

Kjetil Nordesjö,  
Department of Social Work,  
Malmö University, Sweden  
[kjetil.nordesjo@mau.se](mailto:kjetil.nordesjo@mau.se)

Mats Fred,  
Department of Political  
Science, Lund University,  
Sweden  
[mats.fred@svet.lu.se](mailto:mats.fred@svet.lu.se)

## Introduction

In recent decades we have been presented with evidence indicating a growing demand for, and reliance on, various forms of evaluative practices in the public sector around the world (Jacobsson, Pierre, and Sundström, 2019; Pollitt and Bouckaert, 2011; Vedung, 2010). European Union investments, programmes and projects have to comply with evaluation standards; national programmes to reform service delivery are piloted in order to assess whether they are ready to be rolled out nationwide; specific interventions for bespoke target groups are designed to be evaluated at a later stage; and nowadays most policy initiatives are measured and assessed with respect to performance and outcomes. This development has not only generated an increasing number of evaluation procedures (e.g. audits, revisions, performance frameworks, etc.), but also a more fundamental discursive spread of the phenomenon as it becomes embedded, normalised, and institutionalised across organisations, communities, and in everyday lived experience. Unsurprisingly, this change has been described as an evaluation and audit society (Dahler-Larsen, 2012; Power, 1997).

Even though evaluation, “the process of determining the merit, worth or value of something, or the product of that process” (Scriven, 1991:139), is often presented as a neutral, or even technocratic activity, it is pervaded with power relationships and embedded in tensions between stakeholders as well as between values, institutions and belief systems. Evaluation may be initiated in order to instigate specific organisational or behavioural changes. It may also be initiated as a smokescreen or a symbolic gesture, while also having unintended or constitutive effects on people and organisations (Dahler-Larsen, 2014). It is productive and may influence people and organisations at all stages of the evaluation process. Against this background, traditional definitions of evaluation, like Scriven’s cited above, are devoid of politics, history and context. However, evaluation is not contextless, but rather based on certain ideas located at a specific time and place in history and advanced by actors situated in a specific historical context – the power of evaluation is, as such, context dependent.

This is by no means a radical discovery. The idea that power and evaluation have a close relationship has been articulated before (e.g. Karlsson Vestman & Conner, 2006; Triantafillou, 2012). However, because power dynamics are such an integral part of evaluative practices, they are rarely made explicit or occupy a central position in research. Instead, power dynamics are obscured or are lost in broader discussions of governance or outcomes.

The aim of this special issue is to invite scholars in public administration to address issues of power in relation to evaluation practices in the public sector. Power is one of those big, difficult, and inherently nebulous concepts that both scholars and practitioners appear to consciously or unconsciously avoid (see Clegg et al., 2006). Our ambition in this introduction is not to present a comprehensive conceptual review of power, let alone to present the many different aspects of evaluation. Rather, this is a humble attempt to foster a more explicit discussion of the relationship between power and evaluation. The aim of this introduction is, firstly, to introduce a few basic distinctions and conceptualisations of the relationship between power and evaluation. Secondly, we would like to highlight three themes where power may be an especially fruitful lens when analysing, or just trying to make sense of, evaluation procedures. These are themes found in evaluation research, but where power is not always acknowledged or accounted for. We call them (1) *Evaluation and instrumental power*, (2) *Evaluation and contextual power*, and (3) *Evaluation and performative power*.

## Evaluation and Power

The concept of power resists simple characterisation and definition, and different typologies and conceptualisations depend on the field of study, philosophical tradition or historical epoch (Stickl Haugen and Chouinard, 2019). At a general level, one can view power as positive or negative. Although words such as manipulation, violence and domination are often associated with power, power is not necessarily constraining, negative or antagonistic – "...power can [also] be creative, empowering and positive..." (Clegg et al., 2006:2). An example of this dualistic character of power can be found in the literature on the public sector reforms labelled New Public Management (NPM), which, it has been suggested, puts forth a more output-oriented public sector (Parker and Bradley, 2004). As a result, performance measurements, audits and various forms of evaluation procedures have become increasingly prevalent. In terms of negative power, these procedures have been criticised for contributing to administrative burdens on people and organisations (Löfgren et al., 2021; Forssell and Ivarsson-Westberg, 2014), for bureaucratisation (Picciotto, 2016), and an alleged loss of qualitative social values in service delivery in favour of a quantitative focus on numbers (Porter, 2003). At the same time, it has been argued that these same procedures enhance local democratic participation (see Hanberger, 2006), reduce the distance between citizens and political decisions (Fitzgerald et al., 2016), and provide people with tools or data to compare and choose between public sector service providers. Power embedded in evaluation can thus be constraining and repressive, but also a source of creativity and productivity (Triantafillou, 2012:8). It should be noted that much of the critique of evaluation procedures has been in terms of negative power.

Another common conceptualisation of power is power-over, power-to, and power-with. *Power-over* "refers to an asymmetrical relation between two or more actors or groups of actors", *power-to* "consists in the ability of the actor herself to carry out certain specific outcomes", and *power-with* "consists in the ability of a group to act together in view of collective outcomes or goals"

(Pansardi and Bindi, 2021:51). In terms of evaluation, one might think of power-over as evaluation procedures imposed by an authority onto its subordinates, whereas power-to might refer to empowerment evaluation where stakeholders and participants receive tools and resources to evaluate themselves. Power-with could be related to various forms of participatory evaluations. Some scholars also add *power-from-within* to the triad, meaning an individual's awareness of their own capacities that motivate the action (Pansardi, 2011). This line of reasoning would suggest that recent developments of evaluation systems, where many actors are involved, may produce complex patterns of power dynamics consisting of parallel relations of power-over, power-to, power-with, and power-from-within.

A somewhat classical conceptualisation of power is to think of it in terms of different forms, faces, or dimensions. Lukes (1974) identified three dimensions of power where the first one was found in authorships such as Dahl (1957) and Wolfinger (1971), presenting a direct and observable form of power, similar to *power-over* and a negative form of power. This is an intentional and active form of power where “A has power over B to the extent that he can get B to do something that B would not otherwise do” (Dahl, 1957:202-203). The second dimension emphasises that power is also exercised when issues are arranged specifically so that some are kept out of the discussion (Bachrach and Baratz, 1970; Lukes, 1974:16). This involves agenda setting – which topics are discussed, by whom, and how decisions are taken. The third dimension of power is the “supreme and most insidious exercise of power” because it allows rulers to shape the preferences and perception of the masses as well as prevent them from having grievances (Lukes, 1974:23). In terms of evaluation, this third dimension draws attention to the way institutionalised evaluative practices have the power to shape various domains and aspects of individuals, organisations, and society. Here, evaluation practices become social practices that are defined, shaped, and carried out in a social context by actors with interests and values. As an illustrative example one might look at the transformation of “citizens” to “users” or “customers”, as described in the NPM and marketisation literature, where an evaluative way of thinking has played a transformative role (see Lindholm and Balle Hansen, 2020; Clarke et al., 2007). From this perspective, one might also argue that evaluation has contributed to a reconfiguration of democratic societies where evaluation does not need democracy anymore, but legitimises itself.

One of the main concerns with Lukes' conceptualisation of power has been that the dimensions treat power only in terms of domination and miss the productive aspects of power (Morriss, 2002; Haugaard, 2020). Inspired by Foucault, Haugaard (2012) proposes a re-theorisation of the three-dimensional power – with the potential to be emancipating – as well as a proposed fourth dimension. Power, according to Foucault, is *relational*, meaning that “it operates as a social relation of inequality between forces”, something that also brings forth the “possibility of resistance or counterdiscourse” (Cooper, 1994:437).

The fourth dimension of power as proposed by Haugaard (2012) is a more subtle form of power inspired by the *Panopticon* whereby prisoners are made visible in cells surrounding an all-observing eye, which allows them to become objects of knowledge (Foucault, 1991). They become “conscious of their own visibility and, as a consequence, start to observe themselves through the eye of

the observing other” (Haugaard, 2012:48). As such, they subject themselves to the normalising judgement of the observer. In terms of evaluation, one can think of this fourth dimension as the self-disciplinary actions of civil servants operating in a certain way because they are aware *that* (and often even *how*) their practices will be measured and assessed. In the literature this has been formulated as a performance paradox (Meyer and Gupta, 1994) or as “teaching to the test”, resulting in a weak correlation between performance indicators and performance itself and the production of unintended and various constitutive effects (van Thiel and Leeuw, 2002; Dahler-Larsen, 2012). When organisations and individuals learn how their performance is measured (and how it is not), they can put all effort into what is measured, and performance will go up!

To sum up, power can be regarded as something negative or positive, something constraining or emancipating, a capacity for action or as a relational phenomenon. We do not regard the different conceptions as mutually exclusive, but rather as concepts that highlight different views of what is studied and, in our case, help us understand different aspects of the evaluation phenomena. Based on this brief overview and our knowledge in evaluation research, we have identified three themes where power is relevant to understand evaluation procedures. These themes can be read as distinct themes with close relations to each other, but also as a chronology of how evaluation research has developed over the years. The first theme, *evaluation and instrumental power*, represents the most intuitive and dominant form of thinking about evaluation and power: how evaluation, as a product, is intended to improve our organisations and societies by informing decisions. The second theme, *evaluation and contextual power*, includes discussions of the power to define the preconditions in the context and processes of evaluation, such as problem formulation and stakeholder involvement. Finally, in the third theme, *evaluation and performative power*, we discuss the constitutive and performative aspects of power in evaluation, where evaluation ultimately shapes parts of social and organisational life by performing the logic of evaluation.

### Evaluation and Instrumental Power

When we think about the product or the outcome of evaluation procedures, power is easily associated with the evaluation’s ability to directly and openly change the direction of a policy or the course of action in an organisation. This view of power corresponds with a modern ideal of evaluation, where the result of evaluative practices, often in the form of a report, is intended to be used as a basis for decisions to improve the evaluand. In the literature, this is referred to as an *instrumental use* of evaluation – the idea that evaluations are designed on the basis of the needs and routines of an organisation and that the results are intentionally and rationally used by well informed and defined actors (Denvall and Nordesjö, 2021). Research on evaluation use has explored different factors for improving the instrumental use of evaluation as a product or process. For example, it is argued that the evaluation should be executed methodologically correctly, and the evaluator should identify and involve actors who are interested in using the results of the evaluation (Alkin and King 2017). As Patton (1996) argued, the evaluator should aim for intended use for intended users, thus prioritising the usability of the evaluation.

The instrumental model, as a legitimate form of power, rests on the idea of evaluation as a value-neutral and objective scientific procedure that, through evaluative information, guides the improvement of organisations and societies. Here, evaluation can be a form of productive power that supports rather than constrains development. However, the instrumental model has been criticised for many decades because the underlying rationale has proven difficult to validate in empirical research. The literature has instead argued that evaluation (and knowledge gained through research in general) more often is used conceptually, for enlightenment, or to enhance legitimacy (Smith, 2013; Weiss, 1979). Furthermore, ideology, interests, and feasibility are more often prioritised by decision makers than evaluation results (Shadish et al., 1991). Instead of being used at predefined stages in the policy process, knowledge makes discrete and incremental entries into the policy process, and the result of an evaluation might only be one out of several sources of knowledge used (Smith, 2013).

Scholars of evaluation find the instrumental model to be present in contemporary rationales underlying evaluation systems. These scholars relate the model to a linear and rational perspective on knowledge production that adjusts rather than questions an evaluated policy (Leeuw and Furubo, 2008). This perspective ignores more hidden and invisible issues of institutional norms, routines, and belief systems and focuses on the specific evaluation procedures rather than the organisations in which evaluation takes place (Raimondo, 2018). Also, the model builds on intentional and planned use while producing unintentional and constitutive effects in practice (Rijke, et al., 2016). To avoid this mismatch, Hanberger (2011) urges us to look more closely at evaluations' consequences in practice rather than their formal intentions.

Despite the criticism over the years, the instrumental model is found in contemporary ideas of evidence-based policy making and practice, a global movement proclaiming government activities to be based on the best available evidence (Kay, 2011) and the more hackneyed slogan "what matters is what works" (Cairney, 2016). Its proponents argue for the variety of approaches and methods, although there has been an emphasis on randomised controlled trials and systematic reviews as the gold standard (Nutley et al., 2007). This movement can also be associated with a specific science-driven perspective on evaluation (Vedung, 2010) aiming to enhance the role of knowledge produced through evaluation and social science and avoiding decisions based on ideology and prejudice (Stame, 2019).

To sum up, one of the most obvious forms of power in evaluation is power related to power-over, where the result of evaluation processes or products are assigned with the goal to influence action. Although this instrumental form of evaluation power has been criticised from several quarters, it appears to hold a rather firm grip on common ideas of evaluation procedures. If the model's legitimacy rests on the visible and direct capabilities of evaluation practices and systems, its illegitimacy would conversely be related to the neglect and obscuring of more hidden and invisible power dynamics. These play a more important part in our next theme.

## Evaluation and Contextual Power

The second theme focuses on the preconditions for, and context of, evaluation. Here, one can think of evaluation as an arena with boundaries and conditions that are possible to configure in order to affect the process or outcome of an evaluation. For example, Picciotto (2017:313) argues that when evaluation is captured by vested interests, it facilitates surveillance, conformity, and regulated action and defines boundaries and rules. However, when autonomous it can fulfil its public interest responsibility by resisting hegemonic discourse, exposing contradictions, and offering alternative narratives.

Highlighting the borders and conditions of evaluation opens up for a range of power issues within the context of evaluation. An actor who controls the allocated resources for funding the evaluation most likely has influence on the issues that are being investigated and the problems being addressed. For example, designing an evaluation from a gender-sensitive standpoint will draw attention to issues of gender equality. How policy problems are *framed* from different perspectives thus directs attention to evaluative solutions with different underlying rationales (House, 2017).

It is also a prerogative to decide who is to be involved in which phase of an evaluation process (i.e. problem formulation, data collection, analysis, or dissemination). Collaborative evaluation approaches have made this a central issue in order to improve evaluation use, democratic relevance and empowerment among stakeholders. At its broadest, these evaluation approaches suggest partnerships between evaluators and stakeholders in order to increase the relevance of an evaluation. Many of them acknowledge evaluation as a process of conflicting social, contextual, political, institutional, and cultural factors and how relational power dynamics between participants and evaluators can affect the evaluation process (cf. Stickl Haugen and Chouinard, 2019). For example, the literature on practical participatory evaluation argues for increased evaluation use by improving stakeholders' involvement in the ownership and decision making of the evaluation (Cousins and Whitmore, 1998). Empowerment evaluation is intended to help stakeholders gain control of resources, knowledge, and abilities to evaluate on their own (Fetterman, 2001). This draws attention to whether the role of the evaluator is to be a mentor and facilitator who supports self-appraisal or an external assessor of project goals. Also, Macdonald (1976) argues that evaluators influence the changing power relationships of politics, suggesting that a democratic evaluator "recognizes value pluralism and seeks to represent a range of interests in his issue formulation" (p. 134). This approach emphasises the question of the criteria with which evaluative conclusions are drawn. Depending on whether descriptive organisational goals or a prescriptive lens of democratic values such as transparency or equality is used, the evaluation will generate different results (cf. House and Howe, 1999). Similarly, transformative evaluation emphasises anti-discriminatory and social justice themes with a focus on the perspectives of marginalised groups, and critical theory evaluation addresses the macro-political dimension of social phenomena and scrutinises the mental models and ideologies that motivate stakeholders (Picciotto, 2017).

In sum, this theme highlights some of the preconditions for evaluation that affect the emerging knowledge of evaluation procedures. By directing attention

to these, it is possible to understand various power dynamics in evaluation. Although the configuration of these conditions may contribute to a bias towards certain groups of people or organisations, they may also give preference to specific groups and place certain issues on the agenda. In this way, the power of evaluation may not be to produce one truth, but give voice to a plurality of actors previously not heard. This contextual power might not primarily be about gaining power over conditions of evaluation. Rather, it is a way to create power within an evaluation and to give power to stakeholders.

### Evaluation and Performative Power

In our third theme, evaluation procedures are social practices that on the one hand are defined, shaped, and carried out in a social context by actors with interests and values, and on the other hand they shape perceptions, norms, and cognitions among actors, as well as various domains and aspects of organisations and society. This problematises the perception of evaluation as an objective and value-neutral activity and emphasises that evaluative practices should be no less subject to critical scrutiny than the evaluands they are set to investigate in order to understand their contemporary challenges and limits.

Understanding evaluation as a social practice draws attention to the institutionalisation of evaluation procedures. When evaluation is taken for granted as a mandatory and routine-based phenomenon, it obscures the influential actors and driving forces that shape the various aspects of an evaluation context. For example, the notion of an “evaluation society” is in part managed by powerful supranational organisations such as the OECD or the EU (who advocate, fund, and normalise specific forms or evaluative practices), but also managing consultants (pushing and selling the idea of organisational improvement through evaluation) and academic institutions (developing evaluation best practices and educating coming generations of pro-evaluation civil servants and consultants) (see Furubo, 2019). Here, studying how influential actors *translate* evaluative concepts highlights the adaptation of evaluation between institutional contexts (Elvbakken and Hansen 2018, Nordesjö, 2019).

The institutionalisation of evaluation brings constitutive and performative perspectives on evaluation to the fore. Here, evaluation can be understood in terms categorisation, surveillance, and self-monitoring of populations. Such concepts can help investigate how evaluation methods and practices are entangled in issues of social inequity (Rodriguez and Acree, 2021). Also, the sociology of quantification has contributed to understanding numbers in evaluation procedures as constituting the things they measure by “directing attention, persuading, and creating new categories for apprehending the world” (Espeland and Stevens, 2008:404). Standardising and structuring values, experiences, and activities allows them to travel geographically and culturally through the language of numbers. With the help of the legitimacy of scientific reasoning, numbers may be perceived as objective, universal, and impersonal (Porter, 2003).

Consequently, research on evaluation procedures has identified various unintended or constitutive effects that shape, constitute, and perform, rather than objectively represent, social and organisational practices (e.g. van Thiel and

Leeuw, 2002, Leeuw and Furubo, 2008, Rijke et al., 2016, Dahler-Larsen, 2014, Lindgren, 2014). For example, *constitutive effects* capture how measurement and quantification are not objective and simply valid representations of reality, but are constructed through procedures in a social and organisational context. The use of organisational procedures for evaluation and the language of evaluation can thus constitute what is central in work, the relationship between actors, relevant timeframes, and the overall worldview and framing of the problems and evaluative solutions (Dahler-Larsen, 2014). This highlights the *performativity* of evaluation procedures. An evaluation system does not primarily collect data that represents organisational activities. Instead, it productively has the power to shape the purpose of the organisation and its activities in order to adhere to, and *perform*, the logic of the evaluation system.

Indeed, the quantitative and all-encompassing nature of evaluation systems is a relevant object of study when investigating the performative power of evaluation. Such systems collect and analyse streams of continuously flowing data (Rist and Stame 2006). They are permanent and thus part of something ongoing, rather than an ad hoc or a one-time evaluation (Leeuw and Furubo, 2008). The power of such systems has been shown to be far reaching and of various forms. As evaluation systems have become larger, more self-propelling, and able to handle more data, their capacity to shape their surroundings has increased. In this sense, evaluation systems can be said to have adopted almost machine-like characteristics as robotic, predictable, reliable, and automatic systems without human subjectivity, interpretation, or emotion (Dahler-Larsen, 2012). The acceleration of technology supporting such systems also facilitates the development, use, and reproduction of the same systems. Evaluation systems may therefore seem to breed evaluation systems (Segerholm, 2020). The continuous expansion of and subsequent research interest in evaluation systems is reflected in this special issue, where several contributions draw empirically on various aspects of evaluation systems.

To sum up, in our last theme evaluation is a social, constitutive, and performative practice not only shaped by a powerful social and organisational context, but also with the power to shape the very same contexts.

The three themes described here emphasize different ways of thinking about evaluation and power. First, we discussed the intuitive notion of power in evaluation as influencing decisions, which, despite critique, is a strong theme in contemporary society and embedded in the idea of evidence-based policy and practice. The second theme draws attention to power dynamics related to the conditions of the evaluation context. Third, the very fact that evaluation, for example in the shape of evaluation systems, is initiated and put in place, may work as a performative power, and give rise to consequences for various dimensions of social and organizational life. The contributions in this special issue discuss these themes in different ways.

## The Contributions

Through six articles (and six reflections – see below) the goal of this special issue is to contribute to a more explicit discussion of power in relation to the various forms and functions of evaluation, both in research and in practice. We

have not asked the authors to define or relate to power in any certain way. As such, each paper stands on its own. This has led to a diverse set of contributions from authors based in Denmark, Norway, Sweden, and New Zealand, representing different empirical areas with different entry points to power and evaluation. Although some contributions relate to all of the above themes, they all broadly relate to the third theme – the study of evaluation and performative power – and how evaluation can shape (and be shaped by) different parts of an institutionalised evaluation context. When read together, we see how different forms of power are integrated in evaluation practices in public administration, and the implications for how public organisations are funded, organized, and assessed.

The first three articles in this special issue discuss evaluation systems. Peter Dahler-Larsen describes how the *circulation* of evaluative information has become an important factor in how constitutive effects of evaluation are produced. Through the cases of Google Scholar and workplace assessments, he suggests different useful analytical entry points for the study of circulation.

Next, Niklas Andreas Andersen applies the concept of a “contestability differential” to argue that the institutionalisation of an asymmetric power relation between the evaluation system and evaluand creates inherent paradoxes. By investigating an evaluation system within the Danish employment services, he shows how evaluation systems are simultaneously increasing and decreasing the power of evaluations.

The third article on evaluation systems is by Malin Benerdal and Magnus Larsson. They investigate the compatibility of democratic values and the rise of evaluation systems. Drawing on data from school principals’ conceptions of systematic quality work within the Swedish school system, they show how different steering logics affect the integration of democratic values within the evaluation system.

Continuing on the theme of democracy, Karl Löfgren and Annika Agger ask to what extent the increase in citizen-oriented governance practices, and interest in frameworks for evaluating democratic quality, are enhancing the quality of democracy. They seek to identify the underlying rationales behind the evaluation of democracy in a local context by investigating seven evaluative frameworks for the evaluation of democracy.

Next, Jostein Askim, Erik Døving, and Åge Johnsen shed light on the institutionalisation of evaluation. They describe and analyse the Norwegian government’s evaluation practice over a 25-year period by investigating the volume of evaluations, the most active commissioners and providers of evaluations, and the types of evaluations conducted. Among several results, they conclude that the number of evaluations has decreased and that consultants’ share of the performed evaluations has increased.

Finally, Øyunn Syrstad Høydal expands on the concept of framing and its relationship to evaluation and power. It is argued that framing in evaluation involves the power to influence public knowledge production and policy decisions by clarifying perspectives on societal problems and solutions. She uses the Norwegian participation in PISA to demonstrate how framing is not only essential to understanding evaluation, but also contributes to a more power-sensitive use of the framing perspective.

In addition to the six articles, this special issue also includes six reflection pieces – short texts written by “practitioners” experienced in various roles of evaluation in public administration. The idea has been to let them reflect on and contextualise each article from a practice perspective. We write “practitioners” within quotation marks because the line between researchers and practitioners often is blurred, not least in this case and within the field of evaluation in general. Together with the articles and the introduction, the reflection pieces can contribute to an ongoing dialogue on the way in which evaluation as a powerful social phenomenon has implications for organisations and societies today.

Lastly, we would like to extend our gratitude to the authors, “practitioners”, the many anonymous reviewers, and the editorial team of the *Scandinavian Journal of Public Administration* for collectively putting this special issue together. Now, on to the contributions.

## References

- 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): 434-50. <https://doi.org/10.1177/1098214017717015>
- Bachrach, Peter & Morton S. Baratz (1970) *Power and poverty: theory and practice*. New York: Oxford University Press.
- Cairney, Paul (2016) *The politics of evidence-based policy making*. London: Palgrave Springer.
- Clarke, John, Janet Newman, Nick Smith, Elizabeth Vidler & Louise Westmarland (2007) *Creating citizen-consumers: Changing publics and changing public services*. London: SAGE Publications Ltd.
- Clegg, Stewart R., David Courpasson & Nelson Phillips (2006) *Power and organization*. Los Angeles: SAGE.
- Cooper, Davina (1994) Productice, relational and everywhere? Conceptualizing power and resistance within Foucauldian. *Sociology*, 28 (2): 435-454.
- Cousins, J. Bradley & Elizabeth Whitmore (1998) Framing Participatory Evaluation. *New Directions for Evaluation*, 80: 5-23. <http://dx.doi.org/10.1002/ev.1114>
- Dahl, Robert A. (1957) The concept of power. *Behavioral Science*, 20: 201–215.
- Dahler-Larsen, Peter (2012) *The Evaluation Society*. Stanford, Calif.: Stanford Business Books.
- Dahler-Larsen, Peter (2014) Constitutive Effects of Performance Indicators: Getting beyond unintended consequences. *Public Management Review*, 16 (7): 969-986. doi:10.1080/14719037.2013.770058
- Denvall, Verner & Kettil Nordesjö (2021) Bortom det instrumentella idealet – om kunskapsanvändning i utvärderingssystem. In Stig Linde & Kerstin Svensson (eds), *Välfärdens aktörer - Utmaningar för människor, professioner och organisationer*. School of social work, Lund: Social Work Press.
- Elvbakken, Kari Tove & Hanne Foss Hansen (2018) Evidence producing organizations: Organizational translation of travelling evaluation ideas. *Evaluation*, 25 (3): 261-276. <https://doi.org/10.1177/1356389018803965>

- Espeland, Wendy Nelson & Mitchell L. Stevens (2008) A sociology of quantification. *European Journal of Sociology/Archives Européennes de Sociologie*, 49 (3): 401-436.
- Fetterman, David M. (2001) *Foundations of Empowerment Evaluation: Step by Step*. Thousand Oaks, Calif.: SAGE.
- Fitzgerald, Ciara, Stephen McCarthy, Fergal Carton, Yvonne O Connor, Laura Lynch, Frederic Adam (2016) Citizen participation in decision-making: can one make a difference? *Journal of Decision Systems*, 25 (sup1): 248-260.
- Forssell, Anders & Anders Ivarsson Westerberg (2014) *Administrationssamhället*. Lund: Studentlitteratur.
- Foucault, Michel (1991) Questions of Method. In Graham Burchell, Colin Gordon & Peter Miller (eds) *The Foucault Effect*. London: Harvester Wheatsheaf
- Furubo, Jan-Eric (2019) Understanding the evaluation enterprise. In Jan-Eric Furubo & Nicoletta Stame (eds), *The evaluation enterprise - A critical view*. New York: Routledge.
- Hanberger, Anders (2006) Evaluation of and for Democracy. *Evaluation*, 12 (1): 17-37.
- Hanberger, Anders (2011) The real functions of evaluation and response systems. *Evaluation*, 17 (4): 327-349.
- Haugaard, Mark (2020) *The four dimensions of power. Understanding domination, empowerment and democracy*. Manchester: Manchester University Press.
- Haugaard, Mark (2012) Rethinking the four dimensions of power: domination and empowerment. *Journal of Political Power*, 5 (1): 33-54. doi:10.1080/2158379X.2012.660810.
- House, Ernest R. (2017) Evaluation and the framing of race. *American Journal of Evaluation*, 38 (2): 167-189.
- House, Ernest R. & Kenneth. R. Howe (1999). *Values in Evaluation and Social Research*. Thousand Oaks: Sage.
- Jacobsson, Bengt, Jon Pierre, & Göran Sundström (2019) *Granskningsamhället: offentliga verksamheter under lupp*. Lund: Studentlitteratur.
- Karlsson Vestman, Ove & Ross F. Conner (2006) The Relationship between Evaluation and Politics. In I. Shaw, J. C. Greene, & M. M. Mark (Eds.), *Handbook of evaluation: policies, programs and practices*. Thousand Oaks, Calif.: Sage.
- Kay, Adrian (2011) Evidence-based policy-making: The elusive search for rational public administration. *Australian Journal of Public Administration*, 70 (3): 236-245.
- Leeuw, Frans L. & Jan-Eric Furubo (2008) Evaluation Systems. *Evaluation*, 14 (2): 157-169.
- Lindblom, Charles E. (1979) Still Muddling, Not Yet Through. *Public administration review*, 39 (6): 517-526.
- Lindgren, Lena (2014) *Nya utvärderingsmonstret: om kvalitetsmätning i den offentliga sektorn*. Lund: Studentlitteratur.

- Lindholst, Andrej Christian & Morten Balle Hansen (2020) *Marketization in local government diffusion and evolution in Scandinavia and England*. Cham: Palgrave Macmillan
- Löfgren, Karl, Ben Darrah-Morgan, Patrik Hall, Linda Alamaa (2021) The Rise of a New Public Bureaucracy in New Zealand? *Administration & Society*. Online.
- Lukes, Steven (1974) *Power: A Radical View*. London: Macmillan.
- MacDonald, Barry (1976) Evaluation and the Control of Education. In David Tawney (ed), *Curriculum Evaluation Today: Trends and Implications: A Second Collection of Papers from Members of the Schools Council Project Evaluators' Group on Aspects of Their Work*. London: Macmillan
- Meyer, Marshall W. & Vipin Gupta (1994) The performance paradox. *Research in Organizational Behavior*, 16, 309-369.
- Morriss, Peter (2002) *Power: a philosophical analysis. 2nd ed.* Manchester: Manchester University Press.
- Nordesjö, Kettil (2019) Made in Sweden: The translation of a European evaluation approach. *Evaluation*, 25 (2): 189-206. <https://doi.org/10.1177/1356389018803220>
- Nutley, Sandra M, Isabel Walter, & Huw T.O. Davies (2007) *Using evidence: how research can inform public services*. Bristol: Policy Press.
- Pansardi, Pamela (2011). Power to and power over. In: Keith Dowding (ed) *Encyclopedia of power*. Beverly Hills, CA: Sage, 521–524.
- Pansardi, Pamela & Marianna Bindi (2021) The new concepts of power? Power-over, power-to and power-with. *Journal of Political Power*, 14 (1): 51-71. doi:10.1080/2158379X.2021.1877001
- Parker, Rachel & Lisa Bradley (2014) Bureaucracy or Post-Bureaucracy? Public Sector Organisations in a Changing Context. *Asia Pacific Journal of Public Administration*, 26 (2):197-215.
- Patton, Michael Q. (1996) *Utilization-Focused Evaluation: The New Century Text 3rd edition*. Thousand Oaks, Calif.: Sage Publications.
- Picciotto, Robert. (2016) Evaluation and bureaucracy: the tricky rectangle. *Evaluation*, 22 (4): 424-434. doi:10.1177/1356389016657934
- Picciotto, Robert. (2017) Evaluation: Discursive practice or communicative action? *Evaluation*, 23 (3): 312-322. doi:10.1177/1356389017714384
- Pollitt, Christopher & Geert Bouckaert (2011) *Public management reform: a comparative analysis: new public management, governance, and the neo-Weberian state*. Oxford: Oxford University Press
- Porter, Theodore M. (2003) Measurement, Objectivity, and Trust. *Measurement: Interdisciplinary Research and Perspectives*, 1 (4): 241-255. doi:10.1207/S15366359MEA0104\_1
- Power, Michael (1997) *The audit society: rituals of verification*. Oxford: Oxford University Press.
- Raimondo, Estelle (2018) The power and dysfunctions of evaluation systems in international organizations. *Evaluation*, 24 (1): 26-41. doi:10.1177/1356389017749068

- Rijke, Sandra de, Paul F. Wouters, Alex D. Rushforth, Thomas P. Franssen, & Björn Hammarfelt (2016) Evaluation practices and effects of indicator use—a literature review. *Research Evaluation*, 25 (2): 161-169.
- Rist, Ray C. & Nicoletta Stame (2006) *From studies to streams: managing evaluative systems*. New Brunswick, N.J.: Transaction Publishers.
- Rodriguez, Sophia, & Jeremy Acree (2021) Discourses of Belonging: Intersections of Truth, Power, and Ethics in Evaluation With Migrant Youth. *American Journal of Evaluation*, 42 (2): 276-292. doi:10.1177/1098214020963837
- Scriven, Michael (1991) *Evaluation Thesaurus*. Newbury Park, Calif.: Sage.
- Segerholm, Christina (2020). Evaluation systems and the pace of change – The example of Swedish higher education. *Educational Philosophy and Theory*, 52 (6): 613-624. doi:10.1080/00131857.2019.1654372
- Shadish, William R., Thomas D. Cook & Laura C. Leviton (1991) *Foundations of Program Evaluation: Theories of Practice*. Newbury Park, Calif.: Sage.
- Smith, Katherine (2013) *Beyond evidence-based policy in public health: The interplay of ideas*. London: Palgrave Macmillan.
- Stame, Nicoletta (2019) The politics/administration dichotomy and evaluation. In Jan-Eric Furubo & Nicoletta Stame (eds), *The evaluation enterprise: a critical view*. New York: Routledge.
- Stickl Haugen, Jamie & Jill Anne Chouinard (2018) Transparent, Translucent, Opaque: Exploring the Dimensions of Power in Culturally Responsive Evaluation Contexts. *American Journal of Evaluation*, 40 (3): 376-394. doi:10.1177/1098214018796342
- Triantafillou, Peter (2012) *New forms of governing: A foucauldian inspired analysis*. London: Palgrave Macmillan.
- van Thiel, Sandra and Frans L. Leeuw (2002) ‘The Performance Paradox in the Public Sector’, *Public Performance & Management Review*, 25 (3): 267–281.
- Vedung, Evert (2010). Four Waves of Evaluation Diffusion. *Evaluation*, 16 (3): 263-277.
- Weiss, Carol H. (1979) The many meanings of research utilization. *Public administration review*, 39 (5): 426-431.
- Weiss, Carol H. (1980) Knowledge creep and decision accretion. *Knowledge*, 1 (3): 381-404.
- Wolfinger, Raymond E. (1971) Nondecisions and the study of local politics. *American Political Science Review*, 65 (4): 1063-1080.