



Are social mechanisms usable and useful in evaluation research?

Evaluation

2016, Vol. 22(2) 209–227

© The Author(s) 2016

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/1356389016643900

evi.sagepub.com



Erica Melloni

Istituto per la Ricerca Sociale, Milano, Italy

Flavia Pesce

Istituto per la Ricerca Sociale, Bologna, Italy

Cristina Vasilescu

Istituto per la Ricerca Sociale, Milano, Italy

Abstract

This article reflects on the potential contribution of a social mechanism approach to evaluation research, based on the empirical analysis of three evaluations of European policies. A methodological framework for carrying out evaluation based on this approach is put forward that help identify questions to be answered and steps in a working method to be followed. More specifically, the approach focuses on particular types of mechanism: those that stimulate enduring changes in actors' behaviour. Hence the label 'social mechanism'. The underlying hypothesis is that some institutional settings can orientate actors' reasoning in a way conducive to the attainment of policy goals. Moreover, it is assumed that a closer exploration of the workings of mechanisms as part of an evaluation process is useful for both learning and policy design. The article provides examples of evaluations where the proposed approach has been applied.

Keywords

extrapolative case studies, institutional capacity, social mechanism theory, theory of change

Corresponding author:

Erica Melloni, Istituto per la Ricerca Sociale, via XX Settembre 24, 20123 Milano, Italy.

Email: emelloni@irsonline.it

Objectives

Evaluation research has the goal of understanding if, to what extent and why a public policy achieved relevant results. In its role of a supporting process for the improvement of public programmes – i.e. from an *ex-ante* perspective – evaluation has to provide policy makers with usable knowledge that supports this goal.

In order to provide satisfactory explanations of achieved policy results and contribute to policy learning, the analysis must also take into account the conditions specific to a context, the nature of policy actors (their characteristics, goals, values and beliefs) and the way they influence achieved results. This complexity enlarges the policy ‘black box’, requiring evaluators to use new methods to understand how better policy design can contribute to policy success or failure.

While neo-institutionalism sometimes appears over-confident in the explanatory power of institutional configurations (Radaelli et al., 2012), realistic evaluation (Pawson and Tilley, 1997) acknowledges the role of actors as crucial for explaining the success of a policy. However, the general question around policy results is not only whether they have been achieved, and to what degree, but also whether the policy that has stimulated those results can be used as an exemplary case (be it a model, a source of inspiration or something in between) in order to provide lessons for others elsewhere. The point is to understand what works, for whom and in which circumstances. The mechanisms approach has been widely adopted across many disciplines and in evaluation has been diffused through various forms of ‘theory-based’ evaluation. This approach aims to trace the linkages among the complex bundle of elements composing a policy and its context, in order to provide insight on what works in policy making and to enhance the capacity to achieve better results (Barzelay, 2007).

In this article we propose an empirical case study approach to evaluate policies and programmes with the specific goal of producing usable knowledge for policy makers. This working method largely stems from the ‘extrapolative methodology’ proposed by Michael Barzelay in the article ‘Learning from second-hand experience’ (2007). In this article the author critically comments and builds on Eugene Bardach’s Presidential Address to the Association for Public Policy Analysis and Management (Bardach, 2004). Both authors stress the common need, in public policy making, of learning from smart solutions implemented by other policy makers. They suggest working with social mechanisms, a concept that as already observed is certainly not new and in recent times has also been taken up by evaluators.

It is worth noting that, in both Bardach and Barzelay’s view, the goal is quite practical. The main aspiration of ‘extrapolative case studies’ is to provide actors with information that will help them to design policies better. This approach to case studies aims to provide guidance to researchers so that they can ‘effectively investigate practices in source sites to prepare the ground for disciplined and ingenious extrapolation of practices from source to “target sites”’ (Barzelay, 2007: 522). The approach encourages us to expand available tools for evaluation to contribute to policy design and policy innovation. However, the actual use of extrapolation and of its core elements, mechanisms, is not consolidated in policy research nor in practical evaluation. The goal of the article is to present how these suggestions can be applied, recognizing both the strengths and limitations of the approach.

Many definitions have been offered of mechanisms (Mahoney, 2003; Falleti and Lynch, 2009; Hedström and Ylikoski, 2010), which risks making the concept popular but imprecise. We adopt a definition of mechanism narrower than the one given by Bardach (2004: 215;

‘mechanisms are whatsoever advantage deriving from the social or policy environment, from “input substitution” to “nontraditional participants” or “cost-based pricing”’) and by Barzelay (2007: 527; ‘a social mechanism is a fairly general, but only sometimes true, (partial) theorization of complex temporal phenomena in the social world’). More specifically, we are interested in mechanisms triggered by policy arrangements, generating enduring changes in actors’ behaviour. The hypothesis is that in a given context, some institutional and other factors can stimulate actors to change their behaviour in a way that is conducive to the attainment of policy goals. Hence, mechanisms could be considered as triggers for change in actors’ behaviour derived from reflections and new understandings of a situation in which actors find themselves. These changes could take the form of cost/benefit considerations activated by classical ‘sticks and carrots’ mechanisms; or tap into the desire to maintain or increase their own reputation and to realize their values; or derive from responses to peer-group pressure. What is important is that we can recognize some degree of regularity underlying different actors’ changes in behaviour.

The article will first briefly introduce the concept of mechanism, and its use in evaluation research. This is in order to contextualize the evaluative approach proposed later in the article to improve policy understandings by studying successful and failing mechanisms so that lessons can be extrapolated to other contexts. This is followed by three examples of this approach in the evaluation of European policies.

The social mechanism approach: Different approaches and definitions, but a common goal

The concept of ‘mechanism’ is well-established across many disciplines in the social and natural sciences. The concept is now an integral part of the sociological vocabulary (Merton, 1957; Elster, 1998; Hedström and Swedberg, 1998; Tilly, 2001). It also features in European studies (Eberlein and Radaelli, 2010), policy studies (Scharpf, 1997; Mayntz, 2004; Dente, 2014), behavioural economics (Bowles and Gintis, 2011), organizational studies (Bardach, 2004; Barzelay, 2007) and evaluation studies (Suchman, 1967; Chen and Rossi, 1989; Scriven, 1994; Pawson and Tilley, 1997; Weiss, 1997; Donaldson, 2007).

Evaluation research from Theory Based Evaluation to Contribution Analysis (Mayne, 2012), have become particularly fascinated by this fuzzy concept. Chen and Rossi (1989) popularized the notion of mechanisms in theory-driven evaluations, claiming that they provide insights into the causal linkages between a treatment and effects. Promoters of theory-driven evaluation (sometimes known as ‘white box evaluation’; Scriven, 1994) underlined the importance of unpacking the black box so that the inner workings of a programme can be inspected. Scholars such as Suchman (1967), Weiss (1972, 1995, 1997), Wholey (1979), Chen and Rossi (1989) and Donaldson (2007) share the ambition to use mechanisms to understand how and why programmes work, or fail to work, in different contexts and for different programme stakeholders. They have therefore shifted the focus of evaluation from the question ‘Did it work?’ to the question ‘Why and how does it work?’

A renewed interest in mechanisms followed the publication of the seminal book *Realistic Evaluation* by Pawson and Tilley in 1997. According to Pawson (2006), mechanisms:

explain causal relations by describing the ‘powers’ inherent in a system, be those systems substances (like gases and gunpowder) or agents (like examiners or policy makers) or structures (like

bureaucracies or social programmes) . . . The mechanism explains what it is about the system that makes things happen. (p. 23)

Most recently, the updated European Commission's Evalsed guide (2013) identifies two main alternatives for the evaluation of European programmes: counterfactual impact evaluation, aiming at quantifying whether a given intervention produces the desired effects on some pre-established dimension of interest, and theory-based impact evaluation, where mechanisms have a specific role in the explanation of how and why certain outcomes occur. According to Stern et al. (2012), theories based on mechanisms are a specific class of one of the four main approaches to causation, often labelled 'generative' causation. The generative causation framework allows in-depth understanding and fine-grained explanations of complex and context-related causal chains, even though it has potential weaknesses in terms of external validity of the analysis and generalization of results. The core message of all these approaches is that 'causation without explanation is insufficient for policy learning because policy makers need to understand why as well as how they are to use findings from research or evaluation for future policy making' (Stern et al., 2012: 10).

This brings us to the question posed by Bardach and Barzelay: which kind of practical insights useful for policy design can be obtained using mechanisms, and how should we use them in an evaluation exercise? In the remainder of this article we propose an approach to evaluation based on mechanisms. We also provide three examples of evaluations of European policies in which the proposed approach has been applied, together with some conclusions on the strengths and the weaknesses of what is proposed.

A working method to uncover the role of mechanisms in real-world practices

'What works?' is one of the most common questions in evaluation research, aiming at reproducing, ameliorating or mainstreaming desired results in a given policy context. This simple question brings to the fore a bundle of intertwined conceptual elements that need to be unravelled in an evaluation. The working method proposed below sets out to clarify these different elements, by creating a manageable evaluation approach based on mechanisms that can be used in different evaluation contexts. This approach is summarized below in five propositions.

1. Results are changes in actors' behaviours

In a public policy context, results are changes that are judged more desirable than the status quo. In other words, a policy works when it changes (or at least contributes to changing) some undesirable conditions. A corollary is that change does not happen spontaneously: something has to be done in order to achieve desired public policy results (Busetti et al., 2013).

Behaviour change is one of the aims of European policies, promoting convergence around common values and goals. These may include doing more or less of a specific action, taking into consideration someone else's point of view, or incorporating a new technology or competence. This is especially clear in the context of capacity-building policies: how to learn and transfer the complex competencies for managing and implementing effective European policies.¹ Changes in capacity are expected to last longer and are often broader in scope than an initiating policy. This is particularly true when dealing with European policies such as 'European Added Value', a concept often difficult to pin down. It can be narrowly defined

as ‘additional’ value that would not have been achieved without EU support. It can also be understood as the result of enduring and broader changes such as coordination gains, greater legal certainty, greater effectiveness, complementarities, etc. Evaluators have to investigate whether the policy intervention provokes or is capable of achieving lasting changes in actors’ behaviour.

2. *Changes follow actors’ reasoning*

The second proposition deals with what can promote changes in actors’ behaviours. Not all people react to an external stimulus or input in a similar way. People change, or do not change, their behaviour based on personal considerations given the situation to hand, or, as Pawson (2006) put it, on their *reasoning*. A positive or negative reaction could result from different kinds of considerations, such as: anticipations of personal gains or personal losses; the maintenance, increase or possible loss of an actor’s role and reputation; or an actor’s preferences, values and beliefs. Convenience, reputation and values are typical patterns of reasoning, which motivate people to do or not do something, or to do it in a specific way.

Luckily for our powers of prediction, it is often possible to categorize people in groups (e.g. bureaucrats, politicians, experts, etc.) and this classification helps predict sets of actors’ preferences and decisions in a given policy context. For example, bureaucrats are typically deemed to adopt strategies of ‘blame avoidance’ when they do their job. Merton (who described the ‘bureaucratic virtuoso’ in 1957) explained the dangers; and Hood the positive effects of the mechanism of blame avoidance (Hood, 2011). Categorizing policy actors into groups is a useful analytical strategy for questioning who changed their behaviour, and following which kind of reasoning.

3. *Actors’ behaviours are shaped by contextual elements such as institutions and specific policy features*

It is commonplace to say that context matters. Research highlights the particular importance of institutions as part of the policy context. Institutions play a crucial role in explaining why and how people behave in certain ways: sanctioned rules reduce the range of potential actions by specifying required, prohibited, permitted actions/behaviours in a partnership or other setting. Institutions define what is allowed and what is not. Changing institutions to modify actors’ behaviour is possible, but it is often costly in terms of human efforts and other resources. Institutions are therefore usually regarded as a ‘fixed context’ for policy makers. This tends to be the case for European interventions intended to achieve EU 2020 goals. Generally, a policy (or programme) fosters multiple results through a bundle of policy tools directed to different types of actors. In order to assess what part of a policy is responsible for which result, it is necessary to ‘unpack’ this bundle. Tracing the underpinning policy (or programme) processes makes it easier to identify the results (if any), intended and achieved, with regard to a specific category of policy actors.

4. *Mechanisms facilitate the achievement of results*

Social mechanisms can provide convincing explanations for why specific features of the context (e.g. how public employment services are organized within Member States) combined with policy solutions (e.g. apprenticeship or traineeship) shape the behaviours of policy actors

leading to desired results. In our view, mechanisms can be conceived of as triggering change in the motivation and behaviour of actors. These mechanisms are shaped by the situations in which the actors are involved. In other words, mechanisms help explain why a policy led to cooperation from the people involved (be it the implementers or the beneficiaries), and is therefore valuable for policy making and policy learning.

5. *Drawing on literature to identify possible mechanisms*

Very often, in the policy evaluation field and in policymaking practice, there is neither time nor the possibility to conduct experiments or check for success through comparative studies, let alone conduct in-depth ethnographic fieldwork. In order to come close to answering the ‘what works’ question, and, in particular, to detect relevant mechanisms and the conditions that triggered them in a specific case, it is often helpful to draw on the already-considerable literature on mechanisms. This can provide a reference ‘toolbox’ to orientate and inform empirical observations as part of an evaluation.

How we used this working method

Main goals

In many of the evaluation studies we have undertaken, the primary goal was to give convincing explanations of policy implementation, success and failure. The secondary – but even more ambitious – goal was compiling an empirical catalogue of mechanisms and of the way they function in real situations, in order to guide future policy design and resolve future implementation problems.

Extrapolative case studies

For the purpose of our research, we employed the extrapolative case study approach. It is a particular learning-oriented type of case study proposed by Michael Barzelay in 2007. However, in adopting it, we especially emphasized the role and behaviour of policy actors. Extrapolative case studies focus on describing which behaviours and outcomes are explained through which particular combination of contextual conditions and policy features, and on describing which mechanisms facilitate the achievement of ‘outstanding’ outcomes. The elements of an extrapolative case study as we understand them are:

1. Policy outcomes – with particular reference to changes in the actors’ behaviours *congruent* with the policy goal;
2. Policy features – characteristics of the policy and of its implementation process, characteristics, roles and activities of the actors involved;
3. Mechanisms triggering specific policy outcomes;
4. Context features such as institutions, rules and historical events.

Policy outcomes

We broadened Barzelay’s reference to ‘outstanding results’ to include any significant outcome of interest that indicates an enduring change in the behaviours of some of the actors involved

in the policy process. Unfortunately, changes in behaviour are not identified as outcomes measured by systems used to monitor EU policies.² Qualitative interviews with key stakeholders and well-informed witnesses are thus required as part of an evaluation process in order to detect relevant results on the ground. This demands, moreover, a background analysis of the actors involved and affected by the policy.

Policy features

Policy features (or process design features, as Barzelay describes them) are those policy processes that are most strongly connected with the main changes and outcomes. Identifying the most relevant processes depends on the nature of the policy and on the results we are interested in. However, policy makers of complex policies – which European policies usually are – need to take into consideration at least three main types of processes (Busetti et al., 2013):

1. The process of generating and maintaining *engagement* of relevant partners, stakeholders and beneficiaries;
2. The process of enhancing/maintaining/reducing the *role* of certain partners;
3. The *coordination* of participants within a policy network (i.e. those who have some responsibility in implementing these processes).

It is worth noting that attention has to be paid not only to the processes included in the formal design of the policy, but also to implementation processes, including modifications to initial project design and specification that take place over time.

Which mechanisms?

For the purpose of this analysis we have tried to limit ourselves to those mechanisms that might explain change in actors' behaviour deriving from some sort of reflection about the situation they find themselves in. There are many examples of such mechanisms in the literature that we drew on in our research. For instance, we used the mechanisms of:

- attribution of opportunity or threat (Tilly, 2001);
- certification of role and reputation (McAdam et al., 2001)
- blame avoidance (Edelmann, 1977; Hood, 2011);
- performance feedback (Bandura, 1977; Cyert and March, 1992; Barzelay, 2007);
- anticipation of preferences (Scharpf, 1997);
- the 'bandwagon effect' (Granovetter, 1978), etc.

Context

Pawson (2006) points out that uncovering explanatory mechanisms represent only half of the problem. The relation between social mechanisms at work and outcomes is not fixed, but depends also on context. However, identifying what features constitute a context are not necessarily straightforward. Pawson and Tilley (1997) define contexts as the pre-existing institutional, organizational and social conditions that sometimes enable, slow down or block people's choices and interactions. Barzelay (2007) refers to durable situational factors within

Table 1. Guiding questions.

Dimension	Questions	What to search for
Outcomes	What specific changes (if any) were produced in the behaviour of which actors?	Modifications in some actors' behaviour that would not have happened spontaneously.
Mechanisms	Which mechanisms (if any) fostered the change in actors' behaviour?	Implicit or explicit motivations that turned into a change of behaviour.
Policy features	Which parts of the policy (if any) triggered a specific mechanism?	Elements of the policy process that triggered the mechanism (e.g. a granting procedure).
Context features	What context features (if any) sustained/made the triggering of the mechanism possible?	Elements of the context (institutions, rules, historical events), usually not modifiable by the policy, explaining the framework for the action and its constraints.

the undertaking, different from process design features. We would argue that what is key is to identify the relevant factors, which although they usually cannot be altered through policy intervention, nonetheless explain a set of actions and constraints. For example, ways to encourage cooperation among actors could be different depending on whether the organizational context supports prolonged or repeated interactions among actors, or whether the context promotes only occasional interactions, perhaps related to the pursuit of a stand-alone goal.

Explaining linkages

The final part of this working method focuses on the narrative intended to reconstruct the linkages between outcomes, policy features and policy context, and on the explanation of why (i.e. through which mechanisms) the specific outcomes were produced (see also Barzelay et al., 2003).

Four questions presented in Table 1 drove our evaluation studies of specific EU policies.

Applying the social mechanisms approach: Three examples

The following examples are presented to illustrate how this approach can be used in real-world evaluations. The three examples relate to different evaluation studies carried out by the authors. They cover different policy areas (the energy sector; gender equality; and territorial policies), and illustrate different mechanisms at work. Two of these examples refer to ex-post evaluations, while the third deals with an ex-ante evaluation, to illustrate how the approach can be used 'backwards' to support policy design. It is worth noting that the examples presented are not to be regarded as 'best practice' (i.e. more successful practices, compared with similar experiences in the same field). We regard them rather as 'exemplary practice' (i.e. experiences where some interesting outcomes have been produced that exemplify our methodological approach). Examples are presented in order to clarify how specific outcomes, defined as 'change in actors' behaviours', have been favourably affected by different mechanisms, policy and context characteristics. The general intent is to provide insights for policy makers to improve the design and implementation of policies.

The evaluation of the AlpEnergy project

The first example is the evaluation of the AlpEnergy project, a European territorial cooperation project implemented within the Alpine Space programme in 2008.³ AlpEnergy brought together power suppliers, development agencies, research institutes and public administrations from five different regions of ‘Alpine Space’⁴ to address serious energy system challenges of renewable energy supply. These included imbalances between power generation capacities; the load profile arising from patterns and volumes of customer demand; and the performance of the electricity grid even with dispersed power generation.

AlpEnergy aimed to design, develop, evaluate and transfer Virtual Power Systems (VPS) in six different cooperation areas. A VPS is a virtual system for renewable energy management, production and consumption. In particular, the Aosta VPS is seen as representing good practice in saving energy through reductions in domestic energy consumption. The Aosta case fostered energy savings following audits of the energy consumption of 25 private users in Saint Denis, a small mountainous village of around 300 habitants. This was particularly challenging for the VPS policy, as it implied a reorientation of citizens’ energy consumption habits.

Policy outcomes. The VPS test in Saint Denis targeted the reduction and time-shifting the energy consumption of Saint Denis citizens into non-peak periods, as requested by the energy manager of the VPS. The percentage of energy shifted to non-peak periods varied from 45.3 percent to 53.9 percent of the total energy consumed in the test period.

The Allgäu VPS case also focused on energy savings and energy consumption shifting to non-peak periods. In this case the VPS operated mainly through price incentives to private consumers. However, here there were only minor energy savings (2%–3% of the energy saved by participating consumers in comparison to their median consumption during the years 2007–10) recorded in the test phase.

Mechanisms. The main mechanisms triggering the postponement of energy consumption into non-peak periods was performance feedback (Cyert and March, 1992; Barzelay, 2007). The Saint Denis VPS provided consumers with real-time information on their energy consumption and savings through a web-based system and also via SMS (short message service) messages on mobile phone. People were instantly informed in case of peak-energy and were thus able to decide to postpone consumption (e.g. the use of household appliances). As actors who were interviewed underlined (Istituto per la Ricerca Sociale, 2011), feedback on results (i.e. energy saving) increased citizens’ commitment and pushed them towards a continuous reduction in energy consumption, as they had become more aware of their usage, expenditure and waste. Interviews suggested that some participating consumers were disappointed if in a day they had achieved no energy savings.

In addition, competition and a kind of ‘bandwagon effect’ (Granovetter, 1978) contributed to the strong positive outcomes that were achieved. Those interviewed highlighted the triggering of an informal competition among citizens to establish who was better in reducing energy consumption.

Policy features. The Aosta VPS managed both energy production and consumption through a communication network that included the monitoring of energy consumption of all load units in Saint Denis village. The project also monitored energy savings: a website was developed to enable consumers to consult their consumption data; and a computer was installed at the

public library of Saint Denis where consumers could view the graph of their daily electricity consumption. The project also gave the Municipality of Saint Denis a strong role in selecting, engaging and involving consumers in a project perceived as strongly innovative. According to actors interviewed, the local population seemed enthusiastic about being part of an innovative experiment.

Context features. Some characteristics of the Saint Denis context are relevant for positive aspects of project implementation and outcomes. First, Saint Denis has a well-developed sensibility towards renewable energy issues and the role of local public institutions in sustainable development policy. In fact, Saint Denis was chosen as a test location due to its leadership in renewable energy policy among the Aosta mountain municipalities. The municipality had invested considerable resources in sustainability: creation of a centre for the use of a photovoltaic system for public illumination (e.g. street-lights); production of renewable energy (photovoltaic and wind energy plants); sustainable waste and water management; sustainability education; awareness-raising activities to educate citizens about energy saving and energy efficiency, etc. Over time, these actions helped shape citizens' attitudes. For example, recognizing the importance of active participation in sustainable development and supporting the role of public institutions as local leaders in the field of energy and environmental sustainability.

The study on gender mainstreaming in committees and delegations of the European Parliament

The use of the social mechanisms approach in this second example⁵ aimed to provide knowledge about the main drivers favouring or constraining the effectiveness of the Women's Rights and Gender Equality Committee (FEMM) of the European Parliament in promoting gender equality and gender mainstreaming policies.⁶

Policy outcomes. Generally speaking, the FEMM is expected to foster a systematic commitment to women's rights and gender equality within the European Parliament (EP) and mainstream the principle of gender equality in all EP policies. The FEMM is successful when the EP includes gender perspectives in its legislation. This may happen in different ways: when FEMM's own initiatives on relevant gender issues are approved, or when FEMM's policy opinions on major issues or amendments to other EP Committees' legislative drafts are included in the EP final proposals. This latter procedure is intended to mainstream gender perspective in different policies such as transport, environment, regional development, etc.

Between 2011 and 2013 FEMM drafted 18 own-initiative reports, which usually paved the way for the introduction of a gender perspective into policy issues that at first might be thought unrelated to gender equality. Moreover, FEMM delivered 40 formal opinions to legislative proposals relevant for gender issues, which corresponded to a total of 718 legislative amendments or non-legislative suggestions. Only 29 percent of these have been fully or partially integrated in the EP proposals. FEMM also tabled 177 minor gender mainstreaming amendments⁷ to other EP Committee proposals, out of which 85.6 percent were integrated in the EP proposals.

Mechanisms. One of the mechanisms that seem to favour FEMM's effectiveness in mainstreaming gender equality in the EP legislative process is reputation together with actor's certification. The reputation mechanism (Abrahamson and Rosenkopf, 1993) implies that when the source is credible (i.e. perceived as knowledgeable and trustworthy), the recipient will be less suspicious to what is put forward and thus more open and receptive to its details. Actors' certification concerns the validation by specific authorities of actors, of their performance and of their claims (McAdam et al., 2001: 121). Actors' certification is needed when reputation alone is not enough to influence the perceptions of others. Both mechanisms encourage information exchange and decrease the costs of such exchange.

Interviews carried out during the study (Istituto per la Ricerca Sociale, 2014) show that FEMM rapporteurs' expertise, networking capacity, political prestige and interest in gender equality and women's rights issues are essential to successfully pass a gender equality message. When FEMM is not in charge of the procedure, it has to persuade other committees to accept its intervention. The selection of a trustworthy rapporteur is crucial for the success of FEMM's interventions. According to interviews, rapporteurs need to act and be recognized as a role model by other Members of the European Parliament (MEPs) in order to enhance the effectiveness of FEMM opinions and amendments.

The personal prestige can be sustained, or substituted when lacking, by procedural contrivances enhancing role and status certification of FEMM rapporteurs in their relationships with members of other committees.

Policy features. Actor certification can be planned in the initial phase of the legislative procedure, when the Conference of Presidents decides on the role to be assigned to FEMM in the legislative process. It can act as a responsible committee, an associated and joint committee or just as a committee that is asked for an opinion. When FEMM acts as an associated or a joint committee, the possibility of ignoring its concerns is less than when it is completely decoupled from decision-making process. This is why FEMM becoming a joint committee is one of the most important strategies for strengthening the committee's representation and effectiveness.

An exemplary case was the process that led to the drafting of the proposal for a European Directive establishing minimum standards on the rights, support and protection of victims of crime. The drafting of the proposal, prepared under the joint committee procedure between FEMM and the LIBE Committee,⁸ was not only considered by most of the MEPs interviewed as good practice in terms of the inclusion of gender issues but also as best practice in terms of cooperation by LIBE and FEMM. According to interviewees, the process leading to agreement on a joint committee procedure was long and problematic at the outset. LIBE wanted to deal with this matter on its own without involving FEMM, which had requested a joint committee procedure. This disagreement was sent for resolution to the Conference of Committee Chairs who, after discussion, decided to involve FEMM and activate a joint committee procedure. In the event cooperation led to positive results for both committees, each of which was able to bring their own competencies to bear on issues raised by the Directive. While initially LIBE adopted a more legalistic perspective, focusing on equal rights according to their mandate but with no specific focus on gender, the intervention of FEMM resulted in improvements of the text with respect to gender mainstreaming. This example indicates the importance of the role of the FEMM President having the opportunity to access the joint committee procedure through the Conference of Committee Chairs.

The above-mentioned FEMM–LIBE report was highlighted by interviewees as an example of FEMM’s effectiveness in mainstreaming gender equality in the policy process, enhanced by acting as a joint committee.⁹ In comparison, when FEMM acts as an opinion-giving committee, the committee in charge is not obliged nor pushed to take into account all or indeed any of the FEMM suggestions/amendments. In such cases, FEMM has more chance to be influential if the FEMM rapporteur holds a prominent personal position, or if the topic has a well-recognized connection with gender issues.

The context. The specific context in this case is the legislative procedure of the EP. More specifically, according to Rule 43 of the EP’s Rules of Procedure, the EP President refers proposals for legislative acts and other documents of a legislative nature to the committee responsible. Set up in 1984, FEMM is one of the 22 committees participating in the legislative procedure of the EP.¹⁰ FEMM is the main body in charge of promoting gender equality and gender mainstreaming in the EP’s policy and legislative process. It promotes gender equality and gender mainstreaming through: participating in legislative and non-legislative acts; influencing the gender mainstreaming of the work of other EP Committees; preparing and following-up oral questions and plenary resolutions; commissioning studies; organizing public events, hearings, workshops and press releases.

With regard to the EP’s procedures for the implementation of gender mainstreaming in the legislative process of the EP, it is worth stressing that FEMM is in fact a ‘horizontal’ committee with competence on gender equality in all policy fields. On the one hand, this allows FEMM to intervene on every policy issue. However, since FEMM is rarely in charge of the legislative procedure and other committees must therefore accept its intervention, its power is limited. Not only does this status impact on FEMM’s power to promote gender mainstreaming in the legislative process but it may detract from its standing among the members of the EP.

EU strategy has a clear focus on gender equality. Accordingly, the EP put into place different arrangements in order to enhance this principle in European policies. Interviews have brought out a relatively strong commitment on the part of FEMM members and a positive climate leading to effective networking between FEMM’s members irrespective of their specific political affiliations. Despite these efforts, the adoption of a gender perspective in the work of parliamentary committees is still highly variable, voluntary and specific to particular policy areas.

The ex-ante evaluation of the ESPON 2020 Programme

The main ambition of an extrapolative approach is to provide support for policy design. Our third example, the ex-ante evaluation of the ESPON¹¹ 2020 Cooperation Programme (Istituto per la Ricerca Sociale, 2014), in the programming period 2014–20, was theory-based. It was expected to help the managing authority and the 32 States participating in the Programme to strengthen their strategy. The use of mechanisms supported a kind of ‘backward tracking’, from the Programme’s expected outcomes with regard to target beneficiaries, to the mechanisms to be activated in order to encourage behavioural changes and finally to the institutional features intended to influence actors’ reasoning and reflections.

Expected outcomes. The new ESPON 2020 Programme has the general goal of fostering greater awareness, usefulness and use of territorial information for European and regional

policy makers when drafting or monitoring and evaluating programmes. ESPON 2020 also had a host of more specific objectives including: greater focus on responsiveness to policy demands and uptake of ESPON products; more in-house capacity for knowledge production; greater coordination among partners; and reduced administrative burden for stakeholders.

The creation of new institutional arrangements for ESPON based on a single entity (the European Grouping of Territorial Cooperation – EGTC¹²) is intended to act as a catalyst to change the ESPON 2020 Programme implementation process. The new ESPON EGTC is expected to streamline the management of the programme.

Mechanisms. The ex-ante evaluation suggested that ESPON 2020 would achieve better results if it paid attention to specific mechanisms. The first mechanism is the *perception of opportunity* by different types of target groups (such as policy makers, knowledge providers) to participate in ESPON activities. The perception of opportunity depends on the characteristics of the actor and can take the form of grants available (in the case of a knowledge provider), new windows of opportunity for problem-solving (for a policy maker or a technical experts), or increased reputation for actors participating in ESPON activities (for academics and professionals). Basically this mechanism helps to generate and maintain engagement of the target audiences in the Programme.

The second suggested mechanism is the *certification* of the newly established EGTC's role. This could, for example, derive from a perception of EGTC effectiveness. Achieving short-term quality results, together with the communication of those results (feedback on performance), are therefore a priority for the EGTC to be legitimized among Programme partners.

A third mechanism is intended to streamline interactions among ESPON stakeholders, and in particular the authorities formally responsible for implementation, such as a managing authority and monitoring committee, and the new EGTC. This can be achieved through *repeated interactions*, recognized roles and shared rules of cooperation in order to achieve common goals. Repeated interactions favour *trust*, *mutual learning* and *commitment among partners*. For example, direct involvement of the programme's monitoring committee in the assessment of the newly established EGTC should strengthen cooperation in achieving shared goals.

Policy features. These kinds of considerations could, in our view, help policy makers both in the design and in implementation phases. A variety of tools could be used (e.g. for triggering performance feedback goals), as different kinds of solutions could serve a similar goal.

Among the policy features of the ESPON programme intended to foster the above-mentioned outcomes are:

- a new procedure for prioritizing territorial research;
- stronger communication tools; and
- above all, the building up of a new EGTC management structure.

The establishment of EGTC, with a significant increase in human resources and in-house competencies available to the Programme, can be seen as an explicit capacity-building policy. The ex-ante evaluation positively assessed the consistency between the main problems to be overcome and the innovations introduced as part of the ESPON strategy. Moreover, it took a step further in the 'how to' direction, by suggesting a series of programme features that would favour the activation of the most appropriate mechanisms, given the Programme context and its expected outcomes.

The context. For the period 2014–20 the general goal of the ESPON Cooperation Programme is to further support European and local policy makers through the production, dissemination and promotion of territorial or spatial policy evidence. This implies being able to capture the attention of a broad audience of potential stakeholders. These include, all European, national and regional authorities and policy makers, in particular in the field of cohesion policy as well as other relevant sectoral and thematic policies and programmes; regional and local policy makers and practitioners responsible for territorial development and planning; organizations promoting different regional/urban interests at EU level; University academics, both researchers and students as future decision makers.

However, previous evaluation of ESPON programmes highlighted a heavy management burden and procedural complexities, particularly due to the very large size of the ESPON steering committee, made up of 32 state representatives. In order to overcome these weaknesses, the new ESPON 2020 set up the EGTC. The EGTC is a new legal entity, contracted to implement the new programme. It simplifies the former management structure, acting as a single ‘beneficiary’ (instead of the 32 states) for the implementation of the 2014–20 ESPON programme. This has led to several expression of concern about excessive power vested in the EGTC by states that are members of the monitoring committee. ESPON 2020 outcomes could be hampered by a lack of cooperation between the monitoring committee and the new ETGC.

Table 2 presents the main elements of the three exemplar practices presented in the article, together with the different types of contribution made possible by using a mechanism-centred evaluation approach as applied in the cases described.

Conclusions

Knowledge of mechanisms accompanied by a compatible evaluation approach could be highly relevant from two different perspectives: ex-post and ex-ante. From an ex-post or learning point of view, the mechanism-centred approach, focusing on actors and policy processes, allows us to see beyond the formal design of a policy. It promises learning and a deeper understanding of the dynamics of change, adjustment and maintenance that occurs over time. This is often a highly relevant part of a successful policy story. From an ex-ante or design point of view, the mechanism-centred approach raises questions on how to foster behaviour by policy actors coherent with policy goals. As mechanisms can often work similarly across similar contexts, having a good knowledge of examples of mechanisms and of the conditions that triggered them in a specific case is a valuable piece of information for the design of a new policy or programme (Bardach, 2004; Barzelay, 2007). Capacity-building policies could especially benefit from this perspective.

These two aspects are relevant in the European policy setting, where goals are common, but institutions and policy networks differ significantly from one place to another; policy implementation is often characterized by multiple-actors; and context features differ from place to place, which means that policies have to be contextualized in order to produce change. A ‘one size fits all’ policy is hard to implement in different contexts with the same results. It is also worth noting that those in charge of implementing EU policies need high levels of capacity to manage an often complex policy arena. Increasing the capacity of managing authorities and exchange of good practices are two of the most commonly shared intermediate goals of EU policy making.

Table 2. The working method: three examples.

Exemplar case	Policy outcomes	Mechanisms	Policy features	Context features	Contribution of the approach in informing policy
AlpEnergy project	Reduction and postponement of Saint Denis citizens' energy consumption in non-peak periods	Instant feedbacks on energy consumption levels Competition and 'bandwagon effect' among citizens	Public monitoring of energy savings Active engagement of the Municipality in communicating and involving citizens in a innovative project	Very small mountainous area High sensibility towards renewable energy issues Local political agenda focusing on energy projects	Learning from good practices to disseminate findings to different contexts
Gender mainstreaming in committees and delegations of the European Parliament	Effectiveness of FEMM's activity in fostering the mainstreaming of the gender equality principle in the EP legislation	Reputation and/or actor's certification of FEMM's rapporteurs	Planning of FEMM's role in the legislative procedure as an associated or a joint committee Selection of a rapporteur with a high level political prestige Set up of an EGTC.	22 committees participating in the legislative procedure of the EP FEMM committee in charge of promoting gender equality and gender mainstreaming	Providing policy advice to enhance the effectiveness of gender policy
Ex-ante evaluation of the ESPON 2014-2020 Programme	Fostering a greater awareness, usefulness and use of territorial information for European and regional policy makers Streamlining the management of the programme and reducing administrative burden for beneficiaries	Perception of opportunity of different target groups to generate and maintain engagement to the Programme Certification of the newly established EGTC Repeated interactions among stakeholders favouring trust, mutual learning and commitment	Various policy features suggested in order to activate mechanisms	Broad audience of potential stakeholders for the territorial evidence produced by ESPON Large composition of the steering committee of the programme – 32 States Heavy management burden and procedural complexities	Support policy makers in the design and structuring of a new Programme

Nonetheless, the use of social mechanism theory in evaluation research is not easy and faces a number of pitfalls. First, a theory of change is key in order to identify the most relevant outcomes that merit explanation. If a theory of change is not available, it has to be reconstructed by the evaluator. Second, detecting the mechanisms fostering changes in actors' behaviours is a fine-grained exercise that requires qualitative in-depth data, collected mainly through interviews. This may be costly and affect the timetable and deadlines of evaluation studies that are constrained by often tight contractual time-limits. Third, the identification of the mechanism useful for extrapolation is not always crystal clear; as Bardach put it, 'the main point is that there ought to be something worth translating' (2004: 219). In other words, not all experiences can be used to generate new knowledge about mechanisms. Finally, the causal links between mechanisms, policy and contextual features are not always clear and require appropriate narratives that depend on detailed local knowledge to join them up.

The approach we advocate in this article can be seen as consistent with what has been called 'configurational views of causation'. In this, different causal elements coexist and interact and need to be understood both separately and together. Unsurprisingly, therefore, the mechanism-orientated approach we advocate shares both advantages and limits of this broader perspective. Barbara Befani explained it thus:

The black box has become partially transparent and some characteristics of the causal process begin to emerge (we know more than just beginning and end, we do have some in-between elements), but the understanding is not yet fine-grained and combinations need to be interpreted. We have started to 'peek in' but still have not opened the black box. (2012: 15)

More generally, the possibility of a usable methodology to generate practical knowledge based on mechanisms – the 'extrapolative approach' – while promising and even fascinating, is still far from field-ready. Additional primary research is needed, in particular in the public policy domain, to understand the workings of the 'catalogue of mechanisms' suggested by Bardach (2004) – in different contexts and under different policy conditions.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Notes

1. DG Regio, Unit Competence Centre for Administrative Capacity-Building and the Solidarity Fund, Administrative capacity linked to management and implementation of ESI funds: http://www.seio.gov.rs/upload/documents/skrining/eksplanatorni/prezentacije/pg22/22_10.pdf.
2. In this respect, see the final report of the Research ESPON Smart-Ist: http://www.espon.eu/main/Menu_Projects/Menu_TargetedAnalyses/smart-ist.html.
3. The Istituto per la Ricerca Sociale (2011) carried out the final evaluation of the AlpEnergy project, financed within the European territorial cooperation Alpine Space programme on behalf of the Province of Mantova (IT), one of the project partner: <http://alpenergy.net>.
4. AlpEnergy involves the following regions: the Allgäu region in Bavaria (Germany), the City of Mantova in the Lombardia region (Italy), the Province Belluno (Italy), the Autonomous Region Aosta (Italy), the Belledonne chain area in the region Rhône-Alpes (France) and the region of Gorenjska (Slovenia).

5. The study was carried out in 2013–14 by an Istituto per la Ricerca Sociale working group coordinated by Flavia Pesce and Manuela Samek Lodovici with the scientific support of Bruno Dente and composed by Daniela Loi, Erica Melloni and Cristina Vasilescu. See [http://www.europarl.europa.eu/RegData/etudes/etudes/join/2014/493051/IPOL-FEMM_ET\(2014\)493051_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/etudes/join/2014/493051/IPOL-FEMM_ET(2014)493051_EN.pdf).
6. For more details on the decision-making procedure within the EP Committees, see rules 49–56 of the European Parliament: <http://www.europarl.europa.eu/sides/getLastRules.do?language=EN&reference=TOC>.
7. Such amendments enable the FEMM committee to make contributions to the reports of other committees without making formal opinions, particularly in cases when these are not entirely necessary or cannot be drafted due to time constraints. FEMM has already successfully applied this method in several cases.
8. LIBE stands for the Committee on Civil Liberties, Justice and Home Affairs. The proposal for a Directive of the European Parliament and of the Council establishing minimum standards on the rights, support and protection of victims of crime aims to take significant steps forward at the level of protection of victims throughout the Union, in particular within the framework of criminal proceedings. The FEMM – LIBE report 2011/0129(COD) on the proposal for a Directive of the European Parliament and of the Council establishing minimum standards on the rights, support and protection of victims of crime (COM(2011)0275) introduces amendments to it.
9. For further details on the functioning of joint committees, see rule 55 of the European Parliament decision-making procedure: <http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+RULES-EP+20150101+RULE-055+DOC+XML+V0//EN&language=EN&navigationBar=YES>.
10. For more details on the decision-making procedure within the EP Committees, see rules 49–56 of the European Parliament: <http://www.europarl.europa.eu/sides/getLastRules.do?language=EN&reference=TOC>.
11. ESPON stands for European Observation Network, Territorial Development and Cohesion. See <http://www.espon.eu/>.
12. The European Grouping for Territorial Cooperation allows public entities of different Member States to get together under a new entity with full legal status. EGTC was established on 5 July 2006 by Regulation (EC) 1082/2006 of the European Parliament and of the Council and came into force on 1 August 2006.

References

- Abrahamson E and Rosenkopf L (1993) Institutional and competitive bandwagons: Using mathematical modeling as a tool to explore innovation diffusion. *The Academy of Management Review* 18(3): 487–517.
- Bandura A (1977) Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review* 84(2): 191–215.
- Bardach E (2004) Presidential address – The extrapolation problem: How can we learn from the experience of others? *Journal of Policy Analysis and Management* 23(2): 205–20.
- Barzelay M (2007) Learning from second-hand experience: Methodology for extrapolation-oriented case research. *Governance* 20(3): 521–43.
- Barzelay M, Gaetani F, Velarde JCC, et al. (2003) Research on public management policy change in the Latin American region: A conceptual framework and methodological guide. *International Public Management Review* 4(1): 20–42.
- Befani B (2012) Models of Causality and Causal Inference, in Stern et al. (2012) Broadening the range of design and methods for impact evaluation, working paper 38, DFID.
- Bowles S and Gintis H (2011) *A Cooperative Species: Human Reciprocity and its Evolution*. Princeton, New Jersey Oxford: Princeton University Press.

- Busetti S, Dente B and Melloni E (2013) A theory of practice for collaborative governance: The use of social mechanisms in understanding policy success. Paper presented to the 7th ECPR General Conference, Sciences Po Bordeaux, 4–7 September 2013.
- Chen H and Rossi P H (1989) Issues in the theory-driven perspective. *Evaluation and Program Planning* 12: 299–306.
- Cyert R and March JG (1992) *A Behavioral Theory of the Firm*, 2nd edn. Oxford: Wiley-Blackwell.
- Dente B (2014) *Understanding Policy Decisions*. Milan: Springer International Publishing.
- Donaldson SI (2007) *Program Theory-driven Evaluation Science: Strategies and Applications*. Mahwah, NJ: Erlbaum.
- Eberlein B and Radaelli CM (2010) Mechanisms of conflict management in EU regulatory policy. *Public Administration* 88(3): 782–99.
- Edelmann M (1977) *Political Language: Words that Succeed and Policies that Fail*. New York: Academic Press.
- Elster J (1998) A plea for mechanisms. In: Hedström P and Swedberg R (eds) *Social Mechanisms: An Analytical Approach to Social Theory*. Cambridge: Cambridge University Press, 45–73.
- EVALSED (2013) The resource for the evaluation of socio-economic development – Evaluation guide and sourcebooks. Available at: http://ec.europa.eu/regional_policy/en/information/publications/evaluations-guidance-documents/2013/evalsed-the-resource-for-the-evaluation-of-socio-economic-development-evaluation-guide
- Falletti TG and Lynch JF (2009) Context and causal mechanisms in political analysis. *Comparative Political Studies* 42(9): 1143–66.
- Granovetter M (1978) Threshold model of collective behaviour. *American Journal of Sociology* 83(6): 1420–44.
- Hedström P and Ylikoski P (2010) Causal mechanisms in the social sciences. *Annual Review of Sociology* 36: 49–67.
- Hood C (2011) *The Blame Game: Spin, Bureaucracy, and Self-preservation in Government*. Princeton, NJ: Princeton University Press.
- Istituto per la Ricerca Sociale (2011) Final evaluation of the AlpEnergy project, final report.
- Istituto per la Ricerca Sociale (2014) Gender mainstreaming in Committees and Delegations of the European Parliament. Available at: [http://www.europarl.europa.eu/RegData/etudes/etudes/join/2014/493051/IPOL-FEMM_ET\(2014\)493051_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/etudes/join/2014/493051/IPOL-FEMM_ET(2014)493051_EN.pdf)
- McAdam D, Sidney T and Tilly T (2001) *Dynamics of Contention*. Cambridge: Cambridge University Press.
- Mahoney J (2003) Tentative answers to questions about causal mechanisms. Paper presented at the annual meetings of the American Political Science Association, Philadelphia, PA, 28 August.
- Mayne J (2012) Contribution analysis: Coming of age? *Evaluation* 18(3): 270–80.
- Mayntz R (2004) Mechanisms in the analysis of social macro-phenomena. *Philosophy of the Social Sciences* 34(2): 237–59.
- Merton R (1957) *Social Theory and Social Structure*. New York: Free Press.
- Pawson R (2006) *Evidence Based Policy: A Realist Perspective*. London: SAGE.
- Pawson R and Tilley N (1997) *Realistic Evaluation*. Thousand Oaks, CA and London: SAGE.
- Radaelli CM, Dente B and Dossi S (2012) Recasting institutionalism: Institutional analysis and public policy. *European Political Science* 11: 537–50.
- Scharpf FW (1997) *Games Real Actors Play*. Boulder, CO: Westview Press.
- Scriven M (1994) The fine line between evaluation and explanation. *Evaluation Practice* 15(1): 75–7.
- Stern E, Stame N, Mayne J, et al. (2012) *DFID Working Paper 38. Broadening the range of designs and methods for impact evaluations*. London: DFID.
- Suchman E (1967) *Evaluative Research: Principles and Practice in Public Service and Social Programs*. New York: Russell Sage Foundation.
- Tilly C (2001) Mechanisms in the political processes. *American Political Science Review* 4: 21–41.

Weiss CH (1972) *Evaluation*. Upper Saddle River, NJ: Prentice Hall.

Weiss CH (1995) Nothing as practical as good theory. In: Connell J, Kubisch A, Schorr LB, et al. (eds) *New Approaches to Evaluating Community Initiatives*. New York: Aspen Institute.

Weiss CH (1997) How can theory-based evaluation make greater headway? *Evaluation Review* 21(4): 501–24.

Wholey JS (1979) *Evaluation: Promise and Performance*. Washington, DC: Urban Institute.

Erica Melloni is the Director of the Administrative and Institutional Policies Unit of the Istituto per la Ricerca Sociale. Her research interests include institutional capacity building and territorial development.

Flavia Pesce is the Director of the Education, Training and Labour Policies Unit of the Istituto per la Ricerca Sociale. She is a senior expert in evaluating European policies and programmes.

Cristina Vasilescu is a Senior Researcher within IRS. She is an expert in the evaluation of European policies.