



Universiteit
Leiden
The Netherlands

What role can social constructivism play in realist evaluation? The potentials and the limits

Tosun, J.; De Francesco, F.; Pattyn, V.E.; Varone, F.; Jacob, S.; Bundi, P.

Citation

Tosun, J., De Francesco, F., & Pattyn, V. E. (2023). What role can social constructivism play in realist evaluation?: The potentials and the limits. In F. Varone, S. Jacob, & P. Bundi (Eds.), *Handbooks of Research on Public Policy* (pp. 31-46). Cheltenham: Edward Elgar. doi:10.4337/9781800884892

Version: Publisher's Version

License: [Licensed under Article 25fa Copyright Act/Law \(Amendment Taverne\)](#)

Downloaded from: <https://hdl.handle.net/1887/4212640>

Note: To cite this publication please use the final published version (if applicable).

2. What role can social constructivism play in realist evaluation? The potentials and the limits

Jale Tosun, Fabrizio De Francesco and Valérie Pattyn

INTRODUCTION

Ever since Lasswell's (1956) pioneering work on the policy cycle, evaluation has been a vital field of policy research (Hogwood et al., 1984). However, it has become even more important with the 'evidential turn' in public policy (Taylor, 2005, p. 601). This turn can be explained by the re-emergence of governmental emphasis on evidence of 'what works' in public policies and services (Cairney, 2016; Head, 2016; Nutley et al., 2010), rather than relying mainly on policy performance and auditing systems typical of the new public management. Another reason can be seen in the increased complexity of public policies and the problems underlying them. Last, the appraisal of policy impacts is required to obtain or maintain public support for policy measures that entail significant public investment (Pattyn et al., 2018).

From a conceptual viewpoint, policy evaluation is typically, albeit not exclusively, about comparing the intended and actual effects of policy measures, with the goal of generating information or 'evidence' (Pawson, 2002).¹ The evidence can be used for many purposes, such as improving existing policy measures, which indicates a connection between evaluation and policy-oriented learning (Van der Knaap, 1995). And evaluation can produce *ex ante*, *ex durante*, or *ex post* evidence (Scriven, 1991). However, the use of evaluation is not limited to improving a given policy but can also represent a response to political pressure (see, e.g., De Francesco, 2019; Mavrot & Pattyn, 2022; Varone et al., 2020). For example, the policy measures adopted in response to the COVID-19 pandemic have been evaluated meticulously and regularly, not only for the purpose of facilitating policy-oriented learning but also to justify the limitation of freedom rights. Nevertheless, compared to the other stages of the policy cycle, evaluation stands out for its scientific and professional criteria. Insights produced by policy evaluation can produce policy feedbacks (see, e.g., Moynihan & Soss, 2014) and start a new policy cycle (Lasswell, 1956).

While evaluation is a standard tool of policymaking in modern states, and is getting increasingly institutionalized (Jacob et al., 2015; Organisation for Economic Co-operation and Development [OECD], 2020; Stockmann et al., 2020) across countries with different 'evaluation styles' (see De Francesco & Pattyn, 2021), it is seldom a straightforward task. Some of the factors that can complicate evaluation activities include the identification of the goals that a given policy measure pursues, the appropriate definition and measurement of

¹ Alternative evaluation criteria exist, too. For instance, the guidelines by the Development Assistance Committee (DAC) of the Organisation for Economic Co-operation and Development (OECD, 2021) propose to assess the relevance, coherence, impact and sustainability of policies and programmes. However, the effectiveness and the efficiency of policies have been central to the discussion of policy evaluation for several decades.

cause-and-effect indicators, the isolation of a policy's direct impact from other factors, and the political context (Knill & Tosun, 2020). The first three challenges can be resolved with a carefully developed evaluation design and a well-established programme theory (see, e.g., Pawson, 2002). In contrast, the fourth issue, the political context in which evaluation takes place cannot be fully controlled, and requires enhanced attention, especially when reflecting on the insights produced by evaluation activities.

Interestingly, the challenge concerning the political context in which evaluations are carried out represents a unifying element of two otherwise very different approaches to policy evaluation, which were even regarded to engage in a 'paradigm war' within evaluation theory (Taylor, 2005, p. 602). On the one side, realists aim to identify 'what works, for whom in what circumstances', as put concisely by Pawson et al. (2005, p. 21). On the other side, social constructivists (Taylor, 2005, p. 602) emphasize that 'giving meaning' to the individual elements of an evaluation activity is important and that evaluation results are not unequivocal because they do not have the same meaning for all actors involved (Van der Meer, 1999). Despite these differences, both realists and social constructionists agree that 'all public policy evaluations ... are projects in both *political science* and *political science*' (King et al., 2007, p. 480, original emphasis).

This agreement on the political nature of evaluation provokes the question of whether there might be more similarities between realist and social-constructivist evaluation than usually assumed. How much social constructivism is there in realist evaluation theory? How much in practice? And can both approaches be integrated without reducing their ontological, epistemological and methodological integrity? These research questions guide this chapter, which offers a systematic and comparative assessment of the two dominant perspectives on policy evaluation.

The remainder of this chapter unfolds as follows. First, we outline the main characteristics of realist and social-constructivist evaluation. Then, we discuss how much social construction exists in the theory of realist evaluation and how much in a realist evaluation process when it is carried out in practice. Subsequently, we present a proposal for integrating the two perspectives and offer some concluding remarks.

REALIST EVALUATION

In this section, we present the key concepts of realist evaluation as put forth by Pawson and Tilley (1997, 2001). To avoid terminological confusion, first we need to explain that the subject of policy evaluation is public policies, which is a very broad concept. Effectively, this involves what policy scholars would denote as policy instruments (see Hall, 1993), including individual policy instruments as well as sets of them. As Howlett and Rayner (2007) contend, such sets of instruments, also known as 'policy mixes', do not have to be adopted and/or put into practice simultaneously but can evolve gradually over time.

The terminology of realist evaluation deviates from that used in public policy since it refers to 'programmes' rather than policy instruments. For the sake of consistency with the realist evaluation's terminology, in what follows we also use the term 'programme', but it should be noted that we consider a programme to correspond to (a set of) policy measures that entail action by the implementing bodies and/or behaviour change by the target population.

The Role of Theory in Realist Evaluation

Realist evaluation is regarded as ‘theory-driven’ evaluation, as the theoretical dimension is important for understanding this perspective. According to Pawson and Tilley (2004, p. 3), programmes are ‘theories incarnate’: every programme has a theoretical underpinning, which is sometimes made explicit and sometimes not (Pawson, 2002; Pawson & Tilley, 1997, 2001). At the same time, they do not refer to social science theory but to ‘programme theory’. Programme theory is more modest than social science theories (Knill & Tosun, 2012): it comprises assumptions that guide the way a certain programme is implemented and expected to bring about change in social conditions (Donaldson & Lipsey, 2006, p. 64). These assumptions should be developed before a programme is implemented, while it is running, or after it has ended (Astbury & Leeuw, 2010, pp. 364–365).

Programme theory is composed of three interrelated components: impact theory, service utilization plans, and organizational plans. Organizational and service utilization plans together constitute the ‘process theory’ (Knill & Tosun, 2012). To explain these components in detail, we follow the description provided by Rossi et al. (2018).

Impact theory refers to the assumptions made about the process of behaviour change that is triggered by a programme in question. In other words, it is conceived as a ‘change model’. In contrast, service utilization plans and organizational plans correspond to ‘action models’ (Chen, 2012). A service utilization plan is based on assumptions and expectations of how to reach the addressees. In addition, it involves considerations of the provision and sequence of service contacts and the ways of concluding the relationship when services are no longer supplied. Like impact theory, the service utilization plan also describes transactions between the programme and its addressees, though it adopts a different perspective by focusing on the addressees and their social environment.

The organizational plan refers to the necessary resources, personnel, infrastructure, administration and general organization. This plan articulates the functions that a programme is expected to fulfil as well as the resources required for fulfilment. In addition, the organizational plan must include functions that represent preconditions for the service supply. Organizational plans usually include a number of assumptions and propositions.

Impact theory is informed by appropriate theories from social science, while process theory (i.e., service utilization and organizational plans) usually comes almost entirely from the programme concerned (Astbury & Leeuw, 2010; Donaldson & Lipsey, 2006, p. 64). Social science theories are not particularly concerned with policy evaluation but provide a theoretical argument on the basis of which pertinent causal statements (hypotheses) can be formulated. Since realist evaluation is ‘theory driven’, the first activity of an evaluator is to bring the theories on which programmes rest to the surface in order to assess the intended impact.

Pawson and Tilley (1997, 2004) are aware that programmes do not operate in a vacuum but are ‘embedded’ and need to be assessed in the context of the existing social system into which they were introduced. Policy measures are effective if the *individual* capacities of those responsible for policy delivery and their *interpersonal* relationships are considered. Further, other factors determine the quality of programmes such as the *institutional* balance within the given context, and the wider *infra-structural* systems (e.g., political support) that affect the target groups of a policy.

Realist evaluation acknowledges that policy outcomes or impacts stem from the active engagement of individuals, which means that ‘an understanding of the interpretations of

programme participants is integral to evaluating its outcomes' (Pawson & Tilley, 2004, p. 5). Another important feature of realist evaluation is that it accepts that there are externalities (such as unanticipated events, political change, personnel moves, or physical and technological shifts) that have an impact on the delivery of programmes and change their nature. In other words, the conditions for successful interventions can change according to this view.

Key Concepts

As we have seen in the previous section, realist evaluation cannot be carried out without having a theoretical underpinning that has the potential to explain and understand why policy measures work or do not. To achieve this, Pawson and Tilley (1997, 2004) introduce four key concepts.

The first concept, mechanisms, refers to those elements of a programme that potentially or actually bring about changes in human behaviour. In this regard, the interpretation of a programme by the addressees is an important element for explaining whether and how they act upon it. Context is the second concept and refers to those 'conditions in which programmes are introduced that are relevant to the operation of the programme mechanisms' (Pawson & Tilley, 2004, p. 7). Since contexts will vary when a programme is implemented,² so will the working of the mechanisms activated by it, which is likely to result in context-specific consequences and local variations of them. This clashes with the so-called top-down perspective to policy implementation, which assumes that a policy is well implemented if there is little or no variation in the outcomes (Knill & Tosun, 2012, 2020; see also Tosun & Treib, 2018).

The consequences that programmes have for individual or collective actors is the third key concept and is denoted as outcome pattern. A systematic assessment of the outcome patterns can provide many intriguing insights into the functioning of a programme, including implementation variation or deficits, and into the variation that results from the different ways in which different societal groups make sense of a programme.

The overarching goal of realist evaluation is to explain outcome pattern variation by elucidating mechanism variation and context variation, which is termed as the context–mechanism–outcome (CMO) pattern configuration and represents the fourth key concept. These configurations are spelled out and tested empirically in realist evaluation to refine the underlying theory. The configurational thinking underpinning the approach makes realist policy evaluation a sophisticated and challenging task (see Fontaine, Chapter 3 in this *Handbook*; see also Befani et al., 2007; Sager & Andereggen, 2012), in part because it emphasizes the desirability of repeated empirical measurement and testing.

Methodology

Pawson and Tilley (1997, 2004) invite the implementation of mixed methods or multi-method designs (see, e.g., Burch & Heinrich, 2015; Pattyn et al., 2022). As many different types of methods exist, those compatible with realist evaluation are those that test scientific hypotheses explaining why the causal mechanisms – that is, the practical hypotheses – work or do not work. Once the hypotheses are tested using empirical data, realist evaluators would go back to

² Context also matters within a federal system, as subnational units can implement differently the same national legislation.

the theory and assess in which ways it needs to be modified to increase its explanatory power. In other words, this type of policy evaluation emphasizes the research cycle and not only wants it to be implemented in full but also as many times as possible, for the sake of continuously refining the theory. The idea is to use varying data sources for the various testing cycles. Realist evaluation is pragmatic concerning the data to be used for evaluation purposes – it sees value in all sorts of sources, ranging from survey data to interviews and official documents.

In this context, realist evaluators regard stakeholders and affected parties as key sources for eliciting programme theory and providing data on how a given policy measure would work. Stakeholders are involved in such a manner that the evaluator informs them about the programme theory on which a given policy measure rests. As they would ‘learn’ about the theory, stakeholders and affected parties could then ‘teach’ the evaluator about contexts and mechanisms related to a policy measure in a particularly informed way (Pawson & Tilley, 2004, p. 12).

All in all, realist evaluation draws from the logic of natural sciences, as Pawson and Tilley (2004) put it, and is rooted in rationalism and methodological individualism. The latter is suggested by its focus on how programmes can or cannot bring about changes in individual behaviour – a perspective that resonates well with the fact that programmes first and foremost aim at inducing behaviour change (Tosun & Treib, 2018). In terms of explanatory patterns, realist evaluation emphasizes the relevance of the conditions for explaining a phenomenon (for an overview, see, e.g., Cartwright, 1980). Indeed, realist evaluation is keen to provide middle-range theorizing, which means that theory starts with concrete empirical phenomena and then abstracts from them to formulate general statements that can be tested with data (see Merton, 1968). Realist evaluators place strong emphasis on both the theoretical underpinning of a policy measure and the empirical approach to testing the theory. Relying on the generative notion of causality, it especially contributes to middle-range theorizing.

SOCIAL-CONSTRUCTIVIST EVALUATION

Main Features of Social Constructivism

Before outlining the social-constructivist approach to evaluation, it appears reasonable to clarify the main features of social constructivism since it exists in various forms (see, e.g., Barkin, 2003). Moreover, it is important for the subsequent analysis to emphasize the existence of this variation.

Social constructivism is known for contending that social and political realities are, at least in part, constituted by actors through the subjective and inter-subjective understandings they develop to make sense of their experiences (Hay, 2016). In fact, the notion of sensemaking is critical for understanding this paradigm, which has been influential in international relations especially. Policy scholars have known this perspective mostly through the lens of sociological institutionalism (Saurugger, 2013), which emphasizes the importance of norms for the behaviour of human agents (see Hall & Taylor, 1996). As an intermediate perspective between international relations and policy analysis, foreign policy analysis has drawn extensively on social constructivism (see, e.g., Hudson & Day, 2019).

In public policy, the so-called ‘argumentative turn’ started with the work of Fischer and Forester (1993) and produced a body of research that focuses on deliberation, discourse,

rhetoric and semiotics, post-structural policy analysis, and participatory and collaborative policy analysis (Fischer & Gottweis, 2013). Thus, in public policy too, social constructivism has become an influential analytical perspective with numerous empirical applications (Saurugger, 2013).

From an ontological perspective, social constructivism claims that agents do not exist independently of given social structures (Berger & Luckmann, 1966) and the ‘culture’ these structures generate, which corresponds to a ‘shared system of meanings’ (Risse, 2004, p. 160). Put differently, constructivists contend that categories of things exist but that their properties depend on our collective knowledge and understanding of them. Hay (2016) gives a useful example to illustrate this point: a ten euro note in someone’s pocket is only a ten euro note insofar as it is regarded as such. The value of the money is not intrinsic as a piece of paper but is derived from a status bestowed on it from the shared system of meanings or culture. The latter means that actors orient themselves toward their social context through a veil of (inter-) subjective ideas, which includes interpretations, understandings, cognitive templates and normative dispositions.

According to Lombardo and Kantola (2021), from an epistemological perspective, we can differentiate between a ‘rationalist’ (i.e., positivist) version and a ‘reflectivist’ version. Barkin (2003, pp. 326–327) also acknowledges two different extreme ends of social constructivism that differ from each other in epistemological terms, for ‘one argues that an identifiable reality exists out there and can be accessed through empirical research; the other contends that we can never know for sure if what we observe really exists independently of our observation of it and, therefore, no true reality exists for empirical research to find’. He refers to them as ‘neoclassical’ or ‘thin’ constructivism on the one hand and ‘postmodernist’ or ‘thick’ constructivism on the other.

The first perspective attributes interests to actors, though these are idealized perceptions and projections (Hay, 2016). And in line with the basic tenet of constructivism, actors are embedded in social context and their behaviour is shaped by this context. For example, a positive constructivist approach would recognize power structures in politics and assess how these affect normative change in the policy process, and conversely, the way in which norms affect power structures (Barkin, 2003, p. 337).

The second perspective has emphasized the importance of language and discourse (Fischer & Forester, 1993; Fischer & Gottweis, 2013) and practice (Freeman et al., 2011; Freeman & Maybin, 2011; Law & Singleton, 2014). It has been particularly attentive to power and power relationships, as they underlie discursive practices (Risse, 2004). To understand this point, it is useful to emphasize that constructivism pays attention to institutionalization and habitualization (Berger & Luckmann, 1966; Searle & Willis, 1995). From this perspective, language can be regarded as an institution, and institutions have a dual function: they enable but also constrain social practices. However, institutions exist only in and through the practices to which they might be seen to give rise (Hay, 2016).

Despite the existence of different varieties of social constructivism, these share some commonalities. For example, they rest on the assumption that politics is a result of social interactions and is therefore socially constructed, and is therefore subject to a continuous process of change. Another commonality is that constructivists consider agents and structures to mutually shape each other. Last, constructivists of different orientations share a joint interest in assessing the role of norms, ideas, meanings and discourse.

Haas (2004, p. 575) applies the constructivist perspective to the policy process and contends that under conditions of uncertainty, it is impossible to create *ex ante* sufficient information to follow the policy cycle and its individual stages. Instead, the policy process should be conceived as a process in which actors learn about the world and about each other through social interaction. This dynamic process of learning results from the constructivists' ontology of the social world, which regards it as constantly being made and remade in and through practices (Hay, 2016). From this, it follows that constructivists are not interested in describing in detail how programmes change individual behaviour norms but are interested in processes related to the institutionalization of norms and social learning.

Haas's reflection is important, because from the social-constructivist perspective, policy evaluation would not concentrate on its findings, but on the characteristics of these processes. Along these lines, it is important to state that for some social-constructivist evaluators, such as Guba and Lincoln (1989), the product of an evaluation is not a set of conclusions, recommendations or value judgements, but an agenda for the negotiation of claims, concerns and issues.

Main Features of Social-constructivist Policy Evaluation

Starting with the practice of evaluation, the main argument advanced by social constructivists is that evaluation cannot be seen as a purely scientific undertaking but one that is 'socially located' and 'politically contested' (Taylor, 2005, p. 603). Social constructivists reject the scientific objectivity often assumed in policy evaluation (e.g., Guba & Lincoln, 1989). As they argue that policy phenomena can only be understood in context, it appears difficult if not impossible to draw generalizable conclusions from the evaluation of policy measures. Consequently, social constructivists do not see the role of evaluators as producing objective data on how programmes work, but as assessing stakeholders 'critically' with the goal of empowering certain, supposedly disadvantaged, groups in society. In this regard, some scholars, especially in the spirit of 'responsive evaluation', have emphasized the role of evaluators as 'facilitators' who support the organization of the dialogue among stakeholders and help to create the required conditions for dialogue (Abma, 2006).

Social-constructivist evaluation emphasizes the role of the evaluators and their interactions with the programme addressees. Cousins and Whitmore (1998), for example, explicitly acknowledge the critical role of participants' knowledge in evaluation and distinguish between two approaches to it. 'Practical participatory evaluation' strives to increase the impact of the participants' knowledge on the evaluation process. 'Transformative participatory evaluation' aims at enhancing the participants' understanding of how programmes are a manifestation of political power (see Fischer, Chapter 8 in this *Handbook*). This point aligns with the literature on the social construction of target groups, which regards policy measures as 'opportunities' and 'messages' to programme addressees on how the government conceives of them. This perspective allows us to understand how target groups are likely to be treated by the government (Ingram et al., 2019, p. 98). Policymakers use both negative and positive connotations of the social construction of 'target groups' to distribute benefits (Ingram & Schneider, 1990). Research has also shown that the social construction of target groups can shape the findings of policy evaluation. Thus, evaluation should 'reflect on the implications of symbolic messages and distribution of policies toward target populations' (Mendez, 2021, p. 1716).

Another central theme in social-constructivist evaluation is how the evaluation findings are interpreted. In this regard, Van der Meer (1999) contends that to understand when and how

evaluation findings can have an impact, it is important to understand that actors have certain ways of interpreting reality, which he denotes as ‘repertoires’ (also known as ‘frames’), and that understanding these repertoires is key to whether and how the evaluation findings are used. Repertoires, depending on the scenarios that actors construct according to them, can stimulate governmental learning processes.

Interestingly, some constructivist scholars, such as Haas (2004), emphasize the role of institutions for governmental learning. His argument is not limited to the legitimacy of the knowledge produced by policy evaluation but revolves around institutional constraints on knowledge utilization. More precisely, he alludes to the fact that individual governmental agencies may learn from policy evaluation but that others may fail to do so because of the compartmentalized structure preventing communication and exchange. The argument put forth by Haas already suggests that there is room to integrate realist and social-constructivist approaches into policy evaluation, at least to a certain degree.

Methodology

Social constructivist approaches are an antonym of ‘preordinate’ evaluation, where the parameters and boundaries of the evaluation have been defined before its beginning. Further, this approach is characterized by ‘an interactive, negotiated process that involves all stakeholders’ (Guba & Lincoln, 1989, pp. 38–39) as the result of the emergence of more democratic, responsive and participatory-based types of evaluation. Unlike the realist approach, which focuses on how programme theories can be enhanced, the perspectives of users and practitioners on public service and policy quality enhance their legitimacy.

‘Pure’ social-constructivist methodology argues that the object of an evaluation is conceptually constructed as well as emphasizes the need to pay attention to social practices in order to be able to discern normative ideas underlying certain concepts. Social practices refer to actors making normative ideas work by acting in accordance with the concepts’ concerns and their respective beliefs. From this perspective, it is not policies per se that are thought to be evaluated but actors think and act according to the policies concerned. Or in other words, social-constructivist evaluation entails the need to acknowledge that the object of an evaluation is not objective but socially constructed (Liebert, 2010).

SOCIAL CONSTRUCTION IN REALIST EVALUATION

The relationship between realist and social-constructivist evaluation has received some attention in the literature. Interestingly, their relationship has been characterized in different ways. While, for example, Barkin (2003) states that they are seen as two opposing paradigms, Greenhalgh et al. (2009), for example, contend that realist evaluation builds on a constructivist ontology and an interpretivist epistemology and methodology.

How can such strikingly different views on these two approaches and their relationship exist? One reason is that realist philosophy positions itself between positivism and constructivism (Graham & McAleer, 2018). Another, as we argue, is due to the terminology used by Pawson and Tilley (1997, 2004). At first glance, evaluation scholars might be inclined to identify social-constructivist elements in realist evaluation. However, when developing a more thorough understanding of what the terminology refers to, it becomes apparent that from

the theoretical perspective, realist evaluation differs from social-constructivist evaluation. However, this is not because social construction and realism are incompatible (see, e.g., Davis, 2005); rather, it refers to how Pawson and Tilley define their concepts.

An Assessment of the Conceptualization of Pawson and Tilley

Realist evaluation emphasizes the importance of (objective) measurement, but it certainly does not demonstrate a naive understanding of how policies are made and what outcomes they produce. When realist evaluators claim that policy measures are ‘theories incarnate’, they refer to ideas of what programmes can and should bring about. These ‘begin in the heads of policy architects, pass into the hands of practitioners and, sometimes, into the hearts and minds of programme subjects’ (Pawson & Tilley, 2004, p. 3). Effectively, while realist evaluators do not state this, this understanding of theory refers to the concept of policy design (Fontaine, 2020), which is open to both a realist and a constructivist paradigm. Most prominently, Schneider and Ingram (1993) and Ingram et al. (2019) represent the constructivist ontology, which emphasizes the role of power relationships for policy design.

While, in principle, it is possible to associate the realist notion of a theory with a constructivist approach too, the explanations given by Pawson and Tilley suggest that a theory concentrates on the question of how an undesirable condition or situation in society can be overcome. In other words, the authors’ perspective brings considerations in relation to problem-solving to the fore and not those that relate to who has the power to attribute what image to certain groups in society.

Another reason why realist evaluators could be suspected to be social constructivists refers to the realism’s emphasis on programmes being embedded in social systems. At first glance, this suggests that Pawson and Tilley may refer to the ‘shared system of meanings’ that characterizes social constructivism. However, the authors have a different understanding of social embeddedness. Rather than referring to shared rules and norms, they refer to different layers of social reality that constitutes a programme or provides the environment in which it is implemented. This point is important because it refers to all personal and institutional factors that facilitate or hamper a given programme from delivering its intended results.

For example, the social embeddedness of policy measures also includes the welfare systems of the countries in which they are implemented. Similar to the previous discussion, Pawson and Tilley emphasize the resources, infrastructure, and any other factors of relevance necessary to resolve a given social problem. We can think of their reasoning as something similar to the Easton’s (1979) political system theory, which acknowledges the ‘inputs’ that must enter the political process in order to produce ‘outputs’.

Pawson and Tilley emphasize the active engagement of programme addressees. The use of the term ‘interpretation’ could be regarded as a signal that the authors embrace social constructivism here. However, their understanding of ‘interpretation’ is that in most cases programmes entail the need for targeted individuals to change their behaviour, which is not achieved easily as we know from research in psychology. Environmental, genetic and psychological variables affect human behaviour. When developing a programme aimed at behaviour change, the focus is mostly on psychological variables (e.g., health beliefs), since these are the factors that a programme is most likely to be able to change (Crutzen et al., 2017). This is a point that has been at the front and centre of behavioural approaches to public policy (Sunstein, 2020) and public administration (Grimmelikhuijsen et al., 2017).

Although psychological variables can be controlled, they cannot necessarily be controlled in a similar fashion for all individuals. Furthermore, the social environment in which the individuals are embedded is likely to vary strongly. Consequently, when evaluating a programme, both individual-level and context-level factors need to be considered – and this is exactly what the active engagement of individuals means in a realistic context.

Programmes provide individuals with incentives to change their behaviour with a view to solving a howsoever defined social problem. Policy evaluation should not only pay attention to whether a programme is designed to deliver the intended results, but also to the fact that such programmes need to be processed cognitively by the individuals addressed, and that these are embedded in social contexts, which, in turn, affect the processing. Thus, a programme as such could be perfectly designed in terms of its inbuilt mechanism for solving a problem, but it could still fall short in delivering on its promise because of how it is processed – or ‘interpreted’, as Pawson and Tilley put it – by individuals.

The ‘openness’ of programmes and the fact that they can be regarded as self-transformational constitute another element of the conceptualization put forth by Pawson and Tilley that requires attention. This characterization of programmes is interesting because, to some extent, it challenges the notion of (rational) policy design and alludes to the contingency of the policy process. This is indeed a reasoning that aligns with social constructivism, and perhaps it is the dimension that does so most closely. However, this still does not mean that realism embraces the social-constructivist logic here. What matters regarding this argument is that rational considerations shaped the programme when it was first launched. This argument refers to what happens when a programme has been in place for a while. Should the programme deliver the same outcomes over a given time period, or should we expect its outcomes to vary in light of the presence of different types of externalities and the occurrence of learning? With this element, too, the realist framework uses a terminology that hints at social construction, but actually has a different meaning.

Realist evaluation can only be carried out if it is based on explicitly formulated mechanisms that describe how a programme can bring about certain outcomes. In public policy, such mechanisms are, for example, built in the ‘NATO’ scheme, which classifies governance principles. It differentiates between *nodality* (or information), *authority*, *treasure*, and *organizational resources* (Hood & Margetts, 2007). The idea underlying the NATO scheme is that these four types of policy instruments bring about behaviour change in individuals as they respond to the respective incentives. In principle, social constructivism can be one way of defining mechanisms. For example, individuals could change their behaviour because they consider the programme to be legitimate and therefore want to do the right thing, as postulated by the ‘logic of appropriateness’ (March & Olsen, 1989). However, Pawson’s and Tilley’s conceptualization does not mention social construction as a source for identifying mechanisms, but equally it does not explicitly dismiss it. Therefore, it is conceivable that the mechanism through which a programme produces outcomes could be based on social construction.

While context is another key concept in realist evaluation (see Greenhalgh & Manzano, 2022), as we have already discussed above, the authors’ understanding of it differs from what it would mean to social constructivists. Realist evaluators posit that there exist two sources of variation in the outcomes produced by policy measures: those referring to the mechanisms and those referring to contexts. Identifying which of the two sources accounts for how much of the overall variation in outcomes is one of the key objectives of realist evaluation. Another concerns the question of how the sources interact with each other, which corresponds to the

configurational logic. This logic of attempting to explain different types of variation requires a perspective that recognizes the possibility of (objective) measurement. With the interpretative approach to social construction, such a focus on variation and the idea of differentiating between sources of variation and identifying their interaction are very unlikely to resonate. And even with the realist approach to social construction, the notion of analysing variation could be a bridge too far.

As we have shown above, some of the previous components of realist evaluation are termed in a manner that suggests commonalities with social constructivism and its approach to policy evaluation. When turning to the methodological foundation of realist evaluation, the room for seeing possible commonalities diminishes – not necessarily for the sources consulted but for the need to formulate a programme theory and to make explicit how a programme is expected to bring about behaviour change, which is then assessed against the background of the context in which it is being implemented. The data collected for realist evaluation focus on the hypotheses that are derived from the theoretical underpinning of a programme with the goal of testing and refining them in light of the empirical findings (see Manzano, 2016). Interview data are a valid source for realist evaluation, but so are other ‘objective’ sources such as administrative data.

Turning to the generation of data through interviews, these are used in a very specific way when carrying out realist evaluation. Since the goal of the evaluation is to refine the theoretical underpinning of a policy measure, the theories are placed before the interviewee for them to comment on (Nanninga & Glebbeek, 2011). This teacher–learner relationship is the kind of ‘sensemaking’ that takes place in the frame of realist evaluation as an integral part of it, but which follows very explicit rules (for an overview, see Manzano, 2016). And evidently, this approach to sensemaking is significantly different from how pure social constructivists would understand it.

Overall, realist evaluation may appear at instances to be an approach that shares commonalities with social-constructivist policy evaluation, but these disappear when delving more deeply into the meaning of the concept concerned. After all, realist evaluation can be seen as an approach to public policy and policy evaluation that is rooted in scientific realism (which contends that aspects of the world can be described by the sciences), but which at the same time is aware that human behaviour is not only the outcome of incentives provided by policy measures but is also affected by context-specific factors of different kinds. Policy measures are seen as but contributory causes of change in human behaviour that work as part of a causal package in combination with other factors (Stame, 2004).

What this approach calls for is to capture these factors both theoretically and empirically with a view to developing a more meaningful understanding of policy outcomes. As indicated by Fontaine (Chapter 3 in this *Handbook*), the essence of the realist policy evaluation is not easy to grasp because it is such an eclectic approach that comprises an ontology, an epistemology, and a methodology.

An Assessment of the Application of Realist Evaluation

Research has shown that the philosophical foundations that underlie realist evaluation are variably interpreted and applied to different degrees (Marchal et al., 2012). This not only relates to definitional vagueness, but is also the result of some components of realist evaluation being under-researched. Lemire et al. (2020), for example, highlight that more research on the

‘mechanism’ is warranted, while Davis (2005) and Nielsen et al. (2022) contend that this is the case for the notion of ‘context’. Likewise, context lends itself easily to a social-constructivist view, which can be seen as one of the reasons empirical work using realist evaluation has also used constructivist approaches to conceptualize and operationalize it (see Greenhalgh & Manzano, 2022).

Another example of a realist evaluation that adopts a constructivist logic is the analysis of interventions in health care by Greenhalgh et al. (2009). The authors evaluate efforts taken to make health care organizations in inner London more efficient, effective and patient-centred. In so doing, the authors carry out an interpretative case study, which uses different sources of data for the actions carried out and their intended and unintended consequences. The authors state that the financial sponsor of the evaluation asked them ‘not to become tied to predefined milestones or fixed metrics of success’ (Greenhalgh et al., 2009, pp. 394–395). Consequently, the evaluation did not focus on whether the modernization effort worked but to understanding how the modernization initiative was affected by the interaction between the context of the programme and the mechanisms. Based on this, Greenhalgh et al. (2009, p. 414) claim that evaluators must accept the ‘constructionist ontology and interpretivist epistemology’ of realist evaluation. This statement is even more remarkable when considering that one of the inventors of realist evaluation, Ray Pawson, is one of the authors of this study.

The methodological foundation of realist evaluation is another field where the ‘pure’ recommendations, as formulated by Pawson and Tilley, are not always met. Manzano (2016), for example, makes the point that the interviews on which realist policy evaluation are based do not always get the attention they need in the research and are often treated as rather unproblematic, even though there are very specific rules that the evaluators need to follow. In particular, the teacher–learner cycle has been interpreted in a manner as to incorporate social-constructivist methods (see Nanninga & Glebbeek, 2011).

Despite having discussed a limited number of issues and studies, we can state that realist evaluations have incorporated social-constructivist elements in their theoretical models, the operationalization of the individual model components, and the corresponding methodologies. Such a ‘synthetic’ approach aligns with the fundamental philosophical tenet of realism as a middle-range perspective on constructivism and positivism.

INTEGRATION OF BOTH PARADIGMS

As we have seen above, there have been attempts to integrate realist and social-constructivist evaluation, especially in practice. This is most easily feasible with regard to the social science theories that motivate the programme theory underlying realist evaluation. In this regard, and as argued by Fontaine (2020), questions relating to how a programme was designed provide an apt angle for bringing in social construction – for example, in the shape of the target group approach as put forth by Ingram et al. (2019) and Schneider and Ingram (1993). The conceptualization of the context into which a programme is introduced is another element of realist evaluation that could potentially benefit from a more systematic reflection on whether and how social constructivism could come into play here.

Another important question is what the impacts of evaluation findings are, which are usually not taken into consideration by realist evaluators. Adding this aspect to an evaluation exercise could inform realist evaluation interested in continuously improving the knowledge base for

theory testing and therefore the effectiveness of programmes, too. One factor to explain the potential for improving existing policies can be seen in whether the evaluation findings produced a certain impact, and here social constructivism and its attention to power relationships and politics could play an important role (Taylor, 2005; Taylor & Balloch, 2005). In this context, it should be noted that social constructivism has become more compatible with realist evaluation due to recent developments such as the emergence of realist constructivism and the acknowledgement of the important role of actors in political processes (see Saurugger, 2013).

Despite these developments, it should be emphasized that there remain fundamental ontological and epistemological disagreements between realist and social-constructivist evaluators and that the goal cannot be to overcome these. Rather, the two perspectives should be brought together whenever their combination can produce more informative insights, and when they can learn from each other.

CONCLUSION

The policy process and its outputs, outcomes and impacts represent the central research interest of public policy scholars. Policy evaluation represents an integral part of the policy process, and yet this activity can focus on many different aspects of a policy and its implementation. To be able to carry out a policy evaluation that satisfies the requirement of a scientific activity, evaluators need to select a framework, of which there exist several.

The realist evaluation framework is very comprehensive and eclectic in its conceptualization. Some of its components have been applied frequently in evaluation research and others less so. Especially with regard to the latter, overview articles have argued that in practice, scholars have defined certain components, such as context, in a way that incorporates social-constructivist thinking. What is missing from the literature is a more systematic assessment of how much social constructivism is in realist evaluation.

What we attempted to achieve in this chapter was to show how much realist evaluation conceptually and methodologically lends itself to accommodating social constructivism. Our main answer to this question is that it is conceptually more open to social constructivism than methodologically, although real-life applications of them have also been implemented from a more constructivist perspective.

Our reflections in this chapter focus on scientists carrying out evaluations. When shifting the focus to practitioners (e.g., in-house evaluators) carrying out evaluations, we would expect them to be more willing to incorporate social constructivism in their work, for they observe more directly the complex processes that determine ‘what works, for whom in what circumstances’ (Pawson et al., 2005, p. 21). Practitioners are also more likely to regard policies as phenomena that are shaped, at least to some extent, by how they are interpreted by whom and how they are implemented on the ground, which aligns with the basic reasoning of social constructivism.

While here we cannot offer a novel model that combines the two frameworks, we want to end on the note that we consider such effort valuable. We regard any attempts that are carried out in a reflective and systematic manner to contribute to cumulative research and to provide important stimuli for further studies.

REFERENCES

- Abma, T. A. (2006). The practice and politics of responsive evaluation. *American Journal of Evaluation*, 27(1), 31–43.
- Astbury, B., & Leeuw, F. L. (2010). Unpacking black boxes: Mechanisms and theory building in evaluation. *American Journal of Evaluation*, 31(3), 363–381.
- Barkin, J. S. (2003). Realist constructivism. *International Studies Review*, 5(3), 325–342.
- Befani, B., Ledermann, S., & Sager, F. (2007). Realistic evaluation and QCA. *Evaluation*, 13(2), 171–192.
- Berger, P. L., & Luckmann, T. (1966). *The social construction of reality: A treatise in the sociology of knowledge*. Anchor.
- Burch, P., & Heinrich, C. J. (2015). *Mixed methods for policy research and program evaluation*. SAGE.
- Cairney, P. (2016). *The politics of evidence-based policy making*. Springer.
- Cartwright, N. (1980). The truth doesn't explain much. *American Philosophical Quarterly*, 17(2), 159–163.
- Chen, H. T. (2012). Theory-driven evaluation: Conceptual framework, application and advancement. In R. Strobl, O. Lobermeier & W. Heitmeyer (Eds.), *Evaluation von Programmen und Projekten für eine demokratische Kultur* (Vol. 16, pp. 17–40). Springer.
- Cousins, J. B., & Whitmore, E. (1998). Framing participatory evaluation. *New Directions for Evaluation*, 1998(80), 5–23.
- Crutzen, R., Peters, G.-J. Y., & Noijen, J. (2017). Using confidence interval-based estimation of relevance to select social-cognitive determinants for behavior change interventions. *Frontiers in Public Health*, 5, 165.
- Davis, P. (2005). The limits of realist evaluation: Surfacing and exploring assumptions in assessing the best value performance regime. *Evaluation*, 11(3), 275–295.
- De Francesco, F. (2019). Following the evidence practice: An analysis of evaluation studies on EU railway policy. *Politische Vierteljahresschrift*, 60(4), 785–803.
- De Francesco, F., & Pattyn, V. (2021). Policy evaluation styles. In M. Howlett & J. Tosun (Eds.), *The Routledge handbook of policy styles* (pp. 408–421). Routledge.
- Donaldson, S. I., & Lipsey, M. W. (2006). Roles for theory in contemporary evaluation practice: Developing practical knowledge. In I. Shaw, J. C. Greene & M. M. Mark (Eds.), *The handbook of evaluation: policies, programmes, and practices* (pp. 56–75). SAGE.
- Easton, D. (1979). A framework for political analysis.
- Fischer, F., & Forester, J. (1993). *The argumentative turn in policy analysis and planning*. Duke University Press.
- Fischer, F., & Gottweis, H. (2013). The argumentative turn in public policy revisited: Twenty years later. *Critical Policy Studies*, 7(4), 425–433.
- Fontaine, G. (2020). The contribution of policy design to realist evaluation. *Evaluation*, 26(3), 296–314.
- Freeman, R., Griggs, S., & Boaz, A. (2011). The practice of policy making. *Evidence & Policy: A Journal of Research, Debate and Practice*, 7(2), 127–136.
- Freeman, R., & Maybin, J. (2011). Documents, practices and policy. *Evidence & Policy: A Journal of Research, Debate and Practice*, 7(2), 155–170.
- Graham, A. C., & McAleer, S. (2018). An overview of realist evaluation for simulation-based education. *Advances in Simulation*, 3(1), 1–8.
- Greenhalgh, J., & Manzano, A. (2022). Understanding 'context' in realist evaluation and synthesis. *International Journal of Social Research Methodology*, 25(2), 583–595.
- Greenhalgh, T., Humphrey, C., Hughes, J., Macfarlane, F., Butler, C., & Pawson, R. (2009). How do you modernize a health service? A realist evaluation of whole-scale transformation in London. *The Milbank Quarterly*, 87(2), 391–416.
- Grimmelikhuijsen, S., Jilke, S., Olsen, A. L., & Tummers, L. (2017). Behavioral public administration: Combining insights from public administration and psychology. *Public Administration Review*, 77(1), 45–56.
- Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation*. SAGE.
- Haas, P. (2004). When does power listen to truth? A constructivist approach to the policy process. *Journal of European Public Policy*, 11(4), 569–592.

- Hall, P. A. (1993). Policy paradigms, social learning, and the state: The case of economic policymaking in Britain. *Comparative Politics*, 25(3), 275–296.
- Hall, P. A., & Taylor, R. C. R. (1996). Political science and the three new institutionalisms. *Political Studies*, 44(5), 936–957.
- Hay, C. (2016). Good in a crisis: The ontological institutionalism of social constructivism. *New Political Economy*, 21(6), 520–535.
- Head, B. W. (2016). Toward more ‘evidence-informed’ policy making? *Public Administration Review*, 76(3), 472–484.
- Hogwood, B. W., Gunn, L. A., & Archibald, S. (1984). *Policy analysis for the real world*. Oxford University Press.
- Hood, C. C., & Margetts, H. Z. (2007). *The tools of government in the digital age*. Macmillan International Higher Education.
- Howlett, M., & Rayner, J. (2007). Design principles for policy mixes: Cohesion and coherence in ‘new governance arrangements’. *Policy and Society*, 26(4), 1–18.
- Hudson, V. M., & Day, B. S. (2019). *Foreign policy analysis: Classic and contemporary theory*. Rowman & Littlefield.
- Ingram, H., & Schneider, A. L. (1990). Improving implementation through framing smarter statutes. *Journal of Public Policy*, 10(1), 67–88.
- Ingram, H., Schneider, A. L., & DeLeon, P. (2019). Social construction and policy design. In P. Sabatier (Ed.), *Theories of the policy process* (2nd ed., pp. 93–126). Routledge.
- Jacob, S., Speer, S., & Furubo, J.-E. (2015). The institutionalization of evaluation matters: Updating the *International Atlas of Evaluation* 10 years later. *Evaluation*, 21(1), 6–31.
- King, G., Gakidou, E., Ravishankar, N., Moore, R. T., Lakin, J., Vargas, M., Téllez-Rojo, M. M., Hernández Ávila, J. E., Ávila, M. H., & Llamas, H. H. (2007). A ‘politically robust’ experimental design for public policy evaluation, with application to the Mexican universal health insurance program. *Journal of Policy Analysis and Management*, 26(3), 479–506.
- Knill, C., & Tosun, J. (2012). *Public policy: A new introduction*. Palgrave Macmillan.
- Knill, C., & Tosun, J. (2020). *Public policy: A new introduction* (2nd ed.). Red Globe Press.
- Lasswell, H. D. (1956). *The decision process: Seven categories of functional analysis*. University of Maryland Press.
- Law, J., & Singleton, V. (2014). ANT, multiplicity and policy. *Critical Policy Studies*, 8(4), 379–396.
- Lemire, S., Kwako, A., Nielsen, S. B., Christie, C. A., Donaldson, S. I., & Leeuw, F. L. (2020). What is this thing called a mechanism? Findings from a review of realist evaluations. *New Directions for Evaluation*, 2020(167), 73–86.
- Liebert, U. (2010). Challenging the second model: Can a regional supranational form of democratic statehood work? In E. O. Eriksen & J. E. Fossum (Eds.), *Proceedings from the RECON Midterm Conference* (pp. 115–132). IRENA.
- Lombardo, E., & Kantola, J. (2021). Social constructivism. In G. Abels, A. Krizsán, H. MacRae & A. van der Vleuten (Eds.), *The Routledge handbook of gender and EU politics* (pp. 43–55). Routledge.
- Manzano, A. (2016). The craft of interviewing in realist evaluation. *Evaluation*, 22(3), 342–360.
- March, J. G., & Olsen, J. P. (1989). *Rediscovering institutions*. The Free Press.
- Marchal, B., Van Belle, S., Van Olmen, J., Hoérée, T., & Kegels, G. (2012). Is realist evaluation keeping its promise? A review of published empirical studies in the field of health systems research. *Evaluation*, 18(2), 192–212.
- Mavrot, C., & Pattyn, V. (2022). The politics of evaluation. In A. Ladner & F. Sager (Eds.), *Handbook on the politics of public administration* (pp. 244–254). Edward Elgar Publishing.
- Mendez, N. (2021). Do social constructions and gender mainstreaming in policy design affect the evaluation of the policy process? An analysis of a policy intervention in a postconflict society. *Social Science Quarterly*, 102(4), 1699–1718.
- Merton, R. K. (1968). *Social theory and social structure*. Free Press.
- Moynihan, D. P., & Soss, J. (2014). Policy feedback and the politics of administration. *Public Administration Review*, 74(3), 320–332.
- Nanninga, M., & Glebbeek, A. (2011). Employing the teacher–learner cycle in realistic evaluation: A case study of the social benefits of young people’s playing fields. *Evaluation*, 17(1), 73–87.

- Nielsen, S. B., Lemire, S., & Tangsig, S. (2022). Unpacking context in realist evaluations: Findings from a comprehensive review. *Evaluation*, 28(1), 91–112.
- Nutley, S., Morton, S., Jung, T., & Boaz, A. (2010). Evidence and policy in six European countries: Diverse approaches and common challenges. *Evidence & Policy: A Journal of Research, Debate and Practice*, 6(2), 131–144.
- Organisation for Economic Co-operation and Development (OECD). (2020). *Improving governance with policy evaluation*. OECD Publishing.
- Organisation for Economic Co-operation and Development (OECD) (2021). *Applying evaluation criteria thoughtfully*. OECD Publishing.
- Pattyn, V., Álamos-Concha, P., Cambré, B., Rihoux, B., & Schalembier, B. (2022). Policy effectiveness through configurational and mechanistic lenses: Lessons for concept development. *Journal of Comparative Policy Analysis: Research and Practice*, 24(1), 33–50.
- Pattyn, V., Van Voorst, S., Mastenbroek, E., & Dunlop, C. A. (2018). Policy evaluation in Europe. In E. Ongaro & S. van Thiel (Eds.), *The Palgrave handbook of public administration and management in Europe* (pp. 577–593). Springer.
- Pawson, R. (2002). Evidence-based policy: In search of a method. *Evaluation*, 8(2), 157–181.
- Pawson, R., Greenhalgh, T., Harvey, G., & Walshe, K. (2005). Realist review – a new method of systematic review designed for complex policy interventions. *Journal of Health Services Research & Policy*, 10(1_suppl), 21–34.
- Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. SAGE.
- Pawson, R., & Tilley, N. (2001). Realistic evaluation bloodlines. *American Journal of Evaluation*, 22(3), 317–324.
- Pawson, R., & Tilley, N. (2004). *Realist evaluation*. Cabinet Office.
- Risse, T. (2004). Social constructivism and European integration. In A. Wiener & T. Diez (Eds.), *European integration theory* (pp. 159–176). Oxford University Press.
- Rossi, P. H., Lipsey, M. W., & Henry, G. T. (2018). *Evaluation: A systematic approach*. SAGE.
- Sager, F., & Andereggen, C. (2012). Dealing with complex causality in realist synthesis. *American Journal of Evaluation*, 33(1), 60–78.
- Saurugger, S. (2013). Constructivism and public policy approaches in the EU: From ideas to power games. *Journal of European Public Policy*, 20(6), 888–906.
- Schneider, A., & Ingram, H. (1993). Social construction of target populations: Implications for politics and policy. *American Political Science Review*, 87(2), 334–347.
- Scriven, M. (1991). *Evaluation thesaurus*. SAGE.
- Searle, J. R., & Willis, S. (1995). *The construction of social reality*. Simon & Schuster.
- Stame, N. (2004). Theory-based evaluation and types of complexity. *Evaluation*, 10(1), 58–76.
- Stockmann, R., Meyer, W., & Taube, L. (2020). The institutionalisation of evaluation in Europe: A synthesis. In R. Stockmann, W. Meyer, & L. Taube (Eds.), *The institutionalisation of evaluation in Europe* (Vol. 14, pp. 483–522). Springer.
- Sunstein, C. R. (2020). *Behavioral science and public policy*. Cambridge University Press.
- Taylor, D. (2005). Governing through evidence: Participation and power in policy evaluation. *Journal of Social Policy*, 34(4), 601–618.
- Taylor, D., & Balloch, S. (Eds.). (2005). *The politics of evaluation: Participation and policy implementation*. Policy Press.
- Tosun, J., & Treib, O. (2018). Linking policy design and implementation styles. In M. Howlett & I. Mukerjee (Eds.), *Routledge handbook of policy design* (pp. 316–330). Routledge.
- Van der Knaap, P. (1995). Policy evaluation and learning: Feedback, enlightenment or argumentation? *Evaluation*, 1(2), 189–216.
- Van der Meer, F.-B. (1999). Evaluation and the social construction of impacts. *Evaluation*, 5(4), 387–406.
- Varone, F., Bundi, P., & Gava, R. (2020). Policy evaluation in parliament: Interest groups as catalysts. *International Review of Administrative Sciences*, 86(1), 98–114.