
Journal of Public and Nonprofit Affairs

Vol. 11, No. 1

Evaluation and Transformative Social Innovation: A Call for New Habits

Mieke Berghmans - KU Leuven

Joke Vandenabeele - KU Leuven

In this article we argue that dominant measurement and evaluation methods reduce the transformative potential of social innovation initiatives. Taking a pragmatist perspective on transformative social innovation and drawing on a commonly used distinction between simple, complicated and complex intervention logics, we demonstrate that habitually used monitoring and evaluation methods are often supported by a simple or complicated intervention logic. These intervention logics and related understandings of social change and knowledge production are incompatible with the ambition to realise transformation. Less dominant methods, which often are supported by a complex intervention logic, seem to be more apt, especially when they do not focus on adaptation alone, for the monitoring and evaluation of transformative social innovation initiatives. On the basis of this analysis, we plead for new habits of monitoring and evaluation and formulate an agenda for further action and research.

Keywords: Transformation – Social innovation – Monitoring and evaluation
Pragmatism

Introduction

Practices and theories of social innovation are not new. Historically, social innovation can for example be observed in the development of cooperatives as an alternative to rampant 19th century capitalism, the initiation of neighborhood committees in response to the new urban question around the 1970's, or commoning initiatives, like community supported agriculture at the end of the 20th century. People have, in other words, time and again - without necessarily using the term 'social innovation' – deliberated, pursued, and achieved sustainable change (McGowan & Westley, 2015; Moulaert et al., 2017; Moulaert & MacCallum, 2019). While in the past social innovation was predominantly initiated and stimulated at the civil society or local community level, in recent decades government policy makers have become strong supporters of social innovation (Moulaert et al., 2017; Reynolds, et al., 2017). In a world that is confronted with "wicked problems", limited resources and ingenuity gaps" (McGowan & Westley, 2015, p. 54), social innovation has become an attractive pathway for dealing with complex societal challenges to which existing practices, rules and institutions do not seem to find sustainable answers. In this context, government policy makers have started to promote policies, regulatory frameworks, funding mechanisms (Nicholls & Edmiston, 2018) as well as research agendas (Moulaert & Mehmood, 2020) which support and stimulate social innovation.

The adoption of social innovation in governmental policy agendas on regional, national and local levels is met with scholarly scrutiny and critique. Scholars suggest that the transformative potential of social innovation initiatives is likely to weaken when these initiatives become part

and parcel of neoliberal policy strategies. The potential of bottom-up social innovation by civil society easily gets overlooked (Moulaert et al., 2017; Wright & Nyberg, 2019) as many of the contemporary social innovation policies focus on market creation and social business. Social innovation becomes co-opted in a hegemonic market discourse which depoliticizes problem framings and promotes a neoliberal agenda of economic growth (Fougère et al., 2017; Tesfaye & Fougère, 2022; Wittmayer et al., 2019). Moreover, as policy administrators promote dominant accountability mechanisms, which focus on metrics and indicators, social innovation initiatives are forced into dominant cognitive frameworks and institutional structures thus running the risk of inscribing themselves into the dominant order rather than transforming it (Moulaert et al., 2017; Novy, 2017; Oosterlynck et al., 2019; Reynolds et al., 2017).

In this article we delve deeper into this last issue: the idea that the transformative potential of social innovation initiatives gets reduced through the use of dominant monitoring and evaluation tools. This concern is often mentioned in the emerging body of scholarly work on transformative social innovation. Scholars agree that there exists a mismatch between dominant measurement and evaluation methods and the transformative potential of social innovation initiatives (Kok et al., 2023, Moulaert et al., 2017; Novy, 2017; Oosterlynck et al., 2019; Reynolds et al., 2017). In spite of a general consensus that this mismatch exists, little attention has however been paid to explaining precisely how these dominant measurement and evaluation tools reduce or inhibit the transformative potential of social innovation initiatives. Some scholars have pointed out that this mismatch results from an incompatibility between social innovation projects and conventional monitoring and evaluation tools. They indicate that the former consist of uncertain and complex processes that have their own “dynamics and multidimensional impacts” (Antadze & Westley, 2012, p. 143), whereas the latter rely on linear, cause-effect relationships and focus on the realisation of a plannable result (Antadze & Westley, 2012; Cunha & Benneworth, 2020; Oosterlynck et al., 2019). Others have suggested that conventional monitoring and evaluation tools promote a knowledge production which is “external to social innovation dynamics” and thus “reduce[s] the latter’s transformative potential” (Oosterlynck et al. 2019, p. 225).

This article builds further on these observations. It presents a comprehensive analysis of how dominant measurement and evaluation methods may reduce the transformative potential of social innovation initiatives. On the basis of this analysis this article develops a plea for alternative habits of monitoring and evaluation that could contribute to (rather than hinder) transformative change in social innovation projects. It also formulates an action and research agenda for those who engage with monitoring and evaluation in transformative social innovation projects.

This article is structured as follows. We first briefly describe transformative social innovation. We point out the difference between transformative and non-transformative social innovation and identify some key characteristics and conditions for transformative social innovation. Drawing on a pragmatist perspective, we argue that transformative social innovation implies changing social habits and infrastructures, and, as such, requires intricately connected processes of collective action and collective learning. Next, we make a detailed exploration of the many monitoring and evaluation tools that are habitually used to assess the progress and impact made in social intervention projects today. We show how these monitoring and impact evaluation tools support and are supported by three different change logics (simple, complicated, and complex) and can be categorized accordingly. We confront these three types of evaluation methods with the ambition to realise transformative social innovation. In doing so, we show how dominant monitoring and evaluation tools that support and are supported by a simple and complicated change logic, are unlikely to support transformative change. We also suggest that even less dominant monitoring and evaluation tools which are founded on a complexity paradigm, might reduce the transformative potential of social innovation initiatives when these tools focus more on adaptation than on transformation. On the basis of this analysis we conclude with an action and research agenda for those who engage with the monitoring and evaluation of transformative social innovation.

Transformative Social Innovation

Transformative social innovation is “a process of changing social relations that involves the emergence and spread of new knowledge and practices that challenge, alter or replace the established institutions in a specific context” (Haxeltine et al., 2017, p. 61). In adopting this definition, we agree with other scholars (e.g., Dias & Partidario, 2019; Haxeltine et al., 2017) who make a clear distinction between the broader category of social innovation and the more specific notion of transformative social innovation, and who argue that transformation is not an intrinsic quality of social innovation. In and on itself, social innovation - which can for example be defined as “the design and implementation of new solutions that imply conceptual, process, product, or organizational change, which ultimately aim to improve the welfare and wellbeing of individuals and communities” (OECD, n.d.) - does not necessarily aim for change at the institutional level. It pertains to processes of changing social relations that can develop, spread and interact with established institutions. When however social innovation aims for transformation, it intends to alter existing institutions and related practices, organisations, and norms that “both constrain and enable social relations and established patterns of doing, organising, framing and knowing” (Strasser et al., 2019, p. 1,304). These institutions can refer to both formal institutions like organisations, laws, regulations, or policies, as well as to less formal institutions like customs, beliefs, norms, habits, or values (Strasser et al., 2019). Historical examples of transformative social innovation initiatives are – even though they were at the time not necessarily given that name – cooperatives and credit unions, neighbourhood committees, or new models of education, participation, and governance (Moulaert et al., 2017). Contemporary examples of social innovation initiatives with a transformative agenda are hackerspaces, transition towns, community supported agriculture, participatory budgeting, or Via Campesina (Avelino et al., 2019). These examples also demonstrate that contemporary social innovations are developed through a diversity of organizations and collaborations (Unceta et al., 2017). They can take many different forms and may emerge in public, private, and nonprofit organizations as well as in complex cross-sectoral collaborations (Sadabadi & Rahimi Rad, 2022; Shier & Handy, 2016).

There exists a widely shared scholarly consensus that transformative social innovation is the result of intricately connected processes of collective action and collective learning (De Blust et al., 2019; Dumitru et al., 2016; Jessop et al., 2013). Transformative social innovation cannot be achieved by change of individuals alone but requires the emergence of new values, new imaginaries, new practices, new ways of interacting, and new understandings at the collective level so that dominant institutions can be recreated and provided with new meanings (De Blust et al., 2019). In practice this often means that transformative social initiatives:

start when a group of individuals come together to develop a common vision for social and institutional change, responding to perceived deficits or failures in current societal arrangements. Endorsing a set of alternative values, they set out to co-shape a reflexive and experimental space in which their vision may be realized as new social relations and configurations of practices [...]. As the SI initiative develops, it provides a space in which these new values can take root, new interpersonal relations can be shaped and enabled [...] and wherein both individual and collective empowerment can take place [...] (Pel et al., 2020, p. 7).

The intricate connection between collective learning and collective action to achieve transformative change is further elaborated in Kotov and Pedanik (2016) who suggest that transformative social innovation requires “rehabiting.” They argue that transformative social innovation requires changing (sets of) intricately connected conceptual, emotional, and behavioural social habits as well as the infrastructures that shape them.

It is important to underline that, when making this claim, Kotov & Pedanik (2016) do not rely on a “narrow” Cartesian (Pedersen & Dunne, 2020) notion of habit which refers to “repetitive

mindless routine” (Kotov & Pedanik, 2016, p. 139) at the level of the individual. Instead they rely on a Deweyan, pragmatist understanding of reality and identify habits as “general tendencies or predispositions acquired through inquiry and learning or (more commonly) through social interaction manifest in specific action while, typically, the actor is not aware of the predisposition” (Kotov & Pedanik, 2016, p. 143). Habits are in other words “dispositions that coordinate the relationship between human organism and his/ her environment” (Kotov & Pedanik, 2016, p. 144). In this sense habits exist at once at the individual and at the collective level. They are shaped in transaction with the environment, and thus also in transaction with collectively shared and intricately connected conceptual, emotional, and behavioural social habits as well as the infrastructures that support and are supported by these habits. In this, conceptual habits refer to “systems of meanings, beliefs, knowledge, discourses, along with cultural self-descriptions, norms, concepts of what a situation is about, what it means or how it should be dealt with [...], patterns of reasoning and argumentation that, once acquired, become unconscious, unquestioned, and are hard if not outright impossible to change” (Kotov & Pedanik, 2016, p. 150). Emotional habits are “general tendencies to emotionally react in a certain way in specific situations” (Kotov & Pedanik, 2016, p. 151). Behavioural habits are a) “related to the physical body, including the culturally regulated functions of the physical body; and b) skill, a way of doing or handling things and the infrastructure that supports these habits” (Kotov & Pedanik, 2016, p. 151). Infrastructures consist of “tools and artefacts...” that shape our life worlds and environment, which is “structured by a habit that supports the continuation of the habit and its manifestation through individual behaviour” (Kotov & Pedanik, 2016, p. 151).

To transform these intricately connected social (conceptual, emotional and behavioural) habits and infrastructures, which generate “stability and preservation” (Kotov & Pedanik, 2016, p. 145) but can also be changed, it is not sufficient to target the individual alone. Indeed, Kotov and Pedanik (2016) refer to what Dewey has written on the importance of the transactional relationship between individuals and their environment. “No amount of preaching good will or the golden rule or cultivation of sentiments of love and equity will accomplish the results. There must be change in objective arrangements and institutions. We must work on the environment not merely on the hearts of men” (Dewey 1922:22 as cited in Kotov & Pedanik, 2016, p. 145). Nor will it be sufficient to formulate an independent rational argument that is not connected to emotional, behavioural habits or infrastructures; “thought which does not exist within ordinary habits lacks means of execution” (Dewey, 1922, 67).

The desire and need to evaluate and transform habits, which generally remain unnoticed and invisible, emerges in real life situations when a new factor – e.g. an emerging social issue, a new need, or a newly acknowledged problem – arises which disturbs or challenges a habit or generates a conflict between habits. In these situations the values that are embodied by the habit or habit sets, become visible and the question rises whether what is “at present valued – the intentional object of any habit involved – is really valuable” (Kennedy, 1955, p. 90). What is needed for transformative change are then collective processes of value judgment critical inquiry, new imaginations, and experimentation with new hypotheses to be tested in minds and in actions. This process can contribute to the transformation of collectively shared (sets of) conceptual, emotional, behavioural habits as well as the infrastructures that support and are supported by these habits.

Monitoring and Evaluation of Social Initiatives: it’s not Neutral

Just like other social initiatives, transformative social innovation initiatives are confronted with both external and internal requests for monitoring and evaluation. When they are financed or supported by a government or external funders, transformative social innovation initiatives are likely to be confronted with external requests to make their impact and progress demonstrable for accountability reasons. At the same time these initiatives may also have internal drivers to systematically monitor and evaluate. They may for example want to engage

in an evaluation process to stimulate organizational learning or to demonstrate internal accountability. Evaluation, often in a more encompassing manner referred to as ‘monitoring and evaluation’ or ‘monitoring, evaluation, accountability, and learning’ can, in other words be done at the demand of different actors (e.g. funders, local authorities, scientists, Board members, etc.) and with different purposes (e.g. internal or external accountability, evidence building, comparison, fundraising, performance improvement, stimulating organizational learning and reflection processes, etc.) in mind. Tensions exist between these different evaluation goals and purposes (Chelimsky, 2006; Ebrahim, 2005; Regeer et al., 2016; Reinertsen et al., 2022).

Today the field of monitoring and evaluation is filled with a diversity of approaches, methods, and techniques. A selection of dishes on the contemporary monitoring and evaluation menu (Berghmans et al., 2014) are: social return on investment, randomised controlled trial, (quasi-) experimental methods, most significant change (Davies & Dart, 2005), principle-focused evaluation (Patton & Johnson, 2017), success case method (Brinkerhoff, 2003), innovation history (Douthwaite & Ashby, 2005), sensemaker (Cognitive Edge, 2023), process tracing (Beach & Pedersen, 2019), outcome harvesting (Wilson-Grau, 2018), outcome mapping (Earl et al., 2001), and contribution analysis (Mayne, n.d. & 2011). On this menu some dishes are more popular than others. There seems to be a broad societal preference for methods like randomised controlled trial or social return on investment (Archer & Titter, 2000; Benjamin et al., 2023), but also theory-based evaluations have gained popularity (Mackenzie & Blamey, 2005; Van der Knaap, 2004). Other methods, like for example sensemaker or outcome harvesting, seem to have the status of ‘connoisseur dishes’. They are methods that are not widely known (yet) but predominantly enjoyed and used by experts.

Many of these tools are presented as neutral and suitable for any setting or initiative. Consequently, the assessment of progress and impact of a transformative social innovation project might, at first sight, seem to simply be a matter of routinely choosing one tool from this plethora of evaluation tools and implementing it correctly. Evaluations are however not neutral or value-free (House & Howe, 1999). They can affect the social interventions that are the object of evaluation (Benjamin et al., 2023; Ebrahim, 2019; Ruff, 2021). Evaluations can for instance contribute to goal displacement. A strong evaluation focusing on specific outcomes can affect the actions of involved stakeholders who may feel that the quality of their performance is reflected in the evaluated outcome and who may therefore only start working towards the realisation of this outcome whilst overlooking other relevant aspects of the intervention (Benjamin et al. 2023; De Lancer Julnes, 2006). Participation in evaluation may also affect the cognitive understandings of involved stakeholders. The language used in evaluation may affect their perceptions, meanings, and understandings of the evaluated intervention and of social reality (Benjamin et al. 2023; Kirkhart, 2000; Patton, 2000). Participation in evaluation may also have a positive or negative effect on the participants’ individual and collective feelings and affect their views of merit and worth (Benjamin et al., 2023; Froncek & Rohmann, 2019; Kirkhart, 2000). Equally it may impact the power dynamics and perceptions of privilege and power within a program (Gregory, 2000; McKegg, 2019; Weiss, 1993). It is therefore important to carefully consider what kind of monitoring and evaluation methods are suitable and compatible with the evaluated intervention.

Different monitoring and evaluation methods all have their own particularities and characteristics. Nevertheless, drawing on a commonly used distinction between simple, complicated, and complex intervention logics (Glouberman & Zimmerman, 2002; Patton & Johnson, 2017; Rogers, 2008; Stern et al., 2012; Westley et al., 2007), they can arguably, in an ideal type manner, be categorised as monitoring and evaluation methods that support and are supported by a simple, complicated or complex intervention logic. Below we further elaborate on this categorisation. We demonstrate how different methods are supported by different intervention logics and thus reflect different understandings of how social change is established and what kind of knowledge production is required to pursue, monitor, and evaluate this social change.ⁱ We also show how these logics, and the related understanding of

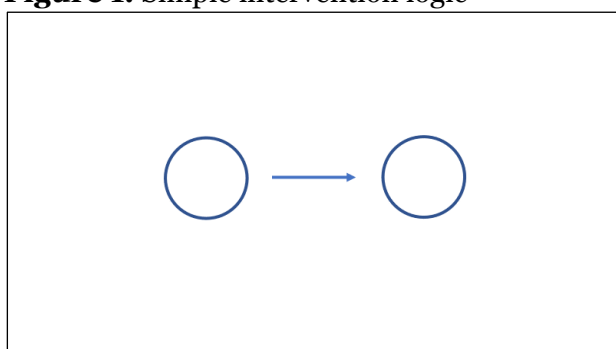
social change and of knowledge production, are not necessarily compatible with the ambition to realise transformative change.

Simple' and 'complicated' monitoring and evaluation methods and social innovation for transformation: not a perfect match

Within the variety of monitoring and evaluation methods, a first group of methods that can be identified are the methods that support and are supported by a simple understanding of social change, such as randomised controlled trials, quasi-experimental methods or (social) return on investment methods. These methods necessitate establishing and measuring specific quantitative performance indicators prior to the project's commencement, enabling before-and-after comparisons to determine if and to what extent the project or program has achieved the anticipated or desired outcomes. Evaluation in these methods means measuring to which extent predefined and expected performance indicators have been realised as an effect of the planned intervention.

These methods start from the assumption that indicators of success can be set beforehand and that there exists a linear and proportional causal relationship between the intervention and the effect (Kurtz & Snowden, 2003; Snowden, 2005). They produce and reproduce an understanding of social change as the result of a single, predictable linear cause-effect relationship. Social change is seen as the result of a simple, 'if A, then B' intervention logic (see figure 1).

Figure 1. Simple intervention logic

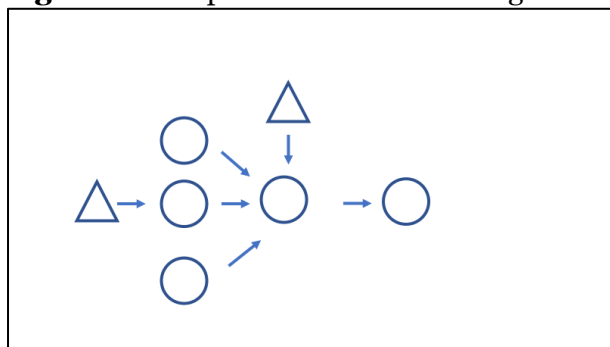


These methods also produce and reproduce a specific understanding of knowledge and knowledge production for social change. The process of measuring, processing, and analysing the extent to which the predefined indicators have been realised is external to the social intervention. Knowledge production about the project happens in other words outside the dynamics of the project. More precisely, knowledge is produced by people, often methodological experts, who take the perspective of a spectator and observes objective facts. If the demonstrated performance on predefined indicators is satisfying, then the social intervention is likely to be understood as a 'recipe' for success. It is in other words assumed that an identical intervention could, in comparable conditions, lead to a similar success. If the measured performance on predefined indicators is not satisfying, then project organizers are invited to make changes and strategically test and evaluate another 'recipe'.

A second group of monitoring and evaluation methods are methods like, for example, mediation analysis and clustered randomised controlled trial (Lemire et al., 2020; Peck, 2012), contribution analysisⁱⁱ (Mayne, 2011), or theory-based evaluations in which the theory underlying to an intervention – i.e. the “causal chain from inputs to outcomes and impact” (White, 2009, p. 271) – is mapped out and underlying assumptions are tested. This type of methods starts from the idea that the purposeful realisation of social change is complicated. They focus on building a theory of change that indicates how different causal mechanisms, conditions, and contextual factors might contribute to a desired outcome and on collecting

evidence about the results and the causal links between the different elements in these cause-effect chains. They are supported by and support an idea of social change as the result of causal chains of different elements that together, under specific conditions, contribute to an intended effect (Rogers, 2008) (see figure 2). Effects are in these methods, at least in principle, understood as predictable and pre-identifiable outcomes of a complicated chain of linear cause-and-effect relationships between different constitutive elements. If unintended effects are realised or if the projected theory of change does not work, one can try to adjust some of the elements in the chain or adjust the existing intervention theory (Berghmans et al., 2014; Poli, 2013; Rogers, 2008).

Figure 2. Complicated intervention logic



Knowledge production is, also in these methods supported by a complicated intervention logic, external to the social intervention itself. Theories are produced before the social intervention takes place and data about the causal chains are collected after implementation of the intervention.ⁱⁱⁱ Knowledge is also produced strategically, in function of developing a more plausible or more effective set of causal chains.

As we already suggested, these two sets of monitoring and evaluation tools, respectively supported by a simple or complicated change logic, are very popular dishes in the monitoring and evaluation kitchen. With a Nobel prize for Duflo, Banerjee and Kremer, the presentation of randomized controlled trials as the ‘gold standard’, and the rise of ‘simple’ monitoring and evaluation methods in evidence based policy making and in non-medical research (e.g. Connolly et al., 2018), methods that are based on a simple intervention logic have become very popular and are increasingly institutionalized in the broader nonprofit sector (Benjamin et al., 2023). At the same time, the habitual use of these methods to measure social change is increasingly subject to criticism. Apart from pointing at the general methodological, ethical, and practical limitations of these methods, critical questions are raised as to whether these methods are really suitable for the evaluation of social interventions as these interventions are in most cases not standardized, have limited predictability, and are strongly influenced by contextual factors (Ebrahim, 2019; Rogers, 2008; van der Meulen Rodgers et al., 2020). In response to these criticisms, the methods which support and are supported by a complicated intervention logic, have become increasingly popular for the monitoring and evaluation of social interventions (Van der Knaap, 2004). It is generally felt that this second group of methods is more apt for the evaluation of social interventions as these methods allow one to work with intended as well as unintended effects and approach change as the result of a combination of activities and factors.

But does this also imply that these methods are more suitable for the evaluation of social interventions that aim for transformative social innovation? On the basis of the above analysis of methods that support or are supported by a complicated causal logic, we are inclined to say no. Of course transformative social innovation initiatives may, in practice and for a variety of reasons and purposes – such as for example to attract financial investment (Antadze & Westley, 2012) or to attain symbolic power (Nicholls, 2015) – choose to adopt a diversity of evaluation tools. This however does not mean that all these methods are fully conducive for

and compatible with the ambition to realise transformative social innovation. In fact, it can be argued that methods founded on a complicated causal logic as well as methods founded on a simple causal logic, are not a well-tuned match for transformative social innovation initiatives. First, whereas transformative social innovation requires processes in which collective action and collectively learning processes are intricately connected, learning is, in these evaluation methods, rather situated outside of the dynamics of the intervention and change processes. It implies a distant look of a spectator who, from a fixed reference point (of a preset theory of change or a preset indicator) neutrally observes the performance on an indicator, a contribution, or the causal link between an intervention and an effect. In this process of learning little room is given to the possibility of questioning existing habits and infrastructures, creating new imaginaries, or experimenting in mind and action with new hypotheses.

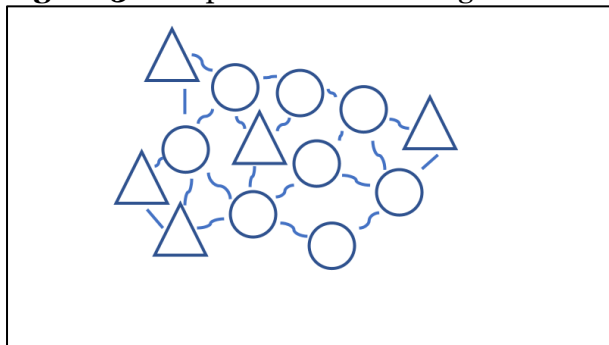
Second, and related to this, these evaluation methods are unlikely to stimulate innovation. They are based, as we demonstrated above, on the assumption that desirable effects and the pathway(s) towards this effect can, at least in principle, be identified beforehand and be predicted. This assumption is at odds with the “innovation uncertainty principle” (Morris, 2011) which refers to the idea that:

the pursuit of innovation necessarily involves a venture into the unknown and if we try to pin these unknowns down too early in our process we may make it more difficult to recognize and realize good opportunities or solutions. If attempts are made to calculate the impact of every idea very early on in the process of its development, the result could be a meaningless and misleading number that may have disproportionate influence on the emergent process at precisely the wrong time (Morris, 2011, p. 3).

The use of ‘simple’ or ‘complicated’ evaluation methods implies pinning down measurable expected outcomes and the expected linear causal pathway(s) towards these outcomes. These predicted outcomes and pathways also acquire prescriptive value, as in these monitoring and evaluation methods success means optimally realising the desired effect through the planned causal pathways. In this manner, these evaluation methods tend to reproduce dominant standards, and dominant ways of thinking, feeling, or behaving through the intervention process. They are unlikely to stimulate risk taking, creativity, and innovation, let alone transformation. In this manner, the mentioned monitoring and evaluation methods might, in spite of their popularity, reduce the transformative potential of social innovation initiatives rather than supporting it.

‘Complex’ monitoring and evaluation methods and social innovation for transformation: the need to safeguard the transformative potential?

The “complexity turn” in the social sciences (Urry, 2005) has contributed to the emergence of a third category of monitoring and evaluation methods which are supported by a complex intervention logic. According to this logic, social interventions – and the wider system in which these interventions are situated – have to be understood as open, dynamic networks of relationships between multiple elements which do not exert a linear influence on each other, but rather interact and adapt to the changing environment (Poli, 2013) (see figure 3).

Figure 3. Complex intervention logic

Key concepts in this complexity perspective on social reality are: adaptiveness and adaptive management, feedback loops, tipping points, iteration, emergence, equilibrium and so forth. According to this complex understanding of change, small inputs can lead to unplanned disproportionate effects. Effects can thus not be predefined, controlled, or managed in a planned way. However, change can be influenced by "learning to dance" with the complex system (Poli, 2013).

Drawing on this complex change logic, authors have developed and used different methods that can support the purposeful pursuit, monitoring, and evaluation of complex change (e.g. Derbyshire, 2019; Kurtz & Snowden, 2003; Rogers, 2009; Van der Merwe et al., 2019; Van Mierlo et al., 2010; Westley et al. 2007). These methods are, for example, most significant change (Davies & Dart, 2005; Ohkubo et al., 2022), sensemaker (Cognitive Edge, 2023; Van der Merwe et al., 2019) or outcome harvesting (INTRAC, 2017; Wilson-Grau, 2015, 2018). These methods are less widely known. It is mostly connoisseurs of the monitoring and evaluation kitchen who know, use, and appreciate these methods. In contrast to the previous two types of methods (which rely on a simple or complicated intervention logic) these monitoring and evaluation methods do not focus on evaluating whether the intervention went as planned or on evaluating whether desired effects were realised. Rather, they start from a given situation, try to identify important past changes, and make sense of how these changes have emerged. This process of real-time monitoring and evaluating the impact of certain interventions can then inform future courses of action.

In spite of their shared inscription into a general complexity logic, these methods seem to give different interpretations to the meaning of 'learning to dance with the complex system'. These interpretations move on a continuum between two poles of interpretation, namely a pole that focuses on adaptation and a pole that stresses transformation.^{iv}

The first pole of interpretation, which stresses adaptation, can for example be observed in the works of Rogers (2009), Derbyshire (2019), and Dotson et al. (2008) but also in the popular examples of "organizing a children's party" (Cognitive Edge, n.d.) or "raising a child" (Glouberman & Zimmerman, 2002). This interpretation starts from the assumption that the system with which one has to learn to dance is relatively stable and knowable. It is, as Funtowicz and Ravetz (1994) would say, an "ordinary complex system" with a relatively stable^v and clear finality or purpose. Though pathways towards this purpose are variable and cannot be predicted nor prescribed, there is a generally shared understanding of what is valued in regard to a successful social intervention. Social change can, according to this interpretation, purposefully be pursued by running small scale experiments. What is needed is regularly gathering feedback about changes caused by these experiments, and also collectively reflecting, together with involved actors and with the support of monitoring and evaluation methods, about emerging patterns and emerging outcomes. If these requirements are fulfilled small adaptations can be made in the direction of the shared finality or purpose. Monitoring and evaluation pertains, in this interpretation, to collectively making sense of and reflecting about

past changes and experimentations in order to make strategic adaptations towards a shared purpose.

A second pole of interpretation, which stresses *transformation*, is reflected in for example Kurtz and Snowden (2003), Lamboll and others (2021), and Fernandez-Gimenez et al. (2008). This orientation starts from the idea that the social intervention is an ‘emergent complex system’ (Funtowicz & Ravetz, 1994). In this interpretation the focus is not (only) on making strategic adaptations towards a shared purpose. The focus is on collaborative dialogue and social learning to further discover what a shared purpose could consist of. In this interpretation, it is assumed that people involved in the complex intervention have different values, different loyalties, and different perspectives; they do not have a shared purpose (yet). There might even exist contradictions and tensions between the interests of the individual people on the one hand and what is good for the emergent complex system on the other. Monitoring and evaluation in this second pole of interpretation is all about collectively making sense of one’s own implicit assumptions and ideas and reflecting the norms and values that underpin these assumption and ideas. This collective reflection may then contribute to developing a new language, new values, new meanings, and new norms, thus creating a new common ground that may further inform and shape the finality of the social intervention.

Methods supported by a complex change logic allow for new discoveries that may emerge throughout the innovation process, and in this manner they stimulate the creativity and risk taking necessary for transformative social innovation. The iterative processes of collective reflection, experimentation, monitoring, and sensemaking, which may involve many stakeholders, may contribute to creating an open space of collective change and learning in which new infrastructures and new (sets of) collective conceptual, emotional, and behavioural habits can emerge.

It might however be a too generalised statement to say that all these methods, which all support social innovation, also support and stimulate transformative change. In fact, it can be wondered to what extent monitoring and evaluation methods that align more with the first pole of interpretation and focus predominantly on strategic adaptations in view of a shared purpose) will stimulate and support transformative change. Indeed, there seems to be a tension between (processes of) adaptation and transformation. Adaption and transformation, do not only imply a different understanding of change but are also supported by a different type of reflexivity (Dias & Partidario, 2019; Moore et al., 2018). Adaptation requires a strategic reflexivity which consists of rapid feedback loops to strategically steer the intervention into the ‘right direction’. This type of reflexivity is more strategic and requires quick responses and is therefore different from the reflexivity that is needed for transformation. Transformation processes, on the other hand, require more and longer time “to understand what aspects of the existing system to break down and which to leverage to build an entirely new alternative system” (Moore et al., 2018, p. 11). They require a deep reflexivity at the institutional level – i.e. a willingness of all involved stakeholders to critically look at and reflect about dominant formal and informal institutions. In doing so they have a focus on developing new imaginaries and testing new collective ideas, emotions, behaviours, and infrastructures in mind and in action. This “deeper systemic reflexivity” (Moore et al., 2018), might not be stimulated in monitoring and evaluation processes that focus on strategic adaptation towards a shared finality. In fact, if such reflection processes are not generating discomfort, they are probably not contributing to transformative change:

In a sense then, a transformative learning space is not a safe space at all. It is actually a training ground to move into these more dangerous spaces. It relies on creating a kind of temporary psychological safety in order for people to do the necessary work of unlearning, crossing scales, confronting diversity, and acknowledging positive and negative dynamics, but it also relies on making people uncomfortable enough to prepare them to move through these contested, unknowable systems with courage, resilience, and grace (Moore et al., 2018, 38).

If one adopts a ‘complex’ monitoring and evaluation method to assess a social innovation initiative that aims for transformation, it is therefore important to be attentive to the nature of the collective reflection, experimentation, and sensemaking processes in this method. It is important to make sure that these processes touch at the institutional level and take a critical stance towards dominant ways of understanding, reacting to, and interacting with and in the world. Only then can these processes, which are likely to also generate some discomfort among the involved stakeholders (Moore et al., 2018), bring to the surface a “re-habiting” of dominant (sets of) habits and infrastructures (Kotov & Pedanik, 2016) and stimulate the transformative potential of the social innovation initiative.

Conclusion and Agenda for Future Action and Research

In this article we have demonstrated how dominant monitoring and evaluation methods may reduce the transformative potential of social innovation initiatives. We have pointed out that many habitually used evaluation methods – such as randomized controlled trial, return on investment, or theory based evaluation – are supported by a simple or complicated intervention logic and rely on the idea of plannable and prescribable change. As such, these methods are at odds with Morris’s (2011) innovation uncertainty principle. They do not give room for uncertainty, risk taking, creativity, or innovation, let alone for transformation. Next to that knowledge production processes in these methods are situated outside of the intervention. Learning in these evaluation processes implies taking the position of a spectator who, from a fixed reference point (of a preset theory of change or a preset indicator), observes the performance on an indicator or on the causal link between an intervention and an effect. This form of ‘spectator learning’ leaves little room for actively and collectively questioning social habits and infrastructures, for creating new imaginaries, and for collectively experimenting in mind and action with new imaginations and pathways to change. The intricate connection between collective learning and collective action, which is a condition for transformation, is inhibited in these (simple and complicated) evaluation methods. In this manner, these habitually used monitoring and evaluation methods tend to contribute to reproducing dominant institutions and habits, rather than challenging, altering, and questioning them.

Methods supported by a complex intervention logic, which are often known and used by ‘connoisseurs’ only, seem to be more suitable for the monitoring and evaluation of transformative social innovation initiatives. When these methods do not focus on strategic adaptation in view of a fixed purpose but generate – often uncomfortable and unsafe – processes of “deeper systemic reflexivity” (Moore et al., 2018), they may have the potential of enhancing rather than reducing the transformative potential of social innovation initiatives. They then hold the possibility of creating an open space of collective action and learning, involving all stakeholders, so that new infrastructures and new (sets of) collective conceptual, emotional and behavioural habits may emerge.

On the basis of the analysis made in this paper, we make a plea for what we call a ‘re-habiting of monitoring and evaluation’ in the field of transformative social innovation. If we really want to stimulate a process of transformative social innovation or “a process of changing social relations that involves the emergence and spread of new knowledge and practices that challenge, alter or replace the established institutions in a specific context” (Haxeltine et al., 2017, p. 61), we should not just adopt ‘old’ monitoring and evaluation habits. It is crucial that we also re-habit the conceptual, emotional, behavioral habits and infrastructures that shape monitoring and evaluation itself.^{vi}

It can however be wondered if the development of new complexity-based evaluation methods and approaches alone will lead to new monitoring and evaluation habits and infrastructures. As we have demonstrated above, a plethora of alternative monitoring and evaluation methods

which rely on a complexity paradigm, already have been developed in order to do more justice to complex interventions. Next to that scholars like Bamberger and colleagues (2015), Williams and Hummelbrunner (2010), and Patton (2019) have already proposed wider evaluation approaches which do more justice to complex interventions. Bamberger et al. (2015) for example propose a complexity-aware approach to evaluation. They claim that complexity-aware evaluations demand unpacking a complex intervention into different evaluable components which have to, in first instance and supported by appropriate evaluation designs, be assessed separately to then, in second instance, be reassembled for the evaluation of the overall program. These authors also plead for the adoption of new technologies, big data, and data analytics in complex evaluations. Williams and Hummelbrunner (2010) propose a systemic approach to evaluation. They claim that evaluators who engage in systemic evaluation need three orientations. They need to be reflective about “assumptions, mental models and values – and how they affect what we see or hear” (Reynolds et al., 2016, p. 668). They also have to pay “due attention to emergence and value differences from plans, as these can provide useful clues for improvements” and they need to adopt “think–act–think circularity” (Reynolds et al., p. 668). Patton (2019), with the Blue Marble evaluation, has, at his turn developed an evaluation approach oriented towards global systemic change and transformation. In fact, Patton’s call to “transform evaluation to evaluate transformation,” seems to be a call for a re-habiting of evaluation habits. His suggestion that the transformation of evaluation requires “moving from a theory of change to a theory of transformation,” moving from “an external, independent stance to a skin-in the game stance,” and moving “beyond the evaluator competence to becoming world savvy through ongoing learning” (Blue Marble Evaluation, n.d.) seems to very much in line with this paper’s plea for ‘re-habiting monitoring and evaluation’.

The mere existence of complexity-aware evaluation methods and approaches might however not be enough to change existing social habits about monitoring and evaluation. It might also require a broader conscientization. Indeed, up until now, evaluation methods that are founded on a complex change logic, seem to be reserved for ‘connoisseurs’. Also, new complexity-aware evaluation approaches, like the ones developed by Bamberger and colleagues (2015), Williams and Hummelbrunner (2010), and Patton (2019), seem to be discussed predominantly by evaluation professionals. At the same time, there continues to be a broad societal preference for classical evaluation methods that are founded on a specific understanding of change and produce and reproduce a specific ‘spectator’ understanding of knowledge and knowledge production. If we really want to re-habit monitoring and evaluation and change dominant ways of thinking, feeling, or behaving in monitoring and evaluation a more general, critical conscientization – of policy makers, practitioners as well as of the general public – about monitoring, evaluation, and transformation may be necessary.

Next to this call for action, or more specifically for critical conscientization, we end this article with two suggestions for future research. A first suggestion pertains to the differences and maybe even incompatibilities (Dias & Partidario, 2019) between adaptation and transformation in complex evaluation methods and approaches. We think that this is an insufficiently explored and under-theorized aspect in complexity-aware monitoring and evaluation methods and approaches. A second suggestion is to fully integrate a pragmatist philosophical perspective into future theory formation about monitoring and evaluation for transformative social innovation. Indeed, it can be observed that scholarly theories about monitoring and evaluation of uncertain and complex interventions, are often grounded in a complexity or systems perspective. As these complex or systems perspectives are rooted in natural sciences, these theories often seem to pay relatively little attention to questions that are also central in transformative social innovation, such as ‘how do we collectively decide what is of value?’, ‘how is collective change and learning established?’, or ‘how can purposeful collective action be made possible when ends and values are unclear or contested?’. Though pragmatist notions like single, double, and triple loop learning and the “reflective practitioner” (Argyris, 1976; Schön, 1983) have already been integrated in reflections about transformative evaluation (e.g. Regeer et al., 2009) we think, in line with the argument formulated by Ansell and Geyer (2017), that a more serious engagement with (Deweyan) pragmatist notions such as

‘habit’, ‘value’, ‘community of inquiry’, ‘intelligent action’, and ‘imagination’ may contribute to a more profound theory formation about evaluation for transformation.

Notes

1. ⁱIn contrast to others (like for example Glouberman & Zimmerman, 2002) we do not see simplicity, complicatedness, or complexity as inherent characteristics of interventions or of social systems. Rather we approach simplicity, complicatedness, and complexity as “pragmatic perspectives” (Petticrew, 2011, p. 397) adopted by evaluation researchers to help describe and understand the interventions in question. Starting from this premise, we then argue that these pragmatic perspectives might not always be compatible with the ambition to realise transformative social innovation.
2. ⁱⁱContribution analysis aims to demonstrate to what extent observed results or outcomes are the result of a specific initiative rather than of other factors. By developing a theory of change, looking for evidence for that theory of change, and paying attention to other factors which may also have contributed to the realisation of specific outcomes, contribution analysis aims to generate evidence about the contribution made by a specific initiative (Mayne, n.d.).
3. ⁱⁱⁱIn some cases, theories are produced in the course of or after an intervention. It could be argued that when these theories-in-use are being constructed, the knowledge production is not external to the social intervention itself.
4. ^{iv}It needs to be stressed that both poles are only poles on a wider continuum of interpretation. Further research is needed to explore whether these poles point at the different ways in which monitoring and evaluation tools are used or if they point at the internal logics of monitoring and evaluation tools.
5. ^vThis goal cannot be exactly defined beforehand and translated in measurable indicators. Nevertheless, it is relatively stable and unlikely to change when the social intervention unfolds. The stability of the goal can emerge in different manners. The goal can for example be stable because it is set and guarded by decision makers who ‘own the project and the program theory’ (e.g. professionals who are determined to pursue a specific goal). But it can also be stable because there exists a general consensus within the system of the social intervention on what precisely has to be achieved through the social intervention.
6. ^{vi}Kotov & Pedanik (2016) make a similar point for habits of problem solving. They argue that transformative social innovation often requires “a rehabilitating also in the ways and methods of problem solving” (Kotov & Pedanik, 2016, p.147).

Disclosure Statement

The authors declare that there are no conflicts of interest that relate to the research, authorship, or publication of this article.

References

- Ansell, C. & Geyer, R. (2017). ‘Pragmatic complexity’ a new foundation for moving beyond ‘evidence-based policy making’? *Policy Studies*, 38(2), 149-167.
<https://doi.org/10.1080/01442872.2016.1219033>

- Antadze, N., & Westley, F. R. (2012). Impact metrics for social innovation: Barriers or bridges to radical change? *Journal of Social Entrepreneurship*, 3(2), 133-150.
<https://doi.org/10.1080/19420676.2012.726005>
- Archer, M. S., & Tritter, J. Q. (Eds.). (2000). *Rational choice theory: Resisting colonization*. Psychology Press.
- Argyris, C. (1976). Single-loop and double-loop models in research on decision making. *Administrative Science Quarterly*, 21(3), 363-375.
- Avelino, F., Wittmayer, J. M., Pel, B., Weaver, P., Dumitru, A., Haxeltine, A., Kemp, R., Jorgensen, M. S., Bauler, T., Ruijsink, S., & O'Riordan, T. (2019). Transformative social innovation and (dis)empowerment. *Technological Forecasting and Social Change*, 145, 195-206. <https://doi.org/10.1016/j.techfore.2017.05.002>
- Bamberger, M., Vaessen, J., & Raimondo, E. (Eds.). (2015). *Dealing with complexity in development evaluation: A practical approach*. Thousand Oaks: Sage Publications.
- Beach, D., & Pedersen, R. B. (2019). *Process-tracing methods: Foundations and guidelines*. Michigan: University of Michigan Press.
- Benjamin, L. M., Ebrahim, A., & Gugerty, M. K. (2023). Nonprofit organizations and the evaluation of social impact: A research program to advance theory and practice. *Nonprofit and Voluntary Sector Quarterly*, 52(1_suppl), 313S-352S.
<https://doi.org/10.1177/08997640221123590>
- Berghmans, M., Deprez, S., Celis, H., Van Hemelrijck, A., & Vandenabeele, J. (2014). *Peilen naar de impact van sociaal-culturele praktijken*. Brussels: Socius.
- Blue Marble Evaluation (n.d.) Transformative alignment principle. Retrieved at <https://bluemarbleeval.org/principles/operating-principles/transformational-alignment-principle/>
- Brinkerhoff, R. (2003). *The success case method: Find out quickly what's working and what's not*. Berrett-Koehler Publishers.
- Chelimsky, E. (2006). The purposes of evaluation in a democratic society. In I. Shaw, J.C. Green & M.M. Mark.(Eds.) *The Sage handbook of evaluation* (33-55). London: Sage.
- Cognitive Edge. (n.d.). How to organise a children's party? Retrieved from <https://www.youtube.com/watch?v=Miw92eZaJg>
- Cognitive Edge. (2023). About sensemaker. Retrieved from <https://thecynefin.co/about-sensemaker/>
- Connolly, P., Keenan, C. & Urbanska, K. (2018). The trials of evidence-based practice in education: A systematic review of randomised controlled trials in education research 1980–2016. *Educational Research*, 60(3), 276-291.
<https://doi.org/10.1080/00131881.2018.1493353>
- Cunha, J. & Benneworth, P. (2020). How to measure the impact of social innovation initiatives? *International Review on Public and Nonprofit Marketing*, 17, 59–75.
- Davies, R. & Dart, J. (2005). The 'Most Significant Change' (MSC) technique: A guide to its use.
- De Blust, S., Schreurs, J., & Devisch, O. (2019). Towards a situational understanding of collective learning: A reflexive framework. *Urban Planning*, 4(1), 19–30.
<https://doi.org/10.17645/up.v4i1.1673>
- De Lancer Julnes, P. (2006). Performance measurement: An effective tool for government accountability? The debate goes on. *Evaluation*, 12(2), 219-235.
<https://doi.org/10.1177/1356389006066973>
- Derbyshire, J. (2019). Use of scenario planning as a theory-driven evaluation tool. *Futures & Foresight Science*, 1(1), e1. <https://doi.org/10.1002/ffo2.1>
- Dewey, J. (1922). *Human nature and conduct*. New York: Henry Holt.
- Dias, J., & Partidário, M. (2019). Mind the gap: The potential transformative capacity of social innovation. *Sustainability*, 11(16), 4465. <https://doi.org/10.3390/su11164465>
- Dotson, Y., Folkman, D., & Syam, D. (2008). Welcome to Cynefin: A framework for learning and action in the face of complexity and chaos. Conference proceedings of the 27th Annual Midwest Research-to-Practice Conference in Adult, Continuing, Community,

- and Extension Education . Retrieved at http://digitalcommons.wku.edu/cebs_pubs/17
- Douthwaite, B. & Ashby, J. (2005). Innovation histories: A method for learning from experience, ILAC Brief 5. Retrieved at https://www.betterevaluation.org/sites/default/files/ILAC_Brief05_Histories_o.pdf
- Dumitru, A., Lema-Blanco, I. & García Mira, R. (2016). Social learning for transformative social innovation, TRANSIT Deliverable 2.3.
- Earl, S., Carden, F., & Smutylo, T. (2001). Outcome mapping: Building learning and reflection into development programmes. Ottawa: International Development Research Centre (IDRC).
- Ebrahim, A. (2005). Accountability myopia: Losing sight of organizational learning. *Nonprofit and Voluntary Sector Quarterly*, 34(1), 56-87.
- Ebrahim, A. (2019). *Measuring social change. Performance and accountability in a complex world*. Stanford: Stanford Business Books.
- Fernandez-Gimenez, M. E., Ballard, H. L., & Sturtevant, V. E. (2008). Adaptive management and social learning in collaborative and community-based monitoring: A study of five community-based forestry organizations in the western USA. *Ecology and Society*, 13(2), 4. <http://www.jstor.org/stable/26267955>
- Fougère, M., Segercrantz, B., & Seeck, H. (2017). A critical reading of the European Union's social innovation policy discourse:(Re)legitimizing neoliberalism. *Organization*, 24(6), 819-843. <https://doi.org/10.1177/135050841668517>
- Froncek, B., & Rohmann, A. (2019). "You get the great feeling that you're being heard but in the end you realize that things will be done differently and in others' favor": An experimental investigation of negative effects of participation in evaluation. *American Journal of Evaluation*, 40(1), 19-34.
- Funtowicz, S. & Ravetz, J.R. (1994). Emergent complex systems. *Futures*, 26(7), 568-582. [https://doi.org/10.1016/0016-3287\(94\)90029-9](https://doi.org/10.1016/0016-3287(94)90029-9)
- Glouberman, S. & Zimmerman, B. (2002). Complicated and complex systems: What would successful reform of Medicare look like. *Discussion Paper*, 8. Commission on the Future of Health Care in Canada.
- Gregory, A. (2000). Problematizing participation: A critical review of approaches to participation in evaluation theory. *Evaluation*, 6(2), 179-199. <https://doi.org/10.1177/13563890022209208>
- Haxeltine, A., Pel, B., Wittmayer, J., Dumitru, A., Kemp, R., & Avelino, F. (2017). Building a middle-range theory of transformative social innovation: Theoretical pitfalls and methodological responses. *European Public & Social Innovation Review*, 2(1), 59-77. <https://doi.org/10.31637/epsir.17-1.5>
- House, E., & Howe, K. R. (1999). *Values in evaluation and social research*. Thousand Oaks: Sage Publications.
- INTRAC (2017). Outcome harvesting. Retrieved at https://www.google.com/url?sa=i&rct=j&q=&esrc=s&source=web&cd=&ved=0CAIQw7AJahcKEwigr_jJ-ooAAxUAAAAAHQAAAAQAg&url=https%3A%2F%2Fwww.intrac.org%2Fwpcms%2Fwp-content%2Fuploads%2F2017%2F01%2FOutcome-harvesting.pdf&psig=AOvVawoq7D-FcoEgQgTtHF2tjRVN&ust=1689175163270577&opi=89978449
- Jessop, B., Moulaert, F., Hulgård, L., & Hamdouch, A. (2013). Social innovation research: A new stage in innovation analysis? In F. Moulaert, D. MacCallum, A. Mehmood, & A. Hamdouch (Eds.), *The international handbook on social innovation. Collective action, social learning and transdisciplinary research* (pp. 110-130). Cheltenham: Edward Elgar.
- Kennedy, G. (1955). The hidden link in Dewey's theory of evaluation. *Teachers College Record*, 56(8), 1-7.
- Kirkhart, K.E. (2000), Reconceptualizing evaluation use: An integrated theory of influence. *New Directions for Evaluation*, 88, 5-23. <https://doi.org/10.1002/ev.1188>

- Kok, K. P. W., van der Meij, M. G., Wagner, P., Cesuroglu, T., Broerse, J. E. W., & Regeer, B. J. (2023). Exploring the practice of Labs for sustainable transformation: The challenge of 'creating impact'. *Journal of Cleaner Production*, 388, 135994. <https://doi.org/10.1016/j.jclepro.2023.135994>.
- Kotov, K. & Pedanik, R. (2016). Social innovation as rehabiting. In C. J. Rodriguez Higuera & T.J. Bennett (Eds.), *Concepts for semiotics* (pp. 133-156). Tartu: University of Tartu Press.
- Kurtz, C.F. & Snowden, D.J. (2003). The new dynamics of strategy. Sense-making in a complex and complicated world. *IBM Systems Journal*, 42(3), 462-448.
- Lemire, S., Peck, L.R., & Porowski, A. (2020). The growth of the evaluation tree in the policy analysis forest: Recent developments in evaluation. *Policy Studies Journal*, 48(S47-S70). <https://doi.org/10.1111/psj.12387>
- Lamboll, R., Nelson, V., Gebreyes, M., Kambewa, D., Chinsinga, B., Karbo, N., Rukonge, A., Sekeleti, M., Litaba Wakun'uma, W. Gutema, T.H., Henjewe, M., Kampanje-Phiri, J., Masikati-Hlanguyo, P., Quaye, W., Duah, S., Kivuyo, M., Nyanga, P., Akuffo-Bea Essilfie, M., Asafu-Adjaye, N.Y., Clottey, V. & Martin, A. (2021). Strengthening decision-making on sustainable agricultural intensification through multi-stakeholder social learning in sub-Saharan Africa. *International Journal of Agricultural Sustainability*, 19(5-6), 609-635. <https://doi.org/10.1080/14735903.2021.1913898>
- Mackenzie, M., & Blamey, A. (2005). The practice and the theory: Lessons from the application of a theories of change approach. *Evaluation*, 11(2), 151–168. <https://doi.org/10.1177/1356389005055538>
- Mayne, J. (2011). Contribution analysis: Addressing cause and effect. In K. Forss, M. Marra & R. Schwartz (Eds.), *Evaluating the complex* (pp. 53-96). Transaction Publishers.
- Mayne, J. (n.d.). Contribution analysis: An approach to exploring cause and effect. ILAC . Retrieved at <https://nonprofitbuilder.org/storage/377/Contribution-analysis-An-approach-to-exploring-cause-and-effect-ILAC.pdf>
- McGowan, K. & Westley, F. (2015). At the root of change: The history of social innovation. In A. Nicholls, J. Simon & M. Gabriel (Eds.), *New frontiers in social innovation research* (pp. 52-68). London: Palgrave Macmillan.
- McKegg, K. (2019). White privilege and the decolonization work needed in evaluation to support indigenous sovereignty and self-determination. *Canadian Journal of Program Evaluation*, 34(2), 357- 367.
- Moore, M.-L., Olsson, P., Nilsson, W., Rose, L., & Westley, F. (2018). Navigating emergence and system reflexivity as key transformative capacities: Experiences from a Global Fellowship program. *Ecology and Society*, 23(2), 38.
- Morris, L., (2011). *The innovation master plan: The CEO's guide to innovation*. Walnut Creek, CA: Innovation Academy.
- Moulaert, F. & MacCallum, D. (2019). *Advanced Introduction to Social Innovation*. Cheltenham: Edward Elgar Publishing.
- Moulaert, F. & Mehmood, A. (2020). Towards a social innovation (SI) based epistemology in local development analysis: lessons from twenty years of EU research. *European Planning Studies*, 28(3), 434-453. <https://doi.org/10.1080/09654313.2019.1639401>
- Moulaert F., Mehmood, A., MacCallum, D., & Leubolt B. (2017). *Social innovation as a trigger for transformations. The role of research*. Luxembourg: Publications Office of the European Union.
- Nicholls, A. (2015). Synthetic grid: A critical framework to inform the development of social innovation metrics. *CRESSI Working papers*, 14. Oxford: Creating Economic Space for Social Innovation (CRESSI) project. Retrieved at <http://www.transitsocialinnovation.eu/resource-hub/synthetic-grid-a-critical-framework-to-inform-the-development-of-social-innovation-metrics-cressi-working-papers-142015>
- Nicholls, A. & Edmiston, D. (2018) Social innovation policy in the European Union. In R. Heiskala & J. Aro (Eds.), *Policy design in the European Union: An empire of shopkeepers in the making?* (pp. 161-190). Palgrave Macmillan.

- Novy, A. (2017). Transformative social innovation. *SRE-Discussion Papers*, 2017/05. WU Vienna University of Economics and Business. Retrieved at <https://EconPapers.repec.org/RePEc:wiw:wus009:5619>
- OECD (n.d.). Social innovation. <https://www.oecd.org/cfe/leed/social-economy/social-innovation.htm>
- Ohkubo, S., Mwaikambo, L., Salem, R. M., Ajijola, L., Nyachae, P., & Sharma, M. K. (2022). Lessons learned from the use of the most significant change technique for adaptive management of complex health interventions. *Global Health: Science and Practice*, 10(1), e2100624. <https://doi.org/10.9745/GHSP-D-21-00624>.
- Oosterlynck, S., Novy, A., & Kazepov, Y. (2019). Conclusion: Local social innovation and welfare reform. In S. Oosterlynck, A. Novy & Y. Kazepov (Eds.), *Local social innovation to combat poverty and exclusion* (pp. 217-228). Bristol: Policy Press.
- Patton, M.Q. (2000). Overview: Language matters. *New Directions for Evaluation*, 86, 5-16.
- Patton, M.Q. & Johnson, N.M. (2017). *Principles-focused evaluation: The GUIDE*. Guilford Press.
- Patton, M.Q. (2019). *Blue marble evaluation. Premises and principles*. Guilford Press.
- Peck, L. R. (2012). Laura R. Peck on “Failure of Intervention or Failure of Evaluation: A Meta-Evaluation of the National Youth Anti-Drug Media Campaign Evaluation.” *Substance Use & Misuse*, 47(13-14), 1425-1426. <https://doi.org/10.3109/10826084.2012.723553>
- Pedersen, M., & Dunne, S. (2020). Dewey’s broad account of habit and its relevance for change management: A conceptual clarification with pragmatic illustrations. *Journal of Change Management*, 20(3), 247-263. <https://doi.org/10.1080/14697017.2020.1755342>
- Pel, B., Haxeltine, A., Avelino, F., Dumitru, A., Kemp, R., Bauler, T., Kunze, I., Dorland, J., Wittmayer, J. & Jørgensen, M. S. (2020). Towards a theory of transformative social innovation: A relational framework and 12 propositions. *Research Policy*, 49(8), 104080. <https://doi.org/10.1016/j.respol.2020.104080>
- Petticrew, M. (2011). When are complex interventions ‘complex’? When are simple interventions ‘simple’? *European Journal of Public Health*, 21(4), 397-398. <https://doi.org/10.1093/eurpub/ckr084>
- Poli, R. (2013). A note on the difference between complicated and complex social systems. *Cadmus Journal*, 2(1), 142-147.
- Regeer, B. J., de Wildt-Liesveld, R., van Mierlo, B., & Bunders, J. F. (2016). Exploring ways to reconcile accountability and learning in the evaluation of niche experiments. *Evaluation*, 22(1), 6-28.
- Regeer, B. J., Hoes, A.-C., van Amstel-van Saane, M., Caron-Flinterman, F. F., & Bunders, J. F. G. (2009). Six guiding principles for evaluating mode-2 strategies for sustainable development. *American Journal of Evaluation*, 30(4), 515-537. <https://doi.org/10.1177/1098214009344618>
- Reinertsen, H., Bjørkdahl, K., & McNeill, D. (2022). Accountability versus learning in aid evaluation: A practice-oriented exploration of persistent dilemmas. *Evaluation*, 28(3), 356-378.
- Reynolds, M., Gates, E., Hummerbrunner, R., Marra, M., & Williams, B. (2016). Towards systemic evaluation. *Systems Research and Behavioral Science*, 33(5), 662-673.
- Reynolds, S. Gabriel, M. & Heales, C. (2017). Social innovation policy in Europe: where next? Retrieved at https://www.siceurope.eu/sites/default/files/field/attachment/social_innovation_policy_in_europe_-_where_next.pdf
- Rogers, P. (2008). Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation*, 14(1), 29-48. <https://doi.org/10.1177/1356389007084674>
- Rogers, P. (2009). Matching impact evaluation design to the nature of the intervention and the purpose of the evaluation. *Journal of Development Effectiveness*, 1(3), 217-226. <https://doi.org/10.1080/19439340903114636>

- Ruff, K. (2021). How impact measurement devices act: The performativity of theory of change, SROI and dashboards. *Qualitative Research in Accounting & Management*, 18(3), 332-360. <https://doi.org/10.1108/QRAM-02-2019-0041>
- Sadabadi, A. & Rahimi Rad, Z. (2022). How can cross-sector partnership promote social innovation? *Systemic Practice and Action Research*, 35, 1-20.
- Schön, D. (1983). *The reflective practitioner: How professionals think in action*. New York: Basic Books.
- Shier, M. L., & Handy, F. (2016). Cross-sector partnerships: Factors supporting social innovation by nonprofits. *Human Service Organizations: Management, Leadership & Governance*, 40(3), 253-266. <https://doi.org/10.1080/23303131.2015.1117556>
- Snowden, D. (2005). Strategy in the context of uncertainty. *Handbook of Business Strategy*, 6(1), 47-54.
- Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R., & Befani, B. (2012). Broadening the range of designs and methods for impact evaluations. *Working Paper*, 38. London: Department for International Development (DFID).
- Strasser, T., de Kraker, J., Kemp, R. (2019). Developing the transformative capacity of social innovation through learning: A conceptual framework and research agenda for the roles of network leadership. *Sustainability*, 11(5), 1304. doi.org/10.3390/su11051304.
- Tesfaye, L.A., & Fougère, M. (2022). Frugal innovation hijacked: The co-optive power of co-creation. *Journal of Business Ethics* 180, 439-454. <https://doi.org/10.1007/s10551-021-04883-4>
- Unceta, A., Castro-Spila, J., & García Fronti, J. (2017). The three governances in social innovation. *Innovation: The European Journal of Social Science Research*, 30(4), 406-420. <https://doi.org/10.1080/13511610.2017.1279537>
- Urry, J. (2005). The complexity turn. *Theory, Culture & Society*, 22(5), 1-14. <https://doi.org/10.1177/0263276405057188>
- Van der Knaap, P. (2004). Theory-based evaluation and learning: Possibilities and challenges. *Evaluation*, 10(1), 16-34. <https://doi.org/10.1177/1356389004042328>
- Van der Merwe, S. E., Biggs, R., Preiser, R., Cunningham, C., Snowden, D. J., O'Brien, K., Jenal, M., Vosloo, M., Blignaut, S., & Goh, Z. (2019). Making sense of complexity: Using SenseMaker as a research tool. *Systems*, 7(2), 25. <http://dx.doi.org/10.3390/systems7020025>
- van der Meulen Rodgers, Y., Bebbington, A., Boone, C., Dell'Angelo, J., Platteau, J. P., & Agrawal, A. (2020). Experimental approaches in development and poverty alleviation. *World Development*, 127, 104807. <https://doi.org/10.1016/j.worlddev.2019.104807>
- Van Mierlo, B., Arkesteijn, M., & Leeuwis, C. (2010). Enhancing the reflexivity of system innovation projects with system analyses. *American Journal of Evaluation*, 31(2), 143-161. <https://doi.org/10.1177/1098214010366046>
- Weiss, C.H. (1993). Where politics and evaluation research meet. *Evaluation Practice*, 14(1), 93-106.
- Westley, F., Zimmerman, B. & Patton, M. (2007). *Getting to maybe. How the world is changed*. Penguin Random House.
- White, H. (2009). Theory-based impact evaluation: principles and practice. *Journal of Development Effectiveness*, 1(3), 271-284. <https://doi.org/10.1080/19439340903114628>
- Williams, B. & Hummelbrunner, R. (2010). *Systems concepts in action: A practitioner's toolkit*. Stanford Business Books.
- Wilson-Grau, R. (2015). Outcome harvesting. Better evaluation. Retrieved from http://betterevaluation.org/plan/approach/outcome_harvesting.
- Wilson-Grau, R. (2018). *Outcome harvesting. Principles, steps, and evaluation application*. Information Age Publishing.
- Wittmayer, J.M., Backhaus, J., Avelino, F., Pel, B., Strasser, T., Kunze, I., & Zuijderwijk, L. (2019). Narratives of change: How social innovation initiatives construct societal transformation. *Futures*, 112, 102433. <https://doi.org/10.1016/j.futures.2019.06.005>.

Wright, C., & Nyberg, D. (2019). Climate change and social innovation. In G. George, T. Baker, P. Tracey & H. Joshi (Eds.), *Handbook of inclusive innovation: the role of organizations, markets and communities in social innovation* (pp. 47-60). Cheltenham: Edward Elgar Publishing.

Author Biographies

Mieke Berghmans

Mieke Berghmans is a postdoctoral assistant affiliated to the research unit Education, Culture, and Society of the KU Leuven (Belgium). She has a broad interest in contemporary governance and performance issues in education and the third sector. She works, teaches and does research on the design and development of innovative educational practices.

Joke Vandenabeele

Joke Vandenabeele is an associate professor social and cultural pedagogy at the Education, Culture and Society department, KU Leuven (Belgium). Her research focuses on how transformative innovation is enacted within practices of non-formal education, community education and public pedagogy.