

Øyunn Syrstad Høydal,  
Oslo Business School, Oslo  
Metropolitan University,  
Norway  
oyhoy@oslomet.no

## Abstract

The concept of framing implies that policy issues can be viewed from different perspectives. The choice of perspective influences the way societal phenomena are understood and the solutions to the perceived problems, as well as the perspectives focused on when evaluating these solutions. Framing thus involves power, and framing in relation to evaluation involves the power to influence production of public knowledge and, subsequently, policy decisions informed by this evidence. While the role of framing in policymaking seems to be a topic of increasing academic interest, thus far evaluation literature has not reflected this trend. Evaluation still tends to be perceived as an objective external voice, providing the political-administrative system with neutral evidence. This paper challenges this dominant image of evaluation and seeks to contribute to clarifying the complex relationship between framing, evaluation and power.

## Introduction

Evaluation has traditionally been perceived as an objective external voice, examining public policy from a professional distance and, from this perspective, providing the political-administrative system with evidence (Pawson and Tilley 1997; Tornes 2013). Research on the use of evaluations in policy has therefore been more or less explicitly based on the perception of the political-administrative system and the academic world as two separate communities (Coleman 1972; Caplan 1979). In theory, knowledge users, such as politicians or civil servants, discuss values and set goals, while knowledge producers find answers to societal challenges through objective facts. However, in practice, evaluation typically involves dialogue, cooperation and mutual influence between knowledge users and producers (Dahler-Larsen 2006; Høydal 2020). As an initiator, the public sector has the power to make choices regarding the aims, topics and methodological approaches that influence subsequent phases of the evaluation and, consequently, the concluding results. This way of steering knowledge production, and the following results, in certain directions through the selection of specific perspectives or designs, could be described as an act of framing (Van Hulst and Yanow 2016). However, framing also occurs at a more general level, in that the dominant values in any given society influence the perspectives and the priorities of policymakers, as well as evaluators and the public (House 2017; Hulst and Yanow 2016).

While the role of framing in policymaking seems to be a topic of increasing academic interest (see for example, Burlone 2020; McIntyre 2020; Park and Lee 2020; Vieira 2020; Willems et al. 2020), thus far the evaluation literature has not reflected this trend (Nordesjö 2019). One exception is House's (2017) work discussing the framing of race in evaluations and how an awareness of racial biases could lead to better evaluations and ultimately to less racism, thus illustrating the significant aspect of power in framing evaluations.

My aims in this paper are to a) demonstrate that framing is essential to understanding evaluation, b) provide evidence of the close relationship between framing and power and thus to contribute to a use of the framing perspective that

is more power-sensitive, and c) argue for a broader perspective on the use of evaluation in the public sector. My paper is theoretically informed by Hulst and Yanow's (2016) conceptualisation of framing, based on Rein and Schön's (1994; 1996) work on frame analyses. Power is discussed in relation to Lukes' (2004) three dimensions, guided by a set of power-sensitive questions, as suggested by Bugge (2002). The theoretical perspectives and the discussion are supplemented by empirical examples from research and public debate. As a recurring example, I use the Norwegian participation in the Programme for International Student Assessment (PISA), which tests the skills and knowledge of 15-year-old students every three years. Due to the rich literature on PISA participation, as well as its international character, influence and relatively long history, I believe that the PISA represents an interesting case.

The global trend of evaluation and evidence-based policymaking (EBPM) occurs in national knowledge regimes, characterised by their respective institutional and cultural features (Campbell and Pedersen 2015). When discussing evaluation in an international context, it is important to bear in mind that this form of knowledge production comes in a wide variety. However, the power in framing is an aspect of all forms of evaluations or knowledge production, and an awareness of these phenomena is consequently relevant across national evaluation or knowledge regimes.

## A Presentation of the Framing Perspective

Use of the frame analysis goes back to the work of Erving Goffman (1974) and symbolic interactionism. According to Goffman, frames are the culturally determined definitions of reality that allow people to make sense of objects and events. Academically, framing has been studied from a wide range of perspectives, including social movement research (Benford and Snow 2000), public opinion research (Chong and Druckman 2007; Øverbye 2006), communication studies (Matthes 2009; Scheufele 1999), environmental studies (Béné et al. 2020; Willems et al. 2020) and public policy analyses (Burlone 2020; Hulst and Yanow 2016; Rein and Schön 1994, 1996). Despite their different focuses, the major premise in this research is the notion that an issue can be viewed from several perspectives and that the preferred perspective will influence further thinking about the issue and the opinions concerning how this issue should be treated (Hulst and Yanow 2016). This implies that policy issues can be viewed from different perspectives. These perspectives will ultimately influence the way societal phenomena are understood and therefore affect the solutions to the perceived problems – as well as the perspectives in focus when evaluating these solutions.

The very idea that policy issues can be understood in multiple ways represents a clear contrast to the rationalistic idea that policymaking is all about finding the best solutions to given problems. Such a perspective has influenced evaluation research for decades. Instead, the framing perspective acknowledges the messy nature of politics and the fact that policymaking entails a battle between different values and the right to define the world (Stone 2012; Bacchi 2009). Framing thus involves power, and framing in relation to evaluation

involves the power to influence production of public knowledge and, subsequently, policy decisions informed by this evidence.

Studying controversies involved in the context of public problems, Rein and Schön (1994) reveal the significance of conflicting perspectives or frames in policymaking. Different perspectives lead actors to argue past each other and make it difficult to establish common ground, which is essential for finding acceptable collective solutions. Building on this work, Hulst and Yanow (2016) distinguish between the dynamic activity of framing (the process in which frames are constructed) and what they refer to as the more strategic and static concept of frames. According to the authors, frames are "...often treated as objects [that] people possess in their heads and develop for strategic purposes" (p. 93). Even though this distinction might not always be clear cut, Hulst and Yanow (2016) claim that framing offers a more dynamic and power-sensitive perspective, better suited for grasping the active processes of policymaking.

Hulst and Yanow (2016) describe framing as sensemaking work, a process of naming (which includes selecting and categorising) and storytelling. Sensemaking refers to a process where the actors involved try to figure out what really occurs. At this stage, framing takes place, in what Hulst and Yanow (2016), inspired by symbolic interactionism as well as the work of Schön, refer to as a "conversation with the situation" (p. 98). Through this interactive process, framing organises prior knowledge and values and provides guidelines for action. According to Hulst and Yanow, the different experiences, values and knowledge of the actors involved lead to different perceptions of the situations and future solutions. To communicate about the situation, boundaries are drawn, and differences are established, such as between natives and immigrants or between normal and abnormal (p. 100). Due to their various perceptions of a situation, different actors will select different elements to be focused on, and their categorising and stories about the same situation will vary. Storytelling binds together the elements of a situation to create a coherent presentation. Such stories define the problems of society and preferable solutions, as well as attribute either blame or praise to the parties concerned.

## Framing and Its Relevance for Evaluation

Hulst and Yanow's (2016) work on framing has been developed in relation to policy issues, but framing – as both a concept and a process – is also relevant for evaluation. **First**, evaluation is by nature an act of framing. Evaluations tell stories about society; they define our problems, as well as tell us whom to blame and how to overcome the trouble. Simply by choosing a programme for evaluation, framing takes place because such a decision involves a process of defining or sensemaking of the current situation. The process of framing becomes more defined as the project takes form through a formal evaluation description and a project announcement. The project description typically includes a selection of focus through the specified aim of the evaluation, often by asking for specific methodological approaches. While this framing would be influenced by the nature of the programme in question, as well as the limited time and budget, it could also involve more political or strategic aspects. For instance, when the Norwegian government ordered a consequence analysis

regarding the permission to dump nearly 6 m tonnes of tailings annually for 50 years into the Norwegian salmon fjord Førdefjorden, the government was interested in the environmental impact of the dumping. The report concluded that the dumping could negatively affect the biosphere in the fjord and therefore was not recommended. Subsequently, a new analysis was ordered. This time, the aim of the assignment was to investigate the economic and societal impacts of a potential mining activity. With this outset, the new consequence analysis recommended the appropriate plans for mining and dumping of tailings in the fjord (Naturvernforbundet 2020; The Guardian 2015). Thus, this example clearly illustrates how the different perspectives of the two ex-ante evaluations led to opposite conclusions and therefore how framing of evaluations could have a huge impact on political decisions. However, the consequences of the framing of evaluations are typically more subtle and less easy to pinpoint. Nonetheless, previous studies have revealed that evaluations that produce knowledge, in line with political regimes and existing budgets, increase their likelihood of being applied and have instrumental impacts (Weiss 1998; Innvær et al. 2002; Oliver et al. 2014).

**Second**, framing occurs through the dominant discourses or stories in any given society. By representing collectively held worldviews, such storytelling has an automated character, providing focus and influencing further thinking about societal issues, for instance, about race, gender or ethnic groups (Bacchi 2009; Espinosa 2013; House 2017). Hence, when evaluating, it is important to be aware of the power of these collectively held ideas and how they unconsciously influence our perspectives and judgements. House (2017) stresses the fact that to be fully able to comprehend societal events, evaluators must understand the culture where these events occur. In North American society, with a history of more than 520 years of racism, it is evident that the white population has been affected by negative images of African Americans. This has shaped the policies that have been evaluated, as well as the evaluators' own perceptions of social issues (House 2017). However, in evaluation research, there has been a limited interest in the meaning of such contextual factors, and the role of culture remains less commonly discussed in the literature (Alkin and King 2017; Højlund 2014; Vo and Christie 2015).

**Third**, categorising is a central part of the framing of evaluations. Støkken (2013) describes how the public quality criteria, as used in ex ante evaluations and certification of child-care institutions, to a very limited degree matched the quality criteria emphasised by the people working or living in these institutions. While the certification criteria were based on measurable factors, such as the educational backgrounds of the staff, formal plans, and so on, the people involved stressed personal and contextual quality factors. In sum, the categories have the power to either certify or not certify child-care units as public institutions, although other quality indicators might have represented the user interests in a better way.

Based on the three points made, the evaluation reports play a significant role in framing societal phenomena. A well-known example is the PISA. Before the PISA, Norwegian authorities assumed that a solid economic foundation would guarantee a high level of student performance (Elstad & Sivesind 2010). However, the first PISA testing (2001) revealed that Norwegian students ranked

below the OECD average and those of the other Scandinavian countries (Baird et al. 2011, p. 24; Norwegian Directorate for Education and Training 2011). This induced a national trauma – “the PISA shock”. In short, Norway had spent a lot of money on education and believed that its students would perform better than they did. The PISA became what Dahler-Larsen (2012) describes as a “magic mirror” (p. 205), where society perceived itself and its realities anew, and overnight, the PISA was given the constitutive power to define Norwegian students as not performing well enough. Constitutive effects could be exemplified by the fact that standards and definitions used in evaluations of public programmes end up as new public standards. Hence, evaluations are not just influenced or framed by the values of society but also frame society.

## Perspectives on Power

As illustrated, the relation between evaluation and framing is multidimensional and complex. When bringing power into the discussion, it therefore becomes evident to apply a perspective that is sensitive to this complexity. Power has traditionally been perceived as an ability to make others act as you would like even though they would have preferred to do otherwise. This reflects classical definitions by Weber and Dahl (1957), and it is based on three central dimensions: power as intentional, relational and causal. This means that power is found in the relation between actors, and it makes sense to define someone having or wielding power as “the one who exercises power [getting one’s] intentions through” (Engelstad 1999). According to Morris (2006), this reflects a power. Have I understood your comment correctly?

In *Power: A Radical View*, Lukes (2004) criticises this classic perception of power for being one-dimensional and solely focusing on situations where there is an observable conflict of subjective interests. Lukes adds the second dimension of power, that is, the ability to set the agenda “... to the extent that [if] a person or group – consciously or unconsciously – creates or reinforces barriers to the public airing of policy conflicts, that person or group has power” (p. 20). However, Lukes claims that the most effective use of power is to prevent conflict from arising in the first place. The third dimension of power is therefore power as domination. Through domination, the powerful transform the powerless in such a way that the latter group behaves as the former group wishes by shaping the latter’s perceptions, cognitions and preferences to ensure its acceptance of a certain role in the existing order. Haugaard (2008) claims the third dimension of power reflects a tacit, interpretative horizon that enables humans to handle their social context, in line with Bourdieu’s reference to *habitus* or Foucault’s use of the concepts of discourses and epistemes.

According to Hulst and Yanow (2016), framing influences the content of policy issues, the actors to be involved and the policy process as such. From this outset, Lukes’ (2004) multidimensional perspective is relevant. According to Lukes, power is an aspect, not just of conflicting interests between A and B, but also in relation to agenda setting and knowledge formation. Hence, his work enables an approach where power is understood in relation to agency and structure, as well as *habitus*. A perspective that recognises that power can be found in both relations and structures.

## The Power in Framing

The many ways that framing is relevant to evaluation and the potential complexity involved when adding the dimension of power to the discussion make evident the necessity to apply a clear structure to the analyses. According to Bugge (2002), an analysis of power must include questions rooted in a general theory of power. Such power-sensitive questions should include the following: 1) What factors make this phenomenon a power relationship? 2) To what extent is power intended by actors who try to promote their special interests? 3) To what extent is it legitimised, supported or actively opposed? 4) By virtue of which mechanisms is this balance of power maintained or changed over time? 5) What are the conditions for changing or abolishing the dominance relationship? In my analyses of the power dimension in the relation between evaluation and framing, I use the emphasised questions as guidelines in the following sections.

### 1. What Factors Makes Framing in Evaluation a Power Relationship?

*Framing* refers to the fact that an issue can be viewed from several perspectives and that the preferred perspective will influence further thinking about these issues, which again could have a clear impact on policymaking. When discussing power in relation to the framing of an evaluation, on one hand, one should be concerned about *the power to* affect evaluation processes and results through *sensemaking, naming and storytelling*. On the other hand, in the process of framing, there is also the *power over* the framing process, exercised by the authority figure who decides which of many possible frames to use in a specific evaluation process, what to evaluate and whom to hire for the project.



While the evaluation initiators have the power to suggest certain ways of framing an evaluation process (i.e., through methodological preferences and focus), evaluators might still challenge these frames. They could do so through their interpretation of the project, their application of the methodological framework and their analyses. Even though the processes of data gathering and analyses are regarded as belonging to the domain of evaluators, the initiators could still interfere if they believe that the project is heading in the wrong direction (Dahler-Larsen 2006; Høydal 2020). For instance, they could do so by suggesting a certain case sample or holding workshops where preliminary findings would be discussed by experts and stakeholders (Høydal 2020). This way, initiators have power over evaluators if they challenge the latter's preferred framing in an undesired direction. In the end, the use of the evaluation results will also be affected by their alignment with the current policy (Weiss 1998).

Another dimension of power and framing in evaluation is the fact that knowledge users might hire specific consultants or firms because they represent a certain framing through the trending management ideas or approaches with which they are associated. According to Røvik (2007), international management ideas hold strong symbolic power – that is, evaluations following certain popular organisational standards or logics might lead to results with a strong standing. Additionally, management consultants offer models, techniques and services that are often packaged, labelled and structured as simple and effective technologies. This way, the framing might lead to a more powerful impact of the final report.

While the evaluation results have the potential power to influence policy and society, the programmes and the activities being evaluated are themselves outcomes of the structural power in the framing of current policies and the dominant discourses in society. In addition, evaluation is itself the result of framing, for several reasons. First, the dominant evidence discourses in the public sector frame policymaking as a process where evaluation is one of the right steps to take to ensure EBPM (Pattyn et al. 2018; Tornes 2012). Second, the way that the public sector chooses to evaluate is just one of many alternatives. Finally, evaluation could have constitutive effects in society. Such effects could be exemplified by the fact that standards and definitions used in evaluations of public programmes end up as new public standards. Hence, evaluations are not just influenced or framed by the values of society but also frame society. In my further discussion concerning evaluation, power and framing, I shall include all the previously mentioned aspects to illustrate the significance of framing and power in relation to evaluation.

## 2. Is the Framing of Evaluations the Result of *Intentions or Special Interests*?

To answer this question, I will use the experience from the Norwegian participation in the PISA evaluations as an illustrative case. Due to its international character, numerous intentions are likely to be at work when the PISA evaluations take place, illustrating the complexity involved. However, as initiator and organizer the OECD has a most central role. The OECD was created in 1961 because of the American influence in postwar Europe and their fight against Communism through support of liberal democracy and market economics (Sellar and Lingard 2013). When education, due to globalization and increasing international competition became a central aspect of the European economy in the mid-nineties, the outcome or results produced by the school system became significant for organizations like EU and the OECD (Grek and Lawn 2012). PISA is based on the idea that economic growth depends on quality educational systems, and that international evaluations can provide information about how well a country's system is performing in comparison with others (Elstad and Sivesind 2010). It is therefore plausible to claim that OECD in line with Lukes' (2004) first and second dimensions, has used its power as initiator to steer PISA in preferred directions through an *intentional* framing reflecting the link between a *certain form* of educational performance and economic growth.

Framing takes place not only by selecting the focus, but also by categorising the elements to exclude from the process and ensuring that the results tell a story in line with the political interest or values of OECD. The PISA tests are not based on national curricula. **Text removed – understood correctly?**  testing has made the Norwegian education system less focused on the value  and ideals traditionally associated with the Nordic school model, such as solidarity and community spirit and more individually and result oriented (Afdal and Afdal 2019). The current political focus on student performance represents a new phenomenon in the history of the Norwegian educational system and has led to a new test-based school system (Elstad and Sivesind 2010). Regarding the dimension of storytelling, the OECD's policy briefs based on PISA results have been criticized for being political documents, defining problems and solutions in

line with the interests of the organization, rather than evidence-based recommendations. The policy briefs have included contested recommendations like “do not increase the schools’ funding” and “close down small schools”, in line with liberal ideas to restrain public spending (Elstad and Sivesind 2010). The fact that the evaluation itself defines acceptable response illustrates one of several constitutive effects of the PISA-testing.

The framing of PISA could be described as a framing not just in line with the values of OECD, but in line with the general liberal economical ideas that have influenced western policy and societies the last decades. This way reflecting internationally taken-for-granted attitudes, hence representing the third dimension of power, power as domination. According to Scott (1990), power as domination could be exercised in a thick or a thin sense. The thick sense refers to a situation where the dominated believe in the values that oppress them, while the thin sense refers to a situation of resignation. The current international position of the PISA-evaluations and the values they represent, are likely to include both forms. In Norway the PISA-participation has changed the educational discourse and established a new way of thinking about schooling and educational quality (Elstad and Sivesind 2010). This development could be described as domination in its thick sense. While domination in its thin sense could be spotted in the critical resignation found among, for instance, Norwegian teachers (Marsdal 2011).

To sum up, PISA has had an enormous influence on the Norwegian educational policy and the general educational discourse. The framing of PISA represents the interests of OECD as well as general liberal economic values dominating Western policy the last decades. A framing reflecting the interests of powerful groups, not the students failing to perform within the test-based system or the different PISA critics. Today, the performance focus introduced by the first PISA evaluation, represents the new normal in Norwegian educational policy and an outset for new evaluations and testing. This way, PISA represents a conservative power, contributing to the reproduction of existing practices, structures and power relations in society.

### 3. To What Extent is the Aspect of Power in the Framing of Evaluations either Legitimised or Actively Opposed?

To answer this question, I will discuss what role evaluations are believed to play in current political-administrative systems. Evaluations take many forms and are conducted in different knowledge regimes (Campbell and Pedersen 2015). According to Duffy (2017), **What do you want me to do here?** the UK policy evaluation has been close to economic traditions, such as auditing and monitoring; in contrast, the Norwegian evaluation tradition is characterised by both theoretical and methodological influences of social science ideals (Dahler-Larsen 2013). According to Lindgren (2014), who writes from a Swedish perspective, by providing information, evaluation has become one of the most central policy instruments in public administration. This **And here?** is reflected in the use of evaluations in the Norwegian context. All parts of the central administration are required to provide information on their efficiency, achievements and results by conducting evaluations. The evaluations must be integrated into plans and strategies, along with other forms of knowledge



production (Norwegian Agency for Public and Financial Management [DFØ] 2016; Tornes 2013). In this perception, there is little room for perspectives that question current practices by introducing aspects such as framing, and following from this, power. Such topics are outside the box and potentially threaten the position of evaluation as a neutral, and useful policy instrument. Regarding PISA, it is administered according to scientific standards and managed by an international organization, and this probably curbs critics at a national level. However, knowledge is never neutral and objective in the evaluation society and the OECD has been criticized for hiding political interests behind the PISA evaluation (Elstad and Sivesind 2010).

When House (2017) claims that evaluators carry a certain baggage regarding race, which frames evaluations in a discriminating way, this is hardly compatible with the perception on the evaluator as a neutral scientist searching for truth or investigating the efficiency of the public sector. Most likely, matters become even more problematic from a rational point of view if predispositions such as racism are in fact implemented in public policy in general and expressed in the public programmes that are being evaluated (see also Hood and Dixon 2015). This phenomenon is also stressed by Rein and Schön (1994).

By defining evaluation as a policy instrument, providing useful information, and enabling control of public funding and efficiency, the aspect of framing is silenced in practice. In this way, the very definition or framing of evaluation curbs any opposition and reflects Lukes' (2004) third dimension of power as domination. Evaluation has become a protected discourse, something taken for granted and not to be criticised or questioned (Dahler-Larsen 2012).

#### 4. Is This Balance of Power either Maintained or Changed Over Time?

When examining the awareness of the framing of evaluations and the dimension of the power involved, the history of evaluation illustrates interesting aspects. Despite the rational dominance in the field of evaluation, the concept of evaluation, due to changes in science and public management ideals, has never been static. In the 1950s and the 1960s, the then-modern evaluation tradition was dominated by science optimism, positivism and the ideal of social engineering. The aim was to build a new and fair society based on evidence. However, when the rational project seemingly failed, user participation, dialogue, consensus and democratic legitimacy came into focus in the 1970s (Vedung 2010). Following the international emergence of new public management (NPM) reforms in the 1980s and the 1990s, goal and result management and subsequent evaluations became the new ways of managing the public sector. When the ideal of EBPM emerged, evaluation shifted from being a political-administrative activity to being associated with large data sets and academic activity (Vedung 2010).

The change between the first wave of science optimism and the dialogue-based ideal in the 1970s referred to a scientific paradigm shift where the rational ideal was not just criticised but partly replaced. The new constructivist paradigm perceived realities as social constructions, denying the idea or the existence of an objective truth on which inquiries could converge and considering it impossible to separate the inquirer from the object of investigation. Regarding methodology, a set of hermeneutic-dialectic processes was elevated as an ideal for evaluation (Vedung 2010). This shift illustrates that evaluation can be and has been

concerned with the core of the framing perspective, that is, the notion that an issue can be viewed from several standpoints and that the preferred perspective will influence further thinking about the issue (Hulst and Yanow 2016). However, as shown, the most recent trends have taken the evaluation tradition in another direction.

Another historic element worth noting is the fact that evaluations were initially perceived as part of a scientific transformation of society. The evaluation activity itself was framed as an instrument for societal change and the improvement of social conditions. The aspiration to include evaluations in societal transformation was also evident in the 1970s, but this perspective was discarded when evaluation became part of the NPM paradigm. In EBPM, the role of evaluations seems limited to the examination of What Works.

The history of evaluation illustrates how the ideological framing of evaluation has changed over time. The dominating idea of what evaluation is and how to evaluate becomes a guideline for knowledge formation in line with Lukes' (2004) third dimension. The different evaluation paradigms follow from larger public management or scientific trends, and this might explain why the users and the producers of evaluations have accepted and adjusted to the shifting ideological framing of evaluation.

## 5. What Are the Conditions for Changing or Abolishing the Dominance Relationship?

Rein and Schön (1994) launched the idea of a reflective policy conversation as a solution to the problems arising from conflicting perspectives or frames in policymaking. The participants of such conversations should “put themselves in the shoes of other actors” (p. 187), as well as examine how their own frames might contribute to the problematic situation. Overall, the actors involved should accept the existence and the influence of frames.

While civil servants, policymakers and evaluators might be aware that the framing of an evaluation has the power to influence the result, there seems to be little room for discussions regarding this phenomenon. The framing of evaluation itself, as a policy instrument providing useful knowledge, is not open to such conversations. As previously illustrated, the meaning of evaluation and the ideas surrounding this phenomenon have changed over the years, and a new evaluation wave might be more sensitive to aspects such as framing. Indeed, such alternative perceptions already exist side by side with the dominant rational concept. According to House (2017), democratic evaluation is one alternative that can hinder the racial biases in the current evaluation activity. Within this tradition, evaluations are designed to include and involve stakeholder groups, promote dialogue among all groups involved and establish the so-called deliberative processes in the evaluation project. Regarding the PISA example, a dialogue including critical stakeholders, could maybe lead to more nuanced use of the evaluation results.

Another alternative that is intended to challenge the power in framing could in fact take place through a deliberate choice of framing. By knowingly emphasising values that tend to be underrated, evaluations could contribute to better knowledge on the impact of public activities. For instance, this could be

done through gender-sensitive evaluations, specifically focusing on the aspect of gender in public policy (Espinosa 2013) or the so-called social impact framing. The latter strategy emphasises a venture's attention to and care for the physical earth, as well as the usual socioeconomic environment. A recent study reveals the positive consequences of social impact framing for the outcomes of business evaluations of female-led ventures (Lee and Huang 2018).

One more way to challenge the power in framing could be to stop defining and thinking about evaluation solely as a neutral policy instrument. Evaluation is also a societal phenomenon, an international trend that changes its character over time and adjusts to other dominant ideas concerning knowledge production or the right way to manage the public sector. Evaluation reflects the values of the society where it takes place and reproduces collectively held worldviews (Dahler-Larsen 2012; Power 1997; Schwandt 2009). By participating in PISA, Norwegian authorities do not only signal their dedication to education and evidence-based policymaking, but also their belief in the values and policy approved by the OECD (Høydal 2019).

According to Røvik (2007), the successful narratives of contemporary organisational trends have laid the foundation for the symbolic meaning of the same ideas. The story of evaluation as a practically useful and necessary tool means that evaluation also becomes a symbol of the same qualities. It may be difficult to determine where the boundaries between the symbolic power and the practical impact of evaluations can be drawn or whether these two sides act as mutual reinforcements of each other. If one manages to combine a sociological or a constructivist view (that acknowledges the symbolic dimension of evaluation) with the notion that evaluation could also be a useful tool for the public sector, the result might be a more realistic and conscious outset of this knowledge production. This could make it more plausible for public administrators, as well as evaluators, to choose alternative and more power-sensitive evaluation models.

By applying a multidimensional power perspective, I illustrate the different aspects of power in the relation between evaluation and framing. Openly conflicting interests regarding the aim, the focus or the methodological approach might be the most obvious examples of the power characterising framing in evaluation. However, power is also a dimension in the self-censorship of evaluators, in reports reflecting the interests and the values of the powerful, in the reproduction of collectively accepted categories and definitions and in the dominant societal discourse. Thus, such a multifaceted perspective illustrates that framing of evaluations includes both intended and unintended aspects. Because evaluations produce knowledge used by the political-administrative system, the categories and the definitions used in evaluation reports tend to end up as new public standards (Dahler-Larsen 2012). Hence, evaluations hold potential constitutive power to change society and are framed by the collectively held ideas and dominant discourses in the societies where they are conducted. Additionally, the current framing of evaluation is itself the result of historic changes in the ontological, epistemic and methodological framing of evaluation. These previous shifts illustrate the potential for new directions for evaluation and a more power-sensitive approach to this activity; in fact, a variety of such evaluation approaches exists and has already been applied with success.

## Concluding Discussion

According to Huberman (1987), discussions and research concerning the relation between evidence and policymaking tend to be normative and unnuanced because such topics are closely related to fundamental scientific questions, such as the rationality of actors, the objectivity of knowledge and the freedom of values. Regarding the limited research interest in the relation between framing and evaluation, I believe that Huberman is right. To question the dominant idea of evaluation as an objective tool, that is, a policy instrument providing neutral and useful knowledge for the public sector, entails to question the dominant rational paradigm in western society. Such a paradigm has influenced the idea of policy, bureaucracy, knowledge production and the relations among these fundamental building blocks of our culture.

Acknowledging the power of framing in relation to evaluation raises the question of the dominant perception on the evaluative activity and the evidence from this knowledge production. Because framing also involves the values of society in general and the dominant discourse in any given culture, the entire discussion takes a fundamental shift from the idea of evaluation as a policy instrument providing policymakers with neutral evidence. However, at a time when large groups have lost their trust in science, in public reports and in evaluations, when people learn from the internet rather than from research and swap newspapers for Facebook, the necessity to prevent the carousel from going wilder has become evident. If evaluations in practice never criticise the establishment, the dominant values or the status quo, what are they worth for the critics of the establishment or for the underprivileged? Why should they trust the politicians referring to such knowledge? They might as well find themselves an online community of kindred spirits. To re-establish trust in academic evidence, such evidence must be trustworthy, and this issue includes far more aspects than choosing the right sampling technique. It must include an awareness regarding framing and subsequently the power in framing, an awareness that is equally important for qualitative and quantitative approaches.

Norwegian authorities have discussed whether knowledge production like evaluation should be undertaken at a greater distance from political interests. However, the current praxis is the opposite (Botheim and Solumsmoen 2009; Christensen et al. 2014). At present, the political-administrative system plays a central role for instance in the intended framing of evaluations (Høydal 2020). The PISA-example illustrates how even international organisations might have the power to influence on a national level through their framing of evaluations. Due to PISA, the Norwegian school system has gone through major changes while the national PISA-results hardly have improved (Høydal 2019; OECD 2019). Given the rational logic behind evaluations, the discrepancy between means and ends should be an incitement for public authorities to take a more critical position to the testing and its consequences. However, this kind of critique is hardly heard of. It might be the apparent international consensus behind PISA and the dominant values and ideas the evaluation represents, which stops participating nations from questioning the role and consequences of PISA - as well as its ideological fundament. Hence, the case of PISA clearly illustrates the complex relationship between evaluation, framing and power.

Framing or power-sensitive evaluation models are available. However, if such models are aimed to be more widespread, evaluation initiators must acknowledge the need for these alternatives. Such sensitivity most likely emerges when the very idea of evaluation as neutral knowledge production at an arm's-length distance is replaced with a more realistic picture. Future research on evaluation should clearly play a role in this regard.

## Acknowledgements

Thanks to colleague and friend Chris Rønningstad for insightful comments – as always.

## References

- Afdal, Hilde W. & Geir Afdal (2019) The making of professional values in the age of accountability. *European Educational Research Journal*, 18(1): 105-124.
- Alkin, Marvin C. & Jean A. King (2017) Definitions of evaluation use and misuse, evaluation influence, and factors affecting use, *American Journal of Evaluation*, 38 (3): 434-450.
- Bacchi, Carol (2009) *Analysing Policy*, Pearson Higher Education, AU. French Forest.
- Baird, Jo-Anne, Talia Isaacs, Sandra Johnson, Gordon Stobart, Guoxing Yu, Terra Sprague & Richard Daugherty (2011) Policy effects of PISA Oxford University Centre for Educational Assessment, Microsoft Word - Policy Effects of PISA OUCEA\_(ox.ac.uk)
- Béné, Christophe, Nozomi Kawarazuka Pham, Steef de Haan, Huynh Tuyen, Duong T. Thi & Chien Dang (2020) *Policy framing and crisis narratives around food safety in Vietnam, Environment and Planning E: Nature and Space*. 0 (0): 1–25 <https://doi.org/10.1177/2514848620941515>
- Benford, Robert D. & David A. Snow (2000) Framing processes and social movements: An overview and assessment, *Annual Review of Sociology*, 26 (1): 611–639.
- Botheim, Ingunn & Dag Solumsmoen (2009) Direktoratene: Rolleblending eller nyttig rollekombinasjon? (Agencies and their roles), *Stat & Styring*, 19 (1): 8–9.
- Bugge, Lars (2002) Den gjennomsiktige makten. Kritikk av Maktutredningens maktforståelse, *Agora*, 20 (3–4): 20–50.
- Burlone, Nathalie (2020) Value-based issues and policy change: Medical assistance in dying in four narratives, *Social Policy & Administration*, Vol.54 (7): 1096-1109
- Campbell, John L. & Ove K. Pedersen (2015) Policy ideas, knowledge regimes and comparative political economy, *Socio-Economic Review*, 13 (4): 679–701.
- Caplan, Nathan (1979) The two-communities theory and knowledge utilization, *American Behavioral Scientist*, 22 (3): 459–470.

- Chong, Dennis & James N. Druckman (2007) A theory of framing and opinion formation in competitive elite environments, *Journal of Communication*, 57 (1): 99–118.
- Christensen, T., M. Egeberg, P. Lægreid & J. Aars (2014) *Forvaltning og politikk (4. utg.). (Public Management and Policy)*, Universitetsforlaget, Bergen.
- Coleman, J. S. (1972) *Policy Research in the Social Sciences*, General Learning Press, Morristown.
- Dahl, Robert A. (1957) The concept of power, *Behavioral Science*, 2 (3): 201–215.
- Dahler-Larsen, Peter (2006) *Evaluering og magt. (Evaluation and Power)*, Magtutredningen, Århus.
- Dahler-Larsen, Peter (2012) *The evaluation Society*, Stanford University Press, Stanford, CA.
- Dahler-Larsen, Peter (2013) ‘Evaluering i Norge set udenfra’, in Anne Halvorsen, Einar Lier-Madsen & Nina Jentoft (eds), *Evaluering. Tradisjoner, praksis, mangfold*, 43-57 Fagbokforlaget, Bergen.
- Duffy, Deidree N. (2017) *Evaluation and Governing in the 21st Century: Disciplinary Measures, Transformative Possibilities*, Palgrave Macmillan UK, London.
- Elstad, Eyvind & Kirsten Sivesind (2010) ‘Innledning’ in Eyvind Elstad & Kirsten Sivesind (eds), *PISA – sannheten om skolen?* 16-19, Universitetsforlaget, Oslo.
- Engelstad, Fredrik (1999) ‘Maktbegrepet etter Max Weber’ in Fredrik Engelstad (ed), *Om makt. Teori og kritikk*, 15-40, Ad Notam Gyldendal, Oslo.
- Espinosa, Julia (2013) Moving towards gender-sensitive evaluation? Practices and challenges in international development evaluation, *Evaluation*, 19 (2): 171–182.
- Furre, Harald & Aase M. J. Horrigmo (2013) ‘Nærhet og distanse: Om forholdet mellom bestiller og utfører av evalueringer’ in I. A. Halvorsen, E. L. Madsen & N. Jentoft (eds), *Evaluering: Tradisjoner, praksis, mangfold*, 180–194, Fagbokforlaget, Bergen.
- Goffman, Erving (1974). *Frame analysis: An essay on the Organization of Experience*, Harvard University Press.
- Haugaard, Mark (2008) Sociological Lukes versus moral Lukes: Reflections upon the second edition of *Power: A Radical View* by Steven Lukes, *Journal of power*, 1 (1): 99–106, DOI: 10.1080/17540290801943471.
- Hood, Christopher & Ruth Dixon (2015) *A Government that Worked Better and Cost Less? Evaluating Three Decades of Reform and Change in UK Central Government*, Oxford University Press, Oxford.
- House, Ernest R. (2017) Evaluation and the framing of race *American Journal of Evaluation* 38 (2): 167-189.
- Huberman, Michael (1987) Steps toward an integrated model of research utilization, *Knowledge*, 8 (4): 586–611.
- Højlund, Steven (2014) Evaluation use in the organizational context—changing focus to improve theory, *Evaluation*, 20 (1): 26–43.

- Høydal, Øyunn S. (2019) Evaluation influence: Method and societal phenomenon, *Evaluation* 25.3 (2019): 381-398.
- Høydal, Øyunn S. (2020) Worlds apart or coproduction? Investigating the relationship between knowledge users and producers, *Evaluation* 26.4 (2020): 404-421
- Innvær, Simon, Mari Trommald Gunn Vist & Andrew Oxman (2002) Health policy-makers' perceptions of their use of evidence: A systematic review, *Journal of Health Services Research & Policy*, 7 (4): 239–244.
- Lawn, Martin & Sotiria Grek (2012) *Europeanizing education: Governing a new policy space*, Symposium Books Ltd, London.
- Lee, Matthew & Laura Huang (2018) Gender bias, social impact framing, and evaluation of entrepreneurial ventures, *Organization Science*, 29 (1): 1–16.
- Lindgren, Lena (2014) *Nya utvärderingsmonstret: om kvalitetsmätning i den offentliga sektorn*, Studentlitteratur AB, Lund.
- Lukes, Steven (2004) *Power: A Radical View*, Macmillan International Higher Education.
- Marsdal, Magnus (2011) *Kunnskapsbløffen*, Forlaget Manifest, Oslo
- Matthes, Jörg (2009) What's in a frame? A content analysis of media framing studies in the world's leading communication journals, 1990–2005, *Journalism & Mass Communication Quarterly*, 86 (2): 349–367.
- McIntyre, Lynn (2020) Framing analysis: Its usefulness as a tool for advocacy on public health nutrition problems, *Public Health Nutrition*, 23 (11): 2053–2055, doi:10.1017/S1368980020000373
- Morris, Peter (2006) Steven Lukes on the concept of power, *Political Studies Review*, 4 (2): 124–135.
- Naturvernforbundet (2020) Førdefjorden - hva har skjedd? <https://naturvernforbundet.no/fordefjorden-hva-har-skjedd/fordefjorden-hva-har-skjedd-article33309-3699.html>
- Nordesjö, Kettil (2019) Made in Sweden: The translation of a European evaluation approach, *Evaluation*, 25 (2): 189–206.
- Norwegian Directorate for Education and Training (2011) Internasjonale studier om norsk skole [International study on Norwegian schooling], [https://www.udir.no/globalassets/filer/tall-og-forskning/rapporter/temanotat/internasjonale\\_studier\\_om\\_norsk\\_skole\\_tema\\_notat.pdf](https://www.udir.no/globalassets/filer/tall-og-forskning/rapporter/temanotat/internasjonale_studier_om_norsk_skole_tema_notat.pdf)
- OECD (2019) PISA results 2018, <https://www.oecd.org/pisa/publications/pisa-2018-results.htm>
- Oliver, Kathryn, Simon Innvar, Theo Lorenc, Jenny Woodman & James Thomas (2014) A systematic review of barriers to and facilitators of the use of evidence by policymakers, *BMC Health Services Research*, 14 (1): 2.
- Øverbye, Einar (2006) 'The Political Economy of Pensions: Politics, Policy Models and Outcomes in Europe' in Camila Arza & Martin Kohli (eds), *The Political Economy of Pension Reform: Politics, Policy Models and Outcomes in Europe*, Routledge,
- Park, Chisung & Jooha Lee (2020) Stakeholder framing, communicative interaction, and policy legitimacy: Anti-smoking policy in South Korea,

- Policy Sciences*, 53(4): 637-665. <https://doi.org/10.1007/s11077-020-09394-z>
- Pattyn, V., S. Van Voorst, E. Mastenbroek & C. Dunlop (2018) 'Policy Evaluation in Europe' in I. E. Ongaro & S. Van Thiel (eds), *The Palgrave Handbook of Public Administration and Management in Europe*, 577-593, Palgrave Macmillan, Hampshire.
- Pawson, Ray & Nicholas Tilley (1997) *Realistic Evaluation*, Sage, London.
- Power, Michael (1997) *The Audit Society: Rituals of Verification*, Oxford University Press, Oxford.
- Rein, Martin & Donald A. Schön (1994) *Frame Reflection*, BasicBooks, New York.
- Rein, Martin & Donald Schön (1996) Frame-critical policy analysis and frame-reflective policy practice, *Knowledge and Policy*, 9: 85-104.
- Røvik, Kjell A. (2007) *Trender og translasjoner. Ideer som former det 21. århundrets organisasjon*, Universitetsforlaget, Oslo.
- Scheufele, Dietram A. (1999) Framing as a theory of media effects, *Journal of Communication*, 49 (1): 103-122.
- Schwandt, Thomas A. (2009) 'Globalizing Influences on the Western Evaluation Imaginary' in I. K. Ryan & J. B. Cousins (eds), *The Sage Handbook of Educational Evaluation*, 19-36, SAGE Publications, Inc., Thousand Oaks, CA.
- Scott, James (1990) *Domination and the Arts of Resistance*, Yale University Press, New Haven.
- Sellar, Sam, and Bob Lingard (2013) The OECD and global governance in education, *Journal of education policy* 28.5: 710-725.
- Stone, Deborah A. (2012) *Policy Paradox: The Art of Political Decision Making*, W.W. Norton & Company, New York.
- Støkken, Anne Marie (2013), 'Evaluering, styring og kontroll av barneverninstitusjoner', in Anne Halvorsen, Einar Lier-Madsen & Nina Jentoft (eds), *Evaluering. Tradisjoner, praksis, mangfold*, 120-133, Fagbokforlaget, Bergen.
- The Guardian (2015) Conservationists slam plans to dump mining waste into Norwegian fjord  
<https://www.theguardian.com/environment/2014/nov/19/conservationists-slam-plans-to-dump-mining-waste-into-Norwegian-fjord>
- The Norwegian Agency for Public and Financial Management (DFØ) (2016, 28 September) *Bruk av evaluering i statlig styring* (The use of evaluations in public administration), Agenda Kaupang, R9392.
- Tornes, Kristin (2012) 'Evalueringstradisjoner' in I. K. Tornes (ed), *Evaluering i teori og praksis (Evaluation in Theory and Practice)*, 71-97, Fagbokforlaget, Bergen.
- Tornes, Kristin (2013) 'Om verdier og kriterier i evaluering' in A. Halvorsen, E. Lier-Madsen & N. Jentoft (eds), *Evaluering. Tradisjoner, praksis, mangfold (Evaluation. Traditions, Practice and Diversity)*, 28-42, Fagbokforlaget, Bergen.



- Van Hulst, Merlijn & Dvora Yanow (2016) From policy “frames” to “framing”: Theorizing a more dynamic, political approach, *American Review of Public Administration*, 46 (1): 92–112, DOI: 10.1177/0275074014533142.
- Vedung, Evert (2010) Four waves of evaluation diffusion, *Evaluation*, 16 (3): 263–277.
- Vieira, Alena (2020) The European Union’s ‘potential we’ between acceptance and contestation: Assessing the positioning of six eastern partnership countries, *Journal of Common Market Studies*. 1:19,
- Vo, Anne T. & Christina Christie (2015) ‘Advancing Research on Evaluation through the Study of Context’ in P. R. Brandon (ed), *Research on Evaluation: New Directions for Evaluation*, 148, 43–55.
- Weiss, Carol H. (1998) Have we learned anything new about the use of evaluation? *American Journal of Evaluation*, 19 (1): 21–33.
- Willems, Jannes J., Anna V. Kenyon, Liz Sharp & Astrid Molenveld (2020) How actors are (dis)integrating policy agendas for multi-functional blue and green infrastructure projects on the ground, *Journal of Environmental Policy & Planning*, 32 (20): 1908–1917, DOI:10.1080/1523908X.2020.1798750