The constitutive effects of social investment evaluation

Kettil Nordesjö
Malmö University, Sweden

Abstract
Evaluation has different uses and impacts. This article aims to describe and analyse the constitutive effects – how evaluation forms and shapes different domains in an evaluation context – of evaluation related to an evidence-based policy and practice, by investigating how the evaluation of social investment fund initiatives in three Swedish municipalities is organized and implemented. Through interviews and evaluation reports, the findings show how this way of evaluating may contribute to constitutive effects by defining worldviews, content, timeframes and evaluation roles. The article discusses how social investment fund evaluation contributes to a linear knowledge transfer model, promotes a relation between costs and evidence and concentrates the power over evaluation at the top of organizations.

Keywords
constitutive effects, EBPP, evaluation use, social investment

Introduction
How social interventions are evaluated, impact and influence individuals and organizations. Within the vast literature on evaluation use, evaluation may contribute to decision-making and accountability, legitimacy of actions taken, knowledge of the quality and results of public services, transparency, democratic dialogue, empowerment and organizational learning. More questionable uses of evaluation have come to include ritual use, in which conducting evaluation shows others that you take action and work in a modern organization (Krogstrup and Dahler-Larsen, 2000). Evaluation systems and performance measurement have also been said to have unintended consequences such as hindering innovation, sub optimization and goal displacement (e.g. Rijcke et al., 2016; Thiel and Leeuw, 2002).

In order to move beyond the intended and unintended use and consequences of evaluation, Dahler-Larsen (2012a, 2014) has proposed the term constitutive effects to capture the way in
which way tests, measurements and indicators help define and constitute the social realities of which they are a part. It directs our attention towards the formative effects of evaluation and relates problems of evaluation use to evaluation itself, rather than to implementation failures outside of evaluation (cf. Furubo, 2019). It underscores that every theory of evaluation has underlying assumptions of how evaluation is supposed to change social programmes and contribute to social change (Shadish et al., 1991). It also directs attention to the notion that the language of evaluation shapes perceptions and affects mutual understanding, which is rarely neutral and requires context in order to be clear about its meaning (Patton, 2000). An implication is that evaluations and their impact may exceed their formal intentions and must be understood in a social, organizational and political context (Dahler-Larsen, 2012b; Schwandt, 2015).

An example of a methodological programme aimed at increasing the utilization of evaluation for policy-making and practice is evidence-based policy and practice (EBPP) (and more recently, evidence informed policy and practice, EIPP, see Boaz and Davies, 2019). Today, it has been integrated into the governance of the welfare sector in the Nordic countries and elsewhere, through evidence-based policy acts and legislation (Jacobsson and Meeuwisse, 2020) evidence-producing organizations (Elvbakken and Hansen, 2018) and standardized assessment tools (Skillmark, 2018). It resonates with government ambitions to govern top-down through standardized knowledge (Martin and Williams, 2019) and have often been united by the emphasis on objective knowledge about ‘what works’ using outcome evaluation based on quantitative data. From this perspective, evaluation within EBPP has been intended to be distanced from and not starting from the viewpoint of professionals, in order to provide trustworthy, objective and scientific knowledge for decision-makers. Overall, this is a far cry from the original bottom-up ‘critical appraisal’ idea of evidence-based practice, in which research and evaluation, professional knowledge and service users’ conceptions are assessed jointly (e.g. Sackett, 1997; Trinder and Reynolds, 2000).

Although there has been an abundance of research on the use of evaluation within EBPP (cf. Boaz and Nutley, 2019), less is known about how evaluation aligned with EBPP – which principally aspires to be objective and separate from the evaluand – forms and shapes different domains in an evaluation context. It is important to explore such consequences because of the continuous emphasis on evidence and outcome evaluation in the welfare sector, globally (cf. Mosley and Smith, 2018). It may improve the understanding of how evaluation in EBPP shapes aspects of an evaluation context beyond the intentions of the intended instrumental knowledge use. Also, in welfare organizations that have to deal with social and wicked problems, in which goals may be vague and contradictory, and it is difficult to achieve consensus on how to operationalize and measure the success of solutions (Hasenfeld, 2010), consequences of evaluation may have implications for how to define the success and long-term goals of welfare interventions and social policy at large.

The empirical case is social investment fund evaluation in Swedish local welfare government. Social investment funds (SI funds) are a materialization of the social investment perspective, which rests on policies that both invest in human capital development and help make efficient use of human capital, while fostering greater social inclusion. Social policies are a productive factor, essential to economic development and employment growth (Morel et al., 2012). In Sweden, SI funds is an answer to a lack of innovation and cooperation between local governments’ departments and a way to work preventively and long term despite short-term budgetary practices (cf. Jonsson and Johansson, 2018). SI funds fund temporary collaborative projects in order to achieve social outcomes and lower future spending. Monitoring, outcome
evaluation and cost calculations are three evaluation types intended to provide evidence for what kind of interventions work and support decision-making.

The aim of this article is to describe and analyse the constitutive effects of evaluation related to EBPP by investigating how evaluation in social investment funds is organized and conducted in three Swedish municipalities. The research questions are as follows:

1. What are the constitutive effects of social investment fund evaluation and in what domains can they be found?
2. What is the significance of the identified constitutive effects for evaluation in EBPP?

The paper is structured in six parts. After this introduction, the literature on evaluation use relevant to EBPP is discussed, followed by the concept of constitutive effects. The case of SI fund evaluation and the qualitative methods used is then described. Finally, the results are presented, and a discussion and conclusion.

**Evaluation use in EBPP**

The study of the formative effects of evaluation can be derived from research on evaluation use, one of the most explored areas of evaluation research. It was often thought that evaluation was used instrumentally and decision-makers implemented the evaluation results, although it was more often used conceptually, for enlightenment or for legitimacy (Weiss, 1979; Weiss, 1977). Instead, ideology, interests and feasibility are more often prioritized (Shadish et al., 1991) and knowledge often makes discrete and incremental entries into the policy process (Lindblom, 1979; Weiss, 1980). Tactical, symbolic, ritual, conceptual, process and anticipat-ing use have since described different outcomes of an evaluation process or product (e.g. Vedung, 2015). An often-debated question concerns the role of the evaluator in the promotion of evaluation use (Alkin et al., 1990; Donaldson et al., 2010). On the one hand, this is not the role of the evaluator since an evaluation should be external, impartial and objective knowledge among other information in policymaking. On the other hand, evaluators can greatly increase the use by identifying the intended users and working with them.

Furthermore, evaluation influence (Henry and Mark, 2003; Kirkhart, 2000) has conceptual-ized how evaluation could affect levels and contexts outside a delimited decision-making con-text. The movement and streams of data between levels and actors became significant through the evolution of evaluation systems (Rist and Stame, 2006), which have been shown to have different functions and consequences (Hanberger, 2011; Hanberger and Lindgren, 2019). The literature on performance measurement has shown how the creation of structures for assessing performance through indicators has led to intended and unintended consequences (Lindgren, 2014; Pollitt, 2013; Sanderson, 2001; Thiel and Leeuw, 2002). The concept of constitutive effects (Dahler-Larsen, 2012a, 2014) emphasizes the formative power of evaluation similar to other literatures, such as the sociology of quantification (Espeland and Stevens, 2008; Porter, 2003), standardization (Timmermans and Epstein, 2010) calculative practices (Miller, 2001) and auditing (Power, 1997). These literatures describe how the use of criteria, numbers and standards is a constitutive feature of modern science and social organization; they deconstruct the often-presumed neutral and objective function of measurement in contested areas of organization. Empirical research on constitutive effects has primarily concerned evaluation systems. Dahler-Larsen (2012a) illustrates the concept through a school testing system that creates a ‘strategic landscape’ in which teachers are locked-in and pragmatically deal with tensions
of measurement. Hanberger and Lindgren (2019) found that evaluation systems in elderly care create different notions of quality, reinforcing certain roles, responsibilities and norms. Andersen (2019) investigates how an evidence-based evaluation system in labour market policy shapes policymaking in regard to a certain understanding of policy content, timeframes and expertise.

Against this background, constitutive effects related to evidence-based evaluation may be a relevant area of investigation due to the focus on instrumental use. Proponents of the ‘evidence wave’ have argued that government activities should be based on ‘what works’, that is, evidence (Vedung, 2010). Although this wave currently encompasses a variety of approaches and methods, there has been an emphasis on experimental and quantitative outcome evaluation as the ‘golden standard’. That certain methods and approaches bring more evidence to decision-makers than others has spurred an intense debate on what counts as credible evidence (Donaldson et al., 2009, 2014). The demand for rigour and experimental evaluation relates the ‘evidence wave’ to a science-driven perspective on evaluation (Vedung, 2010). An aim is to enhance the role of evaluation and social science and contribute to the avoidance of decisions based on ideology and prejudice (Stame, 2019).

With regard to knowledge use, it is possible to discern different generations of EBPP. The first generation relied on a linear model of knowledge use in which evaluation and research generated evidence as a product, disseminated to decision-makers for immediate and instrumental use (Best and Holmes, 2010, cf. Boaz and Nutley, 2019). The model is appealing from a top-down perspective in which powerful governing bodies can support the organizing and synthesizing of evidence in reviews, diffusing recommendations for standardized practice that can be monitored and evaluated. However, the model has met with increased criticism when faced with the realities of politics and professional practice (Martin and Williams, 2019). As Cairney (2016) argues, evidence becomes relevant to decision-makers only when beliefs and values are used to make sense of that information in relation to policy problems and solutions. Thus, the linear top-down model builds on idealized views of policy cycles and comprehensive rationality that attempt to identify the right time and place to infuse policy with evidence. The separation between evaluation and decision-makers also raises the question of whether or not decisions will still be political, or de-politicized, de-democratized and technocratic (Hansen and Rieper, 2010). For professionals, the model may contribute to a narrow, mechanistic, unreflective practice without room for professional adaptation (Martin and Williams, 2019). Part of the criticism has been the insufficient theoretical underpinnings of EBPP implementation in public administration, which has often ignored the rich literature on policy implementation and, consequently, questions of power and street-level behaviour (Johansson, 2010). Still, EBPP has also come to operate within bottom up-approaches operate within relational and system-wide models of knowledge use, which are more inclusive of different forms of evidence and how such evidence is produced (Best and Holmes, 2010). A relational model emphasizes exchanges between actors in research, policy and practice in a wide political context, while a system-wide approach draws on complexity and adaptation in evidence ‘eco-systems’. These kinds of approaches have gained traction but struggle with questions of power assymetries and symbolic participation and may therefore engender a distrust of evaluation (Martin and Williams, 2019).

In light of the influence of the evidence wave and the criticism of top-down EBPP, this study could contribute to exploring how such evaluation shapes and constitutes certain domains of the evaluation context in which it is conducted. In the next section, we turn to these domains of constitutive effects.
Constitutive effects as an analytical concept

Dahler-Larsen’s concept of constitutive effects is used analytically in this article to explore the formative consequences of SI fund evaluation. Measurement, quantification and standardization are not objective, evident and simply valid representations of reality, but are constructed through procedures in a social and organizational context. Measurement have a constitutive relation to the reality it seeks to describe and claims to measure and can only acquire meaning through contexts of use (Dahler-Larsen, 2014). This is in contrast to indicator ‘misuse’ or ‘unintended effects’, which imply that there is a ‘correct’ way of using indicators and that a deviation constitutes perversion (Rijcke et al., 2016).

Indicators are only one part of the evaluation of SI funds. In this article, the concept of constitutive effects is used broadly to cover three interrelated evaluation practices (monitoring, outcome evaluation and cost-benefit analysis) within SI funds under the term SI fund evaluation.

Two factors that shape constitutive effects are organizational procedures and language. Organizational procedures for evaluation such as selection, quantification, retention and forms for presentation of quantitative data make it possible to handle ambiguity and present an orderly view of reality. Evaluation also provides a language in which certain parts of an intervention such as outcomes and indicators are transformed into things and acquire a meaning (Dahler-Larsen, 2014). Four conceptual domains describe areas in which constitutive effects can be explored (Dahler-Larsen, 2012a, 2014).

1. Evaluation defines interpretive frames and worldviews. Evaluation is implicitly or explicitly framed as a solution to a problem with identified causes and normative rationales.
2. Evaluation defines content. It defines what is central in work (processes, skills, outcomes, costs) and determines what the actors should strive to accomplish.
3. Evaluation defines timeframes within which certain results are expected. Pressure to show results may direct attention towards certain criteria while other criteria are neglected.
4. Evaluation defines social relations and identities. Evaluation may create or redefine the identities and roles of professionals, evaluators and organizations.

In this article, the four domains are used as analytical themes to identify and discuss the effects of SI fund evaluation. The term effect refers to the analysis of interviewees’ conceptions of consequences of SI fund evaluation relating to the four domains, and how this is manifested in evaluation reports. The causality between SI fund evaluation and the domains has not been investigated.

Case and methods

In order to describe and analyse the constitutive effects of evaluation in EBPP, the implementation of SI fund evaluation was studies in three Swedish municipalities.

The SI perspective and SI funds

Internationally, the SI perspective has materialized in practices where private capital and entrepreneurs contribute to and profit from solving complex societal challenges by meeting
predetermined targets. For example, through social impact bonds (SIBs), private actors finance welfare interventions and take the risk but also have the possibility to make a profit if goals are met in a logic of ‘payment by results’ (Berndt and Wirth, 2018; Jonsson and Johansson, 2018). Consequently, the SI debate in countries with SIBs such as the United Kingdom has revolved around how to structure legislation, taxes and financing in order to facilitate investments (Balkfors et al., 2019).

In Sweden, the SI perspective has been less about how to finance welfare through the private sector and more a municipal answer to a lack of innovation and public sector silos that complicates complex problem solving that demands the participation of several sectors and levels. Also, the SI perspective seeks to draw attention to rigid short-term budgetary practices in municipalities that hinders preventive and long-term welfare work that is more cost effective (cf. Hultkrantz, 2015; Jonsson and Johansson, 2018). Instead, the SI perspective in Sweden emphasizes coordination of different local bodies and a focus on preventive interventions – act today to lower the costs tomorrow – and thus combines elements of control and efficiency with coordination and trust (Balkfors et al., 2019). It thus relates to the recent Swedish debate on trust-based governance where professional discretion is emphasized to handle the effects of new public management (e.g. SOU, 2019).

More specifically, SI in Sweden is defined by The Swedish Association of Local Authorities and Regions (SALAR), an interest group that has been instrumental to the development of SI in Sweden, as ‘a limited investment which, in relation to the usual working methods, is expected to give a better outcome for the target group and at the same time lead to reduced socio-economic costs in the long run’. (Uph, 2017: 6). This definition includes different SI initiatives, such as SI funds, that is, financial resources that are set aside for certain projects. They exist primarily on a municipal level and are characterized as being (1) implemented through temporary projects, (2) separate from day-to-day public administration, (3) a collaborative effort, (4) evidence- or knowledge-based and possible to (5) monitor and assess from a societal as well as a financial perspective (Uph, 2015). SI funds are mainly carried out on a local level where half of Sweden’s 290 municipalities say they work with social investments in some way (Balkfors et al., 2019).

In this context, SI fund evaluation is given the role to determine whether evidence- or knowledge-based interventions are successful or not and connect them to different cost scenarios. Evaluation then work as a basis for decisions for managers or politicians from two aspects: identifying a choice between specific measures and to identifying these measures’ impact (including costs) for citizens’ wellbeing (cf. Hultkrantz, 2015). However, evaluation is regularly referred to as an area in which municipalities need support structures (e.g. Balkfors et al., 2019; Hultkrantz, 2015). A recent governmentally funded report on the state of SI in Sweden recommends supporting a ‘culture of outcomes’ in public administration and the establishment of a national competence centre to deal with the lack of evaluation skills and evidence-based methods (Balkfors et al., 2019).

Against this background, SI fund evaluation is seen as a critical case of constitutive effects in EBPP, where constitutive effects are likely to occur (cf. Flyvbjerg, 2011). One reason is that decision-making in human services depends as much on moral and ideology as evaluation and research, since it is often difficult to reach consensus on how to operationalize and measure the success of solutions, making human service organizations dependent on societal and political legitimacy (cf. Hasenfeld, 2010). As SI fund evaluation has formally been assigned a decisive role in decision-making in local government, the built-in conflicts make it a case in
which constitutive effects could arise. Furthermore, SI fund evaluation appear to be formative as it seeks to capture the future based on calculations of future costs and benefits for the individual and society.

SI fund evaluation in three municipalities

Three Swedish municipalities (pop. 70,000–150,000) were investigated in a multi-site case study (see Table 1). They had several completed SI fund projects and evaluations and thus had evaluations publicly available and interviewees that could reflect on the evaluation processes. They were also described as at the forefront of SI funds in Sweden but were of different size and with different refunding systems. In particular, M2 stands out as the largest SI fund organization with the largest projects and consequently more people to interview. Another difference is that M1 had external evaluators and thus a smaller organization, while M2 and M3 performed evaluations internally. Otherwise, the municipalities’ SI funds were similarly organized. They set aside financial resources in a fund that finances projects related to one or more specialized departments (e.g. social services). SI staff handle the SI project applications that have been submitted by these departments. An application should describe the social problem to be dealt with, the available evidence or research about the problem and the proposed intervention, the target groups, the intervention itself, the evaluation exercises, the costs and anticipated refunds, and the future implementation. The elected politicians in the municipality ultimately decide on what projects to initiate. A feature of the SI funds is the refund system, in which the projects refund parts of what they had initially received from the fund. In M2, projects refund an amount corresponding to the calculated cost reduction; in M3, refunds only apply if there are evident cost reductions such as a reduced workforce due to new ways of working. In M1, no refund system is in place.

Interventions mainly range from promising professional practices that are getting attention and scaled up within the local government, to more research-based practices that SI staff or politicians picks up from other parts of the country due to good results from research or evaluation. For example, an intervention builds and expands on an earlier project about cooperation between school and social services that has shown relevant results in an evaluation by two university researchers. Some interventions are undergoing randomized controlled trials (RCTs) in other contexts and countries than the Swedish SI funds. No intervention is claimed by the municipalities to be strictly evidence-based, in the sense that it has had multiple RCTs across heterogeneous populations demonstrating its effectiveness (cf. Nutley et al., 2019: 229).

Table 1. Municipalities.

|                        | M1               | M2               | M3               |
|------------------------|------------------|------------------|------------------|
| Number of projects     | 5                | 8                | 10               |
| (Dec 2019)             |                  |                  |                  |
| Budget per project     | 300,000 euros    | 600,000 euros    | 60,000 euros     |
| Refunding system       | No               | Yes              | Partial          |
| Problem profile        | Children and Youth | Children and Youth | Social exclusion |
| Full time SI-staff     | 1                | 2                | 1                |
| SI-fund outcome        | External          | Internal          | Internal          |

SI: social investment.
SI evaluation directs attention towards the outcome side of policy evaluations (Jenson, 2009), although a variety of descriptive quantitative and qualitative data and analyses are common in practice (Vo et al., 2016). In the three municipalities, SI fund evaluation comprises three interrelated evaluation activities that are carried out regardless of intervention. A monitoring system is created by SI staff comprising project goals, activities (with activity indicators), social outcomes (with outcome indicators) and economic outcomes. Project staff are supposed to continuously feed the system with data, which work as a baseline for the outcome evaluation. These evaluations are carried out internally by SI staff or externally and should have control groups, if possible, in order to be counterfactual. Quasi-experimental evaluation is most common, especially in school-related interventions where students primary characteristic is age and students from other grades or schools work as a reference. Where target groups’ characteristics are for example, language or social problem, pre-post analysis tend to be used. Outcome evaluation is conducted at the end of the projects and also possibly after 1–3 years. Some evaluations (especially in M2) are complemented with qualitative interviews to give a broader account of the intervention. Cost calculations such as cost benefit analysis (CBA) are also performed to assess whether or not the interventions have reduced future spending. The final evaluation reports draw on all three activities and provide information for decision-makers such as department managers and the municipal executive committee. If the outcomes are positive, the project or selected parts of the project may be implemented. Thus, predetermined standards, indicators and measurements represent success and may shape organizational practice.

### Data collection

Qualitative data from 16 interviewees and 16 evaluations were collected (see Table 2). Interviews were considered relevant for capturing reflections on how evaluation could affect domains outside of formal intentions. Interviewees were contacted in a combined network and strategic sample. Municipalities were approached through a national network of SI coordinators, after which SI staff, project leaders and department managers were contacted and asked to suggest other interviewees. SI staff manage and oversee SI fund projects in different departments, handle applications, design monitoring systems and may perform outcome evaluations. Department managers are responsible for the acquired funds and are the owners of the SI projects. Project managers work with the interventions. Also, 16 project evaluations describe how the SI projects are valued, evaluative activities are organized and what recommendations

| Interviewees               | M1 | M2 | M3 |
|----------------------------|----|----|----|
| SI-staff                   | 2  | 6  | 1  |
| Department managers        | 2  | 2  |    |
| Project managers           | 1  | 2  | 2  |
| External evaluators        |    |    |    |
| Total                      | 3  | 10 | 3  |

| Evaluations                | M1 | M2 | M3 |
|----------------------------|----|----|----|
| Outcome evaluations        | 5  | 8  | 3  |

SI: social investment.

Table 2. Data collection.
are made to decision-makers. The interviews were held from 2018 to 2019, conducted in person or on the telephone, and lasted 45–120 minutes. They were recorded and transcribed verbatim and organized using NVivo software. Information about the study and interview was sent in advance and verbal consent was received from all interviewees.

The interviews were semi-structured in two themes: governance through SI and evaluation of SI funds. Questions revolved around the need for SI-funds and how the local SI-fund system worked, and interviewees’ experiences of SI-fund evaluation practice – how it was performed, and its consequences for evaluation practice, the organization and different stakeholders. The approach to let interviewees describe the consequences of SI fund evaluation created a first-hand account of how evaluation is perceived to influence different aspects of an organizational context. Still, there are aspects that interviewees may not have been able to assess from their perspective as they may be limited to their role as manager or evaluator and aspects that they presented in certain ways to be morally congruent (e.g. Alvesson, 2003). Also, professionals have not been interviewed which could complement the picture of evaluation consequences in practice. Evaluation reports have complemented interviewees’ accounts as descriptions of SI-fund evaluation to understand interviewees’ statements.

**Data analysis**

An abductive approach was used in a reflective dialogue between the researcher and the previous research, theory, and the empirical material (cf. Alvesson and Kärreman, 2007). The transcribed empirical data directed attention towards the role of evidence and evaluation in decision-making and how they may shape and determine the worldviews of staff and the content of social interventions. Research on EBPP, evaluation use and constitutive effects was consulted in order to make sense of the data (e.g. Boaz and Davies, 2019; Dahler-Larsen, 2014; Vedung, 2010). In a next step, the four domains of constitutive effects provided an analytical lens. For example, the interviewees’ reflections on the significance of evaluation to aspects of time were categorized in the domain of timeframes. Another example was how the domain frame and worldview was studied by connecting problems, solutions and normative rationales (Benford and Snow, 2000). In the same way as numbers, evaluation constitute the things they measure by ‘directing attention, persuading, and creating new categories for apprehending the world’ (Espeland and Stevens, 2008: 404). Thus, the domains were used as themes in a theoretical thematic analysis from the overarching concept of constitutive effects (cf. Maguire and Delahunt, 2017). Still, some inductive sub-themes were developed in order to handle the width of the domains.

**The constitutive effects of SI fund evaluation**

The results are presented in four sections corresponding to the four domains of constitutive effects: frame and worldview, content, timeframes and social relations.

**SI fund evaluation as a frame and worldview**

SI fund evaluation shapes an interpretive frame of how to understand the success of social interventions which is presented in terms of problems, solutions and normative rationales. According to the SI staff in M2, the primary problem that SI fund evaluation was tasked with
solving was the inability to show that preventive social interventions were cost-effective. Preventive work is nothing new in social work, although the ability to demonstrate pedagogically to municipal politicians the monetary consequences of doing nothing or something too late has been lacking (evaluator, M1). Costs and social problems are intimately linked in this problem formulation, requiring a solution that can handle both.

The solution is framed as an evaluation approach that delivers data, which makes it possible to single out which preventive interventions are effective and cost reducing, and which preventive interventions are not. The language of SI fund evaluation supports this solution. Monitoring is said to be important for professional development, but project managers primarily describe it as raw data and a baseline for outcome evaluation or for legitimizing the project, rather than for project improvement. For example, a project manager (M3) finds monitoring crucial for external legitimacy and is ‘working hard for politicians to come here and see the progress’. Outcomes is a term used by all interviewees as something that brings knowledge about which interventions work in their municipality: ‘We want to describe the outcomes that the interventions have contributed to for the target group and for the municipality’s economy’ (SI staff, M2). The outcomes determine the level of success of an intervention which, in a next step, can be related to costs. The term directs attention towards the output side of interventions, away from the processes of professional welfare work. Costs are fundamental to the solution. A typical evaluation summary states that a municipality’s reduced costs amount to EUR 78,000 for pupils no longer in need of attending a certain intervention, and EUR 36,800 each year for reduced social assistance (Evaluation M3, 2019a: 1). An evaluator in M1 argues that this approach is potentially easy to grasp by politicians and decision-makers and the SI staff in M3 say that it is possible to visually illustrate outcomes and costs in order to enhance understanding and comparability between evaluations. Finally, refunding is a term that is primarily used by department managers and SI staff. For the latter group, it connotes the idea that investment in the future results in reduced spending and opportunities for refunding and thus new initiatives. To the former group, it appears to have more negative connotations where ‘re-’ in ‘refunding’ refers to the loss of resources. According to SI logic, future expenditure for a social problem will be lower since a project has reduced the problem. Together, the four terms communicate a solution for decision-makers using influential contemporary terms: the objectivity of the evidence wave (outcomes and monitoring) and the sound financial long-term perspective of the social investment idea (costs and refunding). They facilitate instrumental decision-making by connecting outcomes to costs.

This solution provides normative rationales related to austerity and the scientific rationality of EBPP. Preventive interventions are framed as investments in order to innovate the public sector and solve fiscal challenges. The interviewees describe similar arguments when asked about the framings of SI fund evaluation. Regarding whether or not social workers are comfortable with the economic approach of SI funds, an evaluator in M1 argues that, in recent decades, social workers have become more open to analysing costs:

Social workers work with societal change that provides economic outcomes . . . you must be able to argue for your own role . . . you create benefits in society, dammit, you help people feel better, it’s extremely useful . . . and we live in a capitalistic society, and you have got to learn the rules of the game . . . there is too much ignorance about this.

Thus, the overall frame and worldview appear to resonate broadly with the interviewees, in terms of both how credible and consistent it is perceived, and how relevant it is to their everyday experience (Benford and Snow, 2000).
SI fund evaluation defines the content and what to strive for

Evaluation defines what is central in social interventions. A first subtheme concerns how SI fund projects are decided on, and a second subtheme concerns how lower future expenditure is what to strive for.

First, SI fund projects are ideally thought of as bottom-up organizations in which departments can apply for funding based on both evidence and professional experience. In practice, the SI staff in all municipalities are involved in probing and developing suitable ideas from the existing evaluated interventions. For example, in M2,

[Departments] see the problems or challenges, or what to call them, and identify the target group . . . and we (SI staff) say that from an evidence perspective, this [intervention] should work. (SI staff, M2)

SI projects may also be decided on top-down with minimal participation from the professionals in all municipalities, potentially becoming a burden:

There is much political will in these social investments. . . . Social investments have sometimes been the result of political orders and have been more difficult to implement . . . here is an idea from above that someone wants to try. It is not as successful as when the idea comes from below. And we may have had too many political orders. (Department manager, M2)

Nevertheless, finding evidence-based interventions has proven to be difficult since ‘there are few evidence-based interventions in social work for our problems’, says the SI staff in M3. ‘Instead, we have to experiment, and that’s why evaluation is so important, so we get an understanding of what works and what doesn’t work’.

Second, the SI-funded projects that are initialized strive to achieve lower future costs. Project goals and outcomes are defined in the project application and are related to a cost calculation that entitles the project to funding. Evaluation reports in all municipalities reach conclusions on cost-effectiveness in their executive summaries: ‘The cost of social exclusion on a societal level has decreased by [EUR 9,500] per participant per year, which is above average in [the evaluator’s] evaluations of similar projects’ (Evaluation M1, 2017a: 4). In this case, it is conducted through the monetarization of the project and its clients (i.e. service users) through standardized costs available through public databases, which are compared before and after the intervention. As costs depend on the social problem and intervention, they differ between projects. However, costs for social assistance are analysed in 9/16 evaluations, partly because they are relatively easy to calculate and partly because they are among the municipalities’ largest expenditures. Furthermore, costs require quantification. The objectivity and trustworthiness of quantitative measurements are emphasized in general. For example, a project manager in M3 argues that ‘If you can’t present real results, you can’t say what works. Not teachers’ assessments such as grades, but standardized tests that cannot be questioned’. However, it is not unusual for evaluations to suffer from a lack of data, short time spans and attribution issues. Evaluations in M2 need ‘more data and a longer period of measurement to confirm a relationship’ (2019: 2), ‘sufficient data is not yet available’ (2018: 2) or data ‘have been hard to interpret or shown results that cannot be directly related to the project’s work’ (2017: 2). Evaluations may also lack the possibility for control groups ‘as the target group has not been able to be identified and documented in the social assistance system systematically’ (2017: 13). In M3, ‘objective points of measurement have been hard to find’ (2019b: 3) and,
in M1, some projects were not aware of the monitoring that was needed for outcome evaluation. These problems can make cost calculations difficult: ‘[It is] difficult to demonstrate that the intervention has generated the economic effects that surpass the costs’ (Evaluation M2, 2019: 2). Nevertheless, the interviewees in M2 argue that there is little discussion among decision-makers on methods or results, per se. Instead, department managers want to formulate wide recommendations so as not to lose discretion, and politicians might question the level of refunding. Hence, the results are not questioned.

In all, a constitutive effect in this domain is how SI fund evaluation defines how social interventions should strive for what works and how to lower future expenditure. These aspirations tend to be initialized and organized top-down and are challenged by problems of data accessibility and the lack of evidence-based interventions. However, they may still bring legitimacy for what works.

**SI fund evaluation defines timeframes**

SI fund evaluation contains both long-term and short-term perspectives. On the one hand, the SI perspective rests on the long-term perspective that investing financial resources in preventive welfare work today leads to lower costs, fewer social problems and a decrease in social exclusion among citizens in the future. There is an element of prediction in the evaluations in which the economic outcomes in the form of savings and expenses can be formalized as prognoses on a different year basis related to actors on local, regional and national levels (Evaluation M1, 2017a, 2017b). On the other hand, although the outcomes cannot be measured immediately but generally after several years, attribution becomes more difficult as the years pass. Also, SI fund evaluation must be tailored to a municipal decision-making context in which the evaluation results are in demand at the end of a 1-year budget or a 3-year project:

[The evaluation] should be conducted as quickly as possible for economic reasons. Otherwise, we could have had a longer perspective. And we must also think about when to conduct the evaluation, preferably ten years after the investment. However, that’s difficult for us to decide. So, we have to find indicators of long-term outcomes, for example, school grades for future unemployment. Also, the scaling up must be decided on, you can’t postpone it. (SI staff, M2)

Organizing SI funds in temporary projects also adds to the short-term perspective. This is relevant since the evaluand – the project – is closely related to the evaluation practice (Shadish et al., 1991). Projects are generally characterized by temporality, assignments and transition, in contrast to the permanency, goals and continuous development of bureaucracy (cf. Fred, 2015). Since they should ideally be structured as evidence-based interventions, SI staff and external evaluators generally emphasize the importance of rigid projects, baselines and monitoring systems in order to conduct various forms of outcome evaluation and cost analyses. However, they also entail innovation and experimentation and projects in all the municipalities have tended to differ from the initial plans, challenging the projects’ rigidity and predictability. The SI staff in M2 argue that although projects must adapt at an early implementation stage, rigidity is more important since the outcome evaluation will otherwise suffer. Similarly, an evaluator in M1 argues that the proposed rigidity of the required monitoring and evaluation led to irritation and scepticism towards evaluators who were seen as interfering with social work (evaluator, M1). Finally, the short-term perspective poses an ethical challenge. A project
manager in M2 states that since a municipality ‘goes all in and give society’s most vulnerable children something that you won’t continue with’.

In all, a constitutive effect is how SI fund evaluation tends to be conducted from a short-term perspective, giving rise to tensions such as deciding on long-term interventions based on short-term measurements and keeping temporary and innovative projects rigid due to the requirements of outcome evaluation.

**SI fund evaluation defines social relations**

SI fund evaluation defines social relations and identities for the evaluator and for actors in the evaluation context. The extensive tasks of the SI staff in M2 and M3 make them both internal evaluators and a support function for the project, and sometimes even the ‘informal project owner’ (SI staff, M2). Their involvement is appreciated by the department and project managers who receive various kinds of support. The internal evaluation role of the SI staff in M2 and M3 was meant to keep costs down and evaluation skills within the municipalities. However, providing decision-makers with evaluations while also contributing to organizational development give the SI staff a potentially conflicting role between control and support:

> You can become partial and biased, but it’s also the way we learn. There shouldn’t be a strong distinction [between evaluation and projects]. Then we’d just sit here and evaluate and say this is bad and this is good. It should be a process. (SI staff, M2)

Other actors defined by SI fund evaluation are politicians on the municipal board who are constituted as decision-makers. In cases in which SI projects are also proposed and decided on by politicians without participation from professionals, such a combination represents a top-down version of EBPP. Relatedly, department managers are also constituted as decision-makers since they become SI project owners. As SI projects are relatively small in monetary terms, department managers are involved in decisions about what would otherwise be delegated to other roles. Thus, SI projects receive relatively lot of attention. However, the better the results of an SI project, the larger the refund from the department to the SI fund (in M2). For this reason, the refund model has been criticized by department managers in M2 for being a burden and a loan. On top of refunding, there is also the cost of implementation and upscaling of successful interventions. A department manager (M2) thinks that the refunding mechanism is counterproductive. This raises the question whether SI contributes to incentives to present data in an advantageous way for financial reasons, to which the manager responds that the department’s primary concern has been to implement new SI projects: ‘We’ve been so economically stressed out . . . we’ll pay anyway’. Also, the evaluations’ recommendations are subject to ‘lengthy discussions’ (Department manager, M2) before finalization where they can be negotiated and carefully formulated. SI fund evaluation therefore creates certain incentives and tensions among different decision-makers.

Finally, actors on other levels are present in prognoses in some evaluations as they are generally responsible for health care (regional level) and/or employment (national level). The levels become sectors to which municipal costs can be transferred when interventions succeed (Evaluation M1, 2017a).
Thus, a constitutive effect of SI fund evaluation is the concentration and upwards movement of the power over the formulation and use of evaluation at the top of the municipality, and the lack of professional and client participation.

**Discussion and conclusion**

In summary, SI fund evaluation contributes to constitutive effects in four different domains. It promotes a frame in which it solves the inability to measure the costs of preventive welfare work within scientific rationality. SI fund evaluation also defines the content as top-down evidence-based interventions aimed at gaining knowledge on what works and reducing costs. This creates implementation challenges because of the lack of evidence-based interventions and data accessibility. Furthermore, SI fund evaluation is conceptualized as a long-term commitment but performed in the short term, creating tensions between, for example, the flexibility and rigidity of projects. Finally, the power over evaluation and how to define success is located at the top of the municipal administration.

In this closing section, the limitations of the study, the potential significance of the constitutive effects of evaluation in EBPP and their implications for research and practice are discussed.

A limitation of the study concerns the use of domains as analytical concepts. Dahler-Larsen (2014: 977) describes them as conceptual domains and uses them exploratively to ascertain whether they are ‘good heuristic hooks upon which to hang empirical detail’. Indeed, they have been useful for both exploration and description, although they are broad and have been complemented and operationalized with the help of other concepts in the analysis (e.g. frames). Another limitation concerns data. Project managers are in the minority, and no professionals, politicians or clients were interviewed. This has directed attention towards civil servants and managers, not least since decision-making was seen early on as an essential part of SI fund evaluation. However, even if project managers state that not much is different in an SI fund project compared to other projects (apart from increased administration), professionals might hold different views. Politicians could also provide valuable insights into how evaluations are used. Hence, even if the study provides an account of the production of constitutive effects in an administrative practice, future research could also study the constitutive effects in a professional practice. This would shed light on whether or not various indicators are content to strive for in professional welfare work, or if this is a phenomenon on an administrative level.

A first topic of discussion is how SI fund evaluation promotes an idea of evaluation use built on a linear knowledge transfer model, that is, a one-way process from researchers or evaluators’ production of evidence to end-users who incorporate such evidence into policy and practice (Boaz and Nutley, 2019). It is closely related to a top-down implementation model requiring compliance with knowledge of the best available interventions and assessments based on systematic literature reviews of published research. (Bergmark et al., 2011). Experimental designs are supposed to produce instrumentally usable knowledge for decision-makers about ‘what works’ (Vedung, 2010). It represents a science-driven perspective based on the rationalistic assumption that it is possible to separate goals (politics and decision-making) and means (social science and evaluation) (cf. Furubo, 2019).

At the same time, there are data problems and a lack of evidence-based interventions that make it difficult to perform SI fund evaluation in ways that make it rigorous for decision-making. Nonetheless, the organizational procedures for performing evaluation are upheld and said to be valuable to show externally. Thus, there are signs of loose coupling, that is,
a deliberate separation between organizational structures that enhance legitimacy and the organizational practices within an organization that are believed to be technically efficient (Boxenbaum and Jonsson, 2017: 90). Evaluation may therefore become legitimizing since it is not the substance of results that are used, but what they indicate in terms of their credibility (Boswell, 2012). This is not an unusual phenomenon in human service organizations, which are dependent on legitimacy because of the difficulty of measuring outcomes or even agreeing on the desired outcome (Hasenfeld, 2010). In order to make the organization legitimate in the eyes of voters, through evaluation, senior managers and politicians can show they are able to take action.

Finally, although the idea of a linear knowledge transfer model dominates the SI fund evaluation frame, there are signs in the SI evaluation practice of a relational model of knowledge use in which a variety of actors is included in the production and use of evidence. For example, the requirement for evidence-based interventions is downplayed in the municipalities and may be exchanged for ‘proven experience’. Also, data problems and the lack of evidence-based interventions challenge the linear model and which is partly complemented by professional knowledge. In light of the different pitfalls and criticisms of top-down and linear knowledge use, an implication for evaluation practice in EBPP is how evaluators are challenged to problematize and go beyond a linear knowledge transfer model or otherwise be confronted by several challenges when disseminating evaluation results to decision-makers.

A second finding is how SI fund evaluation as a frame and language promotes a link between evidence and costs. Porter (2003) states that cost calculations such as cost-benefit analysis were an attempt to create decision theory by establishing a rigorous, quantitative basis for public decisions. It was a way of demonstrating honesty and fair-mindedness in the face of corruption, politics and distrust. Quantitative criteria are also authoritarian. They are regarded as accurate or valid representations of something else and perceived as useful for solving problems. They also link actors with relations to the numbers and are associated with rationality and objectivity (Espeland and Stevens, 2008). SI fund evaluation ties such quantitative cost calculations to EBPP. The combination of scientific measurements and decisions based on fair calculations could explain the resonance of SI fund evaluation among decision-makers. It provides them with apparently objective data in order to make decisions about costly welfare interventions in the face of austerity.

The link between evidence and costs has at least two implications for evaluation in EBPP. First, it turns the attention away from qualitative methods, the processes of professional practice and the experiences of clients. Evaluation is not for professional improvement or client use, but for instrumental decision-making. It aims to settle rather than inform EBPP. A narrow focus on demonstrable results may overshadow the significance of missing important information on implementation processes and discourage innovative thinking (Gray et al., 2015; Lehmann, 2015).

Second, the link reveals the costs of social problems and interventions. This is a core argument of the SI perspective in which the financial benefits of early interventions are emphasized, urging decision-makers to act preventively to social problems (cf. Hemerijck, 2017). Instead of looking back, SI thus contributes with a focus on the future by trying to identify preventive interventions that may save resources and decrease human suffering. Hence, the slogan “What works?” may be altered to “what lowers future costs.” However, the link between evidence and costs also raises the concern if citizens, often children, matter instrumentally to economic ends rather than existentially (cf. Lister, 2004). Thus, expenditure might
only be justifiable when there is a demonstrable pay-off, not when it merely contributes to well-being or enjoyment. Future research could explore how SI fund evaluation shapes future local social policies and interventions.

A third topic is how SI evaluation follows a pattern in which the power over evaluation has moved from a professional sphere to a managerial sphere. A professional sphere ideally emphasizes knowledge for improvement, based on experience, to be used conceptually and interactively. However, the responsibility for SI fund evaluation in two of the municipalities has been allocated to internal evaluators on a top management level who have economic rather than professional backgrounds. This is a managerial sphere that typically emphasizes knowledge manifested in standards and used instrumentally to investigate accountability (Evetts, 2009). Thus, the results of this study follow the development in welfare work in recent decades in which the power over evaluation in public sector organizations has moved upwards (e.g. Krogstrup, 2017). Consequently, although professionals’ views are lacking in this study, it is questionable whether trust in professionals, as promoted by trust-based governance, is supported by SI-fund evaluation. Neither professionals or clients may not be represented when discussing criteria for success and how acceptable welfare should be assessed. As research suggests that policymakers have also lost the power of evaluation to administration (e.g. Rosser and Horber-Papazian, 2018), future research could investigate the strategies and justifications for the evaluation jurisdiction of civil servants. Also, economic and statistical competency may increase in demand. However, bearing in mind that the sample is small and that research has pointed to a diverse evaluation practice within social investment initiatives (Vo et al., 2016), this implication needs to be further investigated.

This discussion on the power over evaluation raises the question whether politics can be separated and removed from evaluation. If politics in evaluation refer to actors’ involvement and negotiations on the distribution of societal and organizational resources, it seems difficult. An evaluator ultimately chooses values and criteria against which to value and frame the evaluand, through choices of methods and research questions, with implications for said resources (cf. Karlsson Vestman and Conner, 2006). In this perspective, SI evaluation not only re-locates the power of evaluation to a managerial sphere, but also the politics.

In conclusion, the constitutive effects of SI fund evaluation have significance for evaluation in EBPP in various ways. SI fund evaluation promotes a linear knowledge transfer model and an evaluation frame and language that link evidence with costs. SI fund evaluation relates to a managerial sphere rather to a professional and give legitimacy to welfare organizations.

Acknowledgements
The author would like to thank the two anonymous reviewers for their constructive comments. Mats Fred, Patrik Hall, Dalia Mukhtar-Landgren and Verner Denvall all gave valuable comments on early drafts.

Funding
The author disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Swedish Research Council for Health, Working Life and Welfare under Grant 2017-02151.

ORCID iD
Kettil Nordesjö https://orcid.org/0000-0003-3565-6563
References

Alkin M, Patton MQ and Weiss CH (1990) Debates on Evaluation. Newbury Park, CA: SAGE.

Alvesson M (2003) Beyond neopositivists, romantics, and localists: A reflexive approach to interviews in organizational research. Academy of Management Review 28: 13–33.

Alvesson M and Kärreman D (2007) Constructing mystery: Empirical matters in theory development. Academy of Management Review 32: 1265–81.

Andersen NA (2019) The constitutive effects of evaluation systems: Lessons from the policymaking process of Danish Active Labour Market Policies. Evaluation 26: 257–74.

Balkfors A, Bokström T and Salonen T (2019) Med framtiden för sig: en ESO-rapport om sociala investeringar. Stockholm: Norstedts Juridik.

Benford RD and Snow DA (2000) Framing processes and social movements: An overview and assessment. Annual Review of Sociology 26(1): 611–39.

Bergmark A, Bergmark Å and Lundström T (2011) Evidensbaserat socialt arbete: teori, kritik, praktik. Stockholm: Natur & kultur.

Berndt C and Wirth M (2018) Market, metrics, morals: The social impact bond as an emerging social policy instrument. Geoforum 90(2018): 27–35.

Best A and Holmes B (2010) Systems thinking, knowledge and action: Towards better models and methods. Evidence & Policy: A Journal of Research, Debate and Practice 6: 145–59.

Boaz A and Davies H (2019) What Works Now? Evidence-Informed Policy and Practice. Bristol: Policy Press.

Boaz A and Nutley S (2019) Using evidence. In: Boaz A and Davies H (eds) What Works Now? Evidence-Informed Policy and Practice. Bristol: Policy Press, 418.

Boswell C (2012) The Political Uses of Expert Knowledge: Immigration Policy and Social Research. Cambridge: Cambridge University Press.

Boxenbaum E and Jonsson S (2017) Isomorphism, diffusion and decoupling: Concept evolution and theoretical challenges. The SAGE Handbook of Organizational Institutionalism 2: 79–104.

Cairney P (2016) The Politics of Evidence-Based Policy Making. London: Springer.

Dahler-Larsen P (2012a) Constitutive effects as a social accomplishment: A qualitative study of the political in testing. Education Inquiry 3(2): 171–86.

Dahler-Larsen P (2012b) The Evaluation Society. Stanford, CA: Stanford Business Books.

Dahler-Larsen P (2014) Constitutive effects of performance indicators: Getting beyond unintended consequences. Public Management Review 16: 969–86.

Donaldson SI, Christie CA and Mark MM (2009) What Counts as Credible Evidence in Applied Research and Evaluation Practice? Los Angeles, CA: SAGE.

Donaldson SI, Christie CA and Mark MM (2014) Credible and Actionable Evidence: The Foundation for Rigorous and Influential Evaluations. Thousand Oaks, CA: SAGE.

Donaldson SI, Patton MQ, Fetterman DM, et al. (2010) The 2009 Claremont Debates: The promise and pitfalls of utilization-focused and empowerment evaluation. Journal of MultiDisciplinary Evaluation 6(13): 15–57.

Elvbakken KT and Hansen HF (2018) Evidence producing organizations: Organizational translation of travelling evaluation ideas. Evaluation 25: 261–76.

Espeland WN and Stevens ML (2008) A sociology of quantification. European Journal of Sociology/Archives Européennes de Sociologie 49: 401–36.

Evett J (2009) New professionalism and new public management: Changes, continuities and consequences. Comparative Sociology 8(2): 247–66.

Flyvbjerg B (2011) Case study. In: Denzin NK and Lincoln YS (eds) The SAGE Handbook of Qualitative Research. Thousand Oaks, CA: SAGE, 301–16.

Fred M (2015) Projectification in Swedish municipalities. A case of porous organizations. Scandinavian Journal of Public Administration 19(2): 49–68.
Furubo J-E (2019) Understanding the evaluation enterprise. In: Furubo J-E and Stame N (eds) The Evaluation Enterprise: A Critical View. New York: Routledge, 1–22.

Gray B, Purdy JM and Ansari S (2015) From interactions to institutions: Microprocesses of framing and mechanisms for the structuring of institutional fields. *Academy of Management Review* 40: 115–43.

Hanberger A (2011) The real functions of evaluation and response systems. *Evaluation* 17: 327–49.

Hanberger A and Lindgren L (2019) Evaluation systems in local eldercare governance. *Journal of Social Work* 19: 233–52.

Hansen HF and Rieper O (2010) The politics of evidence-based policy-making: The case of Denmark. *German Policy Studies* 6: 87–112.

Hasenfeld Y (2010) *Human Services as Complex Organizations*. Los Angeles, CA: SAGE.

Hemerijck A (2017) *The Uses of Social Investment*. Oxford: Oxford University Press.

Henry GT and Mark MM (2003) Beyond use: Understanding evaluation’s influence on attitudes and actions. *American Journal of Evaluation* 24(3): 293–314.

Hultkrantz L (2015) *Att utvärdera sociala investeringar*. Stockholm: SNS Förlag.

Jacobsson K and Meeuwisse A (2020) ‘State governing of knowledge’ – Constraining social work research and practice. *European Journal of Social Work* 23(2): 277–89.

Jenson J (2009) Lost in translation: The social investment perspective and gender equality. *Social Politics* 16: 446–83.

Johansson S (2010) Implementing evidence-based practices and programmes in the human services – Lessons from public administration research. *European Journal of Social Work* 13(1): 109–25.

Jonsson AR and Johansson K (2018) Sociala investeringar i en svensk kommunal kontext. *Nordisk Administrativt Tidsskrift* 95: 17–36.

Karlsson Vestman O and Conner RE (2006) The relationship between evaluation and politics. In: Shaw I, Greene JC and Mark MM (eds) *Handbook of Evaluation: Policies, Programs and Practices*. Thousand Oaks, CA: SAGE, 3–65.

Kirkhart KE (2000) Reconceptualizing evaluation use: An integrated theory of influence. *New Directions for Evaluation* 2000(88): 5–23.

Krogstrup HK (2017) *Utvärderingsmodeller*. Lund: Studentlitteratur.

Krogstrup HK and Dahler-Larsen P (2000) Evaluerings og institutionelle standarder: Nyinstitutionelle betragtninger af evaluerings som vor tids ritual. *Politica* 32: 283–99.

Lindblom CE (1979) Still muddling, not yet through. *Public Administration Review* 39(6): 517–26.

Lindgren L (2014) *Nya utvärderingsmonstret: om kvalitetsmätning i den offentliga sektorn*. Lund: Studentlitteratur.

Lister R (2004) The third way’s social investment state. In: Lewis J and Surender R (eds) *Welfare State Change: Towards a Third Way*. Oxford: Oxford University Press, 157–81.

Maguire M and Delahunt B (2017) Doing a thematic analysis: A practical, step-by-step guide for learning and teaching scholars. *All Ireland Journal of Higher Education* 9(3): 335.

Martin GP and Williams O (2019) Evidence and service delivery. In: Boaz A and Davies H (eds) *What Works Now? Evidence-Informed Policy and Practice*. Bristol: Policy Press, 418.

Miller P (2001) Governing by numbers: Why calculative practices matter. *Social Research* 68: 379–96.

Morel N, Palier B and Palme J (eds) (2012) *Towards a Social Investment Welfare State? Ideas, Policies and Challenges*. Bristol: Policy Press, 1–30.

Mosley JE and Smith SR (2018) Human service agencies and the question of impact: Lessons for theory, policy, and practice. *Human Service Organizations: Management, Leadership & Governance* 42: 113–22.

Nutley S, Davies H and Hughes J (2019) Assessing and labelling evidence. In: Boaz A and Davies H (eds) *What Works Now? Evidence-Informed Policy and Practice*. Bristol: Policy Press, 225–50.

Patton MQ (2000) Overview: Language matters. *New Directions for Evaluation* 2000: 5–16.
Pollitt C (2013) The logics of performance management. Evaluation 19(4): 346–63.
Porter TM (2003) Measurement, objectivity, and trust. Measurement: Interdisciplinary Research and Perspectives 1: 241–55.
Power M (1997) The Audit Society: Rituals of Verification. Oxford: Oxford University Press.
Rijcke S, Wouters PF, Rushforth AD, Fransen TP and Hammarfelt B (2016) Evaluation practices and effects of indicator use – a literature review. Research Evaluation 25: 161–9.
Rist RC and Stame N (2006) From Studies to Streams: Managing Evaluative Systems. New Brunswick, NJ: Transaction Publishers.
Rosser C and Horber-Papazian K (2018) From law to reality: A critical view on the institutionalization of evaluation in the Swiss. In: Furubo J-E and Stame N (eds) The Evaluation Enterprise: A Critical View. New York: Routledge, 83–110.
Sackett DL (1997) Evidence-Based Medicine: How to Practice and Teach EBM. New York: Churchill Livingstone.
Sanderson I (2001) Performance management, evaluation and learning in ‘modern’ local government. Public Administration 79: 297–313.
Schwandt T (2015) Evaluation Foundations Revisited: Cultivating a Life of the Mind for Practice. Stanford, CA: Stanford University Press.
Shadish WR, Cook TD and Leviton LC (1991) Foundations of Program Evaluation: Theories of Practice. Newbury Park, CA: SAGE.
Skillmark M (2018) Uppdrag standardisering: införande och användning av manualbaserade utrednings-och bedömningsverktyg i socialtjänsten. Doctoral Dissertation. Växjö: Linnaeus University.
SOU (2019) Med tillit följer bättre resultat: tillitsbaserad styrning och ledning i staten. Stockholm: Norstedts juridik.
Stame N (2019) The politics/administration dichotomy and evaluation. In: Furubo J-E and Stame N (eds) The Evaluation Enterprise: A Critical View. New York: Routledge, 1–22.
Thiel SV and Leeuw FL (2002) The performance paradox in the public sector. Public Performance & Management Review 25(3): 267–81.
Timmermans S and Epstein S (2010) A world of standards but not a standard world: Toward a sociology of standards and standardization. Annual Review of Sociology 36: 69–89.
Trinder L and Reynolds S (2000) Evidence-Based Practice: A Critical Appraisal. Oxford: Blackwell Science.
Uph (2015) Vad är sociala investeringar? 17 frågor och svar. Stockholm: Uppdrag Psykisk hälsa.
Uph (2017) Helhet, hållbarhet och resultat: en guide för sociala investeringar: organisatoriska, strategiska och operativa aspekter. Stockholm: Uppdrag psykisk hälsa.
Vedung E (2010) Four waves of evaluation diffusion. Evaluation 16(3): 263–77.
Vedung E (2015) Six uses of evaluation. In: Nachhaltige Evaluation?: Auftragsforschung Zwischen Praxis und Wissenschaft. Festschrift zum 60. Geburtstag von Reinhard Stockmann / [ed] Hennefeld, Vera; Meyer, Wolfgang; Silvestrini, Stefan, Münster. New York: Waxmann Verlag, 2015, s. 187–210.
Vo AT, Christie CA and Rohanna K (2016) Understanding evaluation practice within the context of social investment. Evaluation 22: 470–88.
Weiss CH (1979) The many meanings of research utilization. Public Administration Review 39(5): 426–31.
Weiss CH (1980) Knowledge creep and decision accretion. Knowledge 1(3): 381–404.
Weiss CH (ed.) (1977) Using Social Research in Public Policy Making. Vol. 11. Lexington, MA: Lexington Books.

Kettil Nordesjö is an Associate Senior Lecturer at the Department of Social Work, Malmö University, Sweden.