Dynamic systems and the role of evaluation: the case of the Green Communities project

Valentina Anzoise
Ca' Foscari University of Venice

Stefania Sardo
BI Norwegian Business School

Dynamic Systems and the role of Evaluation: The case of the Green Communities project

Valentina Anzoise [*][a], Stefania Sardo [b]

[a] European Centre for Living Technology, Ca' Foscari University of Venice
[b] Department of Innovation and Economic Organization, BI Norwegian Business School

Abstract

The crucial role evaluation can play in the co-development of project design and its implementation will be addressed through the analysis of a case study, the Green Communities (GC) project, funded by the Italian Ministry of Environment within the EU Interregional Operational Program (2007-2013) "Renewable Energy and Energy Efficiency". The project's broader goals included an attempt to trigger a change in Italian local development strategies, especially for mountain and inland areas, which would be tailored to the real needs of communities, and based on a sustainable exploitation and management of the territorial assets. The goal was not achieved, and this paper addresses the issues of how GC could have been more effective in fostering a vision of change, and which design adaptations and evaluation procedures would have allowed the project to better cope with the unexpected consequences and resistances it encountered. The conclusions drawn are that projects should be conceived, designed and carried out as *dynamic systems*, inclusive of a dynamic and engaged evaluation enabling the generation of feedbacks loops, iteratively interpreting the narratives and dynamics unfolding within the project, and actively monitoring the potential of various relationships among project participants for generating positive social change.

Keywords: evaluation, innovation, local development, complexity, dynamic systems, narratives, uncertainty

---

[*] Corresponding authors:
Valentina Anzoise | Ca' Foscari University of Venice, European Centre for Living Technology, S. Marco 2940, 30124 Venice (Italy), tel. +39 **Tel.** +39 041 2347594, **Fax:** +39 041 2347589, e-mail: valentina.anzoise@unive.it
Stefania Sardo | BI Norwegian Business School/ BI Handelshøyskolen, Department of Innovation and Economic Organization, Nydalsveien 37, N-0442 Oslo (Norway), **Tel.** +47 464 10 801, e-mail: stefania.sardo@bi.no

1. Introduction

This paper introduces a new evaluation approach called Dynamic Evaluation (DE), aimed at supporting the monitoring, management, and development of projects and programs. DE rests on two theoretical foundations: on one side, ideas in the literature regarding evaluation practices, in particular the perspectives proposed by Participatory Evaluation (Cousins & Earl, 1992; Cousins & Whitmore, 1998; Cousins, 2003; O'Sullivan, 2012), Developmental Evaluation (Patton, 1994, 2011), and by Empowerment Evaluation (Fetterman, 1994, 2001); on the other side, the complexity-based theory of innovation processes, as developed in Arthur, Durlauf, Lane (1997), Lane and Maxfield (1997, 2005) and Lane, Maxfield, Read, and van der Leeuw (2009).

The debate about whether, and how, to apply system thinking (Cabrera, Colosi, & Lobdell, 2008) and complexity theory in the evaluation practices has been raging for a decade. A recent review argued that few contributions to this debate have provided detailed considerations of what would constitute complexity consistent methods (Walton, 2013). Here, we propose one approach to filling this lacuna.

Projects and programs can be considered as temporary organizations in agent-artifact space, which undergo continuous change (Lundin & Soderholm, 1995). These evolving organizations require continuous monitoring and interpretation to define and regulate the system of interactions among the changing set of agents and artifacts that comprise them (Lane & Maxfield, 1997, 2005). Not only do the organization of project and programs lack clear and static boundaries, but the dynamics of the processes through which they evolve are neither linear nor predetermined (Urban, Hargraves, & Trochim, 2014). Conceptualizing, nurturing and implementing projects and programs should be regarded as ongoing experiments, guided by a dynamic evaluation process. The primary aim of this paper is to contribute to the construction of such a process.

The paper is organized as follows. Section 2 presents the theoretical foundations of DE, based upon the two literature streams mentioned above. Section 3 introduces DE principles and methodology. Section 4 discusses the Green Communities project, which served as a test-bed for the design and development of DE, as well as for the ICT tools to support it. Section 5 describes the procedures and tools used to monitor the evolution of the case study, while section 6 analyzes the DE of the GC project. Section 7 presents an overview of the main lessons learnt for evaluation practices and their implications for project planning and program policy-making. The final section offers some concluding remarks.

2. Theoretical foundations

2.1. Innovation and complexity studies

Our complexity theoretical framework relies on theoretical work in innovation dynamics initiated by Lane and Maxfield (1997, 2005) and developed by other authors, including Villani, Bonacini, Ferrari, Serra, and Lane (2007), Read, Lane, and van der Leeuw (2009), and Russo (2000). Even though complexity theory is not a single body of thought (see for example Simon, 1962, 1973; Anderson, 1972; Holland, 1995), the work by Lane and Maxfield has the merit of providing a synthetic theory of innovation processes and a corresponding minimal ontology for them (including a set of entities and their properties, their interaction modalities, and the dynamics through which interactions and their consequences are ordered in time). The theory takes as its primary unit of analysis innovation cascades, which include the construction and modification of entities (agents, artifacts, and attributions of identity) and the kinds of relations among them (Lane et al., 2009). Individuals and organizations participating in innovation cascades usually do not face situations where they can decide which action to take on the basis of a pre-defined set of possible consequences. Instead, the fact that everything is in the process of becoming and change means that participants face ontological uncertainty: they cannot even imagine which kind of consequences may derive from their actions—nor even the entities and interaction modalities that will mediate between these actions and their consequences. In this context, two concepts are central to analyze how agents generate action: generative relationships and the narrative theory of action (Lane & Maxfield, 2005).

Generative relationships among agents are the locus in which new attributions of functionality (for artifacts) and identity (for agents) arise. Even if ontological uncertainty might make it impossible to predict the consequences induced by a particular relationship, one may still "measure" and enhance the potential it has for generating system transformations. This *generative potential* depends upon several elements: *heterogeneity* among agents (with respect to their attributions, competences or relationship structures); *mutual* and *aligned directedness* (that is, reciprocity and mutual willingness to collaborate to transform a common zone of the agent-artifact space); *permissions structures* (which determine what agents can communicate about, with whom, in which illocutionary modes); and *joint action opportunities* to engage in change processes. A project team trying to induce a transformation in a specific zone of agent-artifact space can monitor possible relationships among project participants with respect to these elements and try to create interactions that enhance the generative potential of the more promising of these relationships. Obviously, generative potential is a moving target (as indeed

are the set of agents involved in a project and the relationships among them), so it must be continuously monitored and nurtured.

The narrative theory of action provides an interpretative frame to understand how agents can act amidst the ontological uncertainty that characterizes innovation cascades (Lane, 2014, 2015). According to this theory, agents act out stories they tell themselves: these stories interpret their present contexts by embedding them in narrative structures,[1] which they have "learned" from stories that circulate in the narrative communities to which they belong, augmented by their own past experience.

A narrative can be thought as a sequence of events having a beginning, a middle and an end. It consists of a cast of characters, a plot that serves to structure events temporally, and a denouement centered on a change happening in some of the agents' characteristics and identities (Lane, 2014). Narratives do not just relate, but also explain and constitute reality (Bruner, 1991). Through their sensemaking function they enable human beings to act in the face of ontological uncertainty, because they provide legitimacy and accountability to these actions (Czarniawska, 2004). They help agents to explain the correlation among events in a process, and to encode data that may be relevant for the analysis of a wide range of organizational phenomena (Pentland, 1999). For all these reasons, they provide thick but synthetic descriptions, which make them privileged hermeneutic units and analytic tools for evaluation (Anzoise and Sardo, 2013), although there is still some resistance to the use of this non-standard type of "evidence" in policy-making (Epstein, Farina, & Heidt, 2014).

2.2. Combining evaluation with innovation and complexity studies

In recent years, there have been an increasing number of contributions applying complexity concepts to evaluation theories, practices, and methods (Walton, 2013). These complexity-informed evaluation approaches consider the interactions among components in a system as non-linear, and as giving rise to 'emergent' properties, which cannot be understood just by examining the components separately. Moreover, the interplay of different layers and the high interconnectedness among systems components imply that a change in one of them may have either a negligible or a large effect on the system as a whole (Byrne & Callaghan, 2014), making it difficult to forecast all the possible patterns of transformations that can occur. That is why

---

[1] A narrative structure is a kind of template for narratives. It consists of a set of character types, abstract descriptions of identities; plot elements, which consist of a set of allowable transformations for character identity, as well as the physical and social laws that determine how contexts may change outside of the characters' control and what kinds of coincidences are "normal"" (Lane, 2014).

systems with similar initial conditions may end up developing completely different properties over time (Room, 2011). One consequential proposition is that in evaluation practices one-size-does-not-fit-all projects, since they have to be adapted to the uniqueness of relationships and changes happening in time.

In the next paragraph, we will discuss the Dynamic Evaluation (DE) approach, which has been developed by the authors and other colleagues within the *Emergence by Design* (MD) research project, funded by the European Union.

3. Dynamic Evaluation

The Dynamic Evaluation (DE) methodology is not an all-encompassing evaluation, but can be complementary with others. It shares some features with Responsive, Empowerment, Developmental and Participatory Evaluation (Abma, 2001; Abma, Nierse, & Widdershoven, 2009; Wandersman et al., 2005; Patton, 1994, 2000, 2011; Cousins, 2003; O'Sullivan, 2012; Fetterman, 1994, 1995, 2001). Like these approaches, the principal aims of DE are to enhance reflexivity among project participants, and to foster and support adaptive changes and learning through the continuous generation of feedback loops among the evaluator, the management team, and the other project participants and stakeholders. With respect to evaluation practice, this process generates ideas and suggestions for adaptations and improvements, which proceed in an iterative and cumulative way, taking shape from the agents' interactions, the contexts, and contingencies the project encounters. For DE, fostering and building evaluation capability through participant engagement is a primary form of process use that has an added impact on those involved in it, going beyond the mere use of evaluation findings (Cousins, 2007; Patton, 2000).

What distinguishes DE from the other practices mentioned in the above paragraph is the focus on narratives, both as sources and outcomes of the evaluation, and the fact that its practice and tools rely on complex systems theories and on a specific process ontology. Indeed, DE provides a complexity-consistent approach to understand and monitor the interplay of agency and structure at the local level, which force project participants to think about "a more organic process through which a myriad of different stimuli lead to changing relationships and new sense making across a diverse terrain" (Burns, 2006: 183).

The DE approach aims to provide project participants with structured information regarding what is happening in the project and how the process is interpreted and enacted by the project participants. A Dynamic Evaluator (DE-r) provides the conditions for agents to possibly review

and redefine the projects' objectives, contents, and tasks, and to develop an empathetic understanding of those participating in the cascade of changes the project induces. A principal mission for the DE-r is to enhance the project's *generative potential* for innovation. In order to do so, the following are the main tasks of a DE-r (Lane, 2014, 2015; Anzoise, Gurisatti, & Sardo, 2015):

1) Collect different kind of material (and from different sources) that could be relevant to understand the change processes set in motion by the project;

2) Construct maps of the zones of agent-artifact-attribution space related to the project, to track its dynamics and the change processes stemming from it;

3) Extract and make explicit narratives that describe, from different points of view, how the processes set in motion by the project seem to play out;

4) Alert the project team about divergences as they arise;

5) Provide feedback to project participants and structure consultations among them;

6) Produce project narratives to make retrospective sense out of what happened. These narratives will be used to communicate to others the project experience.

The DE can be conceptualized as an emerging evaluation[2], in the sense that it does not completely adhere to predefined objectives and protocols, but instead it reacts to and elaborates responses based upon the evolution of the project itself.

The development of such a methodology has been the result of a feedback process going from the theory to the case study, following both a deductive and inductive approach. At first, the innovation theory by Lane et al. provided the theoretical frame necessary to start observing the case study. Then, experience on the field served to further design and develop the DE methodology, subsequently applied in the case study.

4. The Green Communities project case study
In order to test and contribute to the theory of innovation dynamics proposed by Lane et al. (1997, 2005, 2009) and to develop the Dynamic Evaluation methodology, the Green Communities project was selected as a case study for several reasons:

---

[2] See also Christie, Montrosse, and Klein (2005) on the notion of (and need for) "emergent design evaluation".

1) It aimed at change the way development projects are usually conducted, reviewing and re-designing both the patterns of growth of mountain economies (and in general of inland areas) and their organizational models;

2) The project designers were motivated by the idea of transforming an existing system of relationships, through the stimulation of new attributions and identities;

3) It was implemented in four territories in Southern Italy, which would provide the occasion to observe how the same project could develop differently in diverse contexts;

4) At the beginning of the project, both the scientific director and the project manager agreed to participating in the evaluation process, providing the evaluators (the authors of the present contribution) access to some relevant information, and introducing them to the other participants (as well as extending permissions to the evaluators and project interactions to engage in close and intensive interactions);

5) The project lasted one year, and this length of time was sufficiently contained to allow the evaluators to monitor the innovation cascades it generated for almost two additional years, within the life-time of the EC Emergence by Design project that was funding the DE-rs work;

6) The project was already equipped with an official ex-post evaluation program that could serve as an object of analysis and comparison.

4.1. Case study history

The Green Communities (GC) project was initiated in Italy in 2012, when the National Association for Municipalities and Mountain Communities (UNCEM) was awarded the contract for a call issued by the Italian Ministry of Environment, under the EU Interregional Operational Program whose funds were specifically dedicated to Convergence Regions, i.e. Campania, Calabria, Sicilia and Puglia.

The official aim of the GC project was to design new development trajectories for self-sustainable communities in mountain and marginalized areas, based on an efficient management of local resources. The following lines were prioritized: 1) To measure the energetic performance of typical public buildings and to produce feasibility studies for further improve it; 2) To valorize forests for energetic and climatic purposes, seen as a crucial asset for communities' transition towards low energy consumption development patterns; 3) To return to each selected area a territorial energy audit. The results of these tasks would have served as guidance for the communities in the decision-making processes regarding local development.

Some members of the project team advocated broadening the project's horizons, to produce an overall *rating system* aimed at assessing territorial sustainability on the basis of twelve dimensions, included mobility, forest and waste management, tourism, welfare, and energy utilization. According to the GC scientific advisor, one of the key proponents of this "enlarged" vision, this tool would have provided community leaders with a new perspective on how to design in a more autonomous way the socio-economic development strategies of their territories, coherent with their endogenous resources, identities, and values. To make this happen, community leaders – particular the mayors – should have become more proactive, changing their identities from policy "executors" to "local development agents". The experience coming from the usage of the rating system in the pilot communities was supposed to enhance the competences of local administrators, civil servants and professionals, and to be seen as a good practice.

For the project implementation, four mountain communities (Fig.1) [3] (i.e. Cilento, Madonie, Pollino, and Titerno-Alto Tammaro) were selected according to specific criteria set by the Ministry of Environment together with Formez[4] and UNCEM. These territories, partially or completely overlap with the administrative boundaries of a National Park or a Local Action Group.

---

[3] Mountain Communities are specific type of administrative organizations that were set in Italy in the early 70s for inland and mountain areas, aimed at providing compensating services for depopulation, which have been recently closed/reformed.

[4] Governmental centre providing services, assistance, studies and training for the modernization of the Italian Public Administration.

Fig. 1 Location of the four pilot communities of the GC project

The project team was organized around the following roles: a managing director, a scientific coordinator, a head of administration, three managers each coordinating one of the three lines prioritized, experts on topics relevant for the rating system development, one operational unit for each pilot community, a steering committee, local political representatives (responsible for channeling the project's goals and engaging other mayors), other internal staff. During the project development, different technical activities had to be carried out. Those pertaining to the rating system development were conducted by experts meeting once every two months to discuss how to structure the tool and how to select indicators and measures. This phase of "reducing the expert knowledge complexity" did not prove to be as simple as expected, so the project management enlisted a third organization, the Green Building Council (GBC), to coordinate this process. Notwithstanding GBC's recognized expertise in developing rating systems, it could not solve the difficulties arising at the level of each single dimension, and the

document finally delivered to the Ministry poorly provided systemic and interdisciplinary guidelines considerably weaker and more sketchy than what the project leaders had anticipated.

The overall project strategy was defined by the management team, while line managers were responsible for carrying on their specific tasks and to coordinate with local technicians. However, all these lines of activity encountered several difficulties along the way, leading them on divergent working paths and causing delays. As an example, project designers had underestimated the fact that most of the forests belonging to the pilot territories were located within National or Regional Parks, whose strict regulations prevented them to be included in the feasibility studies.

Moreover, even if according to the operational plan the involvement of local administrations and of other stakeholders was among the success factors, during the design and implementation phases public authorities took a back seat. As a matter of fact, many of the meetings organized in the communities to meet local representatives and to explain them the project vision and tasks did not involve their active participation. This happened partly because those meetings were not integrated with a more systemic and continuous activity of information and engagement, and partly because some invitations were in some ways "biased" by the local organizers. Consequently, rather than a proactive part in the process, civil servants became to be considered "just" as the expected final users of the rating system and of the feasibility studies. As a confirmation, the formal project evaluation, based on a telephone survey addressed to the mayors of the 111 municipalities formally involved, showed that, after 6 months of project implementation, most of the mayors who replied (the 34,2% of the totality) were only slightly aware of its existence and objectives.

As previously said, beside these technical activities the project designers' vision was to challenge Italian local institutions and authorities' roles, and to demonstrate the need for their reform as a lever for development. Unfortunately, this vision was early undermined by some management deviations: in January 2012, when the promoters received the first tranche of funding and the project started, the president of UNCEM decided to take the formal leadership of the project, and to remove the initial program manager, replacing him with another person. As a consequence, the GC project complexity increased, because the newcomer did not share the same vision of the project initiators. At that time, the DE-rs had just been introduced in the project: their first feedbacks came when it was too late to repair the misalignments, that had already been re-embedded in some participants' narratives, and to steer the project in other directions (see section 6). The project passed the midterm review, but the line concerning the

territorial energy audit was cancelled because of the cumulated delays of the project. As an adjustment, the number of feasibility studies on public buildings was doubled. At the end, the project deliverables were just technical and less ambitious reports and a first draft of the rating system, which has not been further developed.

5. Research strategy and methodology

The GC project was observed and analyzed by DE-rs with a dual objective: on one hand to understand which criticalities an innovation project faces from the moment it is designed, to the period in which it is put into practice, as well as which are the cascades of transformations that it may generate within the space it is intended to effect. On the other hand, the information and observations gathered, as well as the relationships built in time between the DE-rs and the project participants served to the ongoing epistemological and methodological development of the DE, and for the practical application of some of its tasks, procedures and tools in the GC project (sections 6 and 7 below).

The research strategy adopted by the DE-rs to observe and monitor the GC project shares the fundamentals of qualitative research, and was supplemented by the triangulation of data sources and collection procedures, combining quantitative and network analysis with interpretative and narrative approaches (Czarniawska, 2004; Labov, 1972; Plummer, 1995; Bruner, 1990, 1991; Chase, 2005; Webster & Mertova, 2007; Hampton, 2009), and process-tracing methods (Chapman, 2014, Fetterman, 1995, George & Bennet, 2005). In accordance with the ontology of the innovation theory, the units of analysis were *agents* (and their identities), *artifacts* (and their functionalities), *attributions*, *relationships*, *narratives*.

As previously said, to pursue the final research objective of developing the DE methodology, the *case study* has been identified as the most adequate research design, especially useful for the exploration and the understanding of phenomena and issues which are "too complex" for surveys or not controllable or manipulable as needed in (quasi) experimental designs. In particular, a comparative case approach[5] has been used to cope with issues and challenges related with this research design (i.e. in terms of explanation building, data collection, external validity), through the observation of four pilot communities of the GC project and of some

---

[5] The DE was developed within one of the work package of the ''Emergence by Design (MD)'' project (www.emergencebydesign.org), funded by the European Union Seventh Framework Programme [FP7/2007-2013] under grant agreement no 284625. In addition to GC, other case studies were developed by project partners Kennisland (NL) and the University of Warsaw (PL).

additional unexpected follow-on projects (e.g. the Rural Design project in the Titerno-Alto Tammaro mountain community).

The main aim of the information gathering and data analysis were to (1) display the dynamics in the structure of agent-artifact space, especially from the relational point of view; (2) highlight the changes in the alignment or misalignment in agents' attributions; (3) recognize and describe the narratives enacted by the participants in the project. Therefore, the project was monitored through interviews with different agents involved in the project, participation by DE-rs in project meetings and events, and fieldwork in the four pilot territories.

In general, the strategies for data collection have been open, redefined over time and applied to different sources of information, with the aim to continue to expand the empirical base of data. Complex and quantitatively relevant data have been reduced through cross tabulations and cluster and network analysis, in ways that could convey a meaningful depiction of the ongoing processes (at least from the particular zone of the agent-artifact space under observation), showing or suggesting relationships, direct or indirect causal chains (e.g. designed or forced coordination among agents *vs* emergent coalitions[6]). Nonetheless, the design of the DE methodology confirmed that data sources cannot be completely defined *a priori*: they may emerge during the process deployment. Because of the difficulty and feasibility to handle large amounts of data, evaluators have explore those processes that seem to have the greatest potential for discovery and for the generation of (new) hypotheses (or questions). That is why, even if this case study relied on theories, ontologies and questions guiding the observations, the type of data to collect and the way they had to be collected have been revised several times.

The sources of data that have been *produced* and *analyzed* have been field notes, transcripts of meetings, conversations and interviews, together with other material not directly generated by researchers (i.e. emails exchanges and online chats/Skype calls, Facebook threads, participants' notes and diaries, official documents and presentations, media contents, etc.). Audio recordings and photos of fieldwork activities have also enriched the empirical base of data, and they supported the analysis and the cross-checking of the evidence gathered.

In particular, interviews were mostly semi-structured or open, and have been helpful to: contextualize agents' histories; highlight the attributions they assigned to themselves, to other

---

[6] In particular, the Dynamical Cluster Index method (Villani et al., 2013) has been applied to the participation of 175 agents over 101 events occurred in the GC project (Villani, 2015)

agents' identities and functionalities, and to artifacts generated by the project (especially the rating system under development); and detect possible differences in agents' meanings. Furthermore, during the project many of the participants have been interviewed multiple times, in order to gather information about changes in their attributions.

5.1. ICT based toolset to support Dynamic Evaluation

Using multiple sources and methods enhances the scope for insight generation and enables important data to surface. The challenge is to juxtapose these methods in ways that shed more light on the dynamics of interventions (Burns, 2006). Therefore, documents and information gathered have been stored in an archive and a database, designed according to the Lane and Maxfield ontology. A particular attention has been paid to: agents (background), artifacts (functionalities, origins, owners, location, users and their attributions of functionalities), organizations (type of activity, location), relationships among entities (categorized on the basis of their typology, duration, intensity). These tools have been designed also for recording the changes of these variables over time.

As previously said, data have been analyzed both with existing software for social network analysis and for detecting emergent structures and patterns (i.e. Dynamical Cluster Index), and others methods specifically designed for the DE tasks (Anzoise et al. 2015). In particular, a Narrative Modeling Language (namely A4 Language, standing for Agents, Artifacts, Attributions and Actions) graphically visualizes, in a Storyboard, the enacted narratives reconstructed by De-rs or provided by project's participants. This tool maps the stories on entities and relationships contained in the database, in order to identify narrative structures (Anzoise et al. 2015). This continuous shift from the micro to the macro level is useful to avoid the risk of losing the "emergent properties of change processes" which happen if evaluators focus their attention exclusively to lower-level micro-phenomena, or to the bigger picture (Hawe et al., 2009: 97). Moreover, the rules of the A4 language allow users to conform to a coherent writing methodology.

Since data collection and analysis occur together, some further considerations regard the issue of the open nature of complex systems: boundaries are constructs with decisions of inclusion and exclusion reflecting the positions of the actors involved in boundary definitions (Lessard, 2007; Munda, 2004). It is also for this reason that retracing the researchers' thinking is of key importance: the process of comparing data set to data set, data with the theory, and then going back to data coding, initially developed by Grounded Theory (Glaser and Strauss, 1967),

together with writing and organizing memos and notes, help sort data into categories, define their properties, and make sense of them by discovering the relationships among categories.

6. Case study analysis: DE outcomes and feedbacks

From the collection and analysis of the information, the DE-rs noticed a deviation from the linear "expected" deployment of the GC project. The causes of this deviation were captured and structured in a feedback composed of three different narratives, intertwined with one another, and discussed several times with the scientific director and a few other project participants (although, as already stated, the DE was not the formal evaluation, and therefore there was a lack of adequate permission structures to share all the results of the DE with all participants and to generate processes of alignment on the basis of these results).

6.1. The Triumvirate

This narrative highlights the conflict of power emerging among the managing team components (two members of the board vs the project manager), which affected both the long-term vision and the short-term implementation strategy of the project.

The relevant issues raised by this narrative had to do with the team-building process (and its final aim), and with particular constraints on this process. The first constraint derives from the fact that the project manager was "imposed" by the Ministry, as it turned out, and he did not share the longer-term project vision with the other members of the board. In time, his way of acting challenged the systems of relationships and interpretations already in place. The alliance among the three main characters of the narrative, that should have been finalized to better "channel" the project for the achievement of its goals, collapsed.

This story shares many similarities with other well known historical power conflicts, like the one among Caius Julius Caesar, Marco Licinio Crasso and Gneo Pompeo Magno in ancient Rome, or the one involving Georges Jaques Danton, Jean-Paul Marat and Maximilien de Roberspierre in Paris during the French Revolution. These stories are centered on a Triumvirate facing a critical situation, structurally exposed to the risk of a very bad end. By choosing this exemplar story to frame for what was going on in the GC project, the DE-rs wanted to suggest to the project promoters that they reinterpret their actions in the light of the effects they were having on the whole project.

What happened in the GC project was a splitting of participants into two groups, one sharing the "utopic" (the attribution of the other group!) and long-term vision of the restructuration of local development strategies (with the GC project working as the first operative episode in that

direction), and the other conforming to the usual technical and bureaucratic project prescriptions. Even if some of the participants recognized that the Triumvirate narrative was "fitting" the context, and even if they tried to undertake a corrective action, in the meanwhile the project permission structures had changed with considerable cascading effects, which did not leave the triumvirs enough time to re-align, even had they been inclined to do so.

In the end, even if this narrative was explicated in time by the DE-rs, the feedback was not powerful and comprehensive enough to steer the trajectory enacted by Triumvirs away from the negative cascade of consequences induced by the already simmering conflict among them.

6.2. The Magic Artifact

The Magic Artifact narrative was developed to highlight the risk represented by the common and (almost) ideological belief in technical solutions to solve problems. The GC project was relying on the capacity of expert knowledge to provide the "right" frame to select, analyze and assess the complex issue of local development. The DE-rs detected the inconsistencies of this narrative conducting interviews and participating in project meetings, realizing how much faith the experts were placing in the capacity of tools and procedures (such as the rating system) to resolve the socioeconomic development problems addressed.

The DE-rs tried to question this narrative through a constant dialogue with some of the GC experts. As a result of these conversations, some of the experts re-evaluated the role of technical artifacts in resolving the addressed problems, while others did not. The explanation relies on the fact that the narrative guiding the latter group had as its main denouement their change of role into local consultants, applying technical procedures in several territories, spreading social "good", and earning large consultancy fees.

However, the first group was instrumental in leading some of the follow-up activities to the GC project (e.g. the Rural Design project), and discussions around the Magic Artifact narrative were very important to their alignment around the need for alternative narratives that incorporated the lessons from Section 6.3.

6.3 The Missing Bottom

This narrative was constructed to highlight the lack of an adequate engagement of local stakeholders (not only the mayors and the local teams, but also the larger communities involved) in the definition of the social problem. Indeed, a larger participation in the implementation of policies can be an opportunity to ensure the project success, because it may help in reducing the resistance or distance of some members of the community, and in

mitigating potential experts' biases (e.g. concerning needs assessment). This narrative structure refers to familiar stories in which a group of idealistic "leaders" fails to mobilize a mass around their ideas – which the leaders just assume they would be attractive for the "bottom", and sufficient to drag them in playing the pre-assigned roles coherent with their narrative.

In the GC project, the Missing Bottom story is tightly intertwined with the two narratives framed above. The local "productive citizens", which the GC promoters were assuming would have been attracted to participate in the change processes they were initiating, were barely aware of the existence of the project, because the local political leaders were not sharing it to others, counting on to grab its future revenues. Through the discussions with some of the local experts and with other people encountered during fieldwork, the DE-rs detected some "latent" narratives revolved around the everlasting dichotomy between the rich North and the underdeveloped South, as well as between expert and lay knowledge. These narratives, if made explicit, could have been counteracted by a new narrative, which differentiated this project and these experts from seemingly "similar" ones in the past. Such a narrative would have to feature the leadership role that this project and these "foreign" experts wanted local actors and the local communities to assume. The DE-rs tried to shed some light on the "latent" narratives that were blocking GC progress, in order to "persuade" the management team to develop a new narrative based on community engagement, which foresaw the encouragement of agents from the four communities to give shape to the "missing bottom". But even if some attempts in this direction were initiated, the local agents did not align around this possibility.

7. Lessons learnt and conclusions

Focusing on the intended use of evaluation, as described in the first two sections, requires making deliberate and thoughtful choices, which can provide the DE-r with adequate conditions to be effective.

The main aim of conducting a case study was to learn from it and to serve the researchers in the development of an evaluation methodology suitable for tracking and analyzing complex processes. As mentioned in the theoretical section, taking a complexity perspective towards observing a project means to deal with different components that make up a system, and to look at how they relate and therefore influence one another in a dynamic perspective.

In this section, we analyze a set of issues that the DE-rs were able to detect during the Green Communities project, which corresponded to some actions from their part, together with some of the project *cascades of transformations*. As it is shown in the next paragraphs, different

criticalities may characterize a project deployment, such as building the project team, identifying and constructing relationships with the identified beneficiaries and with all the other social groups directly or indirectly affected by it, designing the project's permission structures, enacting the first design and providing opportunities for joint action and further alignment, and more in general dealing with unforeseen issues. All of them relate with the conditions for the potential of relationships to fully deploy, and with the generation and management of change processes.

7.1 Monitoring the generative potential of relationships

Most of the issues that arose during the GC project deployment are related to its top-down, centralized and not-so-flexible organizational structure, which hindered the possibility to re-open spaces of project re-design, and therefore to find any alternative or adjustment to cope with unexpected conditions and drawbacks. This lack of openness and flexibility is in part related to the lack of an evaluation process aimed at continuously questioning the current situation and to give feedback on how best to proceed. The DE-rs were not formally hired by the project, and their presence as observers was accepted but not fully integrated in the project. Therefore, DE-rs could just try to alert the management team about the project deviations. For example, during the course of the project the De-rs came to know that the "gestation phase" of the Green Communities project had begun a couple of years before its formal start, by a group of researchers, professionals and practitioners tied by friendship and previous collaboration experiences. They had sporadically participated in formal and informal discussions, also writing books regarding the dramatic conditions of rural and mountain areas, and on how to develop coherent and sustainable policies (see Borghi and Letta, 2009). After some time, they managed to submit a project centered on these issues to a national call. In order for the proposal to fit the call's formal requirements, it had to be resized in its goals, activities and organizational structure. The project was approved, but it started with some months of delay, and the team had to speed up the planned activities. The managing director and the scientific coordinator asked some of the people who were part of the previous informal planning group, as well as to others, to join the project, but very little space and time were dedicated to check if, and how much, they were all sharing the project vision (i.e. *aligned directedness*). In addition, there was a significant mismatch between the roles assigned a priori by the management to project participants and collaborators, and the roles as perceived and enacted by them.

Beside these weaknesses, the management team underestimated or took for granted some other issues: the commitment of communities and mayors, supposed by the project team to be the

main project beneficiaries; the assessment of the communities' "real" needs (e.g. was the rating system addressing a relevant issue for them?); and the project team's autonomy in organizing meetings and other activities to engage other local stakeholders. If the desired response to the implementation of sustainable local development projects - which are experiments of policy-making by doing - is the change of identities and actions, then the issue of agency has to be introduced as a crucial factor to be explored, monitored and supported. This relates with the selection, recruitment, and support of the "right" agents, especially those who are, at least initially, part of the core team and have the responsibility to frame and translate the project's main objectives into practices and interaction modalities.

Furthermore, the agents' selection has to take into consideration also the *heterogeneity* condition, which in the case of the GC project wasn't given enough attention (e.g. in terms of gender and competencies). Moreover, especially in sustainability-related projects, the soft skills and interactional expertise that are more project-dependent (Collins & Evans, 2002), have to be intertwined with an attitude towards interdisciplinary and inter-sectorial dialogue (rather than a high specialization in one field).

For these reasons, building the formal organization that will run the project - according to certain organizational choices regarding roles, responsibilities and interaction modalities - is a crucial process. These choices shape the meanings of social actions and become, in this way, performative, because they interact with the processes through which individuals express their subjectivity and build group identity (Czarniawska, 2004), which in turn has also influence on the transformative and generative capabilities of the organization itself.

As long as the project develops and other individuals and organizations join it, concepts and dimensions related to the issues the project addresses evolve, and therefore they have to be recalled and questioned, shared with the newcomers and not just mechanically transferred. In some cases, they have to be revised according to the new values these people are bringing with them, because they are contributing to the increase of the overall team heterogeneity. The greatest challenge is not so much to structure the problems, but to make them a matter of discussion among participants and a part of their learning process (Hisschemöller and Hoppe, 1996). This means that the *alignment* of participants' *directedness* over time has to be monitored, and supported by *joint actions*: something that was weak (if not missing) in the GC project.

7.2. The importance of governance and scaffolding structures

A crucial requirement for the successful unfolding of a project or program is the design of an adequate system of governance and of the related *scaffolding structures* (Lane & Maxfield, 2005), able to exploit the human and material resources at disposal, and to cope with unexpected circumstances. In this sense, regulations should be produced, monitored and enforced with the aim of "stabilizing" the project organization when it faces perturbations. At the same time, these regulations should be made for being changed whenever the lack of adaptation may hinder the project survival. Practically, one possibility is to design a system with a certain degree of formalization, but at the same time a partial redundancy in roles, allowing, for instance, the covering of a task when its main responsible is missing, and avoiding as much as possible the costs of information transfer.

The project itself can work as a "scaffolding structure" if it establishes links and allows micro-level entities to learn from one another's experiences. For this to happen, further requirements are needed: the project's members must be adequately trained or experienced in taking care of the structure (i.e. identification, recruitment and monitoring of required roles and functions, roles' coordination, management of the team enlargement, etc.). Moreover, there must be a shared agreement on the basic *permissions structures* regulating the team's interactions and commitment. If these conditions are lacking, the structure can be undermined, paving the way to misunderstandings and drawbacks.

7.3. The role of evaluator and the evaluation practice

Spaces, permissions and timely action opportunities experienced in the GC case study showed that that the Dynamic Evaluator has to be integrated with the project activities from the very beginning (Smith, 1994; Patton, 1994, 2011), and that "evaluative thinking needs to be involved in all stages of the solutions planning and implementation process and needs to be closely coupled with knowledge translation at all stages" (Tannahill & Sridharan, 2012: 164). On the other hand, for the evaluation practice to be effective and useful it has to be conducted with the support of some internal figures belonging to the organizations involved in a project or program. In such an integrated view of evaluation "commitment to intended use by intended users should be the driving force in an evaluation" (Patton, 2000: 436).

Given such a premise, the basic conditions for the DE are:

- Existence of a group (even informal) of project promoters that can define and provide the preliminary narrative structure explaining the overall process of change envisioned (these agents are, in principle, among the primary intended users of the evaluation).

- Consensus on the importance of recruiting (and budget to do it) an evaluator, or an evaluation team, with expertise in narrative analysis and preferably already "familiar" with the narratives and narrative structures (established, prevailing or alternative) characterizing processes similar to those envisaged by the project's promoters.[7]

- Mutual recognition of roles, trust relationships8, clear division of responsibilities, and an adequate permissions structure for the DE-rs and project participants to enable the DE to carry out its tasks.

These basic conditions are necessary (but not sufficient) in order to achieve: a strategic alignment between the project promoters and the DE-rs; a constant verification of the five pre-conditions for generative relationships; and a solid agreement with the participants about the project's directedness and the evaluation aims.

A further lesson learnt concerns the responsibility and tasks of a De-r, which is not to empower vulnerable stakeholders, but rather to enhance a dialogical process (Abma & Widdershoven, 2005) among all the stakeholder groups involved during the process, by gaining mutual understanding of each others' perspectives and mutual learning (Widdershoven, 2001). From this dialogical perspective, empowerment is always relational (VanderPlaat, 1999) and can lead to the generation of (positive) feedback loops. Using participatory techniques and a narrative approach can help to bring out what is going on and what is perceived and expected by participants, considering and activating ''all'' the values and the ''usable knowledge'' available (Lindblom & Cohen, 1979; Schön, 1983; Fareri, 2009), and making them a matter of discussion able to foster action and distributed learning.

Indeed, even when they are not explicitly formulated, projects narratives play a key role in guiding participants' actions. By making them explicit to themselves and to others, the project participants have the possibility to assess the coherence (in the sense of connecting past and present contexts, and contemplated actions with the desired future directionality) and

---

[7] As similarly stated by Czarniawska (2004:5) "to understand a society or some part of a society, it is important to discover its repertoire of legitimate stories and find out how it evolved". Nonetheless, narrative structures are not a fixed a well-defined bunch; and even those that might seem as "the most common" or imagined as "similar", in the end never share exactly the same structure and order.

[8] Trust relationships are a complex issue that would require to be detailed more. Suffice here to say that "trust is an emergent property of generative relationships: it grows as participants come to realize the unforeseen benefits that the relationship is generating" (Lane and Maxfield, 2005)

completeness (i.e. the capacity to incorporate, without the loss of narrative logic, additional information about relevant events and entities) of the narratives that are guiding their project-related interactions (Lane 2014, 2015). Therefore, to succeed in detecting, making explicit and "mapping" dominant, hidden or emerging narratives, the project team and the DE-r have to provide adequate ways (or settings) that allow to uncover and put into questions participants' taken for granted notions, values, patterns and interaction modalities, since they inform the project's orientation and expectations. In this way, DE-r can carry out the activities of uncovering the value of lay participant narratives and making accessible and relevant the "situated knowledge" (Epstein et al. 2014)[9] they can convey, coding and comparing narratives according to the information provided and gathered, so as to detect divergences among agents, transformations occurred, etc. The DE-r does it, "scene by scene", using as starting references the narrative structures the project is embedded in, and those shared by the narrative community the project refers to.

7.4. Limits of this study and challenges of the approach

Further lessons learnt from the GC case study concern the DE limits and challenges. At the time of the events analyzed above, the role DE could play in such complex processes had not been sufficiently worked out. Retrospectively, the problem can be described as: what actions could have the De-rs undertaken in order to use the narratives that were emerging and developing, to enhance the directedness among the project participants? In particular, how could a meta-narrative, inspired by an alternative *narrative logic*, have been developed by the DE-rs together with the project participants to avoid the (inevitable?) unhappy ending of the Triumvirate, steering the project into a diverse direction?

First of all, it should be kept in mind that within the GC case study the DE-rs also paid the price for the fact that the DE tools and methodology were at its first stages of development. Even now, DE does not yet exist as a consolidated practice.

Moreover, with respect to the use of narratives, it must be acknowledged that:

---

[9]Particularly with the growing academic (Claes, van Loon, Vandevelde, & Schalock, 2015) and political (Haskins & Baron, 2011) insistence on evidence-based policy making, policy makers tend to privilege economic, technical, and other quantitative data, statistical analyses, and formal 'premise-argument-conclusion' argumentation when making and explaining their decisions. However, members of the lay public tend to offer context-based reflections on first-hand experiences with the particular problem or general domain that policy makers are proposing to address—a form of contribution we call 'situated knowledge' (cited in Epstein et al., 2014: 243–244).

- Everlasting agent-generating narratives do not exist; they will remain approximately stable (with some re-adjustments in time) only as long as they are accepted as a common frame to position and support coherently participants' particular narratives.

- They cannot be all-encompassing: a narrative is always read and told from a particular point of view, and it is not "enough", alone, to drive changes. Dominant discourses are inscribed in societal institutions, texts, settings, behaviors and material culture, giving them enormous advantages, and alternative discourses could remain marginalized (Witkin, 2010).

- They have a "dark side": each group's narrative privileges some voices and silences others (i.e. they express and are expressions of power). Finding the silent voices and revealing the different values as well as the sources of power constitute one purpose of narratives' deconstruction (Martin, 1990). The evaluator, while unfolding her pro-active role by constantly interacting with the project's initiators and other agents, should always carefully consider the dynamics of power underlying the relationships among agents, "to ensure that the concerns and forms of knowledge held by less powerful people are not excluded—and that these people can mobilize their knowledge to effect change" (Pettit, 2010:820). This condition - inherent in the process of evaluation itself - is frequently overlooked, although it has been specifically addressed by Responsive and Empowerment Evaluation, De Baur, Van Elteren, Nierse, and Abma (2010).

8. Final remarks

We strongly agree with the point of view articulated in the constructivist approach to evaluation (Guba & Lincoln, 1989), that the relevant outcome for an evaluation is the evaluation process itself, rather than its specific ex-post findings (Patton, 1996). Indeed, if performed (and unfortunately it is common to do so) as a mere assessment or "judgment", with a steep learning curve all on the part of the evaluator rather than on the participants and the project team (Regeer, Hoes, van Amstel-van Saane, Caron-Flinterman, & Bunders, 2009), evaluation does not offer much added value for the project implementation and for those directly and indirectly involved in it. On the contrary, an evaluation process should understand and reconstruct the plurality of the points of view it involves in time, and should be viewed through its intrinsic meaning of *giving value*. This happens when the evaluation take on the task of developing and enhancing the reflexivity of project designers, project managers and participants, by questioning and alerting, whenever necessary, the project's directionality and its coherence with its current set of procedures and configurations. The individuals and organizations involved in the evaluation

should conceive it as a not-standardizable-process, which can re-open the space of the project's design (Anzoise and Sardo, 2015). In particular the organizations, through the generation of feedback loops, can become what Nooteboom (2001) has called *learning organizations,* which nurture and modify their actions as they accumulate experience and knowledge. Therefore, facilitating, co-designing, and setting the conditions for learning and reflective action correspond to an interpretation of the evaluation knowledge and of the evaluation practice as having the potential to provide a bridge from the "problem space" into a richer and more useful "solution space" (Tannhill and Sridharan, 2012). Achieving this, however, requires the expansion of the evaluation theory and method, and a stretching of the role of the evaluator (Patton, 2010:58).

The present contribution is intended to enrich the debate on evaluation methods and on how to plan sustainable development projects, thus adding a new perspective on process evaluation and management. Not only it is necessary to better coordinate and intertwine project evaluation with project management activities, as well as to adapt them to the specific contexts, but also to reflect on the way projects and programs are currently framed. A change has to be carried on in the way calls for public funding are written, projects and policy-making are conceived, and in particular in on which kind of resources (material, immaterial and of time) are necessary for project design, implementation and evaluation. This issue is anything but trivial: in many cases the formal structures in which projects have to fit (e.g. budget, contract and time constraints) prevent project participants from capturing, understanding and taking full advantage of the cascades of consequences that can arise from the processes the project sets in motion.

References

Abma, T. A. (2001). Reflexive Dialogues: A Story about the Development of Injury Prevention in Two Performing-Arts Schools, *Evaluation* 7(2), 238–52.

Abma, T. A., & Widdershoven, G. A. M. (2005). Sharing stories: narrative and dialogue in responsive nursing evaluation. *Evaluation & the Health Professions*, 28(1), 90–109.

Abma, T. A., C. J. Nierse and G. A. M. Widdershoven (2009). Clients as Partners in Responsive Research: Methodological Notions for Collaborations in Mixed Research Teams, *Qualitative Health Research* 19(3): 401–15.

Anderson, P. (1972). More is different: Broken symmetry and the hierarchical nature of science, *Science*, 177, 393-396.

Anzoise, V. & Sardo S. (2013). Dynamic Evaluation process: the co-design of a vision. Deliverable 3.1. *MD project.*

Anzoise, V. Gurisatti P. and Sardo S. (2015). Innovation policy and projects' narratives. Deliverable 3.3. *MD project*.

Arthur, W.B., S.N. Durlauf, and D.A. Lane (1997). Introduction. In *The economy as an evolving complex system II*, ed. W.B. Arthur, S.N. Durlauf, and D.A. Lane (pp. 1-14). Reading, MA: Addison-Wesley.

Borghi E. & Letta E. (2009) (eds.). *La sfida dei territori nella Green Economy*. Bologna: Il Mulino.

Bruner, J. S. (1990). *Acts of meaning*. Cambridge, MA: Harvard University Press.

Bruner, J. (1991). The Narrative Construction of Reality. *Critical Inquiry*, 18 (1), 1-21.

Burns, D. (2006). Evaluation in complex governance arenas: The potential of large system action research. In B. Williams & I. Imam (Eds.), *Systems concepts in evaluation: An expert anthology* (pp. 181–196). Point Reyes, CA: American Evaluation Association.

Byrne, D., & Callaghan, G. (2014). Complexity theory and the social sciences: The state of the art. Oxon: Routledge.

Chase, S. (2005). Narrative inquiry: Multiple lenses, approaches, voices. In Denzin, N.K. & Lincoln, Y.S. (eds) *The SAGE handbook of qualitative research*. 3rd edition (pp. 651-679). Thousand Oaks, London, & New Delhi: Sage Publications.

Christie, C.A., Montrosse B. & Klein, B.M., (2005). Emergent Design Evaluation: a case study. Evaluation and Program Planning, 28, 271-277.

Claes, C., van Loon, J., Vandevelde, S., & Schalock, R. (2015). An integrative approach to evidence based practices. *Evaluation and program planning*, 48, 132-136.

Collins, H. M. & Evans R. (2002). The third wave of science studies studies of expertise and experience. *Social studies of science* 32(2): 235-296.

Cousins, J. B. and Earl. L. (1992). The case for participatory evaluation. *Educational evaluation and Policy Analysis*, 14, 397-418.

Cousins, J. B., & Whitmore, E. (1998). Framing participatory evaluation. *New directions for evaluation*, 1998(80), 5-23.

Cousins, J. B. (2003). Utilization effects of participatory evaluation. In T. Kellaghan & D. Stufflebeam (Eds.), *International handbook of educational evaluation* (pp. 245–266). Boston, MA: Klewer.

Cousins, J. B. (Ed.). (2007). Process use. *New Directions for Evaluation*, No. 116.

Czarniawska, B. (2004). *Narratives in Social Science Research*. London: Sage.

Cabrera, D., Colosi, L., & Lobdell, C. (2008). Systems thinking. Evaluation and program planning, 31(3), 299-310.

Chapman, S. (2014). A framework for monitoring social process and outcomes in environmental programs. Evaluation and program planning, 47, 45-53.

De Baur, V.E., A.H.G. Van Elteren, C.J. Nierse & T.A. Abma (2010). Dealing with distrust and power dynamics: asymmetric relations among stakeholders in responsive evaluation. *Evaluation*, 16(3): 233-248.

Epstein, D., Farina, C.R. and Heidt, J. (2014). The Value of Words: Narrative as Evidence in Policy Making. *Evidence and Policy*, 10(2); Cornell Legal Studies Research Paper No. 14-23.

Fareri, P. (2009). *Rallentare. Il disegno delle politiche urbane*, Milano: Franco Angeli.

Fetterman, D.M. (1994). Empowerment evaluation, *Evaluation Practice* 15(1), 1–15.

Fetterman, D.M. (1995). In Response to Dr. Daniel Stufflebeam's Empowerment evaluation, objectivist evaluation, and evaluation standards: where the future of evaluation should not go, where it needs to go. October, 1994, 321-338. *American Journal of Evaluation*, 16, 179-199.

Fetterman, D. (2001). *Foundations of empowerment evaluation: Step by step*. Thousand Oaks, CA: Sage.

George, A. and Bennet, A. (2005). Process- tracing and Historical Explanation. In George A., and Bennet A., *Case studies and Theory development in the Social Sciences* (pp. 205-232). Cambridge, Mass.: MIT Press.

Glaser, B. and Strauss, A. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*, Chicago: Aldine Publishing Company.

Guba, E. and Y. Lincoln (1989) Fourth Generation Evaluation. Newbury Park, CA: Sage

Haskins, R. & Baron, J. (2011). Building the connection between policy and evidence: The Obama evidence-based initiatives, London: Nesta.

Hampton, G. R. (2009). Narrative policy analysis and the integration of public involvement in decision making. *Policy Sciences: an international journal devoted to the improvement of policy making*, 42 (3), 227-242.

Hawe, P., Bond, L., & Butler, H. (2009). Knowledge theories can inform evaluation practice: What can a complexity lens add? *New Directions for Evaluation*, 124, 89–100

Hisschemöller, M. & Hoppe R. (1996). Coping with Intractable Controversies: The Case for Problem Structuring in Policy Design and Analysis. *Knowledge and Policy*, 4(8):40-60.

Holland, J. (1995). *Hidden Order: How Adaptation Builds Complexity*. Reading: Addison-Wesley.

Labov, W. (1972) *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.

Lane, D.A. & Maxfield, R. (1997). Foresight, complexity and strategy, in Arthur B., S. Durlauf and D. Lane, *The Economy as an Evolving Complex System II*, Addison-Wesley.

Lane, D. A & Maxfield R. (2005). Ontological uncertainty and innovation, *Journal of Evolutionary Economics*, 15, 3-50.

Lane, D.A. & Maxfield R., Read D. and van der Leeuw S. (2009). From Population to Organization Thinking. In Lane et al. (eds.), *Complexity Perspectives in Innovation and Social Change*, Methodos Series 7, Springer.

Lane, D. A. (2014). *The Innovation Society and Social Innovation*, Deliverable 2.2. INSITE Project.

Lane, D. A. (2015). *Innovation Dynamics, Innovation Society, Social Innovation*, Deliverable 2.3. *MD project*.

Lessard, C. (2007). Complexity and reflexivity: Two important issues for economic evaluation in health care. *Social Science and Medicine*, 64(8), 1754–1765.

Liebman, J. L. (2013). *Building on recent advances in evidence-based policymaking*, Washington, DC: Brookings Institution.

Lindblom, C. E., & Cohen, D. K. (1979). Usable knowledge: Social science and social *problem solving*. New Haven, CT: Yale University Press.

Lundin, R. A., & Söderholm, A. (1995). A theory of the temporary organization. Scandinavian Journal of management, 11(4), 437-455.

Martin, J. (1990). Deconstructing Organizational Taboos: The Suppression of Gender Conflict in Organizations. *Organization Science*, 1, 339-359.

Munda, G. (2004). Social multi-criteria evaluation: Methodological foundations and operational consequences. *European Journal of Operational Research*, 158(3), 662–677.

Nooteboom, B (2001). *Learning and Innovation in Organizations and Economies*. New York: Oxford University Press.

O'Sullivan, R. G. (2012). Collaborative evaluation within a framework of stakeholder-oriented evaluation approaches. *Evaluation and program planning*, 35(4), 518-522.

Patton, M. Q. (1996). A world larger than formative and summative. *Evaluation Practice*, 17,131–144.

Patton, M.Q. (1994). Developmental Evaluation, *Evaluation Practice*, 15(3), 311-319.

Patton, M. Q. (2000). *Utilization-focused evaluation*. Springer Netherlands.

Patton, M. Q. (2011). Developmental Evaluation. Applying complexity concepts to enhance innovation and use. New York: The Guildford Press.

Pentland, B. (1999). Building Process Theory with Narrative: From Description to Explanation, *Academy of Management Review*, 24(4).

Pettit, J. (2010). Learning to do Action Research for Social Change. *International Journal of Communication* 4, 820-827.

Plummer, K. (1995). Life story research. In J. Smith, R. Harré, & L. Langenhove (Eds.), *Rethinking methods in psychology*. (pp. 50-64). London: SAGE Publications.

Read, D., Lane, D., van der Leeuw, S. (2009) The innovation innovation, in Lane et al., *Complexity Perspectives on Innovation and Social Change*, 43-84.

Regeer, B. J. & Hoes, A. C., van Amstel-van Saane, M., Caron-Flinterman, F. F., & Bunders, J. F. (2009). Six guiding principles for evaluating mode-2 strategies for sustainable development. *American Journal of Evaluation*, 30(4), 515-537.

Room, G. (2011). Complexity. Institutions and public policy. Cheltenham: Edward Elgar Publishing.

Russo, M. (2000). Complementary innovations and generative relationships: an ethnographic study. *Economics of Innovation and New Technology*, 9, 517-557.

Schön D. (1983). The reflective practitioner - How professionals think in action. New York: Basic Books.

Simon, H. (1962). The architecture of complexity: Hierarchic systems, *Proceedings of the American Philosophical Society*, 106, 467-482.

Simon, H. (1973). The organization of complex systems, in H. Pattee (ed.), *Hierarchy Theory: The Challenge of Complex Systems*, NY: George Braziller.

Smith, M. F. (1994). Evaluation: Review of the past, preview of the future. *Evaluation Practice* 15(3), 215-227.

Sunstein, C.R. (2011). Empirically informed regulation, *University of Chicago Law Review* 78 (4), 1349-1429.

Tannahill, C. & Sridharan S. (2012). Getting real about policy and practice needs: Evaluation as a bridge between problem and solution space, *Evaluation and program planning*, 36, 157-164.

Urban, J. B., Hargraves, M., & Trochim, W. M. (2014). Evolutionary Evaluation: Implications for evaluators, researchers, practitioners, funders and the evidence-based program mandate. *Evaluation and program planning*, 45, 127-139.

VanderPlaat, M. (1999). Locating the feminist scholar: Relational empowerment and social activism. *Qualitative Health Research*, 9(6), 773–785.

Villani, M., et al. (2007). An agent-based model of exaptive processes. *European Management Review,* 4 141-151.

Villani M. (2015). Modeling Directed Emergence, Deliverable 2.1. *MD project*.

Villani, M., Filisetti, A., Benedettini, S., Roli, A., Lane D.A. & Serra R. (2013). *Advances in Artificial Life, ECAL 2013*. MIT Press, 372–378.

Walton, M. (2013). Applying complexity theory: A review to inform evaluation design, *Evaluation and program planning*, 45, 119-126.

Wandersman, A., Snell-Johns, J., Lentz, B. E., Fetterman, D. M., Keener, D. C., Livet, M., Imm, P.S., & Flaspohler, P. (2005). The Principles of Empowerment Evaluation. In D. M. Fetterman & A. Wandersman (Eds.), *Empowerment Evaluation Principles in Practice*, 27 – 41. New York: Guilford Press.

Webster, L., Mertova, P. (2007). Using Narrative Inquiry as a Research Method: An introduction to using critical event narrative analysis in research on learning and teaching, London, UK: Routledge.

Widdershoven, G. A. M. (2001). Dialogue in Evaluation: A Hermeneutic Perspective. *Evaluation* 7(2), 253–63.

Witkin, S.L. (2010). A discourse perspective on evaluation, *Culture & Society*, 1(1), 109-122.