Inspection of social services in Sweden: a comparative analysis of the use and adjustment of standards

Eva Hämberg & Thomas Sedelius

To cite this article: Eva Hämberg & Thomas Sedelius (2016) Inspection of social services in Sweden: a comparative analysis of the use and adjustment of standards, Nordic Social Work Research, 6:2, 138-151, DOI: 10.1080/2156857X.2016.1156015

To link to this article: http://dx.doi.org/10.1080/2156857X.2016.1156015

© 2016 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

Published online: 08 Mar 2016.

Submit your article to this journal

Article views: 126

View related articles

View Crossmark data
Inspection of social services in Sweden: a comparative analysis of the use and adjustment of standards

Eva Hämberg and Thomas Sedelius

School of Education, Health and Social Studies, Dalarna University, Falun, Sweden; School of Law, Psychology and Social Work, Örebro University, Örebro, Sweden

ABSTRACT

State supervisory instruments in the social services sector have increasingly been regarded as one of the most important safeguards for promoting the quality of, and public confidence in state-regulated activities. In several countries, there have been a number of organizational changes and changes in instrument design aimed at improving the effectiveness of supervision within the social service sector. However, little attention has been given to the impact of different standard designs on supervision. The aim of this study is to empirically contribute to a broadened understanding of how standards in supervision are designed, and what implications the design has on the effectiveness of social service supervision. In this study we employ a systematic and comparative analysis of the content (in terms of input, process, output and outcome aspects) and precision (low, medium and high) of 186 standards used in completed state supervisory decisions within two different types of social work: Investigation, Assessment and Decision-Making (IAD); and Treatment Work (TW) in Sweden during 2012. Contrary to expectations based on a supervision system ideal, the findings show that output aspects, i.e. factors related to long-term client effects, are rarely used and not more so in TW than in IAD supervision. In addition, the analysis of standard precision indicates that supervision of TW cannot identify deficiencies related to complex and subjectively experienced aspects to any greater extent than IAD. The overall results indicate that the supervisory instrument is primarily adapted to IAD-supervision, but is less suited to optimally covering central aspects of TW.

Introduction

State inspections have increasingly been seen as one of the most important methods for ensuring quality in state-regulated activities and for promoting public confidence in them (Bevan and Hood 2006). This method has two features worth underlining here: the standards for such inspections are derived from legislation; and those subject to such inspections must comply with the demands made therein (Bengtsson and Ek 2013).

In countries such as Sweden, critics have charged that mainly non-disciplinary or ‘soft’ inspections are ineffective at preventing maltreatment and scandals within the social services. This has led to
government investigations and reforms. The purport of these reforms has been to replace ‘soft’ inspections with standards-based and mainly disciplinary control systems (SOU 2002:14; SOU 2004:100; SOU 2007:82).

There is now a quite substantial body of research across different disciplines on the general pros and cons of control system based policy instruments. Hood et al. (1999) depict such instruments as ideally based on three interrelated components: a standard-setting component, an information-gathering one, and a behaviour-changing one. The major function of the standard-setting component is to set out demands on the controlled phenomenon. Standards help ensure transparency, predictability, and a focus, enabling policy-makers to develop a communicable vision and to formulate demands about what is to be achieved and how (de Lancer Julnes 2006; van der Knaap 2011). However, these presumably positive attributes of control-based instruments also present logical grounds for criticism. Instruments of this kind, several scholars contend, are inappropriate for many situations. They are based on single-loop learning, and they are overly simplistic, excessively static, and potentially misleading. They are accordingly inadequate for capturing critical dimensions of programme quality or for suggesting improvements (Argyris 1976; Greene 1999; Abma and Noordegraaf 2003; Bevan and Hood 2006; de Lancer Julnes 2006; Martin et al. 2010; van der Knaap 2011). Such drawbacks are particularly salient when the knowledge base is weak, when subjective experiences are key factors, or when complex problems have to be solved (Greene 1999). In contexts such as these, other methods offer more promising avenues for improvement. Examples of such instruments include double-loop or dialogue-based approaches, as well as methods based on power-sharing among competent participants (Argyris 1976; Hämberg 2013).

The social services are a field where the use of control-based policy instruments – inspections, audits, performance measurements, etc. – is increasingly common (Munro 2004; Tilbury 2004; Davies and Gregory 2010; Ek 2012; Lindgren 2014). It is also a field where professionals must carry out complex tasks, where solutions often have to be found on a weak knowledge base, and where relational aspects are of great importance for outcomes (Munro 2004; Beadle-Brown, Hutchinson, and Mansell 2008). These conditions have raised questions about how inspections in areas like social work can be done in more appropriate and adequate ways (Daniel 2003; SOU 2007:82; Nielsen and Ejler 2008). One suggestion proffered by several analysts is to adjust the design of standards to the character of the phenomenon being inspected (Helsby and Saunders 1993; de Lancer Julnes 2006; van der Knaap 2011; Abma and Noordegraaf 2003). When such adjustments are made, the content and precision of the standards employed are crucial for the functioning of the policy instrument (Braithwaite, Makkai, and Braithwaite 2007).

Several studies have addressed the design of inspection standards in the field of social work, but they have done so in general and mainly theoretical terms (Munro 2004; Clegg 2008; Davies and Gregory 2010). Few empirical studies have been done on the standards used in social-service inspections, and we still know very little about the extent to which such standards are actually adjusted to the differing nature of different social services.

The overall aim of this study is to contribute to an improved empirical understanding of how standards are designed in the area of social-service inspections, and to throw light on the extent to which such standards are adjusted to the differing character of different social services.

Our study is based on unique empirical data consisting of 189 inspection reports by the Swedish National Board of Health and Welfare (NBHW). The focus of these reports is on Personal Social Services (PSSs), which are provided under the aegis of Swedish municipalities to individuals and families suffering psycho-social problems of various kinds. We compare – in terms of their content and precision – a number of paragraphs in laws and regulations that set out standards for the inspection of two types of social services: Investigation, Assessment, and Decision-making procedures (IADs); and institutionally provided Treatment Interventions (TIs). In the case of IADs, the individual’s need for and legal right to PSSs are investigated, and a package of benefits and services is selected for him/her: Is a TI required? Is child protection needed? Is compulsory care advisable? The TI as such includes the provision of treatment aimed at helping individuals solve their psycho-social problems and improve
their social functioning. This can involve programmes for parents with reduced parenting capability, for youths with a pattern of criminal behaviour, and for adults with a history of drug addiction.

We raise the following research questions: What kinds of legal standards are used in the inspection of the two types of social-service inspection (IADs and TIs)? Are there any differences in design between the standards used in the two types of inspection? To what extent is the design of these standards adjusted in order to capture the complexity and unpredictability of social work?

The remainder of the article is divided into four main parts. In the first we discuss the function of standards, and the relevance of content and precision in the standards used in the inspection process. In the second, we explain the differing character of IADs and TIs, and we provide a brief description of inspection as a regulatory instrument in the Swedish context. In the third section we present our methods and materials, followed by our empirical results. The latter are structured according to the two main stages of the inspection process: i.e. the information-gathering stage and the evaluative stage. We conclude the article, finally, with a discussion of the adaptability of control-based policy instruments as such, and their capacity to capture different aspects of social work.

**Control-based policy instruments**

We see control systems as based on four interrelated components that serve different functions: a standard-setting component, an information-gathering one, an evaluative one, and a behaviour-changing one (Hood et al. 1999; Boyne, Day, and Walker 2002; Hämberg 2013). The standard-setting component serves as a norm. The information-gathering component provides the system with information about the controlled phenomenon. The evaluative component makes judgments about the phenomenon in question. The behaviour-changing component, finally, creates incentives for action.

Control systems are found in a variety of contexts, government regulation being but one. Policy instruments may be described as tools or ‘techniques by which governmental authorities wield their power in attempting to ensure support and effect social change’ (Bemelmans-Videc 1998, 3). Inspections, audits, and certification processes are three different policy instruments which are also designed as control systems (Hood et al. 1999). How the four components mentioned above are empirically processed within these instruments may vary. In the case of inspections, the standard-setting component largely emanates from legislative sources, including laws and regulations (Bengtsson and Ek 2013).

**The content and precision of standards**

On the basis of theories of control systems, we can identify two distinct normative functions for the standard-setting component. First, it specifies what to gather information about; second, it forms the basis for comparison and judgment. The theoretical concept of standards relates here to the functioning of this component, but without presupposing any restrictions on how it is designed. Thus, the content of a standard can be described more precisely or less, and it can be formulated in qualitative or in quantitative terms. As several researchers have shown, the content and precision of standards are of decisive importance for the functioning of control systems (Hood et al. 1999; Braithwaite, Makkai, and Braithwaite 2007; Hämberg 2013). The content of standards specifies the aspects on which the control system is to focus. Researchers have described the content of standards in terms of input, process, output, and outcome (Helsby and Saunders 1993; Scheirer 2000; de Lancer Julnes 2006; Winter 2006). These aspects do not completely exclude others; nor do they cohere with each other fully. They are all based, however, on a common logic in which they form the basic components of a process (Scheirer 2000). Input standards can be defined as ‘the structures and support which need to be set up or provided to better enable the desired processes to take place’ (Helsby and Saunders 1993, 69). Process standards describe ‘what needs to happen within institutions […] in order to achieve the desired outcomes’. (69) Output standards specify what is being delivered in terms of ‘performance of the implementers’ (Winter 2006, 159). Outcome standards, finally, describe the ‘effects on [the] target population’ (159).
Outcomes can be divided in turn into short-term and long-term ones (Scheirer 2000). By using different standards that cover different aspects, the control system is better able to reconstruct the logic of the programme itself and to evaluate the success of said programme (Scheirer 2000).

In terms of how well its content is described, each standard can be more precise or less. Hämberg (2013) has argued that the precision of standards makes a great difference for the information-gathering and behaviour-changing components of a control system. More precise standards provide more exact guidance on what kind of information the system is to gather, as well as on what requirements are to be imposed upon the controlled subject. Where the behaviour-changing component is concerned, a higher degree of precision in standards legitimates the use of more coercive means; a lower degree of precision legitimates the use of less coercive and more voluntary means.

Quality and validity of standards

The quality of public services is a specific dimension of interest to evaluate and control by means of standards-based instruments, but it is also a dimension which exposes such instruments to certain validity problems (Greene 1999; Munro 2004; Tilbury 2004; de Lancer Julnes 2006; Beadle-Brown, Hutchinson, and Mansell 2008; Malley and Netten 2008). The quality of public services is specifically mentioned in the Social Service Act (SSA). Quality is something experienced, and it is based on value judgments; accordingly, a complete consensus can hardly be reached on which content aspects should be covered by standards (Greene 1999; Donabedian 2005). Limitations in the ability of standards to capture complex and subjective experiences have led many scholars to call for complementary evaluative instruments (de Lancer Julnes 2006; Nielsen and Ejler 2008; Martin et al. 2010; van der Knaap 2011). Several ways to improve the validity of standards-based instruments – and thus their effectiveness – have been suggested. One way is to use a broad or balanced array of standards that cover different aspects of the evaluated phenomenon (Tilbury 2004; Donabedian 2005; de Lancer Julnes 2006). Another way is to combine the use of precise standards and of vague ones (Braithwaite, Makkai, and Braithwaite 2007). A third way is to involve target groups in the definition of standards or the setting of goals (Davies and Gregory 2010). In this study, we investigate the extent to which the first two strategies – i.e. the use of an array of standards covering different aspects of the supervised phenomenon on the one hand, and the use of a mixture of precise and of vague standards on the other – are applied to the inspection of social services in Sweden. We spell out our assumptions below about how this logic works in connection with the two kinds of social work under study.

Investigation, Assessment, and Decision Procedures; and Treatment Interventions

Social services are handled by human-service organisations (HSOs), the three main tasks of which are to protect, to maintain, and to enhance the well-being of their clients (Johansson, Dellgran, and Höjer 2015). The technologies used in the performance of these tasks can be grouped into three ideal-types: people-processing technologies, people-sustaining ones, and people-changing ones. The aim of the first-mentioned is to label and classify clients, thereby making it possible for HSOs to link them to other persons ‘whose anticipated responses to the labels are presumably needed for transforming the clients’ attributes’ (Hasenfeld 1983, 135). The function of people-sustaining technologies is to prevent or to retard the deterioration of personal welfare or well-being. The purpose of people-changing technologies, finally, is to alter the psychological or social status of clients in order to enhance their well-being (Hasenfeld 1983). All three service technologies are used in the provision of PSSs; however, different ones predominate in the different services (Sallnäs 2015).

In Sweden, responsibility for investigating the needs of individuals and families rests with the municipalities, which are enjoined to carry out various interventions. IADs have to be carried out by the municipalities directly. TIs can be done by the municipalities themselves, or else outsourced to private actors.
In the case of IADs, the different cases are mainly dealt with by classifying clients through the use of people-processing technologies. TIs, by contrast, typically aim at changing the persons in question, through the use of people-changing technologies. In comparison with IADs and people-processing technologies, TIs and people-changing technologies are more focused on long-term client outcomes.

HSO tasks and technologies differ, and not just in respect of their aims; they also vary in respect of their complexity (Perrow 1967; Dornbusch and Scott 1975). TIs involve a kind of problem-solving task, often with a weak or incomplete knowledge base. Thus the results of TIs depend, to a greater extent than those of IADs, on the quality of the relationship between the professional and the client (Hasenfeld 1983; Munro 2004; Davies and Gregory 2010). TIs are more complex than IADs, and their results are less predictable. These differences have implications for how social work is organised and controlled (cf. Perrow 1967; Dornbusch and Scott 1975). In the case of less complex tasks for which service technologies with a relatively high degree of outcome predictability can be used – such as IADs – standard procedures and detailed routines are fit for purpose. In the case, however, of more complex tasks for which technologies with a low degree of outcome predictability must be employed – like TIs – standard procedures and detailed routines are less appropriate. Cases of the latter kind involve, by necessity, a greater reliance on dialogues founded on professional knowledge and competence.

Bearing these differences in mind, we can formulate two general presumptions regarding the adjustment of standards used in the inspection of IADs and TIs. In an ideal inspection system (in terms of the validity of its standards), we would expect (1) standards used in the inspection of IADs to be more precise in general but to focus less on long-term outcomes. By contrast, we would expect (2) standards used in the inspection of TIs to be less precise in general but to focus more on long-term outcomes.

Inspection of social services in Sweden

The inspection of social services in Sweden has been subject in recent decades to several governmental investigations aimed at improving its efficiency. In 2010, social-service inspection was transferred from the 21 county councils to a single national authority, the NBHW; and then in June 2013 to a newly established authority, the Health and Social Care Inspectorate (Proposition 2012/2013:20). These organisational changes have led to more centralised responsibility and to greater specialisation within the authorities. During the period reviewed in this study, the NBHW was responsible for social-service inspection. A stronger emphasis has also been put, due to the government’s investigations, on inspection as an instrument of sanction and control SOU (2002:14). Attempts were made, for example, to limit the advisory ingredient in inspections; and the government demanded that compulsory on-site inspections be conducted twice a year at all social-service institutions that provide TIs for children, youths, and families (Proposition 2008/2009:160). The latter represented a response to investigations that uncovered severe cases of neglect and abuse in institutional child care in Sweden over several decades (SOU 2009:99).

Inspections of PSSs are either initiated by the authority itself or else carried out after complaints from service users or others. Such inspections may concern a wide variety of issues. Examples include how individuals have been treated by social workers, how their cases have been handled, and whether the interventions undertaken have been of sufficiently high quality. Inspections of TIs for children, youths, and families are mainly initiated by the NBHW, reflecting the more stringent demands imposed by the government in this area; by contrast, inspections of IADs are mainly undertaken in response to complaints from service users. Most of the information in TI inspections is gathered through on-site examinations by one or two inspectors. A variety of sources and methods is used in such inspections. The actual on-site inspection normally lasts from a few hours up to one day, and it may be either announced or unannounced. IAD inspections, by contrast, are for the most part so-called desk inspections, which do not include on-site visits. They entail gathering information about the inspected service mainly from documents, including case files. The information thus acquired is then assembled into a report put out by the NBHW. The report describes the inspected service, reviews how the inspection was conducted and the information gathered, and lists the legal and regulatory
provisions underpinning the inspection. The report then presents an evaluation of the inspected service. In cases where no deficiencies have been discovered, the inspection is declared complete. In cases where deficiencies have been found, the report normally describes said deficiencies in terms of the applicable legal and regulatory paragraphs. In these cases, the report concludes with a list of actions that must be taken by the inspected service.

**Data and methods**

This study is designed as a focused comparison of the standards used in inspections at the information-gathering stage and subsequent evaluative stage of two different social services (IADs and TIs). Each social service represents a distinct ideal-type of service technology. Our empirical data consist of 186 paragraphs from twelve different laws, five governmental and NBHW regulations, and one international convention. These paragraphs, which feature in 189 NBHW reports on inspections of IADs and TIs, all relate to standards for social services. The SSA, which is the dominant regulatory source, contains a broad array of rules, ranging from quite general to highly specific ones. Paragraph 3 in chapter 3 of the SSA exemplifies how general demands are set forth in these provisions:

Social services shall maintain high quality. Staff with appropriate training and experience must be available to perform tasks in this area. The quality of social services shall be developed and secured in a systematic and continuous manner. [Authors’ translation of SSA 3:3].

Paragraph 8 in chapter 6 illustrates how more specific demands are laid out:

If a child, in accordance with the provisions of this law, is furnished with foster care outside of his or her own home, the Social Welfare Committee must consider anew at least every six months whether such an arrangement is still needed, and if so how it is to be designed and provided. When the child has resided in the same foster home for three years since his or her initial placement there, the Committee shall consider whether there is cause to apply for a transfer of custody in accordance with Chapter 6, Section 8 of the Parental Code. [Authors’ translation of SSA 8:6].

The 189 reports reviewed in this study constitute a special subset of the 1171 reports completed by the NBHW during the 01/01/2012–23/10/2012 period on PSSs for children, youths and families and for adults suffering substance abuse. The 189 reports are the ones in which deficiencies in the inspected services were identified and demands for remedial action made. Out of these 189 reports, 112 concern the inspection of IADs and 77 the inspection of TIs. Most of the IAD inspections were of the desktop variety, while most of the TI inspections included site visits. It should be noted that the TI reports display a certain bias, as a majority of them bear on institutions for refugee children.

Each of the 189 reports describes the service inspected, evaluates its merits and deficiencies, and lays out demands for remedial action. With few exceptions the reports explicitly cite – either by number or with a quotation – the specific paragraphs to which, in the judgment of the inspectors, the inspected social service failed to adhere.

The paragraphs in question come to 186 in all. The least frequently cited of these paragraphs is mentioned in a single report; the most oft-cited is mentioned in 90 (mean 9.6, median 3.0). The operationalization and subsequent coding of the content of each paragraph has been done in terms of the

---

**Table 1. Operationalization of paragraph content.**

<table>
<thead>
<tr>
<th>Content</th>
<th>Process</th>
<th>Output</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>External and internal conditions necessary to pursue the social work.</td>
<td>How the services should be provided.</td>
<td>What services should be provided.</td>
<td>Long-term client effects</td>
</tr>
<tr>
<td>Legal status of clients and other actors.</td>
<td></td>
<td>Short-term results e.g. service properties,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>experienced quality,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>and rule of law</td>
<td></td>
</tr>
</tbody>
</table>

---
four categories of content aspects presented earlier: input, process, output and outcome. Table 1 shows how we have operationalised these aspects.

Among the paragraphs cited, 176 present legal demands which the inspected service must meet, while 13 set out non-mandatory recommendations and demands.

The input category concerns the necessary conditions for social work, while the process category captures how the work should be conducted. The output and outcome categories both concern the results of the work in question, but they differ with respect to time: the output category relates to short-term results; the outcome category captures long-term effects.

To describe the degree of precision of each paragraph, we have employed an ordinal scale with three levels: low, medium and high. A low degree of precision indicates that the paragraph's content is vaguely defined, leaving a significant degree of discretion to the inspector. A medium degree of precision indicates that the content is described in concrete but not detailed terms. A high degree of precision indicates that the content is described in detail and/or contains measurable indicators.

We did the coding in two consecutive steps. In the first, we coded the type of content and level of precision for each paragraph sentence by sentence. In the second, we did the same for each full paragraph, according to the most frequent content and precision coding accorded to the sentences within it. Specific paragraphs are cited more extensively in the information-gathering section than in the evaluative section of NBHS decisions, so an underestimate of the number of paragraphs cited at the evaluative stage is to be expected. An example from the paragraphs most commonly cited at the evaluative stage of IAD inspections may serve to demonstrate how we made these coding decisions.

Paragraph 1 and 1a1 of chapter 11 in the SSA contain six sentences:

If a child, in accordance with the provisions of this law, is furnished with foster care outside of his or her own home, the Social Welfare Committee must consider anew at least every six months whether such an arrangement is still needed, and if so how it is to be designed and provided. When the child has resided in the same foster home for three years since his or her initial placement there, the Committee shall consider whether there is cause to apply for a transfer of custody in accordance with Chapter 6, Section 8 of the Parental Code. [Authors’ translation of SSA 8:6].

When a report regarding a child or young person is submitted to the NBHW, the Board shall immediately make an assessment of whether the child or young person is in need of immediate protection [3]. This assessment shall be documented [4]. A decision to open or not to open an investigation shall, unless exceptional reasons require otherwise, be taken within a period of fourteen days after the report has been received [5]. Such a decision is not necessary if an investigation concerning the child or young person is already ongoing [6]. [Authors’ translation of SSA 11:1 and 11:1a]

The first three sentences (1–3) focus on how a task is to be handled by the Board; they are thus coded as ‘process’ standards. The following three sentences (4–6) concern the immediate result of an activity, and so are coded as ‘output’ standards. Since process and output sentences are equally common in this case, and since we judge sentence (1) to be the one most commonly used as a standard in the empirical data, we have coded the paragraph as a whole as a ‘process’ standard.

Where its level of precision is concerned, we have classified sentence (1) as ‘high’, because it describes in detail when and under what circumstances an investigation is to be initiated. Sentence (2) contains concrete but not very detailed phrases: i.e. ‘[t]hat which has come to light in such an investigation’; ‘of significance’; and ‘recorded in a secure manner’. This is typical of sentences of the ‘medium’ category. However, since all the other sentences – (1), (3), (4), (5) and (6) – contain detailed and rather exact measures, we have chosen a coding of ‘high’ for the precision of the paragraph overall.

We have analysed the standards for the information-gathering and evaluative stages of the inspection process separately. The distinction between the two stages is highly relevant, because they perform different functions.

**Paragraphs used as standards at the information-gathering stage**

In our analysis of the 186 paragraphs cited as standards at the information-gathering stage of the inspection, we found that 0–34 paragraphs are cited per report (mean 8.7, median 8.0). At the evaluative stage,
0–18 paragraphs are cited per report (mean 4.0, median 3.0). This discrepancy between the two stages in the number of paragraphs cited is quite predictable. At the information-gathering stage – where the standards actually decide the type and amount of information to be collected by the inspectors – a large number of paragraphs cited is an indication that information about a broad array of aspects has been gathered. By contrast, a smaller number of paragraphs cited at the evaluative stage shows that deficiencies have only been detected in relation to a small number of standards.

Taking the information-gathering stage of the inspection process separately, what aspects of IADs and TIs are covered? A first analysis discloses a considerable difference between the two types of social work in the number of paragraphs being used as standards. In IAD-inspections, 0–18 paragraphs are cited per report (mean 6.3, median 5.0); the corresponding number in TI inspections is 0–24 (mean 12.0, median 12.0). The larger number of paragraphs cited in TI inspections at this stage accords with our expectation, and allows for both a broader array of aspects and a greater number of deficiencies to be identified in TIs than in IADs.

As Figure 1 shows, there is a relatively equal distribution in TI inspections between paragraphs that describe input, process, and output aspects. By contrast, IAD inspections at the information-gathering stage cite paragraphs that define process and output aspects more often. Our most important finding, however, is that the inspection of long-term client outcomes is the least emphasised aspect

Figure 1. Content of paragraphs referred to as standards at the information-gathering stage of IAD and TI inspection, percentages. Note. The no. of cases (n) for IAD is 112 and for TI 77. P-value < 0.001. e no. of cases (n) for IAD is 112 and for TI 77.

Figure 2. Precision of paragraphs referred to as standards at the information-gathering stage of IAD and TI inspection, percentages. Note. The no. of cases (n) for IAD is 112 and for TI 77. P-value < 0.001.
at the information-gathering stage, and even less so in TI inspections than in IAD ones. This finding challenges our initial thesis about what an ideal inspection system would entail. We presumed that, since TIs involve a heavier use of people-changing technologies that have long-term client improvement as their primary goal, paragraphs which cover long-term outcome aspects will be cited more frequently in the inspection of TIs than in that of IADs. At this stage, however, such a presumption finds no support in our data.

Turning now to the level of precision in paragraphs used at the information-gathering stage of inspections, we see in Figure 2 that – quite as expected – the paragraphs used in TI inspections are less precise than those used in IAD inspections. At the information-gathering stage, our data indicate, TI inspections are governed to a greater extent than are IAD inspections by paragraphs with a low or medium level of precision. This accords with our thesis that the inspection of complex tasks – represented by TIs in this study – must be based on more flexible or vague standards in order to be effective.

Paragraphs used as standards at the evaluative stage

As an inspection process enters the evaluative stage, there is a fall in the average number of paragraphs being used as standards (see above). This is to be expected, for inspections at this stage are focused on making judgments about deficiencies. Our data reveal too that there is a greater drop in the average number of paragraphs used at the evaluative stage of TI inspections than at the same stage of IAD inspections. In TI inspections, the number of standards falls from an average of 12.0 paragraphs per report (median 12.0) at the information-gathering stage to 3.7 (median 3.0) at the evaluative stage. In IAD inspections, the number of paragraphs also drops from the one stage to the other, but only from an average of 6.3 per report (median 5.0) to one of 4.2 per report (median 4.0). This sharper drop cannot be explained simply by the fact that inspections generally involve a narrower approach at the evaluative stage. The data do not allow for far-reaching conclusions here, but the pattern suggests either that our

Figure 3. Content of paragraphs referred to as standards at the evaluative stage of IAD and TI inspection, percentages.

Note. The no. of cases (n) for IAD is 112 and for TI 77. P-value < 0.001.

Table 2. Content of paragraphs referred to as standards at the information-gathering and evaluative stage of IAD and TW inspection, percentages.

<table>
<thead>
<tr>
<th></th>
<th>Input</th>
<th>Process</th>
<th>Output</th>
<th>Outcome</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information-gathering stage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IAD</td>
<td>15</td>
<td>38</td>
<td>39</td>
<td>8</td>
<td>100</td>
</tr>
<tr>
<td>Ti</td>
<td>30</td>
<td>35</td>
<td>28</td>
<td>6</td>
<td>100</td>
</tr>
<tr>
<td>Evaluative stage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IAD</td>
<td>14</td>
<td>38</td>
<td>44</td>
<td>3</td>
<td>100</td>
</tr>
<tr>
<td>Ti</td>
<td>27</td>
<td>37</td>
<td>33</td>
<td>2</td>
<td>100</td>
</tr>
</tbody>
</table>

Note. The no. of cases (n) for IAD is 112 and for TW 77.
TI cases have fewer deficiencies on average than our IAD cases, or that inspections are generally less effective at identifying deficiencies in cases of the former type than in cases of the latter type.

Figure 3 shows the distribution of content of the paragraphs used as standards at the evaluative and information-gathering stages. As can be seen, the distribution in the two cases is very similar. Process and output aspects predominate in both IAD inspections and TI inspections, while aspects focused on long-term outcomes have almost entirely disappeared.

In Table 2, the distribution of paragraphs used as standards in the two types of inspection is summarised for each stage. Input, process and output aspects predominate at both the information-gathering stage and the evaluative stage of both IAD and TI inspections. The distribution between the three content categories is somewhat more balanced in the case of TI inspections than in that of IAD inspections. While process aspects predominate at both stages of the two types of inspection, input aspects are covered less in the case of IAD inspections. Regardless, furthermore, of the type of social service being inspected, and regardless of the stage in the inspection process, our data suggest that long-term client outcomes are by far the least emphasised aspect of social-service inspections. The paragraphs used as standards at the information-gathering stage indicate that at least some information about outcome aspects is collected; however, a comparison of the distribution of paragraphs between the stages suggests that deficiencies related to outcome aspects are harder to identify than deficiencies related to other aspects. At the information-gathering stage, 8% of the paragraphs used in IAD inspections and 6% of those used in TI inspections concern outcome aspects, but the figures drop to just 3 and 2% respectively at the evaluative stage. The opposite is true for output aspects. At the information-gathering stage, 39% of the paragraphs used as standards in IAD inspections and 28% of those used in TI inspections cover this aspect, with 44 and 33% respectively do so at the evaluative stage. These differences between the two stages in the use of paragraphs focused on output and on outcome aspects may be attributed to one of two circumstances: either inspections as such are better able to
identify deficiencies related to output aspects than to identify those related to outcome aspects, or else deficiencies related to output aspects are more common in social services than are deficiencies related to outcome aspects. We will discuss these two possible interpretations further in our conclusions.

Turning again to the level of precision, we find some interesting tendencies toward convergence between the two types of social-service inspection, as well as between the two stages of the inspection process under examination here.

First, as Figure 4 indicates, the differences are very small between IAD inspections and TI inspections when it comes to the precision of standards used at the evaluative stage. Although a somewhat larger proportion of standards with low precision are used in the case of TI inspections, by far the greater number of standards used in both types of inspection are detailed or measurable, i.e. high.

Table 3 summarises the distribution of levels of precision in standards as between the two stages. The differences observed between IAD inspections and TI inspections at the information-gathering stage of the process virtually disappear at the evaluative stage. In the case of TI inspections, paragraphs with a high degree of precision increase from 36% at the information-gathering stage to 57% at the evaluative stage. In the case of IAD inspections, on the other hand, the proportion of high-precision paragraphs remains at similar levels. Again, these findings suggest that the adjustment of standards in TI inspections – so as to capture greater complexity and to recognise subjective experiences – holds at the information-gathering stage of the process but does not survive the evaluative stage. Our analysis of both the content and the precision of standards suggests that TI inspections are not able to any greater extent than IAD inspections to capture complexity and subjective experiences, or to identify deficiencies related to long-term client outcomes. As a policy instrument, it would appear, inspections are primarily adapted to IADs; they are less able to address central aspects of TIs in any optimal fashion.

Conclusions

The purpose of this study has been to contribute to an improved understanding of how standards in social-service inspections are designed, and to what extent this design is adjusted in relation to different social services. In an ideal inspection system (in terms of the validity of its standards), we would expect TI inspections to cover all four content aspects, but to focus on long-term outcomes to a greater extent than IAD inspections would do. Similarly, we would expect the paragraphs used as standards in TI inspections to be less precise than those used in IAD inspections.

Where the content of standards is concerned, our study shows that paragraphs used at both the information-gathering and evaluative stages of both IAD and TI inspections deal largely with aspects related to input, process and output. Paragraphs focused on long-term outcomes are rarely used – almost, in fact, as rarely in the case of TI inspections as in that of IAD inspections. This certainly contradicts our expectations about how an ideal system of inspections would function, since it implies that only a small amount of information is gathered about long-term client outcomes. An overall aim of inspections is to ensure high quality in social services, by identifying deficiencies and demanding improvements. If the standards covering central qualitative aspects of the social services are applied only rarely (or not at all), then the efficacy of inspections in promoting this aim is undoubtedly reduced.

When it comes to the precision of standards, our study has turned up some expected differences between IAD inspections and TI inspections at the information-gathering stage. A larger proportion of vague and less precise paragraphs are used at the information-gathering stage of TI inspections than at the same stage of IAD inspections. This confirms our initial expectations about an ideal inspection system: we hypothesised, namely, that TI inspections will show a greater capacity to collect information of a complex nature and to recognise subjective experiences. As we move from the information-gathering stage of the process to the evaluative one, however, there is a convergence between IAD and TI inspections, and the difference almost disappears: i.e. a sharp increase takes place in the proportion of high-precision paragraphs used as standards in the case of TI inspections as well. This is certainly at odds with what would expect in an ideal inspection system. What plausible explanations might account for these findings? One approach would be to accept in some measure
the government’s strategies and ideas behind the design of inspections as a policy instrument. These emphasise centralization, distant bureaucratic control, and an increased number of inspections as ways to improve the social services (Proposition 2008/2009:160; SOU 2004:100). We would then assume that the discrepancies in precision observed between the paragraphs used as standards at the information-gathering stage and at the evaluative stage are actually indications of high quality within the TI services. Since information which is complex and based on subjective experience is taken in at the information-gathering stage of TI inspections, but then disappears at the evaluative stage, we would assume that few deficiencies related to such aspects are actually present. This line of argumentation leads to the conclusion that inspections are indeed capable of overseeing TI services and guaranteeing their quality. Some researchers have turned up findings that support such a conclusion: the inspection of local government services, according to their data, promotes improvements indirectly by identifying deficiencies (Downe and Martin 2007).

Another and quite contrary explanation is less comforting. According to this view, the fact that few or no deficiencies have been identified cannot be taken as an indication that deficiencies do not exist. It could just as well indicate that existing deficiencies are not revealed for some reason, and so are less unaddressed within the present framework of formal authority and bureaucratic control. This explanation is supported by research showing that inspectors use variable strategies or styles of enforcement, and that they may choose to handle deficiencies relating to complexity and subjective experience in a more inter-professional and dialogue-based manner (Hood et al. 1999; Johansson 2006; Downe and Martin 2007). By this line of reasoning, dialogue-based inspections make for a more appropriate control system in the case of certain social services.

A third explanation – and arguably a more challenging one – is that the effectiveness of inspections should be questioned. According to this account, standards have indeed been designed in such a way as to support the gathering of complex information and the recognition of subjective experiences, but for other reasons it has not been possible to gather needed information or to identify deficiencies. Important aspects of the social services can neither be supervised nor improved on the basis of a standards-based control system. This explanation harkens back to Argyris (1976) theories about the need for double-loop-based decision-making, and to Greene’s (1999) contention that complex phenomena cannot be meaningfully defined by simple endpoints. The findings of our study are insufficient for making strong claims about which of these three explanations is the most plausible. We believe, however, that our findings are of interest in relation to the question of whether inspections are effective as an instrument for regulating human-service organisations. In an effort to throw light on this question, we have examined how standards underlying the information-gathering and evaluative stages of social-service inspection in Sweden are designed and actually used. We certainly need more in-depth studies on the extent to which inspections in this area actually succeed in gathering adequate information, as well as on the extent to which they really are able to identify deficiencies connected with complexity and subjective experience. Future research may also profit by going beyond the context of Swedish social services.

Note

1. The reason why articles 11:1 and 11:1a in the SSA are put together is that the latter is never explicitly cited in the material by its number, whether at the information-gathering stage or at the evaluative stage. It is clear from the text, however, that this paragraph is used as a standard, but by reference to SSA §11:1 as the overall paragraph number.

Disclosure statement

No potential conflict of interest was reported by the authors.
Funding

This work is financially supported by the The Swedish Research Council for Health, Working life and Welfare (FORTE) under the granted project [2012-0234].

References


