<http://www.diva-portal.org>

This is the published version of a paper published in *American Journal of Evaluation*.

Citation for the original published paper (version of record):

Groth Andersson, S., Denvall, V. (2017)

Data Recording in Performance Management: Trouble With the Logics.

American Journal of Evaluation, 38(2): 190-204

<https://doi.org/10.1177/1098214016681510>

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:lnu:diva-59307>

Data Recording in Performance Management: Trouble With the Logics

American Journal of Evaluation
1-15

© The Author(s) 2016

Reprints and permission:

sagepub.com/journalsPermissions.nav

DOI: 10.1177/1098214016681510

journals.sagepub.com/home/aje

Signe Groth Andersson¹ and Verner Denvall²

Abstract

In recent years, performance management (PM) has become a buzzword in public sector organizations. Well-functioning PM systems rely on valid performance data, but critics point out that conflicting rationale or logic among professional staff in recording information can undermine the quality of the data. Based on a case study of social service staff members, the authors explore three recording logics. The findings reveal a complexity of recording behavior and show how frontline staff shift between recording logics according to the situation. The actual data recordings depend not only on the overall logic but also on factors such as attitudes, assumptions, and motives. The authors suggest that shifting recording logics weaken the validity of performance data. These shifts undermine the idea of PM as a trustworthy strategy to bridge the gap between professional and managerial staff, as well as the possibility of a well-informed management.

Keywords

performance measurement (goal attainment), logics, performance data, occupational professionalism, organizational professionalism

The use of performance management (PM) systems has become commonplace in social programs as a means to measure program results. The dominant belief is that today's challenges—be they poverty, customer satisfaction, or leadership—cannot be managed if they cannot be measured (Chopra & Kanji, 2011). This quest for measurement has given rise to many systems that monitor performance and results (Lynch-Cerullo & Cooney, 2011). Management strategies such as PM, result-based management (RBM), and outcome measurement (OM) have gradually gained popularity (de Bruijn, 2007; Krogstrup, 2011; Pollitt, 2006). These management and evaluation strategies are now regarded as a major wave flooding the public sector (Albæk, 2003; Dahler-Larsen & Schwandt, 2012; Vedung, 2010).

However, sometimes the system or strategy is used as a tool to manage the program. In this study, we focus on the use of performance measures used by social service workers in the course of their

¹ Social Development Centre SUS, Copenhagen, Denmark

² Department of Social Work, Linneaus University, Växjö, Sweden

Corresponding Author:

Verner Denvall, Department of Social Work, Linneaus University, 351 95 Växjö, Sweden.

Email: verner.denvall@lnu.se

services to and interactions with the disadvantaged clients they serve through a social services agency. We found that this dual purpose—measurement and case management—coupled with the attitudes and values of the caseworkers, affects the accuracy and reliability of the measures. This may be at odds with the needs of central office managers, who need reliable information about the program's overall effectiveness and efficiency to satisfy their own management and accountability responsibilities.

Management and evaluation strategies originate from a need for greater accountability and transparency as well as a need to continuously evaluate and adjust employee and overall organizational performance and results (Lynch-Cerullo & Cooney, 2011; Mayne & Rist, 2006). Moreover, ongoing measuring and monitoring of performance and outcome must be undertaken to ensure the quality of public interventions (Nielsen, Jacobsen, & Pedersen, 2005; Pollitt, 2013; Rist, 2009).

We argue that this trend is challenged by conflicting and shifting perspectives within an organization. Employees' individual perceptions and assumptions regarding the performance measurement and management system as well as the specific performance being measured influence the validity of performance data. Based on individual assumptions and motives, employees apply different recording reasoning as conditions change. Thus, we take our point of departure in research that suggests organizations should consider alternative logics and sublogics (Pollitt, 2006). Pollitt notes that "any set of management practices must include—consciously or unconsciously, hidden or explicit—some assumptions about how people think, and what motivates them" (p. 347); these logics offer a chain of reasoning that drives performances. He argues that these logics and sublogics can affect the use, and thus the value, of the PM system and that many of the problems with PM stem from "divergences between this assumed logic and a series of what might be termed 'alternative' logics" (Pollitt, 2006, p. 347).

The major focus on, and use of, performance and OM of public interventions has led to a proliferation of types of PM systems that measure output and outcomes using different structures, competencies, and tools (de Bruijn, 2007). In this article, we use the term *performance management* (PM). Many researchers have noted that successful PM requires organizations to build adequate skills, systems, and structures (Hunter & Nielsen, 2013). Furthermore, conflicting logics and rationales within the organization, among other things, can influence the quality of the performance and outcome data generated by these systems (Mayne, 2007). Some critics even argue that the data produced as the basis for PM risk being worthless or directly misleading (Arnaboldi, Lapsley, & Steccolini, 2015; de Bruijn, 2007; Pollitt, 2013). Pollitt also stresses that only a small fraction of the literature on public-sector PM systems directly addresses the behaviors and rationales of the staff who record performance data (Pollitt, 2006).

In this article, we explore the logic used by caseworkers when recording outcome data. This study contributes to our understanding of the nuances of the logics and sublogics in play among employees and the impact these logics may have on the validity of the performance and outcome data generated. The empirical data are based on a case study of a Danish municipality that examined the use of a PM system in the social services as part of the organization's overall PM strategy.

Alternative Logics in PM

Initially used by the private sector and later adopted by the public sector, PM defines objectives and how to accomplish them uniformly through services and procedures (Brignall & Modell, 2000; Wandersman, Imm, Chinnman, & Kaftarian, 2000). Gradually, PM has expanded to social services and education, where it is characterized by less uniformity in task performance, multiple values, and less-defined targets (Krogstrup, 2011). Organizations providing social services to people typically require external funding, appeal to moral standards, and deal with outcomes that are not easy to estimate (Hasenfeld, 2010; Lipsky, 2010). This complexity makes it difficult to measure outcome

and performance accurately (Arnaboldi et al., 2015; Bliss, 2007). Conflicting interests and logics among management and staff may cause problems not only in the welfare sector but also among organizations in general (de Bruijn, 2007).

Two ideal types of rationales typically coexist in professional organizations: one oriented toward the interests of the organization (often represented by management) and the other oriented toward the interests of the profession, such as the frontline staff who provide the organization's services and also record performance and outcome data (Evetts, 2009). A given management practice will always build on certain assumptions about how people think and what motivates them—the rationales of the actors involved (Pollitt, 2013). These rationales depend not only on cognitive processes such as thinking and calculation but also on motives, emotions, and values. Recent Scandinavian research has suggested that employees' recording behavior is typically influenced by their assumptions about how the data they collect and record will be handled and used (Krogstrup, 2011). Thus, staff's responses to a given management practice are likely to build on assumptions about managers' values and how they think and act in combination with the staff's own personal motives and values. In this regard, the employees' assumptions of the existence of two divergent rationales between management and employees can influence their recording behavior, impeding successful implementation of the PM system. Although, in this article, we focus on the reasoning applied by social service employees in frontline positions, it should be noted that diverse reasoning at the managerial level might also play an important role (cf. Hansen & Vedung, 2010).

PM relies on the competence to define objectives and develop operationalized indicators of goal achievement, along with continuous collection of performance and outcome data. Therefore, a crucial part of any PM system consists of a recording tool with a number of operationalized, and often quantified, indicators of how close people and organizations come to achieving certain goals. It is through this recording tool that performance and outcome data are collected (Dahler-Larsen, 2001; de Bruijn, 2007; Pollitt, 2013). The data will then form the basis for evaluation, assessment, adaptation, and development of initiatives and interventions in relation to defined objectives (Hunter & Nielsen, 2013; Nielsen, Bojsen, & Ejler, 2009). Data are most often collected by staff performing the organization's core tasks; the data are then gathered, aggregated, and used or disseminated by managers and leaders in charge of administering and developing the organization.

Ideally, PM systems ensure the relevance of objectives and indicators by involving staff when defining the pertinent measures and supporting the interpretation of performance data with evaluation knowledge (Hunter & Nielsen, 2013). Research finds unequal focus on managing the challenge of ensuring the validity of the data recorded. Several studies have identified the risk of "gaming the system" and cheating in connection with data recording and have presented various factors that may contribute to this behavior (de Bruijn, 2007; Pollitt, 2013). The tendency to cheat or to misuse may be influenced by loose or tight coupling between performance and outcome data and incentives as well as by different logics and conflicts of interest among the various stakeholders in the organization (de Bruijn, 2007). A stronger focus on those responding to PM and on the context may help generate a better understanding of the little-known world of these alternative logics in PM (Pollitt, 2013).

Theoretical Framework

The analytical framework in this study is inspired by Dahler-Larsen's (2001) constructivist approach to program theory and Hansen and Vedung's (2010) concept of theory-based stakeholder evaluation. Several theories of change may be in play, which might affect the behavior and actions of employees involved and thereby the outcome of an intervention (Dahler-Larsen, 2001; Hansen & Vedung, 2010; Vedung, 2011). From a social constructivist perspective, a theory of change represents one of many perceptions of reality among stakeholders in the organization (Wenneberg, 2000). That is, the theory of change of a particular intervention will depend on the mediator (Dahler-Larsen, 2001;

Hunter & Nielsen, 2013). The various theories of change can be regarded as different logical models based on specific rationales and assumptions about the present situation and how a specific effort or intervention will affect it (Hansen & Vedung, 2010). In this theoretical viewpoint, actors' differing logical models for particular actions, as well as their interrelation, are crucial to the process and outcome for a specific effort or intervention. PM can, in this context, be regarded as a comprehensive governance system, in which central office managers undertake specific efforts or interventions (such as setting targets and collecting and interpreting performance and outcome data) to achieve specific goals (such as obtaining a more efficient service provision). This will be the assumed theory of change underlying PM—its “core logic” (Pollitt, 2013).

To understand the main rationales at stake, we have used Evetts's concepts of occupational and organizational professionalism (Evetts, 2009). The rationale of organizational professionalism is usually articulated by the organization's managers and leaders and is expressed as a management-oriented discourse of control. This rationale is based on decision-making structure, rational/legal forms of authority, standardized procedures, and objectives. Occupational professionalism, on the other hand, is typically represented by the professional groups that carry out the organization's core tasks. This rationale is based on peer authority and confidence in employees' ability to make judgments and assessments. Autonomy is given high priority and is based on education, experience, and professional identity. This perspective stresses ensuring the quality of work through peer control and professional ethics. When applying these ideal types of professionalism to modern social welfare organizations, the interests of the organization can be understood as administering public funds in a transparent and accountable way so as to provide the best possible service to the largest possible number of clients. The interests of staff such as frontline workers, on the other hand, are generally perceived as providing the best service to the individual client in a given situation (Lipsky, 2010). Tension often occurs between the two forms of professionalism and may be strong in organizations that deal with multiple values and less uniform tasks and targets (Björk, 2013; Liljegen & Parding, 2010).

Evetts (2009) describes how variations in these two ideal types of professional rationales can emerge and how both management staff and frontline staff can be representatives of them. She also introduces new public management professionalism as a form of organizational professionalism. This variation of organizational professionalism relies on the rationale that overall organizational performance is best sought by having employees conduct performance reviews and consider solutions for self-improvement. This rationale implies recreating them as managers, maneuvering within a predefined frame of PM (Evetts, 2009). In our case study, we use these concepts to understand how the existence of different professional rationales in the organization influences the preconditions for successfully implementing PM.

Method and Data

In this study, we focus on the part of the PM system that deals with data recording—specifically, the use of an OM tool. Readers might recognize it as the “Outcomes Star,” a measurement tool that has a similar structure and scoring system; however, we do not rely on the United Kingdom version of the Outcomes Star (<https://www.outcomesstar.org.uk>), which has been translated and modified in many ways from the original. The measurement tool in focus had been in use for about 2 years, and outcome data were systematically recorded and collected. The tool consists of 10 defined success indicators in the form of changes in clients' ability to master specific areas of living. They serve as operationalized indicators of the extent to which the client is able to manage a specific area. Change in each indicator is illustrated using a Likert-type scale of 1–10, so it is possible to detect how far an individual has progressed for each indicator (see Figure 1). A score of 1 means that clients have very poor mastery of that dimension of life or do not recognize they have a problem even though it seems evident from the outside. A score of 10 reflects very strong mastery or experiencing no problems.

Figure 1. The structure of the performance measurement tool. *Source:* Authors' illustration.

Several versions exist in the city in focus. The structure is the same, but the 10 dimensions are adapted to different target groups such as children and youth, persons with disabilities, and homeless populations. In this study, we focused on how the tool was used by providers of social services to homeless persons in shelters. Those providers are, in this article, named as “social/caseworkers” or “frontline staff,” not to be mixed up with the top officials, “management staff” or “managers.” The clients were dealing with additional challenges such as substance or alcohol abuse, mental health issues, and other severe social problems.

Frontline workers' daily tasks consist of providing assistance, practical help, and social support to shelter residents. They are required to use the performance measurement tool to report the progress of each of their assigned clients and to reassess their clients every 3 months. Scores are electronically recorded and used in ongoing evaluations of client well-being. Recording guidelines are provided. Clients' progress on the scale is related to the interventions provided by the social workers in contact with the client during the period in question. The recorded data are considered outcome data in the sense that they represent the result of the efforts and assistance provided. The frontline staff members are encouraged to schedule meetings to discuss the results of their local unit in relation to their efforts and methods in use and to draw inspiration from other units showing good results. Leaders and managers use the aggregated data to identify and investigate the interventions and methods of units, showing better or poorer results as a basis for engaging in dialogues with the specific units concerning targets, methods, and results, and to prioritize future strategies for improvement. The aggregated data should enable managers and caseworkers to share the results of their work with others (such as politicians and citizens) to ensure transparency and accountability. In this study, we do not look further into how the aggregated data are managed and used. The focus is on the rationales caseworkers use when recording outcome data and how these choices affect the validity of data (cf. Arnaboldi et al., 2015, p. 18).

The empirical material consisted of interviews with central office managers and caseworkers as well as analyses of official documents. The management office and four shelters agreed to participate. No clients were contacted. The first author interviewed two employees from central management who had been involved in developing and implementing the PM system and analyzed internal and external documents that referenced the PM system and strategy. In addition, the first author conducted seven interviews with 17 frontline staff from three shelters and one drop-in center for homeless persons. (In this article, all four are referred to as shelters.)

Through initial interviews and document reviews, we sought to establish general theories of change for the performance measurement tool and the management strategy at the central and local

levels. The analyses were based on Hansen and Vedung's (2010) tripartite model for theories of change. The result was two overarching logic models/theories of change: one formulated by leaders and managers and the other formulated by frontline staff, both representing understanding and usage of the performance measurement tool and underlying rationales. These two logic models served as the basis for additional interviews that focused more specifically on the informants' individual actions and use of the measurement tool. The interviewees were asked if they recognized the overall logic model and to elaborate on and discuss the actions and assumptions it described. These interviews were processed using a more open approach, and emerging themes were coded. From this analysis, the initial logic models were further developed and nuanced, drawing the picture of coexisting and shifting logic models among the frontline staff with specific regard to their recording behavior. Quotations are kept as close to the original Danish as possible.

Findings

The PM system in this case not only features a control and steering discourse, especially in official documents authored by managers and leaders, but also a discourse of organizational learning and self-regulation, especially expressed in the interviews with central management. In the quotation below, the employee from central management emphasizes the learning discourse:

We won't be assessing the specific units' efforts and results *so very* much. The way we are doing this is supposed to allow for continuously learning on a local level. So that practitioners find it helpful [. . .] for instance to discuss what happens when they apply a certain intervention or method to a client in a certain situation. (Employee from central management)

Despite the strong focus on learning, the manager starts by suggesting that central assessment and control of the specific units will occur to some extent, although *not very much*. This dichotomy characterizes the interview as a whole. The same manager later argues that "it is a political organization, so I would say that it is fair enough that the politicians are able to see the results of the money being spent." The control discourse is especially evident in the official documents:

By looking at the progress of several clients from one social service center or within one specific area, central management can assess whether the efforts and interventions are leading to good overall results or whether they may need adjustment [. . .] the ten dimensions ensure that the employee remembers to get around all the relevant aspects about the client, when talking to him. (Official documents on the city website)

Transparency, accountability, and standardized procedures play a key role in the rationale of the PM system. Management regards the outcome management tool as a way to bridge the local frontline level with the central managerial level, creating a path for sharing knowledge. At the same time, learning and self-regulation discourses are emphasized. Thus, management's perceptions of the PM tool adhere closely to a rationale of classical organizational professionalism and the perception of PM as a control strategy and a structure for communication, on the one hand, and as a rationale of new public management professionalism, emphasizing the perspective of organizational learning and self-regulation, on the other hand. According to Evetts (2009), both are expressions of an organizational rationale but with clear differences in their sublogics.

The interviews with frontline workers in the shelters revealed another rationale. These case-workers spoke about close relationships with clients and the ability to accommodate, recognize, and involve clients as fundamental elements and prerequisites to success in their efforts to provide help and support. Several frontline workers expressed skepticism toward the PM system. For example, one frontline worker felt that the PM system was management's attempt to standardize their work:

For me, it is about making something happen that the client can benefit from—on his own terms. So it does not necessarily have to do with some centrally outlined guidelines. (Oliver, Shelter 3)

In the interviews with caseworkers, the values of professional autonomy and discretion and the importance of individual consideration for the client are expressed as contrary to the values represented by central management and the PM tool:

We just had a discussion about our values versus the values of the central management [...] and [our value of] taking the individual client into account is to say, “We respect you and understand you and recognize your situation. It will be at your pace and when you are ready.” (Louise, Shelter 2)

Overall, the frontline workers regarded the OM tool as an expression of hierarchical managerial control and many did not believe that the conversion of their daily work into numbers and discussions based on outcome data ensured progress and quality in the social services provided. In their opinion, progress and quality can be provided and evaluated through discussions and education based on the values and knowledge of their profession:

Those discussions that this tool and its 10 dimensions are supposed to facilitate [...] we manage them better and are more qualified in other ways and by other means. So if we use the scales on the measurement tool in this way, that would be a step backward in terms of qualified discussions about the clients. (Oliver, Shelter 3)

Three Recording Logics Among Frontline Staff

The different and sometimes conflicting perceptions of the merits and impacts of the PM system give rise to various recording logics and shifting recording behavior among the social workers. Our analysis identified three recording logics applied when they record performance data. The first is a *tactical* recording logic, a pragmatic and calculating approach in which the actual recording process is regarded as a game to be played in order to obtain certain outcomes. Second, we find a *client-centered* recording logic based on valuing considerations of occupationally professional moral standards and relationships with clients. The final logic is presented as *by the book*, a recording logic in concordance with the organizational rationale and the guidelines of the PM system. Recording logics for the same staff member may include variations with dissimilar outcomes depending on their dominant attitudes, assumptions, and motives when they are recording the data as shown in Table 1.

The recording logics in play are not conclusive. Although based on a specific logical rationale, they cannot be expected to have a particular outcome in terms of intentional or random higher or lower scores or scores that accommodate the guidelines provided. According to our findings, the complexity of factors that may affect professionals' recording behavior makes it challenging to predict and assess the quality of data produced through this PM system.

Tactical

The tactical recording logic is based on a pragmatic rationale from a perspective of consistency, where records are expected to have specific consequences for the client, the workplace unit, or the employee. Recordings are based not on the given guidelines or moral or occupational standards, but rather on the intent to send specific signals to the user of the data, according to the employee's personal judgments of what is most advantageous. This recording logic can be based on mistrust of either the PM system's ability to convey meaningful data or the manager's ability to interpret them properly. It can also be based on lack of confidence in management's intentions and a belief that there is a hidden agenda regarding how the data will be used. The logic may also reflect a

Table 1. Recording Logics Applied by Frontline Staff.

Recording Logic	Rationale	Attitude/Assumption	Motive	Recording Behavior
Tactical	Pragmatic and calculating	Lack of confidence in the PM system's ability to convey meaningful data	Consideration for the unit's position	Calculating recording: Intentionally higher or lower scores than the guidelines prescribe, if it is believed to accommodate a beneficial communication strategy; "cheating" and "gaming"
		Lack of confidence in manager's ability to interpret data	Consideration for the professional's position	
		Distrust of manager's intentions	Consideration for the client's well-being/ position	
Client centered	Occupational professionalism	Lack of confidence in the PM system's ability to convey meaningful data	Consideration for the relationship and the client's well-being	Symbolic recording: Purposely indifferent, higher, or lower scores than the guidelines prescribe, if it is believed to accommodate the well-being of the client or the values of the profession
		Resistance to the values of the PM system	Loyalty to occupational principles and values	
		Indifference to the PM system		
By the book	Organizational professionalism	Confidence in the PM system and management's overall intentions with implementing it	Consideration for the organization	Dutiful recording: Scores that accommodate the recording guidelines provided by management
		Confidence in the PM system's ability to convey meaningful data	Consideration for the larger population of clients	
			Desire to comply with the given guidelines and procedures	

Note. PM = performance management.

combination of these considerations. The outcome of this recording logic depends on the caseworker's more specific assumptions about how, for what, and by whom the data will be interpreted and applied. Some staff members distrust the system's ability to communicate valid and accurate information or the management's ability to understand the complexity behind the numbers.

The following quotation shows how Mike, a caseworker in one of the shelters, is concerned that it is unclear how the data he records will be used. Furthermore, he is concerned that the data present an incomplete picture: "When you choose to look at the numbers alone, it doesn't show the whole picture of the client!" (Mike, Shelter 3). Similarly, Eric distrusts the tool's ability to present a fair and holistic picture of the client's situation and needs:

If you hit a good period 2 or 3 times in a row, then it may very well look like the client is ready to move out [of the shelter]. A little more support and off you go. But it wouldn't be someone I would like to have as a neighbor, because every three months he "tumbles" for 3 weeks and breaks everything into pieces. So therefore [...], well, I'm not so enthusiastic about this tool. (Eric, Shelter 4)

Some employees in the shelters report their suspicion that management holds back information about the use of the data. The idea of a hidden agenda gives rise to numerous speculations about consequences of the recordings:

It's a little strange that [...] at the time we all were taught about it [the recording tool], there were some meetings. We asked many, many times, "What will it be used for?" We never got an answer! [...] And then you sit back and say, well, does it have something to do with financial resources or something? What is it? No one really wants to say anything. Everything is very secretive and confusing. (Marcus, Shelter 2)

Parts of these speculations are based on the belief that the collected data provide the basis for prioritizing or allocating resources among workplaces:

It would indeed be very, very easy for [management and politicians] to go out and say, well, this shelter does better than this one, so we can allocate resources from one to the other. (Daniel, Shelter 4)

Some employees gave scores lower than their actual assessment because they feared their workplace would be deprived of resources if clients seemed to be functioning too well. Other employees gave scores higher than their actual assessment because they expected that the workplace would receive resources if they demonstrated good results. Some gave scores lower than their actual assessment to demonstrate that their clients needed further support. That is, they expected their clients to be allocated more help if they had lower scores. In other cases, employees were inclined to give higher scores because they believed that clients who did not show progress would not receive as much help as those who showed potential for change. Finally, some mentioned the possibility of people giving higher scores in order to prove their own skill.

The following quotations demonstrate the tactical rationale that intentionally lowered scores. In one of the shelters, the social workers collectively decided to give low scores in order to prevent staffing cuts, even though the low scores were not always consistent with their assessments:

We have agreed to keep the scorings low. That we should score low, and that we have to look at [the outcome measurement tool] as a political strategic tool. [...] We experience ongoing problems with staffing cuts, so if we score everyone staying here high, then it may well be that management continues to cut staff because "it is not so bad after all—the clients are not doing so poorly." (Peter, Shelter 4)

At another shelter, an employee tried to score as accurately as possible, following a solution-focused approach that sought to demonstrate the client's positive changes. But her supervisor told her that she ought to reconsider this practice since high scores may indicate that her work is superfluous:

And so my boss says to me, "The higher the scores you give, the more redundant you have made yourself." [...] Well, then you give it some thought. [...] So while before I would have scored a client an 8, I think now I would be inclined to score 7 instead, and then in three months it can climb to an 8. Even though I myself am thinking, "It's 8 now—and next time it's actually 9." (Louise, Shelter 2)

Depending on the employee's attitudes, assumptions, and motives, this logic leads to a tactical and strategic recording that does not take into account the guidelines for recording data and results in a different score than the actual assessment. That is, the recording is intended to send a signal to the central management rather than to reflect accurately the client's situation and needs.

Client Centered

This recording logic reflects a rationale in which the occupational values and moral concerns, on an individual level, override the inclination to follow the organizational rationale (represented by the PM system), including the given guidelines and procedures. When this recording logic is applied, guidelines for recording performance data are bypassed, based on the assumption that they stand in contrast to occupational values, the professional relationship to the individual client,

or the well-being of the individual client. For example, frontline staff members record a number, as they are instructed, but do not put time or energy into the assessment because they do not think the procedure matches their professional values or in any way benefits the client. That is, the use of the measurement tool and the performance data does not match the perception of doing good professional social work:

It would be a shame to spend time discussing numbers [...] [W]hen we deal with people, we do not discuss simple dimensions—we deal with relationships and all the many nuances of a person and what have you! These are complex problems where the various dimensions are interconnected and where you cannot separate one part from the other. (John, Shelter 3)

This recording logic can result from several attitudes or assumptions. The caseworkers may lack confidence in the PM system and the recording tool's ability to convey accurate and useful data, or they may resist the assumed values of the PM system as seen in relation to the tactical recording logic. Or this logic can be a result of mere indifference to the PM system and the recording tool:

We are resigned a bit when it comes to these recordings. [...] [W]e don't really talk about it anymore. We do the recordings in 5 minutes—because we have to. We are not really interested in it. [...] I guess we have realized that it isn't really useful in this context. (Chad, Shelter 2)

Several motives are at play when considering the client's well-being: consolidation of a trusting relationship, respectful involvement with the client, and loyalty to perceived occupational values such as acknowledging the individual, not putting people into boxes, and feeling it is disrespectful to describe unique human beings using numbers. For example, one caseworker adamantly resisted quantified evaluations of clients: "In my world, a description of another human being will always be a qualitative description!" (John, Shelter 3).

The employees may record numbers with such indifference that the scores are relatively random. Or for the sake of the client, they may deliberately put down a higher or lower score than their professional assessment based on the guidelines. Several caseworkers claimed it was disrespectful or inappropriate to the client relationship to "overrule" the client's own assessment when scoring. This tension might occur if the employee assesses a client low on the scoring scale, but the client thinks he or she is doing very well. Eric, for example, recorded a score higher than his own initial assessment indicated due to his client's protestations:

If I can see that I exceeded a clear line where the client just doesn't agree with the score or feels really uncomfortable with it, I make the score higher. I do this for the simple reason that I value the relationship with the client highly. The relationship you build up with the client on the basis of these conversations [...] to me that is more important than if the registered score is 2 or 6. (Eric, Shelter 4)

The social workers might shift from one recording logic to another, depending on the situation. Alma describes how she took over responsibility for a client from a colleague. She didn't agree with the colleague's latest assessment and score of that particular client. Still, she elected to make a similar record, considering that it would be an uncomfortable and demotivating experience for the client to drop suddenly to a lower score without having made any noticeable change in behavior. In this situation, the interest in the client's well-being wins out over the inclination to follow the guidelines:

It really is incredible how differently two people can assess the scores! I had to "lie" when making the score, because I couldn't just drop the client 6 scores, which would be the right thing to do according to my assessment. I thought, "Wow, the latest score is really different from my point of view!" But I cannot allow a client to drop from a 9 to a 3 just because I'm suddenly in charge of recording. (Alma, Shelter 3)

The client-centered recording logic may reflect various attitudes, motives, and sublogics. Regardless, it is a symbolic recording with low validity. The employee makes a recording, but the number will not communicate accurate information. It can be either a swift and completely random score or a higher or lower score, if deemed appropriate in relation to the interests of the client.

By the Book

The recording logic “by the book” is the least complex one. It is based on allegiance to the organizational rationale—the PM system, the managers, and the leaders. It does not contain sublogics to the same degree as the other two recording logics. Like the PM system and the measurement tool, it builds on an organizational professionalism rationale. The employee complies with the terms of standardized procedures and the objectives the PM system represents. When this recording logic is applied, the employee’s attitude is typically characterized by trust in the PM system and management’s intentions. Or, at the very least, it indicates confidence in the PM system—a belief that the recordings and measurements do not do any harm and that it is best for everyone to follow instructions. It is based on the assumption that the managers and leaders, through the use of aggregate data, are able to manage the organization and make decisions in a way that favors the largest possible number of clients and/or that the aggregate data are useful to frontline workers themselves in improving and regulating their own work. This is well in line with the managers’ ambitions and the city’s public statements:

In the future, the ambition is that clients must have an outcomes-star that comes with his or her journal.
(Official document on the city website)

The guiding assumption will thus be confidence in the PM system, the measurement tool, and the guiding motive consideration for the overall interests of the organization (and the larger population of clients). Applying this logic results in a recording in which the employee does not speculate on what potential negative or positive consequences an individual recording might have. Below, Louise describes how she does her best to follow the instructions and make a fair assessment:

When I go through the client’s individual plan with the client, I get around to the various dimensions. Then afterward, I interpret and evaluate on the 10 dimensions and where on the scale the client fits, based on the conversation we have had. Then I’ll do the same again 3 months later when we have a follow-up conversation. (Louise, Shelter 2)

Yet Louise had previously stated that she had changed her recording behavior to give lower scores after her supervisor approached her. This caseworker basically follows the guidelines but is inclined to shift to a tactical recording logic under influence of others. For employees who tend toward this recording logic, assumptions about how the data may be used do not generally play a major role:

There was a lot of talk about it in the beginning, that there might be cuts in resources [...] or additions [...] or whatever there was talk about. There were many such conspiracy theories, but [...] I really do not think about it. I think about making the recording as accurate as possible and avoiding major fluctuations over time. (Alma, Shelter 3)

The OM tool is viewed as an intention to improve how management operates and is not perceived as directly related to employees’ own professional work or values. In the next quote, the same caseworker argues for sticking to the instructions and trusting in management’s ability to handle wisely the data collected. Her point is that if they begin to record data on the basis of other logics and reasons—for instance, with consideration for the professional’s own position—they will fail to

predict the consequences and might risk creating problems for the client. This is also an indirect recognition of management's rationale:

But let's say that you score 9 for one client, because you want yourself and your work to look good, and then the client is granted less or no support on that basis; that would be . . . that would be pretty bad. (Alma, Shelter 3)

The frontline staff who follow this recording logic generally express respect for and confidence in management's approach to steering the organization and the tools in use. Nevertheless, Alma earlier described how in one particular situation, when she took over responsibility for recording for a client from another colleague, she gave a score higher than her actual assessment out of consideration for the client's well-being. This social worker in general applies an organization-focused recording logic but tends to apply a recording logic focused on the assumed best interests of the individual client under certain circumstances. Thus, the organizational rationale doesn't exclude the existence of an occupational professional rationale, rather they exist side by side, depending on the situation.

The application of a recording logic by the book leads to recordings of performance data, in which the employee does his or her best both to follow the instructions and to register accurate data. From management's perspective, valid and accurate data will be produced. Nevertheless, this study shows that the nature of the situation and specific influencing factors may determine whether employees tend to apply a recording logic based upon a rationale of organizational professionalism.

Discussion

Our research corroborates Pollitt's awareness that the same employee or group of employees can apply different logics over time depending on both rational and irrational factors. A plethora of attitudes and considerations is in play.

As shown above, a given assumption or motive is not synonymous with the application of a certain logical rationale and does not predetermine a particular recording behavior. The specific recording behavior will instead reflect a combination of logical rationales, assumptions, attitudes, and motives that predominate in the recording situation. These influencing factors will come together differently. In some situations, the consideration of a client's immediate well-being or the appreciative relationship will prevail, while in other situations the tactical rationale will win out over the organizational rationale. The same employee or group of employees may juggle different recording logics, which ultimately can result in shifting recording behavior. This makes it challenging to determine the validity of the performance data produced.

This case illustrates how the various recording logics applied comprise some of the same attitudes and motives. Consideration for the individual client can be a determining factor for applying both a tactical and according to instructions correct logic, depending on the employees' individual assumptions or attitudes. And an attitude of distrust in the PM system can lead to the application of either a tactical or a recording logic by the book. Thus, it is not a specific rationale, attitude, or motive in itself, but rather a combination of these that is decisive in the actual recording behavior of the professional.

Overall, the social service workers might express a strong affiliation with an occupational rationale in which professional standards, autonomy, and the respect and concerns of the individual client are dominant. However, they will not always make recordings based on this rationale. If the implemented PM system is perceived as representing an organizational rationale of control, resistance or disregard of the measurement tool is likely to occur among the professionals. And if employees experience a considerable gap between their own values and interests and those of management, or uncertainty about how data are used, this may also lead to distrust and concern regarding management's intentions in the implementation of the PM system. All of this can give rise to various alternative recording

logics. Additionally, there may be situations in which employees hold their own occupational/professional standards separate from management's PM tools, and the two rationales exist concurrently.

The study also hints about the factors that actually contribute to the occurrence of alternative recording logics. We suggest that the larger the perceived gap between management and employee values and interests, the greater the difference employees will experience between recording data according to the guidelines and recording data according to the occupational professional rationale, consideration for the individual client, and consideration for the professional unit. This makes it more likely that frontline staff will record data not based upon the guidelines. The study also indicates that trust is a factor that plays an important role when employees record performance data. Lack of confidence in either the PM system or management's intentions has a significant impact on recording behavior. The greater the distrust, the more likely that staff are to not apply an intended recording logic. Instead, employees' tactical or moral concerns will lead to scores either higher or lower than the actual assessment or random recordings without proper assessments. The study thus shows that a lack of trust between local actors and central management can give rise to several different alternative logics that weaken the validity of the data produced.

The management's ability to communicate clearly the rationale for and the purpose of the PM strategy is crucial to both the issue of trust and how the gap between management and employee values and interests is perceived as was suggested by Hunter and Nielsen (2013). In this case, the communicated rationale behind implementing the PM system shifts between a discourse of top-down control and a discourse of bottom-up learning and self-regulation. At the same time, communication concerning how the data are being used is unclear. Altogether, it makes the core logic of the PM system seem vague and, to some extent, ambiguous.

As Pollitt (2013) argues, the logic that guides recording scores in the PM system depends on assumptions of how people think and what motivates them. As long as it is unclear who will be looking at—and using—the data, and as long as employees do not embrace the measurement tool and the produced data as useful in their daily professional work, there is plenty of room for speculation and interpretation of management's motives, intentions, and rationales. In this regard, the character of the implementation process and the PM system's ability to accommodate values, professional standards, and perspectives of occupational staff play a major role in how the PM system is perceived and received—and, thus, the validity of the data produced.

Implications

It is vital to acknowledge and deal with seemingly irrational components influencing the PM system, such as feelings of trust or distrust, assumptions of contrasting values, or perceptions of hidden agendas. The risk of emerging gaps is apparent, and implementation of PM might be undermined by the social service professionals if they regard it as an attempt by management professionals to impose its values and organizational professionalism on their professional work. The gap, or perceived gap, between local and central levels, and local professionals' distrust of the PM system and management's intentions, may thus jeopardize the PM system.

It is wise to bear in mind the importance of this field of tension when formulating and communicating the core logic of the implementation of PM in organizations. It is essential to relate PM to local issues, such as the extent to which employees experience that the specific PM system represents their values and rationales, as well as the extent to which there is a relationship of trust between employees and management. In addition, leaders and managers should be conscious of how their own understanding and communication about the PM system are likely to affect the behavior of the employees making the recordings. It behooves management to take local conditions into account, realizing that PM is dependent not only on properly functioning internal structures but also on relationships of trust within the organization.

It is also relevant to consider whether it is possible to develop a PM system based on a common professional rationale that accommodates the perspectives and interests of both frontline employees and managers. Under most circumstances, it is important to work on reducing the tension that might exist between employees and management. Since it is usually management that defines and implements PM systems, it is advisable to make an effort to approach and understand employees' occupational professional points of view, such as inviting them to participate in defining, planning, and disseminating the PM system. Of course, this implies waiving the desire for central control.

In summary, there are certainly ways to go and precautions to take in order to meet contextual and relational issues and thereby improve the quality of performance data. On the other hand, various unforeseen factors may come into play and affect individual recording behavior. Therefore, the validity of recorded performance data will always be subject to uncertainty. Our research demonstrates the complexity in this issue and outlines some of the logics that present challenges to the idea of PM. We hope it also provides opportunities to take meaningful action.

Acknowledgments

The authors wish to thank two anonymous referees and the editors of *American Journal of Evaluation* for important suggestions on previous versions of this article. Thanks also to Lise Greene and Evert Vedung.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

- Albæk, E. (2003). Evalueringens Fremtid—Fremtidens Evalueringer [The future of evaluation—Evaluations of the future]. In P. Dahler-Larsen & H. Krogstrup (Eds.), *Tendenser i evaluering [Tendencies in evaluation]* (pp. 260–270). Odense: Syddansk Universitetsforlag.
- Arnaboldi, M., Lapsley, I., & Steccolini, I. (2015). Performance management in the public sector: The ultimate challenge. *Financial Accountability & Management*, 31, 1–22.
- Björk, A. (2013). Working with different logics: A case study on the use of the Addiction Severity Index in addiction treatment practice. *Nordic Studies on Alcohol and Drugs*, 30, 179–199.
- Bliss, D. L. (2007). Implementing an outcomes measurement system in substance abuse treatment programs. *Administration in Social Work*, 31, 83–101.
- Brignall, S., & Modell, S. (2000). An institutional perspective on performance measurement and management in the “new public sector.” *Management Accounting Research*, 11, 281–306.
- Chopra, P. K., & Kanji, G. K. (2011). On the science of management with measurement. *Total Quality Management & Business Excellence*, 22, 63–81.
- Dahler-Larsen, P. (2001). From programme theory to constructivism: On tragic, magic and competing programmes. *Evaluation*, 7, 331–349.
- Dahler-Larsen, P., & Schwandt, T. A. (2012). Political culture as context for evaluation. *New Directions for Evaluation*, 135, 75–87.
- de Bruijn, H. (2007). *Managing performance in the public sector*. New York, NY: Routledge.
- Evetts, J. (2009). New professionalism and new public management: Changes, continuities and consequences. *Comparative Sociology*, 8, 247–266.
- Hansen, M. B., & Vedung, E. (2010). Theory-based stakeholder evaluation. *American Journal of Evaluation*, 31, 295–313.

- Hasenfeld, Y. (2010). The attributes of human service organizations. In Y. Hasenfeld (Ed.), *Human services as complex organizations* (2nd ed., pp. 9–32). Thousand Oaks, CA: Sage.
- Hunter, D. E., & Nielsen, S. B. (2013). Performance management and evaluation: Exploring complementarities. *New Directions for Evaluation*, 137, 7–17.
- Krogstrup, H. K. (2011). *Kampen om evidens: Resultatmåling, effektevaluering og evidens [The battle of evidence: Performance measurement, impact evaluation, and evidence]*. Copenhagen, Denmark: Hans Reitzels.
- Liljegren, A., & Parding, K. (2010). Ändrad styrning av välfärdsprofessioner: Exemplet evidensbaserad i socialt arbete [Changed management of welfare professions: Example of evidence-based social work]. *Socialvetenskaplig Tidskrift*, 17, 270–288.
- Lipsky, M. (2010). *Street-level bureaucracy: Dilemmas of the individual in public services*. New York, NY: Sage.
- Lynch-Cerullo, K., & Cooney, K. (2011). Moving from outputs to outcomes: A review of the evolution of performance measurement in the human service nonprofit sector. *Administration in Social Work*, 35, 364–388.
- Mayne, J. (2007). Challenges and lessons in implementing results-based management. *Evaluation*, 13, 87–109.
- Mayne, J., & Rist, R. C. (2006). Studies are not enough: The necessary transformation of evaluation. *Canadian Journal of Program Evaluation*, 21, 93–120.
- Nielsen, S. B., Jacobsen, M. N., & Pedersen, M. (2005). Øje for effekterne—resultatbaseret styring kan styrke offentlige indsatser [An eye for impacts—Result-based management can improve public services]. *Nordisk Administrativ Tidsskrift*, 86, 276–295.
- Nielsen, S. T., Bojsen, D. S., & Ejler, N. (2009). Introduktion til resultatbaseret styring [Introduction to result-based management]. In N. Ejler (Ed.), *Når måling giver mening: Resultatbaseret styring og dansk velfærdspolitik i forvandling [When measuring makes sense: Result-based management and Danish welfare policies in transformation]* (pp. 40–70). Copenhagen, Denmark: Djøf Forlag.
- Pollitt, C. (2006). Performance management in practice: A comparative study of executive agencies. *Journal of Public Administration Research and Theory*, 16, 25–44.
- Pollitt, C. (2013). The logics of performance management. *Evaluation*, 19, 346–363.
- Rist, R. C. (2009). På jagt efter troværdig evidens: At konstruere monitorerings- og evalueringssystemer indenfor områder med knappe ressourcer [On the hunt for trustworthy evidence: To construct monitoring and evaluation systems within areas of scarce resources]. In N. Ejler (Ed.), *Når måling giver mening: Resultatbaseret styring og dansk velfærdspolitik i forvandling [When measuring makes sense: Result-based management and Danish welfare policies in transformation]* (pp. 259–267). Copenhagen, Denmark: Djøf Forlag.
- Vedung, E. (2010). Four waves of evaluation diffusion. *Evaluation*, 16, 263–277.
- Vedung, E. (2011). Spridning, användning och implementering av utvärdering [Diffusion, impact, and implementation of evaluation]. In B. Blom, L. Nygren, & S. Morén (Eds.), *Utvärdering i socialt arbete: Utgångspunkter, modeller och användning [Evaluation in social work: Starting points, models, and use]* (pp. 285–299). Stockholm, Sweden: Natur & Kultur.
- Wandersman, A., Imm, P., Chinnman, M., & Kaftarian, F. (2000). Getting to outcomes: A results-based approach to accountability. *Evaluation and Program Planning*, 23, 389–395.
- Wenneberg, S. B. (2000). *Socialkonstruktivisme: Positioner, problemer og perspektiver [Social constructivism: Positions, problems, and perspectives]*. Fredriksberg: Samfundslitteratur.