knowing the factors that influence a system does not mean we can control these factors and expect predictable outcomes (Parsons, 2007). In the field of evaluation too, the analysis of (causal) links between policy interventions and their outcomes is considered challenging (e.g., see Perrin, 2002), particularly when complex problems or programs are concerned (e.g., Dye-house, Bennet, Harbor, Childress, & Dark, 2009; Forrest, 2007; Patton, 2008; Williams & Imam, 2007).

Third, as mode-2 strategies are fundamentally experimental, academic theory on systems change cannot be straightforwardly translated into practice. Friedman (2001) refers to Argyris and Schön (1974, 1978) in observing that, particularly under conditions of uncertainty, ambiguity, and goal conflict (such as is the case with mode-2 strategies for sustainable development), there are gaps between program theory and program practice. Actors are frequently unable to behave in ways consistent with their espoused theories. Thus, even if program participants' espoused change theory corresponds to academic theory on complex, multilevel transitions, we can expect that in practice actors will not act accordingly. This has consequences for the use of program theory in evaluation of mode-2 strategies.

Let us illustrate by means of the current Dutch discourse. As described above, a mode-2 understanding of sustainable development has led to a search for new intervention strategies (mode-2 approaches) better suited to the management of new kinds of problems. In the Netherlands, especially since the "transition terminology" was embraced by the fourth National Environmental Policy plan (VROM, 2001), a range of intervention theories and tools has been developed and research programs set up. The hard and pioneering work done in this area has probably resulted in the disposition to approach "program theory evaluation" primarily from a *prescriptive* perspective. In Dutch discourse surrounding the evaluation of transition programs, we observe an inclination to assess whether the "program theory" (i.e., the expressed and/or enacted relationship between objectives, interventions, and outcomes) corresponds to the theoretical "blueprints" of transition management. However, in interviews and informal conversations that we held with a range of program managers and project leaders, they have expressed agitation and annoyance with being confronted repeatedly with the gaps between program theory and their practice. They argue that they know about the theory but struggle with the implementation and have expressed the need for help and guidance. Thus, extending Friedman's argumentation (2001), the challenges for program theory evaluation in the context of highly complex issues may rather be described as (a) learning why in practice the interventions may not adhere to some of the fundamental aspects of the intended program theory and (b) finding and monitoring ways to overcome these difficulties.

Challenges for Evaluating Mode-2 Strategies

To recapitulate, a distinction can be made between three types of monitoring and evaluating sustainable development, addressing the following questions: (a) *Where* do we stand with respect to sustainable development (generally or in relation to an intervention program)? (b) *What* interventions do we plan and execute and what are their outcomes? and (c) *Why* are certain interventions developed and implemented? (see Table 2). We have described the different approaches and their potential contribution to sustainable development briefly and elaborated on the challenges they meet when confronted with the increasing complexity, multilevel nature of the issues, and actor pluralism that a mode-2 view on sustainable development entails (see Table 2).

Below we will look at the implications of these challenges for mode-2 evaluation approaches. We do this by articulating guiding principles that emphasize essential aspects of mode-2 evaluation approaches. The guiding principles are based on our experience with one such approach: the ILA monitoring approach. The guiding principles modestly aim to contribute to the widely expressed

Table 2. Prevailing Approaches to Evaluating Sustainable Development

Focus	Approach	Features	Challenge
1. <i>Where do we stand?</i>	Progress assessment	Indicator-based modeling	Participative development of indicators Including social, institutional indicators
2. <i>What effects do interventions have?</i>	Performance measurement	Goal oriented, ex post	Acknowledging changing and unstable nature of declared sustainability goals Contribute to intervention strategy
3. <i>Why are certain interventions chosen?</i>	Program theory evaluation	Logic models	Accommodating differences in intended intervention theory and practice

need to develop innovative indicator frameworks. Both the nature and the content change in innovative indicator frameworks, henceforth called mode-2 evaluation frameworks.

ILA Monitoring in Practice: Focus, Tools, and Role of Evaluator

The ILA monitoring approach is an example of a mode-2 evaluation approach and is based on our most recent experiences with evaluating 11 projects with mode-2 aspirations, in the context of two Dutch system innovation programs on the sustainable development of agriculture. The transition program TransForum is set up as a 6-year program to support the sustainable development of agriculture in the Netherlands by bringing local practices (of farmers, growers, civil servants) together with scientists with the aim to induce changes in both science and practice (see Veldkamp et al., 2009). Different members of the Athena monitoring and evaluation team were involved in evaluating eight of TransForum's innovation projects as well as the TransForum program itself. The second context in which our research was carried out is that of the program "monitoring and evaluating networks" supported by the Dutch Ministry of Agriculture, Nature, and Food Quality. Four innovation projects (or practice networks in the language of the program) were monitored, two of which used ILA as primary approach. In addition, an independent project on regional innovation in sustainable development of agriculture supplements the set of 11 projects. In another section below, we will discuss the criteria we used for case selection. The 11 projects together have formed the "laboratory" in which we developed and tested our tools and articulated our focus. In the current section, we will use examples from the ILA monitoring approach to illustrate ways to address the challenges to evaluating sustainable development that were put forward in the previous section.

Focus: the Development and Practice of Interventions by Mode-2 Intermediaries

The first guiding principle for a mode-2 evaluation framework concerns the focus of analysis. Acknowledging the relevance of program theory evaluation, we state that a mode-2 evaluation approach focuses on the "inside" of efforts to stimulate sustainable development, that is, on the practices and intervention strategies developed and conducted by mode-2 intermediaries rather than the outside (e.g., the effects of these efforts or the cause of the problems). The reason for this is that developing and implementing mode-2 approaches to sustainable development is challenging. Writings on mode-2 approaches contain precautions on the complexity of the task, due to the context-dependent nature of sustainable development, the diversity of actors involved, and the many levels at which action is needed. A mode-2 approach not only requires a flexible and context sensitive methodology but also an interdisciplinary team with interactional expertise of the fields it operates in (Collins & Evans, 2002), competences in project, process, knowledge, and network management, and support from its commissionaires. What makes it particularly difficult, however, is the power of prevailing

modes of operation, institutional structures, and individual and organizational competence that constrains the required change. It is because of hampering conditions, changing circumstances, and unexpected effects that a mode-2 intervention strategy is inherently open to change—it cannot follow a blueprint. Similar notions are apparent in the literature on systems thinking in the context of evaluation (Cabrera, Colosi, & Lobdell, 2008; Churchman, 1970; Dyehouse et al., 2009) and CAS (Eoyang, 2007; Parsons, 2007). Reflecting and learning are therefore essential features of mode-2 approaches to sustainable development (see, e.g., Armitage, Marschke, & Plummer, 2008; Hendriks & Grin, 2007) and can be strengthened by corresponding approaches to evaluation (e.g., Burns, 2007; Imam, LaGoy, & Williams, 2007; Midgley, 2007).

Guiding Principle 1: on the Focus of Evaluation

Mode-2 evaluation focuses on the inner workings of intermediaries, in particular by supporting continuous learning about intervention strategies for sustainable development.

The ILA monitoring approach aims to help mode-2 intermediaries guide and refine their interventions for sustainable development in response to changing situations and observations. Thus, although the used mode-2 approaches to sustainable development by intermediaries constitute the focus of evaluation, we do not assess whether the program theory of the intermediary adheres to the main features of mode-2 approaches but how they are used, where it gets difficult, and why and how these obstacles are resolved. As Friedman (2001) observed, the mechanisms that govern the *implementation* of a program theory are not addressed in regular program theory evaluation and are a potential blind spot in the evaluator's field of vision. Guiding Principle 1 thus addresses the challenge faced by program theory evaluation when applied to mode-2 strategies (see Table 2).

Action research approaches (Burns, 2007; Hoes, Regeer, & Bunders, 2008; Reason & Bradbury, 1990; Zweekhorst, 2004; Kemmis & McTaggart, 1988) have provided evidence that following a cyclical intervention strategy, whereby tailor-made interventions follow observation of and reflection on previous interventions, is particularly suitable for dealing with the types of problems associated with sustainable development (see Figure 1). This leads to the second guiding principle.

Guiding Principle 2: on the Role of Evaluation

Monitoring and evaluation are intrinsically linked to the intervention process, by being part of the iterative process of defining, implementing, and adjusting interventions.

As such, mode-2 evaluation meets the challenge, faced by performance measurement approaches, to contribute to intervention strategies. Our experience has shown that evaluation contributes to this spiral of activities in at least three different ways, indicated by the grey marks in Figure 1. First, the evaluator may contribute to investigating the current state of affairs with respect to the challenge at stake (reconnaissance), by performing desk research, stakeholder interviews, or focus groups. Interviews with key players in mode-2 evaluation should be prepared not solely from a research perspective (gathering data) but from an intervention perspective (affecting actor views and relations) and should preferably be conducted in collaboration with the project team. Unlike Figure 1 suggests, we have found that suggesting to start with reconnaissance activities at the beginning is often not well received. Soon after the start of a project, teams often feel a much greater need to start executing part of the project plan, than to reflect on underlying assumptions. Reconnaissance activities positioned between the first and second action research cycle are rich in terms of experiences and observations brought in and are of great use for subsequent phases of plan and action. Thus, for productive evaluator-project team relations to develop, careful observation of and responding to the dynamics and expressed needs of the project team is important.

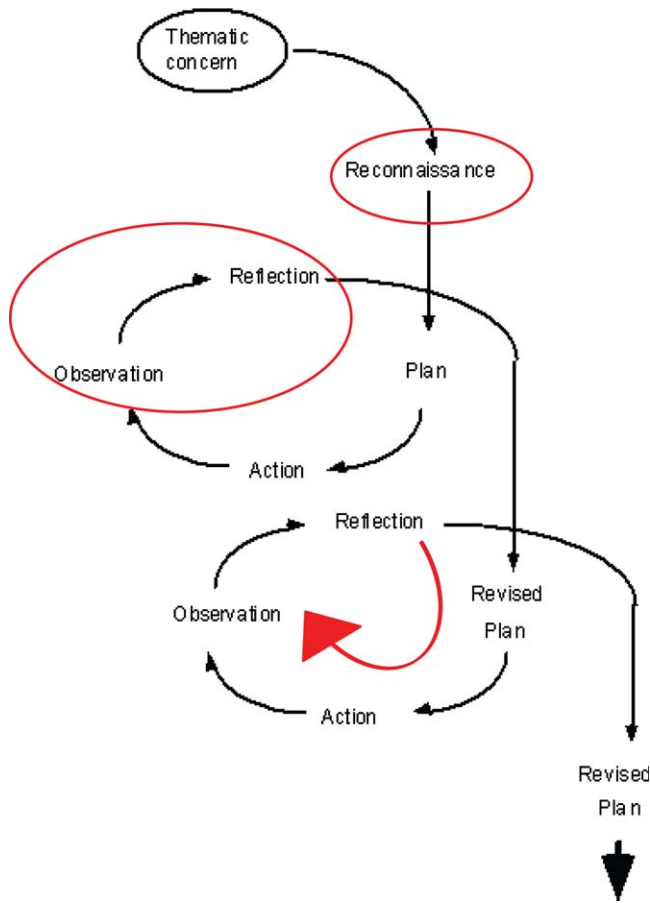


Figure 1. Adopted from Kemmis and McTaggart, 1988, and Adjusted

A second, commonly used, role for evaluation in the spiral of activities of a mode-2 intermediary specifically concerns the activities of observation and reflection (see Figure 1). As reflection and learning are recognized as essential aspects of any mode-2 strategy toward sustainable development, sponsors may request allocation of part of the budget to reflection activities, which is subsequently subcontracted. In practice, this strategy risks insufficient use of results, by considering observation and reflection as separable, and hence separate, activities. Emphasis, in terms of time and effort, is on gathering and analyzing data, with a steep learning curve on the part of the evaluator rather than the project team. Time spent on reporting back results and mutual reflection is often limited to one or two occasions.

The third way in which evaluation may contribute to the spiral of activities in a mode-2 project involves including the evaluator in all parts of the cycle. Especially in projects where participant observation is part of the methodology, regular interaction between team members (of which evaluator is one) enhances the possibilities for connecting observation and reflection to plan on one hand and action on the other hand. The emphasis in terms of time and effort is on interaction between evaluator and other project team members and the mutual rearticulation of observations, challenges, and strategies. Strong communication skills are needed for successful role sharing between evaluators and project teams, as the empirical study by Cartland, Ruch-Ross, Mason, and Donohue (2008) confirms.

Table 3. Examples of Terminologies that Indicate Mode-2 Aspirations

Mode-2 Aspirations: Projects State Their Focus on	Variety in Terminology
- Knowledge cocreation	“Building a transdisciplinary network,” “connecting the worlds of knowledge and practice”
- Different way of working	“Learning-by-doing,” “community of practice,” “new role for government”
- Bringing together a variety of actors	“Apply KOMBI-approach,” “bring together 5 O’s”
- Creating impact on system level	“Change knowledge infrastructure,” “change climate for decision making”

Creating a Framework for Evaluation

Thus, the scope of a mode-2 evaluation framework is set by guiding Principles 1 and 2. Focusing on the intervention strategies of mode-2 intermediaries helps them with their difficult task of stimulating sustainable development in a way that incorporates the increased complexity that a mode-2 view on sustainable development entails. Now that we have set the scope of the mode-2 evaluation framework, we can turn to the contents. What exactly is being evaluated in this reflexive and iterative manner, and how do we determine this? In other words, what are appropriate mode-2 indicators and how are they developed? We start with the latter, by formulating guiding Principle 3, explicitly addressing the first challenge faced by progress assessment approaches.

Guiding Principle 3: on the Process of Creating a Framework for Evaluation

Use a participative approach to the development of a mode-2 evaluation framework.

If evaluation aims to enable actors to learn about and enhance their practices, their learning objectives should be central to the framing of evaluation issues. Thus, the issues and concerns of stakeholders and end users constitute the start of the inquiry (see also Greene & Abma, 2001; Guba & Lincoln, 1989; Patton, 2008; Stake, 1967 for a comparable rationale). To show how this may work out in practice, we describe the way in which cases are selected in the ILA monitoring approach.

Case Selection: Mode-2 Aspirations. In selecting our cases, we find it important that they aim to contribute to a more sustainable way of producing food, organizing care, developing urban and regional spaces, or transporting food and flowers; the cases on which our examples are based are all in the area of the sustainable development of agriculture. Second, projects express an interest in the added value of evaluating the project. Third, the ILA monitoring approach is especially conducive to projects that express mode-2 aspirations, albeit not necessarily in those terms. Thus, although recent writings about systems-based evaluation particularly concern the use of systems concepts to broaden the practice of evaluation, whether applied to large programs or to the design of a questionnaire (Imam et al., 2007: 10), our focus is on evaluating projects or programs that express their belief in a systems-based approach to tough problems. Examples are given in Table 3.

In introductory interviews with project leaders, we go to great lengths at trying to understand their main concerns and current struggles. Throughout the discussion, we are looking for common ground and a terminology that suits the project. The terminology thus developed forms the starting point of the inquiry and shapes the preliminary focus of the evaluation. The concept of a mode-2 approach functions not as a framework for assessment, which would make it deficit oriented, but is used instead as a guide to inform our appreciative inquiry (see also Preskill & Catsambas, 2006).

We would like to note here the subtle but important difference between involving stakeholders and end users in the research process of the evaluator on one hand and a collaborative learning effort between evaluator and practitioners on the other. In the first case, the primacy is with the inquiry of the evaluator where including stakeholder perspectives leads to more complete and integral research results. In the second case, evaluation is instrumental to (and at the same time an integral part of) the learning process of the actors; the primacy is with the process of realizing sustainable development. The ILA monitoring approach aims for the latter.

The Content of a Mode-2 Evaluation Framework: from Indicators to Sensitizing Concepts. As established above, mode-2 evaluation needs to acknowledge the changing and unstable nature of declared goals. Sustainability projects or programs may even benefit from an unfixed, changeable formulation of goals so as to accommodate the multiple interests of concerned parties. Two earlier sustainability programs (Dutch National Initiative for Sustainable Development and Real Prosperity) evaluated by one of the authors came to similar conclusions: going into the field with a set of goals and definitions defined in advance did not generate the aspired movement toward sustainable development. In both cases, strategies were adjusted; the questions of, respectively, “What is sustainable urban development?” and “What is Real Prosperity?” became leading in interactions with actors concerned; they generated energy and aided the development of agency (Regeer & Bunders, 2007/2009). Thus, through an interactive process of deliberation and dialogue, new understandings of sustainable development were developed and aligned with local contexts and institutional contexts of participants. In our monitoring, we have found that these new understandings contain *sensitizing concepts*, capturing a specific quality of mode-2 indicators. Their function is not to assess but rather to perceive, or make visible, aspects that are or seem relevant to sustainable development. Moreover, the local articulation of mode-2 indicators stimulates sustainable development through discursive activity. Similarly, in the context of systems inquiry and evaluation, it is observed that “deeper meaning-making is more likely to promote valuable action than better data” (Imam et al., 2007: 8).

Guiding Principle 4: on the Nature of Indicators

Mode-2 indicators do not assess but sensitize.

Mode-2 indicators are context dependent and dynamic.

Note that articulating the local meaning of sustainable development is not limited to interaction between the evaluator and the mode-2 intermediary. Entering the dialogue with relevant stakeholders on their understanding of sustainable development is part of a mode-2 intervention strategy to sustainable development. As a result of evolving insights on the part of the mode-2 intermediary and due to an increased number of stakeholders and perspectives entering the interpretative space over time, mode-2 indicators necessarily change during the course of the project.

To further our understanding on the content of mode-2 indicators, in the next section, we will describe a specific tool we developed in the context of ILA monitoring: the dynamic learning agenda. It contains sensitizing concepts and aims to contribute to both the *learning by* the mode-2 intermediary and *learning about* intervention strategies for sustainable development. Two more guiding principles for mode-2 evaluation will be formulated.

Supporting Learning: the Dynamic Learning Agenda

As described above, the *focus* of our research is on intervention strategies of intermediaries with mode-2 aspirations; that is, temporary project teams or program organizations aiming for systems

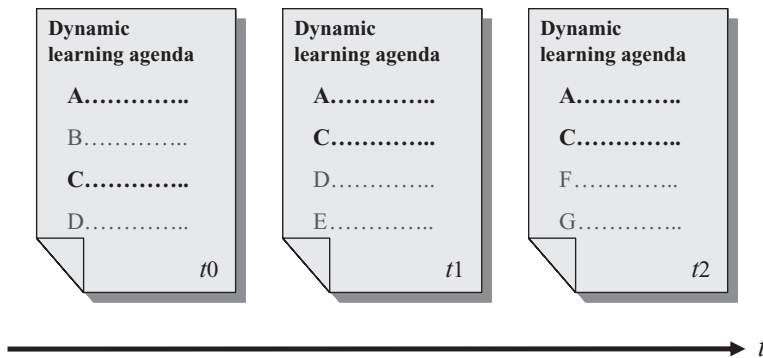


Figure 2. Dynamic Learning Agenda

change. The *goal* of our research is to strengthen the intervention strategy by articulating and generating feedback loops that enable actors to learn. A crucial question on the part of the evaluator is, “How can we capture learning experiences in such a way that the learning process of project participants is enhanced?”

Capturing learning experiences implies making an intangible process visible. The intervention process is a continuous and ongoing flow of decisions, observations, actions, thoughts, reflections, interactions, adjustments, and so on. Much like water; when you try to get hold of it, it disappears. At the same time, the reification of the learning experiences can serve the process (Czarniawska & Joerges, 1996). A (temporary) materialization of ideas and experiences may act as a point of reference, and on collective reflection, it can act as a tool for mutual sense making. In the context of ILA monitoring, we developed the dynamic learning agenda: a tool that both enhances the learning process and makes it tangible. Another technique that seeks to contribute to program development as well as communication is the most significant change (MSC) technique that involves the regular collection and participatory interpretations of “stories” of change (Dart & Davies, 2003).

The Dynamic Learning Agenda

Learning starts with articulating questions. The learning agenda contains the issues (formulated as questions) that a mode-2 intermediary struggles with, in the development and implementation of an intervention strategy for realizing sustainable development, at a specific moment in time. By constructing a sequence of learning agendas, the agenda becomes dynamic. Although we have used several variations of the dynamic learning agenda, the basic idea is depicted in Figure 2. Before we show and discuss some variations drawn from our cases, we will clarify how evaluators may approach creating a dynamic learning agenda.

Dynamic learning agendas can be constructed in different ways. Often, the first learning agenda is constructed by explicating the challenges, as they are implicitly discussed in a meeting on the progress of the project. The learning agenda then forms input for the next meeting and after some meetings may become a device to structure the meeting and discussions about progress. Learning agendas can also be constructed in interactive working sessions, using interactive group methods such as focus groups, open space, and so on. Furthermore, specific working sessions are organized addressing specific questions that remain on the agenda for a considerable amount of time. Evaluators take different roles in these activities and need corresponding skills, ranging from analytical capacities to facilitation skills. In terms of visualization, we have found that the full learning agenda at a particular moment in time usually looks like a low tech, plain list of questions, possibly clustered in categories (see Figure 4). A dynamic learning agenda, showing changes over time, is often organized

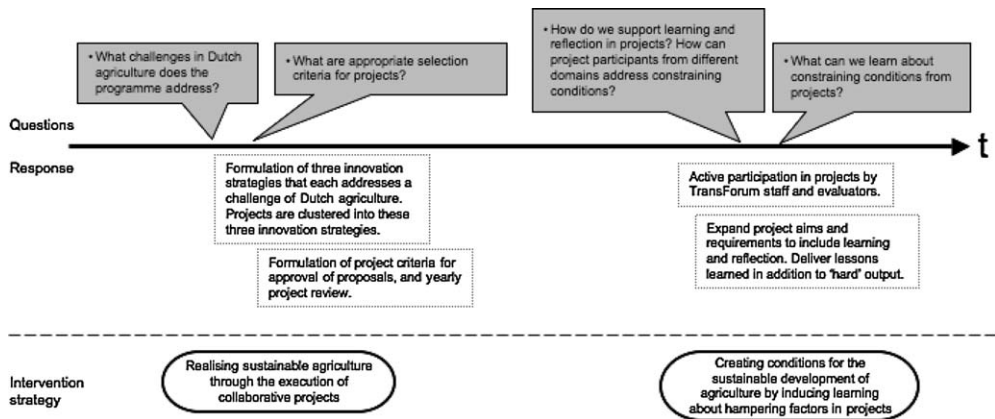


Figure 3. Selection of the Dynamic Learning Agenda of TransForum in the Period 2005–2006

around a single question or a set of connected questions and usually shows the actions taken in response to the questions (see Figure 3 for a selection of a dynamic learning agenda). Depending on the intended use of the dynamic learning agenda, it can be more or less comprehensive; possibly including major events in the environment of the project, the formulation and appropriation of questions by different stakeholders over time, or significant project outcomes.

As a result of our experimentation with learning histories (developed by Kleiner and Roth at the Center for Organizational Learning of the MIT in the 90s of the last century) in earlier stages of the ILA approach to evaluating sustainable development, the dynamic learning agenda shares some of the epistemological and methodological assumptions underlying learning histories. Dynamic learning agendas start from the specific and situational, and are constructed in a participative way. These features are in accordance with guiding Principle 3. Like learning histories, they should be “judged by the quality of the conversation they provoke” (Kleiner & Roth, 1996, p. 20). As such they are intended to contribute to the *sensitization* of participants to the issues that emerge as relevant, which corresponds to guiding Principle 4. By including perspectives of a wide range of participants (initiators, followers, and opponents of the sustainable development at stake), no one perspective on sustainable development is excluded in advance. Furthermore, the dynamic learning agenda is devised not as an end in itself but rather as a means toward learning and mutual sense making and is thus inextricably linked to the intervention process (guiding Principle 2). Finally, dynamic learning agendas are particularly aimed at bringing to light the difficult, tough issues that are normally “swept under the rug” (Kleiner and Roth, 1996: 14). We will elaborate on these persistent questions next.

A Living Archive of Challenges to Realizing Sustainable Development

The dynamic learning agenda helps to construct and discuss the main challenges ahead on a regular basis and thereby support the learning process of the mode-2 intermediary. In addition, they function as living archives of those challenges. Preliminary analysis of dynamic learning agendas shows that some issues remain on learning agendas for a considerably longer time than others. In Figure 2, questions A and C depict such persistent, tough questions. As the resolution of single-loop learning questions lies within the capacities of the practitioners (single-loop learning involves doing things better through incremental improvements of existing routines), they disappear from the agenda relatively easily. Double-loop learning questions, however, are particularly persistent (they involve change in underlying beliefs, norms, and assumptions) and exceptionally relevant to the challenges of sustainable development (for the original account of single and double-loop learning see Argyris & Schön,

1974, 1978). In a similar vein, humanistic philosopher Kunneman (2006) noted in his account of the existential state of contemporary societies that although tough questions (“trage vragen” in Dutch) may be shoved away under the table, from this subordinate position, they will continue to give importunate signals. According to Kunneman, these signals can become visible when there is room for exploration and even acceptance of differences between people and positions. Reflecting on the dynamic learning agenda in a safe environment aims exactly at that. Likewise, critical systems heuristics (e.g., Reynolds, 2007) and soft systems methodology (e.g., Attenborough, 2007) may help to get the “undiscussibles” in an evaluation (Imam et al., 2007).

On analyzing dynamic learning agendas, we have found that persistent questions appear in two forms. First, we often see a discrepancy between mode-2 intentions and practice. Armitage et al. (2008) have called this the learning paradox. They have observed that despite increased attention for the importance of learning in resource and environmental management literature, in practice it is used in an unspecific, vague sense. They contribute to resolving the paradox by a careful examination of literature on learning of emerging governance models in the context of environmental and resource management. The dynamic learning agenda is yet another answer to the paradox; it addresses the discrepancies when they arise and where they arise. Thus, the fifth guiding principle.

Guiding Principle 5: on the Contribution of Indicators to Learning

Mode-2 indicators (e.g., in the form of questions on a dynamic learning agenda) are conducive to finding strategies to accommodate the discrepancy between mode-2 intentions and mode-1 practice.

An example can be provided using the case of the TransForum program that is being evaluated using the ILA approach since 2005. The questions on the dynamic learning agenda of TransForum reveal changes in the intervention strategies over time that indicate an increased understanding of what it means to be a mode-2 intermediary (see Figure 3). To start with, TransForum-by-design adheres to the principles of a mode-2 intermediary: it aims to realize a sustainable agricultural sector, not through linear knowledge transfer but through collaboration among scientists, entrepreneurs, and government in real-life experiments (for more details about the program see Veldkamp et al., 2009). In its 1st year of operation much energy was put in setting the scope, creating coherence, and formulating criteria. Questions on the learning agenda were related to program coherence (what challenges in Dutch agriculture does the program address? What relevant links can we make between the projects and the scientific program?) and project criteria (what are appropriate criteria that can be used to approve project proposals?). Soon, however, it was found that setting appropriate criteria for projects (e.g., requesting application by consortia of actors) does not automatically lead to the desired results (i.e., actors were not able to overcome their respective interests and remained in their original domains). TransForum-in-action in its 1st year appears to adhere more to a regular mode-1 program in terms of its relation to projects. Tracking the formulation and reformulation of questions on the dynamic learning agenda shows that the more analytical “what” questions are replaced by posing “how” questions that are necessary to develop the ability to act in a meaningful way. The assertion that collaboration is guaranteed through project criteria changes over time to a series of questions addressing the process behind realizing collaboration between a diversity of actors (How can we support collaboration between different actors? How can we foster learning in projects to build the capacities necessary to deal with constraining conditions?). In terms of action, this meant that TransForum staff began to actively participate in projects to be part of the cocreation process. In the course of the program, the intervention strategy thus changed from “contributing to the sustainable development of agriculture through the execution of projects” to “creating conditions for the sustainable development of agriculture by inducing learning about

Figure 4. Learning Agenda Showing Constraining Conditions for Sustainable Regional Development**Learning agenda “project name” at “date”**

Project challenges

- How do we involve a broad range of actors, for example, including leisure or care professionals besides farmers and nature conservationists, each with their own cultures, language, interests, objectives?
- How do we create room to incorporate diverse views while at the same time maintaining focus and energy on the initial inspiration?

Contractor context

- How to deal with the tension between originally formulated project deliverables and a process of emergent design?

Academic context

- How can scientists adopt a more interventionist role whereas academic culture demands publications?
- If relevant knowledge questions are collectively defined, how can they be made relevant to the scientists field of research?

Political context

- How to align the long-term transition intended by the project with the yearly funding cycles of regional/local governments?
- How to align the long-term process with the 4-yearly political cycles of elections?

Entrepreneurial context

- How can the unique qualities of the region be capitalized on the market notwithstanding the farmers' long tradition of operating in highly subsidised markets and perhaps therefore limited perception among farmers of potential new value propositions?
- How to deal with the situation that the innovative business idea (e.g., combining sectors such as animal husbandry and care, or horticulture and energy, or agriculture and tourism) might be at odds with legal principles in these different domains?

hampering factors in projects” (see Regeer, Mager, Beekman, & Bunders, submitted for publication, for a more extensive description and current questions).

A second form in which persistent questions appear on dynamic learning agendas is as tension between a mode-2 approach and a mode-1 context. Funders, local politicians, or research managers may not understand the need for discursively shaping strategies for sustainable development in collaboration with previously unfamiliar actors. Moreover, the institutional structures in which they operate do not always allow for changing goals and shifting perspectives. Articulating these tensions carefully and reflecting on them gives the intermediary the opportunity to move beyond the initial aggravation (“they just don’t understand”) and beyond accepting the constraints as they are (“we could have been successful, but alas, the conditions weren’t right”). It opens up opportunities to develop strategies that create alignment between a mode-2 approach and a mode-1 context (“How can our project contribute to the needs of stakeholder X? What are their needs? And why?”).

The emerging body of literature on systems-based evaluation (e.g., Williams & Imam, 2007) has introduced the important notions that we must attend to the deeply held principles, values, and beliefs of people in systems and that we must broaden our inquiry from the particular situation of interest toward the broader *system* of interest, whose boundaries derive from human purpose. Critical systems heuristics provides a set of questions that helps describe the boundaries of a system, distinguish what “is” the case from what “ought” to be the case, and thereby inform new understandings and practices (see Reynolds, 2007). Similarly, we have found that the questions on the learning agendas of our cases reveal relevant boundaries (in terms of constraining conditions in the environment of the system) and help identify possible courses for action. Figure 4 depicts a typical learning agenda featuring questions that we found in three regional development projects in the area of sustainable agriculture. Various political, economical, and institutional factors appear to inhibit

Table 4. The Guiding Principles of a Mode-2 Evaluation Approach Linked to Challenges Identified

Guiding Principle	Challenge Met
<p><i>Guiding Principle 1: on the focus of evaluation</i> Mode-2 evaluation focuses on the inner workings of intermediaries, in particular by supporting learning about intervention strategies for sustainable development</p>	<p>Contributing to intervention strategy Accommodating differences in intended intervention theory and practice</p>
<p><i>Guiding Principle 2: on the role of evaluation</i> Evaluation is intrinsically linked to the intervention process, by being part of the iterative process of defining, implementing and adjusting interventions</p>	<p>Contributing to intervention strategy</p>
<p><i>Guiding Principle 3: on the process of creating a framework for evaluation</i> Use a participative approach to the development of a mode-2 evaluation framework</p>	<p>Participative development of indicators Including social, institutional indicators</p>
<p><i>Guiding Principle 4: on the nature of indicators</i> Mode-2 indicators do not assess but sensitize Mode-2 indicators are context dependent and dynamic</p>	<p>Acknowledging changing and unstable nature of declared sustainability goals</p>
<p><i>Guiding Principle 5: on the contribution of indicators to learning</i> Mode-2 indicators are conducive to finding strategies to accommodate the discrepancy between mode-2 intentions and mode-1 practice</p>	<p>Contributing to intervention strategy</p>
<p><i>Guiding Principle 6: on the contribution of indicators to learning</i> Mode-2 indicators make visible the different conditions that constrain sustainable development, for example, social, political, economical, and institutional factors, and aid their inclusion in strategies for sustainable development</p>	<p>Contributing to intervention strategy Including social, institutional indicators</p>

the intended course of action, the collective formulation of which allows project participants to broaden the scope and start working on these conditions (see also Broerse, 1998). This leads to the final guiding principle for a mode-2 evaluation framework.

Guiding Principle 6: on the Contribution of Indicators to Learning

Mode-2 indicators make visible the different conditions that constrain sustainable development, for example, social, political, economical, and institutional factors, and aid their inclusion in strategies for sustainable development.

Thus, we have shown that the dynamic learning agenda may contribute to the intervention strategy during the process, revealing and accommodating differences in the intended interventions strategy and actual practices. We also saw that it brings to the fore the conditions that constrain sustainable development as well as strategies to cope with these constraints.

In Table 4, the guiding principles of the mode-2 evaluation approach proposed in this article are summarized and related to the earlier mentioned challenges that are faced by common monitoring and evaluation approaches when applied to mode-2 strategies for sustainable development.

Conclusions

Global carbon emissions continue to rise and ecological footprint analyses show that we have exceeded the biosphere's natural carrying capacity. Moreover, a fair distribution of welfare and

quality of life across the world is far from being realized; rather developing countries are suffering from the effects of climate change to an unrepresentative degree. Policy intervention programs, grass roots movements, and multinational policy agreements have not been sufficiently successful in changing societal development substantially into a more sustainable direction. This urges the question: *If realizing sustainable development is so difficult, how can evaluation contribute to understanding and realizing sustainable development?*

In these final conclusions, we would like to identify the ways in which mode-2 evaluation as developed here complements prevailing types of monitoring and evaluation and accommodates some of its main challenges. A combination of elements from prevailing types of monitoring and evaluation, including addressing their challenges, along with elements from interactive approaches such as fourth generation evaluation (Guba & Lincoln, 1989), responsive evaluation (Greene & Abma, 2001), and appreciative inquiry (Preskill & Catsambas, 2006) results in a type of evaluation that is particularly appropriate for problems that arise in the context of the complexity of sustainability endeavor as we understand it now.

In this article, we have focused our attention on evaluating the intervention strategies for sustainable development that are currently emerging under the banner of mode-2 strategies. The guiding principles for a mode-2 evaluation framework as developed in this article start with a seemingly narrow focus, comparable to program theory evaluation: intervention strategies for sustainable development. Although strategies that incorporate a dynamic view on sustainable development are well described in literature, and programs are set up to stimulate mode-2 approaches to sustainable development, we still observe discrepancies between intentions and actual practice and between practice and context.

A mode-2 evaluation approach addresses this challenge by assisting mode-2 intermediaries in their efforts to stimulate sustainable development and is as such intrinsically linked to the intervention strategy. Mode-2 evaluation aims to enhance learning by elucidating discrepancies between intentions, practice, and context and encouraging the development of bridging strategies. Potentially, it also results in generic knowledge about appropriate and successful strategies for realizing sustainable development.

Furthermore, the mode-2 evaluation framework can complement the valuable information provided by current forms of progress assessment. Using the dynamic learning agenda tool in the context of ILA monitoring, we found that it is possible to track the spreading of persistent questions by regularly interviewing broader groups of stakeholders surrounding a mode-2 project. If mode-2 intervention strategies aim to induce learning (especially double loop) in actors at different levels, in different domains and places, the appropriation of persistent questions by a growing group of actors can be identified as a success. Moreover, monitoring the degree to which constraints in institutions are mitigated, adds to the body of knowledge on the “state” of sustainable development. However, above all, with the evaluation framework described in this article, we aim to enhance the efforts of the many practitioners who contribute to complex societal issues such as sustainable development in the face of continuing change.

Authors' Note

The strength of the ILA monitoring approach lies in its direct relation to ongoing intervention programs. Hence, further developing the approach was only possible due to the rare opportunity given to us by the TransForum program and by all participants of the 11 cases to monitor them “live” and “real-time,” for which we express great gratitude. We also thank our research partners at the Communication and Innovation Studies Group at Wageningen University and Research Center for valuable and inspiring discussions. The writing of this article would not have been possible without the partnership and support (financially and otherwise) of TransForum (particularly through the projects “Networked Learning and Learning from Networks” [KV-050] and Streamlining Greenport Venlo

[RO-023]) and the Dutch Ministry of Agriculture, Nature, and Food Quality (through the project “Monitoring and Evaluating Networks” [ond/2006/16/01]).

Declaration of Conflict Interests

The authors declared no conflicts of interest with respect to the authorship and/or publication of this article.

Funding

The Athena Institute received funds from TransForum and the Dutch Ministry of Agriculture, Nature and Food Quality for the execution of the research.

References

- Abma, T. (2001). Variaties in beleidsevaluaties: Meten, beoordelen en onderhandelen. In T. Abma & R. J. In 't Veld (Eds.), *Handboek beleidswetenschap* (pp. 311-320). Amsterdam: Boom.
- Argyris, C., & Schön, D. (1974). *Theory in practice: Increasing professional effectiveness*. San Francisco, CA: Jossey-Bass.
- Argyris, C., & Schön, D. A. (1978). *Organizational learning: A theory of action perspective*. Reading, MA: Addison Wesley.
- Armitage, D., Marschke, M., & Plummer, R. (2008). Adaptive co-management and the paradox of learning. *Global Environmental Change, 18*, 86-98.
- Attenborough, K. (2007). Soft systems in a hardening world: Evaluating urban regeneration. In B. Williams & I. Imam (Eds.), *Systems concepts in evaluation. An expert anthology* (pp. 75-88). Point Reyes, CA: EdgePress/American Evaluation Association.
- Beck, U. (1986). *Risikogesellschaft – Auf dem Weg eine andere Moderne*. Frankfurt: Suhrkamp Verlag.
- Becker, E., Jahn, T., Stiess, I., & Wehling, P. (1997). *Sustainability: A cross-disciplinary concept for social transformations* (Policy Papers No. 6). Paris: UNESCO.
- Bell, S., & Morse, S. (2004). Experiences with sustainability indicators and stakeholder participation: A case study relating to a ‘blue plan’ project in Malta. *Sustainable Development, 12*, 1-14.
- Broerse, J. E. W. (1998). *Towards a new development strategy. How to include small-scale farmers in the biotechnological innovation process*. Delft: Eburon.
- Bunders, J. F. G. (Ed.). (1990). *Biotechnology for small-scale farmers in developing countries. Analysis and assessment procedures*. Amsterdam: VU University Press.
- Bunders, J. F. G., & Broerse, J. E. W. (1991). *Appropriate biotechnology in small-scale agriculture: How to orient research and development*. Wallingford: CAB International.
- Burns, D. (2007). Evaluation in complex governance arenas: The potential of large system action research. In B. Williams & I. Imam (Eds.), *Systems concepts in evaluation. An expert anthology* (pp. 181-196). Point Reyes, CA: EdgePress/American Evaluation Association.
- Cabrera, D., Colosi, L., & Lobdell, C. (2008). Systems thinking. *Evaluation and Program Planning, 31*, 299-310.
- Cartland, J., Ruch-Ross, H. S., Mason, M., & Donohue, W. (2008). Role sharing between evaluators and stakeholders in practice. *American Journal of Evaluation, 29*, 460-477.
- Chen, H. (1990). Issues in constructing program theory. In L. Bickman (Ed.), *Advances in program theory: New directions for evaluation* (Vol. 47, pp. 7-18). San Francisco, CA: Jossey-Bass.
- Christie, C. A., & Alkin, M. C. (2003). The user-oriented evaluator’s role in formulating a program theory: Using a theory-driven approach. *American Journal of Evaluation, 24*, 373-385.
- Churchman, C. W. (1970). Operations research as a profession. *Management Science, 17*, B37-B53.
- Collins, H. M., & Evans, R. (2002). The third wave of science studies: Studies of expertise and experience. *Social Studies of Science, 32*, 235-296.

- Czarniawska, B., & Joerges, B. (1996). Travels of ideas. In B. Czarniawska & G. Sevón (Eds.), *Translating organizational change* (pp. 13-48). Berlin, New York: Walter de Gruyter.
- Dart, J., & Davies, R. (2003). A dialogical, story-based evaluation tool: The most significant change technique. *American Journal of Evaluation*, 24, 137-155.
- De Lancer Julnes, P. (2006). Performance measurement: An effective tool for government accountability? *Evaluation*, 12, 219-235.
- Dyehouse, M., Bennet, D., Harbor, J., Childress, A., & Dark, M. (2009). A comparison of linear and systems thinking approaches for program evaluation illustrated using the Indiana Interdisciplinary GK-12. *Evaluation and Program Planning*, doi:10.1016/j.evalprogplan.2009.03.001.
- Eoyang, G. H. (2007). Human systems dynamics: Complexity-based approach to a complex evaluation. In B. Williams & I. Imam (Eds.), *Systems concepts in evaluation. An expert anthology* (pp. 123-140). Point Reyes, CA: EdgePress/American Evaluation Association.
- Forrest, J. (2007). Evolutionary and behavioral characteristics of systems. In B. Williams & I. Imam (Eds.), *Systems concepts in evaluation. An expert anthology* (pp. 197-210). Point Reyes, CA: EdgePress/American Evaluation Association.
- Friedman, V. J. (2001). Designed blindness: An action science perspective on program theory evaluation. *American Journal of Evaluation*, 22, 161-181.
- Friedman, V. J., Rothman, J., & Withers, B. (2006). The power of why: Engaging the goal paradox in program evaluation. *American Journal of Evaluation*, 27, 201-218.
- Funtowicz, S. O., & Ravetz, J. R. (1993). Science for the post-normal age. *Futures*, 25, 739-755.
- Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., & Trow, M. (1994). *The new production of knowledge: The dynamics of science and research in contemporary societies*. London: Sage.
- Giddens, A. (1984). *The constitution of society. Outline of the theory of structuration*. Cambridge: Polity.
- Gore, A. (2006). *An Inconvenient Truth: The Planetary Emergency of Global Warming and What We Can Do About It*. Rodale Press in Emmaus, Pennsylvania.
- Greene, J. C., & Abma, T. A. (2001). Editors' notes, special issue: Responsive evaluation. *New Directions for Evaluation*, 2001, 1-6.
- Guba, E. G., & Lincoln, Y. (1989). *Fourth generation evaluation*. Newbury Park, London, New Delhi: Sage.
- Hajer, M. A. (2003). Policy without polity? Policy analysis and the institutional void. *Policy Sciences*, 36, 175-195.
- Hendriks, C. M., & Grin, J. (2007). Contextualizing reflexive governance: The politics of dutch transitions to sustainability. *Journal of Environmental Policy & Planning*, 9, 333-350.
- Hildén, M., & Rosenström, U. (2008). The use of indicators for sustainable development. *Sustainable Development*, 16, 237-240.
- Hisschemöller, M., & Hoppe, R. (1996). Coping with intractable controversies: The case for problem structuring in policy design and analysis. *Knowledge and Policy*, 8, 40-60.
- Hoes, A.-C., Regeer, B. J., & Bunders, J. F. G. (2008). Transformers in knowledge production: Building science-practice collaborations. *Action Learning: Research & Practice*, 5, 207-220.
- Imam, I., LaGoy, A., & Williams, B. (2007). Introduction. In B. Williams & I. Imam (Eds.), *Systems concepts in evaluation. An expert anthology* (pp. 3-10). Point Reyes, CA: EdgePress/American Evaluation Association.
- Jasanoff, S. (Ed.). (2004). *States of knowledge. The co-production of science and social order*. London, New York: Routledge.
- Kemmis, S., & McTaggart, R. (1988). *The action research planner*. Victoria: Deakin University Press.
- Kemp, R., Parto, S., & Gibson, R. B. (2005). Governance for sustainable development: Moving from theory to practice. *International Journal of Sustainable Development*, 8, 12-30.
- King, J. A. (2008). Bringing evaluative learning to life. *American Journal of Evaluation*, 29, 151-155.
- Kleiner, A., & Roth, G. (1996). *Field manual for the learning historian*. MIT, Center for Organizational Learning.

- Kunneman, H. (2006). *Voorbij het dikke-ik. Bouwstenen voor een kritisch humanisme*. Humanistics University Press.
- Lehtonen, M. (2005). OECD environmental performance review programme. *Evaluation*, 11, 169-188.
- Loorbach, D. (2007). *Transition management. New mode of governance for sustainable development*. Utrecht: International Books.
- Marra, M. (2004). The contribution of evaluation to socialization and externalization of tacit knowledge. *Evaluation*, 10, 263-283.
- McCoy, M., & Hargie, O. D. W. (2001). Evaluating evaluation: Implications for assessing quality. *International Journal of Health Care Quality Assurance*, 14, 317-327.
- Meadows, D., Meadows, D., & Randers, J. (2004). *Limits to growth. The 30-year update*. White River Jct., Vermont: Chelsea Green Publishing.
- Meadows, D. H., Meadows, D. L., Randers, J., & Behrens, W. W., III. (1972). *The limits to growth*. New York: Universe Books.
- Midgley, G. (2007). Systems thinking for evaluation. In B. Williams & I. Imam (Eds.), *Systems concepts in evaluation. An expert anthology* (pp. 11-34). Point Reyes, CA: EdgePress/American Evaluation Association.
- Nowotny, H., Scott, P., & Gibbons, M. (2001). *Re-thinking science. Knowledge and the public in an age of uncertainty*. Oxford: Polity.
- Parsons, B. A. (2007). The state of methods and tools for social systems change. *American Journal of Community Psychology*, 39, 405-409.
- Patton, M. (2008). *Utilization-focused evaluation* (4th ed.). Thousand Oaks, CA: Sage.
- Perrin, B. (2002). How to – and how not to – evaluate innovation. *Evaluation*, 8, 13-28.
- Preskill, H., & Catsambas, T. T. (2006). *Reframing evaluation through appreciative inquiry*. Thousand Oaks, CA, London, New Delhi: Sage.
- Reason, P., & Bradbury, H. (1990). *Handbook of action research*. London: Sage.
- Regeer, B. J., & Bunders, J. F. G. (2007/2009). *Knowledge cocreation: Interaction between science and society. A transdisciplinary approach to complex societal issues* [translated from Dutch]. Den Haag: RMNO/COS.
- Regeer, B. J., Mager, S., Beekman, V., & Bunders, J. F. G.. New approaches to the sustainable development of agriculture. TransForum – The case of a mode-2 intermediary. Manuscript submitted for publication 2009.
- Reynolds, M. (2007). Evaluation based on critical systems heuristics. In B. Williams & I. Imam (Eds.), *Systems concepts in evaluation. An expert anthology* (pp. 101-122). Point Reyes, CA: EdgePress/American Evaluation Association.
- Rosas, S. R. (2005). Concept mapping as a technique for program theory development: An illustration using family support programs. *American Journal of Evaluation*, 26, 389-401.
- Rotmans, J., Kemp, R., & Asselt, M. v. (2001). More evolution than revolution. Transition management in public policy. *Foresight*, 3, 15-31.
- SenterNovem. (2007, June 05). Bsik: Investing in knowledge infrastructure. Retrieved October 15, 2008, from http://www.senternovem.nl/english/products_services/encouraging_innovation/bsik_investing_in_knowledge_infrastructure.asp
- Stake, R. E. (1967). The countenance of educational evaluation. *Teachers College Record*, 68, 523-540.
- Veldkamp, A., Altvorst, A. C. v., Eweg, R., Jacobsen, E., Kleef, A. v., Latesteijn, H. v., et al. (2009). Triggering transitions towards sustainable development of the Dutch agricultural sector: TransForum's approach. *Agronomy for Sustainable Development*, 29, 87-96.
- Voß, J.-P., Bauknecht, D., & Kemp, R. (2006). *Reflexive governance for sustainable development*. Cheltenham UK: Edward Elgar.
- Voß, J.-P., Kemp, R., & Bauknecht, D. (2006). Reflexive governance: A view on an emerging path. In J.-P. Voss, D. Bauknecht & R. Kemp (Eds.), *Reflexive governance for sustainable development*. Cheltenham, UK, Northampton, MA, USA: Edward Elgar.

- VROM. (2001). *Een wereld en een wil: Werken aan duurzaamheid: Nationaal milieubeleidsplan 4*. Den Haag: Ministerie Volkshuisvesting, Ruimtelijke Ordening en Milieubeheer, Centrale Directie Communicatie.
- Walker, G., & Kubisch, A. C. (2008). Evaluating complex systems-building initiatives: A work in progress. *American Journal of Evaluation, 29*, 494-499.
- WCED. (1987). *Our common future*. UK: Oxford University Press.
- Williams, B., & Imam, I. (Eds.). (2007). *Systems concepts in evaluation. An expert anthology*. Point Reyes, CA: EdgePress/American Evaluation Association.
- Yampolskaya, S., Nesman, T. M., Hernandez, M., & Koch, D. (2004). Using concept mapping to develop a logic model and articulate a program theory: A case example. *American Journal of Evaluation, 25*, 191-207.
- Zweekhorst, M. B. M. (2004). *Institutionalising an interactive approach to technological innovation. The case of the Grameen Krishi Foundation*. Amsterdam: Vrije Universiteit.