Six guiding principles for evaluating mode-2 strategies for sustainable development

B.J. Regeer*, A-C Hoes*, M. van Amstel-van Saane*, J.F. Caron-Flinterman* and J.F.G. Bunders*
* Athena Institute, VU University Amsterdam, The Netherlands
º Animal Science Group, Wageningen UR, The Netherlands

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

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 CO$_2$ levels or by measuring effects on student learning as a result of educational reform. As Jean 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 programmes and thereby foster societal change. Evaluators can choose from, and skill themselves in, a wide range of available methods and techniques. But what if the problems to be addressed, and the corresponding programmes 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 programmes 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 than 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 organisations; 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. Realising 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 programmes 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) en Bunders en Broerse (1991), aimed at supporting more inclusive (agricultural, health and (bio)technological) 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 programme evaluation and programme theory evaluation) and articulate the challenges they meet when confronted with the complexity of designing and conducting mode-2 intervention programmes 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 programme activities at particular moments in time. The specific approach we developed however, has been referred to as the Interaction Learning and Action (ILA) monitoring approach, which, for the sake of clarity, we will maintain.

In order 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 (D. Meadows, Meadows, & Randers, 2004; D. H. Meadows, Meadows, Randers, & Behrens III, 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, 1987). Whilst there is a widespread understanding 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 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, Stiess, & 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, realised 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 employing 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 programmes 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 emphasised 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 programme 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 multi-level 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 co-creation, and developing problem solving strategies for sustainable development. Mode-2 intermediaries can be individual actors or institutionalised organisations, but most often we mean project teams or
temporary programme organisations set up to create an impact on sustainable
development. We also use the mode-2 prefix when we refer to emerging approaches to
evaluation.

Table 1. Features of mode-1 and mode-2 strategies

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

Some characteristics of the mode-2 strategies, in relation to characteristics of mode-1 strategies, are summarised in Table 1. Whilst mode-1 strategies would focus on seeking solutions for relatively concrete problems employing well-established methodologies, mode-2 strategies emphasise the creation of conditions for structuring problems and supporting the search for possible directions for solutions (Regeer & Bunders, 2009). Particularly in the case of complex problems, characterised 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 offers diverse experiences to learn from and evaluate (J-P. 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 in order 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 of the realm of the usual behaviour (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 behaviour of agents (see also writings on Complex Adaptive Systems in relation to evaluation, e.g. Eoyang, 2007; Parsons, 2007). While fundamental changes are required, existing structures cannot be ignored, overridden or simply by-passed. 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 optimising intervention strategies, the difficult route towards 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 programmes 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, multi-levelness 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 & 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 programmes were set up to assess national or global trends and facts related to sustainable development (progress assessment). Other evaluation projects and programmes have been launched with the purpose to assess the goals and results of interventions aimed at contributing to sustainable development (goal-oriented programme evaluation). Finally, there are studies that aim to investigate and assess the (often implicit) theories that underlie those interventions (programme 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 programmes 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. IPCC, Millennium Ecosystem Assessment, Biodiversity Outlook). They can be deployed in general or in relation to an intervention programme. 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 behaviour. Al Gore has been able to draw on many different studies, models and empirical results to support his ‘Inconvenient Truth’.

While 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, Jahn, Stiess, & Wehling, 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. (1997) 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, Parto, & Gibson, 2005). The challenge to developing innovative indicators thus has a process and a content dimension.

**Goal-oriented Programme 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 programmes (e.g. Millennium goals, Kyoto protocol), citizen initiatives (e.g. Local Agenda 21, Ecoteams), NGO collaborations (e.g. Anti globalisation 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 lead to a growth of performance measurement and (goal-oriented) programme 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 programmes and/or corporate initiatives, in terms of achievement of predetermined goals and objectives, after the programme has been completed (i.e. summative); resulting in amongst others annual sustainability reports of companies based on the GRI indicators, or the Environmental Performance Review.
The dominant view on programme 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 programme 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 non-intended and non-expected 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 programme; more emphasis is placed on incorporating evaluation in the intervention process (see also Friedman, 2001). Programme evaluation that focuses on evaluating the programme itself with the purpose of assisting in making timely adjustments (formative) rather than focusing on the end result of the programme (summative) seems expedient in cases of mode-2 approaches to sustainable development and other complex problems alike.

Another important challenge of programme 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 recognised 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 programme itself, rather than settling down with an incomplete or inadequate set of goals and measuring the outcomes of the programme 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 et al. (2006) advocate to make goal setting a process of systematic and participative inquiry into goals, particularly in cases of controversy or dissent.

**Programme theory evaluation**

In contrast to goal-oriented programme evaluation, 'programme theory evaluation' focuses on understanding how programmes 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 programme is specified in terms of its inputs, its expected outcomes, the assumed relationships between them, and the underlying mechanisms relating expected programme ends and means to each other (Chen, 1990). These mechanisms, explaining how programme 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). Programme 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 programme teams (e.g. Christie & Alkin, 2003; Patton, 2008).

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

First, programme theory evaluation presupposes a relatively stable programme, 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 characterised 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 towards sustainable development, (characterised 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 programme. As Complex Adaptive Systems (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 programmes are concerned (e.g. Dyehouse, Bennet, Harbor, Childress, & Dark, 2009; Forrest, 2007; Patton, 2008; e.g. 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 programme theory and programme practice. Actors are frequently unable to behave in ways consistent with their espoused theories. Thus, even if programme participants’ espoused change theory corresponds to academic theory on complex, multi-level transitions, we can expect that in practice actors will not act accordingly. This has consequences for the use of programme theory in evaluation of mode-2 strategies.

Let us illustrate by means of the current Dutch discourse. As described before, 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 programmes set up. The hard and pioneering work done in this area has probably resulted in the disposition to approach 'programme theory evaluation' primarily from a prescriptive perspective. In Dutch discourse surrounding the evaluation of transition programmes we observe an inclination to assess whether or not the 'programme theory' (that is, 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 we held with a range of programme managers and project leaders, they have expressed agitation and annoyance with being confronted repeatedly with the gaps between programme theory and their practice. They argue 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 programme 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 programme 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 questions: 1) *Where* do we stand with respect to sustainable development (generally or in relation to an intervention programme)?; 2) *What* interventions do we plan and execute and what are their outcomes?; 3) *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, multi-level nature of the issues and actor pluralism that a mode-2 view on sustainable development entails (see Table 2).

<table>
<thead>
<tr>
<th>Focus</th>
<th>M&amp;E</th>
<th>Features</th>
<th>Challenge</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. <em>Where</em> do we stand?</td>
<td>Progress assessment</td>
<td>Indicator-based, modelling</td>
<td>- Participative development of indicators - Including social, institutional indicators</td>
</tr>
<tr>
<td>2. <em>What</em> effects do interventions have?</td>
<td>Performance measurement</td>
<td>Goal oriented, ex-post</td>
<td>- Acknowledging changing and unstable nature of declared sustainability goals - Contribute to intervention strategy</td>
</tr>
</tbody>
</table>

Below we will look at the implications of these challenges for mode-2 evaluation approaches. We do this by articulating guiding principles that emphasise essential aspects of mode-2 evaluation approaches. The guiding principles are based on our experience.
with one such approach: the Interactive Learning & Action (ILA) monitoring approach. The guiding principles modestly aim to contribute to the widely expressed 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 programmes on the sustainable development of agriculture. The transition programme TransForum is set up as a six year programme 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., 2008). Different members of the Athena monitoring & evaluation team were involved in evaluating eight of TransForum’s Innovation projects as well as the TransForum programme itself. The second context in which our research was carried out is that of the programme ‘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 programme) 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 programme 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, it also requires 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 organisational 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,
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, while the employed mode-2 approaches to sustainable development by intermediaries constitute the focus of evaluation, we do not assess whether the programme theory of the intermediary adheres to the main features of mode-2 approaches, but how they are employed, where it gets difficult, and why, and how these obstacles are resolved. As Friedman (2001) observed: the mechanisms that govern the implementation of a programme theory are not addressed in regular programme theory evaluation and are a potential blind spot in the evaluator’s field of vision. Guiding principle 1 thus addresses the challenge faced by programme theory evaluation when applied to mode-2 strategies (see Table 2).

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 red 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 upon underlying assumptions. Reconnaissance activities positioned between the
first and second action-research cycle are rich in terms of experiences and observations brought in and of great use for subsequent phases of plan and action. Thus, in order 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.

A second, commonly employed, role for evaluation in the spiral of activities of a mode-2 intermediary, specifically concerns the activities of observation and reflection (see fig. 1). As reflection and learning are recognised as essential aspects of any mode-2 strategy towards 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 analysing 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 on the one hand to plan 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 (re)articulation 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 et al. (2008) confirms.

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 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*

Employ 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, organising 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. Secondly, projects express an interest in the added value of evaluating the project. And thirdly, the ILA monitoring approach is especially conducive to projects that express mode-2 aspirations, albeit not necessarily in those terms. Thus, whilst recent writings about systems based evaluation particularly concern the use of systems concepts to broaden the practice of evaluation, whether applied to large programmes or to the design of a questionnaire (Imam, LaGoy & Williams 2007: 10), our focus is on evaluating projects or programmes that express their belief in a systems based approach to tough problems. Examples are given in Table 3.

*Table 3. Examples of terminologies that indicate mode-2 aspirations*

<table>
<thead>
<tr>
<th>Mode-2 aspirations: projects state their focus on:</th>
<th>Variety in terminology</th>
</tr>
</thead>
<tbody>
<tr>
<td>- knowledge co-creation</td>
<td>‘building a transdisciplinary network’, ‘connecting the worlds of knowledge and practice’</td>
</tr>
<tr>
<td>- different way of working</td>
<td>‘learning-by-doing’, ‘community of practice’, ‘new role for government’</td>
</tr>
<tr>
<td>- bringing together a variety of actors</td>
<td>‘apply KOMBI-approach’, ‘bring together 5 O’s’</td>
</tr>
<tr>
<td>- creating impact on system level</td>
<td>‘change knowledge infrastructure’, ‘change climate for decision making’</td>
</tr>
</tbody>
</table>

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 the 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 realising sustainable development. The ILA monitoring approach aims for the latter.

*The content of a mode-2 evaluation framework: from indicators to sensitising concepts*

As established above, mode-2 evaluation needs to acknowledge the changing and unstable nature of declared goals. Sustainability projects or programmes may even benefit from an unfixed, changeable formulation of goals so as to accommodate the multiple interests of concerned parties. Two earlier sustainability programmes (NIDO-DSV 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 towards 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, 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 *sensitising 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, LaGoy & Williams 2007: 8).

*Guiding principle 4: on the nature of indicators*

- Mode-2 indicators do not assess but sensitise
- 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 sensitising 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; i.e. temporary project teams or programme organisations aiming for systems 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, etc. 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) materialisation of ideas and experiences may act as a point of reference, and upon 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 programme development as well as communication is the Most Significant Change (MSC) technique, which involves the regular collection and participatory interpretations of “stories” of change (Dart, 2003).

The Dynamic Learning Agenda

Learning starts with articulating questions. The learning agenda contains the issues (formulated as questions) a mode-2 intermediary struggles with, in the development and implementation of an intervention strategy for realising sustainable development, at a specific moment in time. By constructing a sequence of learning agendas, the agenda becomes dynamic. While we have used several variations of the Dynamic Learning Agenda, the basic idea is depicted in Figure 2. Before we will 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, etc. Furthermore, specific working sessions are organised 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 visualisation 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 organised 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, 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 sensitisation 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 towards 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 realising 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 & Schon, 1974; Argyris & Schon, 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 at exactly 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, LaGoy & Williams, 2007).

On analysing 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 employed 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 programme, which is 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 realise 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 programme see Veldkamp et al. 2008). In its first year of operation much energy was put in setting the scope, creating coherence and formulating criteria. Questions on the learning agenda were related to programme coherence (What challenges in Dutch agriculture does the programme address? What relevant links can we make between the projects and the scientific programme?) 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 first year appears to adhere more to a regular mode-1 programme 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 realising collaboration between a diversity of actors (How can we support collaboration between different actors? How can we foster learning in projects in order 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 in order to be part of the cocreation process. In the course of the programme 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 hampering factors in projects’ (see Regeer, Mager, Beekman, & Bunders, subm for a more extensive description and current questions).

![Dynamic Learning Agenda of TransForum](image)

**Figure 3.** Selection of the Dynamic Learning Agenda of TransForum in the period 2005-2006.

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 upon 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?").

<table>
<thead>
<tr>
<th>Learning Agenda “Project name” at “date”</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Project challenges</strong></td>
</tr>
<tr>
<td>- How do we involve a broad range of actors, e.g. including leisure or care professionals besides farmers and nature conservationists, each with their own cultures, language, interests, objectives?</td>
</tr>
<tr>
<td>- How do we create room to incorporate diverse views whilst at the same time maintaining focus and energy on the initial inspiration?</td>
</tr>
<tr>
<td><strong>Contractor context</strong></td>
</tr>
</tbody>
</table>
• 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 four-yearly political cycles of elections?

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

Figure 4. Learning agenda showing constraining conditions for sustainable regional development

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 towards 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 & Vince, 2004). 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 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, e.g., 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 summarised 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.

Table 4: The guiding principles of a Mode-2 evaluation approach linked to challenges identified.

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

**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 realised; rather developing countries are suffering from the effects of climate change to an unrepresentative degree. Policy intervention programmes, grass roots movements, and multi-national policy agreements have not been sufficiently successful in changing societal development substantially into a more sustainable direction. This urges the question: *If realising sustainable development is so difficult, how can evaluation contribute to understanding and realising 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 endeavour 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 programme theory evaluation: intervention strategies for sustainable development. Although strategies that incorporate a dynamic view on sustainable development are well described in literature, and programmes 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 realising sustainable development.

Furthermore, the mode-2 evaluation framework can complement the valuable information provided by current forms of progress assessment. Employing the Dynamic Learning Agenda tool in the context of ILA monitoring we found 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. But 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.

References


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.
complex evaluation. In B. Williams & I. Imam (Eds.), Systems concepts in
evaluation. An expert anthology (pp. 123-140). Point Reyes, CA:
& I. Imam (Eds.), Systems concepts in evaluation. An expert anthology (pp. 197-
Friedman, V. J. (2001). Designed Blindness: An Action Science Perspective on Program
201-218.
755.
The new production of knowledge: The dynamics of science and research in
Cambridge: Polity.
Policy Sciences, 36, 175-195.
of Dutch Transitions to Sustainability. Journal of Environmental Policy &
Planning, 9(3), 333-350.
Sustainable Development, 16(4), 237-240.
for problem structuring in policy design and analysis. Knowledge and Policy,
8(4), 40-60.
Practice, 5(3), 207-220.
(Eds.), Systems concepts in evaluation. An expert anthology (pp. 3-10). Point
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(1/2), 12-30.


