
Annika Hasselblad, Sheila Zimic & Leif Sundberg


To link to this article: https://doi.org/10.1080/15228835.2023.2287241

© 2023 The Author(s). Published with license by Taylor & Francis Group, LLC.

Published online: 29 Nov 2023.

Submit your article to this journal

View related articles

View Crossmark data

Annika Hasselblad\textsuperscript{a}, Sheila Zimic\textsuperscript{a} and Leif Sundberg\textsuperscript{b}

\textsuperscript{a}Mid Sweden University, Sundsvall, Sweden; \textsuperscript{b}Umeå University, Umeå, Sweden

**ABSTRACT**
Contemporary welfare organizations engage in various evaluation practices to assess the quality of their services. In this paper we report a qualitative exploration of how technology-enabled evaluations are understood by organizational members who participate in quality assurance activities in Swedish social services. The study contributes to critical information systems literature, focusing on the tensions professionals experience in relation to the digital systems they use for evaluations. For example, “quantities” take precedence over the qualities of such work, as information systems constrain ambitions to realize knowledge-based social services. The results reveal three tensions in professionals’ evaluation-related activities arising from conflicting uses or desires. One is between desires for flexible systems that enable reflection and standardized digital support systems. Another is between uses or desires for indicators that are meaningful at the operational level and for general, comparable measures at the management level. The third is between desires to use evaluation procedures for learning and control. The study contributes to both theory and practice related to technology-enabled evaluation of welfare services, and critical perspectives on information systems.

**Introduction**

The development of the welfare state during the 20th and early 21st centuries has been accompanied by various kinds of evaluations of the activities and services involved (Bergmark, Bergmark, & Lundstrom, 2012; Power, 1997). No evaluations are performed in isolation as they reflect contemporary societal trends and discourses. They also inevitably involve use of, and interaction with, contemporary technologies. Thus, as with any narrative construct, we argue that it is important to critically examine welfare services’ evaluation from perspectives of the people who engage in evaluations in their everyday practice. Moreover, as technology plays an increasingly salient role in welfare organizations’ systematic evaluation,
it is important to scrutinize technology-related dimensions of evaluation practices. Such analysis is essential for understanding the types of knowledge that are considered important and reproduced in, for instance, digital evaluation systems and what is omitted. Examples of deployed systems include standardized databases for comparing different organizations based on a common set of indicators (Trygged, 2017). In line with the increasing use of information and communication technologies (ICT, hereafter “digital” systems) in evaluations for such purposes, scholars have referred to a trajectory toward an “audit society” rooted in political demands for accountability and control (Bowerman et al., 2000; Power, 1997) or an “evaluation society” for quality assurance (Dahler-Larsen, 2014).

Evaluation practices are not only enabled by the technologies operatively used in welfare programs, but also constrained and interactively affected by them. Further, the interactions may have much broader organizational and societal effects. Various authors have argued that ranking systems and calculative practices shift power away from professionals and toward managers and administrators as new assignments and titles are created to accommodate evaluation tasks (Kurunmäki & Miller, 2006; Miller & Rose, 1990; Shore & Wright, 2015). Examples include “controllers” and “quality managers” (Baines, 2006; Hjärpe, 2020). Their administrative and accounting tasks can be understood in the light of what Osborne (2006) described as the key elements of new public management (NPM), including strong attention to input and output control, evaluation, performance management, and auditing. These elements have been criticized for claimed negative effects on the public sector arising from associated foci on economic rationality, excessive trust in market-based discipline, managerial control, competition, and results assessment (Deakin & Michie, 1997, p. 1; Verbeeten & Speklé, 2015; Walsh, 1995). In social work, claimed negative consequences include a reduction in values, partly due to a focus on short-term service outputs rather than long-term outcomes (see, for example, Munro, 2004). Thus, there are several tensions between (techno)-economic rationalist views of evaluation and the values and practices that guide welfare professions, such as social services. However, few studies have focused on the micro level sensemaking (as Ek Österberg & de Fine Licht, 2021) and situated practices (Suchman, 1987) of actors who are engaged in quality management and have responsibilities to report numbers to managers and politicians.

Against this backdrop, the purpose of this paper is to explore how evaluation practices are perceived by respondents drawn from members of social service organizations (in Sweden) engaged in quality assurance activities. We specifically address the following research question (RQ): What tensions do quality management professions experience during technology-enabled evaluations of social services?
The rest of the paper is organized as follows. In “Background” section, the paper is positioned in relation to previous research on the evaluation of welfare organizations and critical information systems. Then, methodological considerations are presented in “Methods” section, followed by results in “Results” section. The results are discussed in “Discussion” section, and finally conclusions are offered in “Conclusions” section.

**Background**

Previous studies have clearly shown that evaluation systems influence diverse domains, often with “unintended consequences,” via constitutive effects (Dahler-Larsen, 2011, 2014; Nordesjö, 2021; see also Gillingham, 2019) as they become institutionalized (Andersen, 2021; Dahler-Larsen, 2014; Leeuw & Furubo, 2008). An example of such unintended consequences mentioned by Dahler-Larsen (2014) is the use of "waiting time" as an indicator to assess quality in healthcare. A narrow definition of waiting time that only includes time spent in emergency waiting rooms (and thereby excludes other locations where people wait) led to patients being relocated to corridors, in an effort to reach good quality regarding this indicator. Such effects generate tensions, manifested in the digital technologies applied and associated practices. Some of these tensions have been explored in previous research, including conflicts between different “epistemic cultures” (Wagner & Newell, 2004), and issues related to efforts to implement standardization and flexibility (Hanseth, Monteiro, & Hatling, 1996), with possibilities for customization (Sinsky, Bavafa, Roberts, & Beasley, 2021). Such tensions are highly relevant here, as technology-enabled evaluations require decisions about what should be measured, who should measure it, and designs of evaluations involving use of digital systems.

In the following “Evaluation as technology” and “Critical information systems” subsections, we outline previous research on welfare services’ evaluation then position the paper in critical information systems research. The objectives are to provide solid foundations for discussing how (digital) technology enables and constrains evaluation practices based on the understandings of selected practitioners.

**Evaluation as technology**

The quantification, measurement, and compilation of statistics in public welfare organizations are not new phenomena (Foucault, 1980). However, digital technology is increasingly used for these purposes, in modern welfare organizations generally (Dencik, 2022; Dencik & Kaun, 2020) and
in social services particularly in countries such as Canada (Baines, 2006), the United Kingdom (Kurunmäki & Miller, 2006; Munro, 2004) and Sweden (Hjärpe, 2020). According to Kurunmäki and Miller (2006), this increasing reliance on technology is a fundamental element of the modernizing government agenda, which creates expectations for organizations to adjust to more calculative practices.

In line with (Høydal, 2021), we challenge the notion of evaluation as an “objective” activity or practice that provides “neutral evidence,” recognizing that any data generated or compiled may not be objective and neutral (Dalton, Taylor, & Thatcher, 2016). Both the practices and perceived objectivity of the data may also vary substantially and be influenced by numerous factors. Accordingly, we aimed to obtain rich insights into how evaluation practices are framed in sociotechnical settings through the compilation or use of datasets that are always situated or framed to achieve certain goals (Gitelman, 2013; Mayer-Schönberger & Cukier, 2013). (Rose and Miller, 1992) argue that governmental programs require “technologies” to be operable, so the technologies are deployed in ways deemed appropriate to reach political goals (Stone, 2012). Thus, we understand the use of technology for evaluation as inherently political (Berg, 1998), value-laden, and socio-technical (i.e., situated in social contexts). Previous research has identified a contrast between techno-deterministic reasoning in government policies concerning technology and the daily practices in, for example, social care settings (Lindberg, Kvist, & Lindgren, 2022). These types of narratives in policy may lead to the materialization and stabilization of certain technological discourses (Akrich, 1992; Bijker & Law, 1994; Bowker & Star, 1999; Latour, 1987), and prevent the establishment of more flexible systems, and reflexive system designs (Gidlund, 2010).

More than 20 years ago Henfridsson (2000) noted the importance of participants’ meaning-making and perceptions of information technology in Swedish social work organizations. More recently, Lagsten and Andersson (2018) have highlighted missing links between information systems (IS) research and social work, identifying several important areas that require more research, notably mismatches between “social” and “system” conceptualizations, and the creation of data for accountability and quality assurance purposes. Building on and extending these authors’ work, here we explore quality management professionals’ perceptions of their daily work involving use of digital technologies for evaluating social services. Tensions are identified that strongly influence the meanings associated with the deployment of information technology, which may “ambiguously” both facilitate and constrain the acquisition, compilation and use of data in multiple ways.
Other studies suggest that the emphasis on implementation of technology has led to a greater focus on recording data rather than utilizing it for professional purposes, thus reducing social work to a technical practice rather than a service that is responsive to the needs of children and families (Devlieghere & Roose, 2018). In the same vein, Ylönen (2023) argues that the use of information systems in social work changes priorities and foci, and according to other previous studies (Huuskonen & Vakkari, 2015; Parton, 2006; White, Hall, & Peckover, 2009) it raises risks of de-contextualizing information, which can have adverse consequences for the clients. Although social workers try to resist these changes of priorities (Ylönen, 2023), several studies have found that they result in social workers spending most (60–80%) of their available time with the systems, for example managing digital records (White, Wastell, Broadhurst, & Hall, 2010). This may be partly due to clear deficiencies. For example, Seniutis, Petružytė, Baltrūnaitė, Vainauskaitė, and Petkevičius (2021) found that the absence of functionality in the IS used led to data duplication, resulting in additional time consumption in child welfare agencies. Such deficiencies (and tensions like those explored here) may clearly affect the workplace ecology, reduce the reliability of information, challenge data protection, exposing practitioners to excessive managerial control and ultimately limit the time they can directly work with clients.

**Critical information systems**

Critical information systems (CIS) literature is concerned with the contradictions that arise during the implementation and use of technology in social settings, and seeks to offer an alternative to more common managerial approaches (Richardson & Robinson, 2007). As our study includes a focus on the tensions that arise during quality management professionals’ encounters with digital systems in their evaluations, we apply the following three key elements of CIS as an analytical lens to facilitate discussion of our empirical materials.

First, we pay attention to context by questioning taken-for-granted stories about digital technologies’ effectiveness, acknowledging the importance of avoiding limiting consideration of system goals to narrowly defined cost-effectiveness versus human values (Lyytinen & Klein, 1985). Within organizations, the values and institutional characteristics of actors (Orlikowski, 1992) and their contextual practices (Suchman, 1987) influence individuals’ perceptions of technology. Thus, we recognize the importance of social context and political behaviors inherent in the development and use of technology (Stone, 2012). Accordingly, a major objective was to grasp how the respondents in our empirical study (responsible for quality management in social service
organizations) comprehended and utilized evaluation practices in their organizations.

Second, we challenge techno-rational (or “techno-economic,” see Avgerou & McGrath, 2007) management perspectives (Cecez-Kecmanovic, Janson, & Brown, 2002), for instance, by recognizing the limitations of purely managerial or engineering perspectives of IS innovation outcomes (Gidlund & Sundberg, 2021; Klein & Hirschheim, 1991; Kling, 1980; Kumar, Van Dissel, & Bielli, 1998; Ramiller, 2001; Robey and Markus, 1984). In alignment with perspectives rooted in science and technology studies (STS), CIS researchers have contended that ingrained techno-rational values stem from the positivist tradition and a "focus on the exclusive validity of objectified, systematized knowledge coupled with a distinct separation of facts from values" (Lee, 1991). Various researchers have pointed out that techno-economic rationality is too restricted or incongruent for comprehending the processes of IS innovation (Ciborra, 2002; Introna, 1997; Walsham, 2000). It is equally important to understand behavioral aspects of both the individuals and organizations, as well as the intricate socio-technical interactions involved.

Third, we acknowledge that designers inevitably inscribe certain values into technologies during the design process (Akrich, 1992). Examples include the optimization of functionality (Cecez-Kecmanovic et al., 2002), efficiency, rational planning, assumptions about political neutrality, technical objectivity, and subject-object dualism (Lake, 1993). As the term implies, CIS is concerned with unveiling and questioning the status quo, and what is taken for granted, to open possibilities for alternative realities, paths, interpretations, and in the case of technology, designs. Our CIS-inspired analytical lens is explained in more detail in the following summary of our methodological approach.

In sum, these three elements allow us to analyze the stories of quality management professionals on three levels: system (by questioning narrowly defined or taken for granted goals associated with these systems), organization (by recognizing limitations of only including managerial and engineering perspectives on technology use), and knowledge regime/strategy (by opening up for alternative realities, paths and designs)

Methods

In this section, we outline the research context, the methodological approach, and the materials created and used in the empirical study to gain an understanding of the perceptions of technology-enabled evaluations of participants in quality assurance activities of social service organizations. It draws on a qualitative study revolving around six focus group interviews (Stewart and Shamdasan, 2014) with officials engaged in quality assurance in social service organizations in a county in Sweden.
Research context

Social services are provided in Swedish municipalities by public organizations with political leadership mediated through one or more committees (most often a Social Welfare Board) that have responsibility for overseeing municipal activities, within legislative frameworks set by the national government. Municipal officials’ work is guided by the committees’ decisions. Evaluations in social services have several purposes, including assurance that operations: have sufficient quality in relation to politically decided goals, comply with relevant laws and regulations (e.g., the Social Service Act), and meet internal operational objectives and criteria regarding the quality of social care. The internal evaluations include statutory follow-up of authority decisions, e.g., the interventions for each individual or family receiving help from the municipal social services organizations.

The evaluations, particularly those of interventions, are regarded as important in efforts to ensure that social services have high quality and support evidence-based practice. However, a systematic approach must be applied in evaluation to contribute to evidence-based practice and organizational learning. Most of the evaluations depend on information documented in an information system. In social services a case management system, which contains detailed information that can be used for various evaluation purposes, is used in assessments with clients (most commonly Procapita/Lifecare, and Treserva in current Swedish settings). Other information systems that assist evaluation include, for example, registers (e.g., Palliative Care Register, and Register of dementia symptoms and interventions in elderly care), decision support systems (e.g., Stratsys) and larger evaluation systems in the form of performance measurement systems (e.g., Open Comparisons). During the past decade, standardization of inter-organizational evaluations has become prioritized in Swedish social services on local, regional, and national levels. The Swedish government and Swedish Association of Local Authorities (SALAR) established an agreement regarding support for evidence-based practice intended to ensure that social services have high quality, in which systematic evaluation practices play an important role. This study was part of an initiative to develop evaluation practices that support evidence-based practice more robustly, through collaboration enabled by a network of officials engaged in quality assurance activities in social services in municipalities of the focal Swedish county.

Empirical materials

In total, six municipalities in the county participated in the study over a period of six months. In each municipality, a focus group interview was conducted with participants selected by a responsible manager or
representatives from the quality network. Participants were selected because they were deemed competent for answering questions regarding quality assurance evaluation practices and information systems used in them. The number of participants in the interviews varied from three to nine, depending on how the municipalities were organized. Table 1 shows job titles of the participants in each focus group interview.

Focus group interviews were chosen as group interactions have known ability to generate richer empirical data regarding focal phenomena than individual interviews (Finch, Lewis, & Turley, 2003). The participants were comfortable with each other as each focus group consisted of officials who worked together on quality assurance in their organization. Discussions in the groups revolved around the topic of systematic evaluations and formed a reflexive process in which participants discussed different perceptions of the concept, validity and reliability issues when reporting data in information systems, and other challenges they typically faced in their evaluation-related work. The focus group interviews were semi-structured, with open-ended questions, enabling examination of the everyday ways that participants made sense of systematic evaluations (Lunt & Livingstone, 1996). The questions concerned:

- the participants' professional role
- how they would describe systematic evaluation (i.e., what it means to them),
- what types of evaluations were conducted and how they were used,
- the questions that evaluations were intended to answer, or knowledge needs they were intended to meet,
- whether evaluation results contribute to organizational development, and
- participants' hopes for evaluation practices in the future.

<table>
<thead>
<tr>
<th>Focus group</th>
<th>Job title</th>
</tr>
</thead>
</table>
| 1           | Activity strategist (1)  
             | Quality manager (3)     |
| 2           | Quality manager (4)     
             | Coordinator responsible for social services (1) 
             | Quality leader (1)     |
| 3           | Operational controller (1)  
             | Quality manager (2)     
             | Medically responsible nurse (1) |
| 4           | Board secretary (1)     
             | Specialist social worker (1)  
             | Operations manager (2)     
             | Assistant operations manager (social and elderly care) (1)  
             | First-line manager (2)     
             | Quality manager (2)     |
| 5           | Systems administrator (3)  
             | Quality manager (1)     |
| 6           | Systems administrator (1)  
             | Medically responsible nurse (1)  
             | Quality manager (1)     |

Table 1. Job Titles of Study Participants and Number of Participants in Each Job Title Category.

<table>
<thead>
<tr>
<th>Focus group</th>
<th>Job title</th>
</tr>
</thead>
</table>
| 1           | Activity strategist (1)  
             | Quality manager (3)     |
| 2           | Quality manager (4)     
             | Coordinator responsible for social services (1) 
             | Quality leader (1)     |
| 3           | Operational controller (1)  
             | Quality manager (2)     
             | Medically responsible nurse (1) |
| 4           | Board secretary (1)     
             | Specialist social worker (1)  
             | Operations manager (2)     
             | Assistant operations manager (social and elderly care) (1)  
             | First-line manager (2)     
             | Quality manager (2)     |
| 5           | Systems administrator (3)  
             | Quality manager (1)     |
| 6           | Systems administrator (1)  
             | Medically responsible nurse (1)  
             | Quality manager (1)     |

Total in all focus groups (30)
Before the focus group interviews, all participants received a table with questions to be used for guidance regarding types of evaluations in each field of activity (see Appendix B).

**Analytical approach**

Our analytical approach has been inspired by “critical grounded theory” (Charmaz, 2005; Hense & McFerran, 2016). Regarding our study participants as “knowledgeable agents” who could explain their thoughts, intentions and actions (Gioia, Corley, & Hamilton, 2013, p. 3), our initial analytical approach involved heavily engaging with the data using coding procedures described below. To broaden the focus, we then engaged in a critical inquiry, based on the CIS literature summarized in “Critical information systems” subsection. This provided both a sharper analytical lens rooted in grounded theory, and opportunities for deeper analysis based on CIS, or in the words of Charmaz (2005) “Combining the two approaches enhances the power of each.”

All interviews were recorded and transcribed. To draw rich insights (Walsham, 1995) from the qualitative data, we first approached the semistructured empirical materials using open axial coding (Corbin & Strauss, 1990; Gioia et al., 2013). This allowed us to build an understanding of the participants’ experiences of their encounters with technology during their daily work. We then generated second-order themes and aggregated constructs from the first-order codes derived from the transcripts.

The coding involved the following steps. Initially, the participants’ quotations were converted into first-order “codes.” Related codes were then grouped together in a table in Word. This enabled us to generate second-order themes based on the groups. Finally, three aggregated constructs (“tensions”) were generated based on the themes. The coding required several iterations, as expected in interpretative work.

During the research process we were sensitized (Timmermans & Tavory, 2012) by the literature presented in “Background” section. Thus, although we sought to apply and axial coding, we do not regard ourselves (or any researchers) as “empty vessels.” As noted by Cutcliffe (2000): “the researcher and all his/her knowledge and prior experience is bound up with the interactive processes of data collection and analysis” (see also Dunne, 2011).

Thus, our study resides in the intersection between approaches commonly associated with inductive stances and critical approaches that regard research as value-laden and advocates use of a priori codes. For more detailed outlines of combinations of these paradigms, e.g., “critical grounded

During the coding we paid specific attention to tensions between the participants’ practices and the action spaces provided by the information systems they used when creating data for evaluations. While outlining these tensions, we simultaneously and abductively (Dubois & Gadde, 2002) developed the theoretical framework presented in “Critical information systems” subsection with the three focal areas of concern from CIS literature to enable discussion of the implications of the tensions on three levels of abstraction (system, organization, knowledge regime/strategy). The mentioned tensions form the main results (“Results” section) are the foci of the discussion in “Discussion” section.

Results

The axial coding of the empirical material resulted in identification of three aggregate constructs (Figure 1):

i. Reflexive spaces versus standardized digital support systems

ii. Operational versus managerial indicators

iii. Learning versus control

We refer to these three constructs as “tensions” because they encompass ambivalences and differences in how evaluation practices were understood and applied by organizational members who engaged in quality assurance activities. These tensions are further outlined in the following subsections.

<table>
<thead>
<tr>
<th>1st Order Concepts</th>
<th>2nd Order Themes</th>
<th>Aggregated Constructs</th>
</tr>
</thead>
<tbody>
<tr>
<td>* The impossibility to compile information based on text feeding data into systems.</td>
<td>Wishes for more flexible systems</td>
<td>Reflexive spaces vs. standardized digital support systems</td>
</tr>
<tr>
<td>* Participants wish to be able to devote time for reflection and analysis instead of feeding data to systems.</td>
<td>Ambivalent attitudes to standardization</td>
<td></td>
</tr>
<tr>
<td>* &quot;Duty changes [...] So it would be good if we could fix a little bit ourselves and add our own parameters and build up parameters in the information system.&quot; P30</td>
<td>Aspects regarding individuals becoming invisible</td>
<td></td>
</tr>
<tr>
<td>* Numbers are perceived as irrelevant and meaningless.</td>
<td>Goals are perceived too high level to be useful</td>
<td></td>
</tr>
<tr>
<td>* Participant expressing needs to align organizational goals with the goals for individual care takers.</td>
<td>Emphasis on organizational learning</td>
<td></td>
</tr>
<tr>
<td>* Numbers are perceived as irrelevant and meaningless.</td>
<td>Focus on using technology as a means for control</td>
<td></td>
</tr>
<tr>
<td>* Formal execution plans do not encapsulate operational practices.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>* Wishes to systematically evaluate individuals’ satisfaction.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>* &quot;Everything we do should benefit individuals that we are helping&quot; P8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>* Participants want to generate knowledge for practitioners who are working closest to the people we are here for.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>* Participants want to use the evaluations to improve the entire social services organization.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>* Participants emphasize the importance of providing feedback to clients and practitioners to foster meaningfulness and engagement.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>* Participants refer to using evaluations to detect errors to know how legal requirements are being followed.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>* Participants refer to using evaluations as a means for error detection.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>* &quot;It feels like we are delivering [information] only to satisfy a system.&quot; P17</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1. Coding Results.
Tension 1: reflexive spaces versus standardized digital support systems

The first tension is related to the functionalities of digital systems used for evaluative practices. The participants expressed views that digital systems should support the extraction of data that could be used to compile statistics and make comparisons with other units and/or municipalities. They expected results of some key, required evaluation tasks to be reported in, for example, a system for performance measurements, such as Open Comparisons (Swedish: Öppna Jämförelser). However, instead of easily extracting information from the case management system, the participants experienced difficulties matching datasets with the information requested by external actors. A lack of compliance between the requested information and available data led the quality management professionals to manually gather information from the system. Sometimes they even had to consult specific documentation stored in dossiers. The participants described this manual gathering of the data as time-consuming and were concerned about having to choose which variables to include when reporting data in national surveys or systems for performance measurements. Hence, they expressed views that the manual procedures reduced trust in the numbers since elements of interpretation and subjectivity were embedded in the practice. Moreover, the participants expressed discomfort about acting as “interpreters,” using phrases such as “This does not feel good” and “I don’t want to hand [the data output] over.” They also questioned the process of creating indicators at a national level, which they described as lacking the involvement of professionals with information system competence:

To know what to extract you have to have knowledge about details, you have to know exactly what is stored in the system. (Focus group 1, Quality manager)

The lack of trust in numbers also meant that the statistics compiled from national surveys such as Open Comparisons were not perceived as reliable. Some participants even described using the data for comparisons as “dangerous” and “risky.” In other cases, the results were perceived as “a number no one cares about.” To avoid these risks, the participants desired a common structure. They all agreed that they should have the same codes in their case management systems, so that the extracted data would be comparable between units or municipalities, thereby enhancing the reliability (and hence utility) of the end results. For example, the National Board of Health and Welfare requested information that social service organizations in local governments were obliged to report on a regular basis. Participants described this data as structured (standardized) as it specified exactly which variables to include.
Conversely, they expressed a desire to be able to extract data “on demand” and easily “dig out more” at any time from their information systems, such as case management systems. One participant emphasized the impossibility of predicting future needs, and expressed a desire for social services to have the flexibility to include their own parameters in systems they use to meet new challenges they face as society changes.

Society changes [...] so it would be good if we could fix the system a bit ourselves and add our own parameters and build up [parameters in the information system].

(Focus group 6, Quality manager)

Furthermore, participants frequently mentioned that an optimal system should have the potential to include “everything,” i.e., all types of information that could be of use for any relevant actors and organizational units. Thus, the participants experienced tensions between the static systems available to conform to standardized evaluations, and a longing for more flexible systems that would enable more reflexive approaches (which we refer to as “reflexive spaces”).

**Tension 2: operational versus managerial indicators**

The second tension concerns the use (or non-use) of results of the evaluations, and in whose interest they are produced. The participants frequently discussed a desire for their organization to act on the results of evaluations, and generally expressed a perception that significant effort was expended on reporting information, with limited utility as an outcome. While the actors could account for the type of information they had to report, for example, to the National Board of Health and Welfare, they often could not explain what the information was used for. Indeed, some participants said that they were curious about what central actors did with all of this information. The type of information they reported made little sense to organizational members at the operational level, as the results—the knowledge products—were not new knowledge to them. The compiled results were perceived as “checkpoints” or “measurements of the temperature” of the organization. They felt that to be useful and make sense to organizational members at the operational level the measures should reflect the organization’s goals. Much of the discussion during the focus groups revolved around the organization’s aims, purposes and work performed to produce benefit or meaning for the individuals that social services are intended to help and how evaluations could be expressed and aligned with organizational goals. However, the social services’ goals were not perceived as self-evident. One participant referred to differences in goals related to organizational levels. For example, a local welfare board’s goal (e.g., “to base operations on stable core
values”) clearly differs from that of an individual (e.g., “to learn to walk independently with the help of a walker”). Different actors were identified as having different interests, and the tension in this context was to distinguish what was relevant to measure in relation to goals of different levels. The participants expressed a belief that indicators could have a direct or indirect relationship with quality of work. They repeatedly returned to the question “What is it that we want to know?” A common topic was an interest in knowing whether social service efforts were helpful for recipients of the services. However, these efforts are not easy to measure. Some discussions concerned individuals’ action plans, which should capture elements such as the individuals’ needs and goals of the assessed interventions. The participants agreed that the quality of action plans varied. While some could be detailed and specific, others were abstract and difficult to evaluate. Furthermore, some indicated that action plans did not capture everything, such as human interactions and the quality of the relationship between a practitioner and a client:

The action plan is the single most important document in operational work. But it doesn’t capture everything. It doesn’t capture how you respond to a client. How should we measure that? How should we measure that I am kind and warm, do what I’m supposed to, and respect you and your integrity? How do I measure that? It [The action plan] captures a lot, and it serves as a base, but there are things that do not fit in there. Like, treatment of people, the base values, well, my perspective on human beings, and my personality—all this doesn’t fit in there [in the document]. (Focus group 2, Quality manager)

This quotation illustrates the tension between aspects of work quality that can and cannot be measured. The participants highlighted a central value conflict between ensuring that work was explicitly described and measurable in information systems (or documents) and the difficulty of capturing the quality of professional work at the operational level.

The participants also argued that individuals’ experiences should be systematically captured at the end of an intervention. Although clients’ experiences appeared to be the most relevant type of information, they were often excluded or unreliably captured (e.g., through user surveys with people suffering from dementia). Participants believed that clients’ voices, if systematically captured, could be important sources of knowledge for social workers, and thus useful operational-level indicators.

**Tension 3: learning versus control**

The third tension concerns the purpose of evaluations. Many discussions focused on the challenges of reporting information in a system for performance measurements (such as Open Comparisons). Moreover, the
reporting itself was depicted as a cumbersome task, with immense workloads associated with reporting information, and physical reactions such as “one gets sweaty when it [the reporting duty] comes.” This was closely related to a lack of clarity regarding the purpose of evaluations. One participant stated, “It feels like we’re delivering [information] only to satisfy a system.” These experiences were coupled with the struggles of making the results of evaluations meaningful and useful within the organization and reporting information.

The participants also repeatedly returned to the issue of meaning, asserting that each evaluation should reflect the work quality and benefit for the clients. They also emphasized the importance of providing feedback to clients and practitioners to foster meaningfulness and engagement, as they expressed “detachment” from evaluands in terms of loss of meaning:

For those of us who extract statistics, it takes a lot of time, and then you’d want to know it’s something sensible. Otherwise, we’re just doing things in vain. It doesn’t feel good. We have other things to do, so to say. (Focus group 5, Systems administrator)

The lack of clear meaning and purpose of evaluations created tension over usability of the evaluation systems within the organization. In some cases, the participants called for a clearer relationship between a quality management system and evaluations that could be fed into the system to embed them in a more organizationally meaningful way. Regarding the purpose of comparisons, the participants unanimously expressed a need for more standardized requirements that could provide more equal comparisons. However, as mentioned, they also wished for more flexible systems from which they could “dig out” locally pertinent data. As standardization and alignment with evaluation systems invoke a sense of “control,” the urge for flexibility increasingly corresponds to “learning.” Interestingly, the participants did not express a position for or against one or the other: rather, they called for greater clarity regarding the purposes of evaluations and more stringent alignment between expressed and actual uses. They also wanted evaluations to serve as analytic tools (some for the purposes of control and others to enable learning) that could support the improvement of social services.

The politicians may be more interested in one thing, caretakers another, and the National Board of Health and Welfare a third. Some issues may be shared but there are probably also different interests in play. (Focus group 5, Systems administrator)

Discussion

Figure 2 presents the three tensions identified in our in-depth analysis of quality management professionals’ experiences of evaluation practices,
distinguishing between two positions of welfare and techno-economic practices, divided into three levels: system, organization and knowledge regime/strategy. These three tensions are discussed in more detail in the following text.

The results of this study provide insights into tensions that may arise in the wake of technologically enabled evaluations, particularly evaluations conducted for quality assurance purposes in government welfare services. As digital systems rely on formalized knowledge in the form of data created for comparing welfare services, they are subject to quantification and standardization. In line with Hjärpe (2020), we identified a tension between the pursuit of measurability and calculative practices and perceptions that these practices fail to incorporate the richness of important aspects of work in sectors such as social care. We also detected a paradox as the interviewees expressed desires for more flexibility in their reporting systems to support more reflexive approaches (Gidlund, 2010), but also called for more standardization to enable relevant comparisons. Thus, there were both resistances to standardized systems (Timmermans and Epstein, 2010) and a call for greater standardization.

Evaluation systems fail to fully capture rich practices in the provision of care and support for human beings in the welfare environment, as they favor the generation and compilation of comparable “numbers” (Noordegraaf, 2008). At the same time, what may appear to be comparable, fixed, and stable is always preceded with negotiations (Akrich, 1992; Bijker & Law, 1994; Bowker & Star, 1999; Latour, 1987). In addition, the actors responsible for reporting data in technology-enabled evaluations sometimes describe their recordings as dependent on their own attitudes, assumptions, and motives (Denvall, 2015). Since evaluations require technologies (Rose

---

**Figure 2.** Tensions Between Welfare Practices and Techno-Economic Practices.
and contemporary technologies are digital, they also reflect the conditions in which they were created. As noted in our case, the digital systems used in evaluations encourage techno-rational management practices (Avgerou & McGrath, 2007; Cecez-Kecmanovic et al., 2002), which may result in them being a poor fit for welfare systems (Gillingham, 2019; Hasselblad & Sundberg, 2020).

Facts, in the form of “numbers” and “hard data,” are the preferred forms of information entered into databases used for evaluations as they enable objective comparisons between different units of welfare services. However, they are also deemed irrelevant due to their failure to capture operational-level practices, which have strongly subjective elements. Thus, prioritizing the acquisition and recording of “objective facts” rather than subjective “values” leads to measures that lack meaning according to the interviewees, who questioned the utility of knowledge that evaluations are expected to provide for learning. Therefore, in line with Andersen (2021), we observed that epistemological coherence, as manifested in digital evaluation systems, may lead to increased use but decreased usability of evaluations. Thus, we argue that “epistemological fixation” (Andersen, 2021, p. 39) of techno-rational practices may lead to managers receiving data that are relatively easy to report, but have limited relevance in social work practices.

In addition, the tensions and struggles presented in this paper suggest that evaluation practices are not—or, rather, should not be—seen as static but rather dynamic and not monolithic but multi-dimensional. However, unless these practices are accompanied by relevant and more symmetrical representations, e.g., between the evaluator and evaluand (Andersen, 2021, p. 43), in the mediating technologies the results may provide poor indications of “quality” as they may hinder (or at least fail to promote) learning processes within an organization. Moreover, information systems in their current forms cannot support infinite logics, as they are built on notions of digital data structured in databases. Thus, technology has its own constraints, and essentially different logics must probably be accompanied by equally different technological systems. As previous research in critical information systems has indicated, the values embedded in technological systems tend to serve the techno-rational management logic (Avgerou & McGrath, 2007; Cecez-Kecmanovic et al., 2002) and can thus be combined with NPM. Moreover, the ideal of rational knowledge following structural natural science-based methods and rendering quantitative information objective and neutral (Iliadis and Russo, 2016) influences what is inscribed in technological systems and the ways in which systems enable or constrain possibilities for professional judgments and learning. Thus, it may be necessary to dissolve closures of control knowledge regimes according to techno-economic practices in evaluation systems to provide more openings...
for reflections in practice and raise the prominence of a learning-oriented knowledge regime, thereby potentially balancing the tensions between welfare and techno-economic practices.

**Conclusions**

By rooting our study in previous studies on the evaluation of welfare services and CIS literature we drew rich insights concerning our RQ (What tensions do quality management professions experience during technology-enabled evaluations of social services?). The results, based on first-hand experiences recounted by professionals engaged in evaluation practices, reveal gaps between participants’ everyday practices and functionalities of available technological systems. Currently applied technology strongly promotes a focus on quantifiable measurements and statistics, while strongly hindering inclusion of quality-related aspects of work in evaluations, according to the participants. This gap is manifested through three tensions:

i. Reflexive spaces versus standardized digital support systems
ii. Operational versus managerial indicators
iii. Learning versus control

Our findings provide both theoretical and practical contributions to CIS literature and social work practice, as summarized in the following subsections.

*Theoretical contribution.* As we empirically investigated evaluation practices from a micro-level perspective, we identified tensions that arise when quality management professionals interact with digital technologies intended to enable them to perform their evaluations. By doing so we unveil power asymmetries between welfare and techno-economic practices embedded in current evaluation technologies on three levels of abstraction: system, organization and knowledge regime. This contributes to CIS literature in the form of a structure to facilitate critical inquiries about the evaluation of sociotechnical settings such as social services. In line with previous research, we argue that there is a need for much closer collaboration between experts in social work and information systems. Our study paves the path for a much closer linkage between these previously separate domains. By opening up and unveiling the tensions we challenge the techno-rational practices and open avenues for alternative ways to understand, talk about and make conscious choices regarding evaluation systems.

*Practical contribution.* Evaluations have a sociotechnical nature, so the information systems involved can (and should) be re-negotiated and refined in practice, through ongoing interaction between social workers and system
designers. Social work practitioners could gain insights and learn about their own practices, for instance by scrutinizing the content of standard systems by translating and expressing their practices through technical code. Meanwhile systems designers could explore limitations of their technical systems by increasing their understanding of the complex practice of social work. Such mutual shaping could, as a boundary spanner, create possibilities for more relevant and innovative systems through pursuit of more responsive technical design, and thus understanding of both the complexity of social work practice and technical code. However, if the dominant knowledge regime is too heavily aligned with standardized systems for managerial control there seems to be limited action space to gain knowledge by learning. This may hamper or even eliminate possible incentives for organizational change.

**Limitations and strengths**

This research was not without limitations. As with all qualitative studies, boundary conditions include contextual and temporal framing of the involved organizations. As we relied on focus group methodology, the results depend on the interactions between the participants where certain individuals and narratives may become dominant and others (consciously or unconsciously) suppressed (see, e.g., Smithson, 2000). Meanwhile, our sample of participants relied on dialogue with managers and is thus not representative of all workers in the organizations. Thus, we strongly encourage future researchers to conduct similar studies in other contexts and from other professional perspectives. Moreover, our methodological approach entailed the creation of a model including constructs linked to three dimensions (system, organization, knowledge regime). These levels should not be understood as separate entities, but as interrelated with more blurry boundaries than our graphical representation suggests.

Despite these limitations we obtained rich insights into common but understudied practices in contemporary welfare societies. Our study also contributes an important structure for analyzing and critically scrutinizing technologically supported welfare evaluation, which could be applied and extended in other settings.

**Directions for further research**

By critically examining the underlying logics of technology-enabled welfare evaluation, we believe that this study paves at least three fruitful paths for future research. First, we have supplied CIS scholars with a structure, or framework, to unveil the rationality behind evaluations. We encourage scholars to build upon and extend this framework. Second, we welcome studies focused on specific emerging technologies, such as AI and ChatGPT,
as their use is introduced into welfare services. This is highly important to prevent the un-reflexive use of new technology in these settings. Third, we look forward to additional studies in other national settings, to broaden the knowledge foundation of technology-enabled welfare evaluations.

**Disclosure statement**

No potential conflict of interest was reported by the author(s).

**References**


**Appendix A**

<table>
<thead>
<tr>
<th>Focus group</th>
<th>Job title</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Activity strategist (P1)</td>
</tr>
<tr>
<td></td>
<td>Quality manager (elderly care) (P2)</td>
</tr>
<tr>
<td></td>
<td>Quality manager (individuals and family care) (P3)</td>
</tr>
<tr>
<td></td>
<td>Quality manager, information technology (P4)</td>
</tr>
<tr>
<td></td>
<td>Quality manager (home care service) (P6)</td>
</tr>
<tr>
<td></td>
<td>Quality manager (non-institutional care) (P7)</td>
</tr>
<tr>
<td></td>
<td>Quality manager (retirement/ care homes) (P8)</td>
</tr>
<tr>
<td></td>
<td>Coordinator responsible for social services (P9)</td>
</tr>
<tr>
<td></td>
<td>Quality leader (P10)</td>
</tr>
<tr>
<td>2</td>
<td>Operational controller (P11)</td>
</tr>
<tr>
<td></td>
<td>Quality manager (social and disability care) (P12)</td>
</tr>
<tr>
<td></td>
<td>Quality manager (social and elderly care) (P13)</td>
</tr>
<tr>
<td></td>
<td>Medically responsible nurse (P14)</td>
</tr>
<tr>
<td>3</td>
<td>Board secretary (P15)</td>
</tr>
<tr>
<td></td>
<td>Specialist social worker (P16)</td>
</tr>
<tr>
<td></td>
<td>Operations manager (individuals and family care) (P17)</td>
</tr>
<tr>
<td></td>
<td>First-line manager (children, young people and families) (P18)</td>
</tr>
<tr>
<td></td>
<td>First-line manager (children and young people) (P19)</td>
</tr>
<tr>
<td></td>
<td>Operations manager (social and elderly care) (P20)</td>
</tr>
<tr>
<td></td>
<td>Assistant operations manager (social and elderly care) (P21)</td>
</tr>
<tr>
<td></td>
<td>Quality manager, information technology (P22)</td>
</tr>
<tr>
<td></td>
<td>Quality manager (social services and school) (P23)</td>
</tr>
<tr>
<td>4</td>
<td>Systems administrator (individuals and family care) (P24)</td>
</tr>
<tr>
<td></td>
<td>Systems administrator (social care) (P25)</td>
</tr>
<tr>
<td></td>
<td>Systems administrator (social care) (P26)</td>
</tr>
<tr>
<td></td>
<td>Quality manager (social and disability care) (P27)</td>
</tr>
<tr>
<td>5</td>
<td>Systems administrator (individuals and family care) (P28)</td>
</tr>
<tr>
<td></td>
<td>Medically responsible nurse (P29)</td>
</tr>
<tr>
<td></td>
<td>Quality manager (P30)</td>
</tr>
</tbody>
</table>
Appendix B

<table>
<thead>
<tr>
<th></th>
<th>What kind of evaluations are conducted in the area? (local, national etc.)?</th>
<th>What kind of tools are used to conduct evaluations?</th>
<th>How are compilations of information conducted in the systems?</th>
<th>How are results presented?</th>
<th>What kinds of report are possible to conduct?</th>
<th>How are reports used?</th>
<th>Who receives and takes part in the reports?</th>
<th>What kind of information or knowledge is desirable?</th>
<th>Who wants to know?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Care and support for children and young people</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial aid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drug and dependent care</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social psychiatry and disability care</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Violence in close relationships</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elderly care</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interdisciplinary (unit-wide)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>