

Pedagogy of regulation: Strategies and instruments to supervise learning from adverse events

Josje Kok 

Erasmus School of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, The Netherlands

Ian Leistikow

Dutch Healthcare Inspectorate, Utrecht, The Netherlands

Roland Bal

Erasmus School of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, The Netherlands

Abstract

Diverse scholars have argued that standards and performance measurements are “instruments of control” that have a profound influence on the day-to-day lives of individuals and organizations, causing constitutive effects. Regulatory bodies increasingly use standards to oversee and monitor the regulated. This paper discusses the Dutch Health and Youth Care Inspectorate’s use of both standards and a performance measurement system introduced to monitor how Dutch hospitals investigate and learn from serious adverse events. Rather than focusing on how standards affect regulated practices and organizations, our study examines how the use of these instruments affects the standard maker, that is, the Inspectorate. We explore how the Inspectorate’s work practices, standards, and coupled performance measurement system influence its regulatory pedagogy, reviewing practices, and decisionmaking. We conclude that standards and performance measurement systems are not by definition “instruments of control” as their constitutive effects are (under)determined by the relationships in which they are enacted.

Keywords: constitutive effect, healthcare regulation, pedagogy, performance measurement, standards.

“Making mistakes is inevitable, not learning from them is unacceptable.”

Long-term policy plan, Dutch Health Care Inspectorate 2016–2019, citing Sir Liam Donaldson, England’s Chief Medical Officer, 2004 (Inspectie voor de Gezondheidszorg 2016).

1. Introduction

A growing number of social scientists argue that both standards and performance measurements are highly relational (Porter 1996; Lampland & Star 2009), shaped in a process of collective sense-making (Garfinkel 1967; Weick 2001). Although widely used as instruments of objectification, standards and performance measurements are in themselves interpretations that set their own political and normative effects in motion and demarcate the way in which the world is defined (Bowker & Star 1999; Dahler-Larsen 2014). As a consequence, the use of standards and performance measurements has constitutive effects in that they constitute the very practices that they measure.

Originally introduced by Dahler-Larsen (2012, 2014), the concept of constitutive effects refers to the many subtle and not-so-subtle ways in which evaluation machines steer certain values, orientations, interpretations, and practices in the direction of a particular construction of social reality (Dahler-Larsen 2012). These constitutive effects result from the profound influence of standards on the day-to-day lives of individuals and organizations

Correspondence: Josje Kok, Erasmus School of Health Policy and Management, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands. Email: kok@eshpm.eur.nl

Accepted for publication 12 January 2019.

(Bowker & Star 1999; Lampland & Star 2009). Performance measurements and standards are therefore increasingly recognized as “instruments of control,” holding the power to regulate behavior (Brunsson & Jacobsson 2000; Lascoumes & Le Gales 2007). But as Slager *et al.* (2012) have illustrated, little is known about the ways in which the regulatory power of standards is created and maintained over time. Also, most studies interested in the power and effects of standards have focused on the receiving end of the line: the outcome of standards on the actors who are the objects of regulation. What is missing is an analysis of how standards discipline and control the standard makers and have constitutive effects on the regulators themselves. To understand these effects, just looking at the standard or performance measurement system and examining its content is insufficient. Rather, we must look at the way in which they are shaped and used in practice.

In this paper we explore the ways in which standards and performance measurements are “instruments of control” (Brunsson & Jacobsson 2000), and how internal constitutive effects in regulatory agencies are enacted in a real-life empirical setting. We do so by focusing on the daily work practices of the Dutch Health and Youth Care Inspectorate; a regulatory body responsible for – among numerous other tasks – overseeing that Dutch healthcare organizations “learn” from serious adverse events. To be precise, we use the term adverse event (AE) to denote “an unintended and/or unexpected event related to the quality of care, having caused the death of or serious harm to the patient or client” (Buijsen 2014, Page 388).

The Dutch Health and Youth Care Inspectorate (hereafter “Inspectorate”) has only recently attempted to benchmark “learning” via self-developed standards and a coupled performance measurement system (Leistikow *et al.* 2017). As well as regulating and monitoring healthcare organizations, forcing organizations to manage risks and be accountable for mistakes “from the inside” (Power 2007), these instruments were also introduced after a public crisis at the Inspectorate (see “Contextual backdrop” section) to objectify the Inspectorate’s internal work practices, limit individual inspector’s “street-level” discretion (Lipsky 2010), and improve efficiency and traceability; an intervention in the relationship between Inspectorate and inspectees. This dual role of the introduced standards and performance management system makes the Dutch setting particularly interesting to examine.

In the Netherlands, as in other Western countries, a national AE reporting system is in place; in the Netherlands this system is overseen by the Inspectorate. Dutch law mandates healthcare organizations to report AEs to the Inspectorate and investigate the cause(s). The required investigations usually consist of a root-cause, or similar form of analysis, complemented with proposed improvement measures to minimize the risk of reoccurrence. The Inspectorate evaluates these inquiry reports, provides feedback, and uses the data to monitor risk trends – both nationally and at the level of individual healthcare organizations.

However, how and what healthcare organizations can “learn” from AEs is debated, as we shall explain in our theoretical framework (Tucker & Edmondson 2003; Tamuz *et al.* 2011; Anderson *et al.* 2013; Mitchell *et al.* 2015). Thus it is relevant to explore exactly what the Inspectorate wants healthcare organizations to learn – what, in other words, is its pedagogy? We hypothesize that this “theory of learning” (Kenklies 2012) is embedded in the above-mentioned infrastructure of the Inspectorate’s standards and work routines. That is, by studying this infrastructure in practice, we expect to find the pedagogy in use by the Inspectorate. The following research questions guide our exploration: How do the Inspectorate’s standards and performance measurement system manage to control the Inspectorate’s work routine? and Which pedagogy can be distilled from this work routine?

To answer these questions we foreground the backstage elements of work practices (Lampland & Star 2009) at the Inspectorate. Hereby we not only build on the theoretical notion of standards and performance measurement systems but also shed light on some of the many unknowns surrounding the practical execution of daily work within regulatory institutions (Sparrow 2000), globally faced with an ever more critical public and political arena demanding the management of risks and uncertainty (Power 2007).

This article consists of six sections. Recognizing the influence of the institutional constellation on regulatory processes (Jordana & Levi-Faur 2004), we start by presenting a brief historical and political overview and discuss the subsequent influence on the Inspectorate’s current work routine, to serve as a contextual backdrop. We then introduce the theoretical framework, discussing the concepts of “learning” and pedagogy, as well as the notion of standards and performance measurement systems as instruments of control. The framework is followed by a description of the research methods used. In Sections 4 and 5, we present the research findings and discuss our results. In the concluding section, we reflect on the implications of our analysis.

1.1. Contextual backdrop: A new “sharpened” work routine

The Inspectorate has a long history of overseeing the quality of Dutch healthcare services (Boot 2013; Robben et al. 2015). It does so by inspecting, advising, and stimulating organizations and sanctions in the case of non-compliance; an approach based on Ayres and Braithwaite's (1992) responsive regulation model. The Inspectorate's current approach to monitoring AEs is relatively new; a regulatory approach that has been shaped by complex configuration of external and internal drivers, including societal, political, legal, and institutional changes (Jordana & Levi-Faur 2004; Baldwin & Black 2016).

In line with the increasing global political and societal demands for public institutions and private organizations to control and manage risks (Power 2000, 2007), the current routine was set in motion in 2012 in the wake of wide media coverage of several high-profile incidents that had impaired the Inspectorate's reputation. In the aftermath of these incidents, the public complained that the Inspectorate's method of dealing with AEs was careless, inadequate, and favored the medical professionals under regulation (Ombudsman 2011, 2012; Dute 2015). Moreover, the organization was accused of passivity, lacking vision and transparency (Dute 2015). Alongside the public turmoil and political accusations of regulatory capture (Bardach & Kagan 1982), internally the Inspectorate was struggling to cope with an ever-increasing administrative workload (Sorgdrager 2012; Legemaate et al. 2013; Schippers 2013), as well as continual reorganizations (Vries 2011). To come to terms with the public's mistrust, deal with the increasing workload, and put the political unease to rest, Edith Schippers, then Minister of Health, Welfare & Sports, geared the Inspectorate toward a new way of working, formulated as a “sharpened work method” (Volksgezondheid, Welzijn en Sport [VWS] 2012; Schippers 2013). The sharpened routine introduced the work practices, standards, and performance management system this paper explores (see Fig. 1 for a concise overview).

First, the formerly contested regional focus that held individual inspectors responsible for assessing inquiry reports sent in by local healthcare organizations was reorganized into national incident reporting consultation teams, called National Incident Reporting Consultation (NIRC) teams (Legemaate et al. 2013; Inspectie voor de Gezondheidszorg [IGZ] 2013b). These comprised a mix of “general” inspectors, including former nurses, lab technicians, and biomedical scientists, and “specialized” inspectors, predominately non-practicing medical specialists. Each LMO team was made responsible for a different part of the healthcare sector, for example, hospitals, long-term elderly care, and general practitioners. Henceforth, assessments were a group effort performed at the national level, limiting individual inspector's regulatory discretion.

Following this shift, the guideline for reporting AEs to the Inspectorate (IGZ 2013a) was updated, specifying how AEs should be reported to the Inspectorate and what procedure the Inspectorate follows once a report has been made (Vries 2011; Legemaate et al. 2013). In addition, a uniform protocol for inquiry reports (IGZ 2014) replaced all earlier guidelines, specifically instructing healthcare organizations as to what their inquiry reports need to contain.

Last, and perhaps most significant, the Inspectorate's problem definition changed: not learning from mistakes was deemed a greater risk than making mistakes. The regulatory focus therefore switched from the medical content of an AE (what went wrong?) to primarily an assessment of the quality of the learning process reflected in the inquiry report (how has the organization learned from the event?) (Legemaate et al. 2013; Leistikow et al. 2017). To track and document these learning processes, the LMO team responsible for watching over hospitals created a standardized scoring system (“scoring analysis (SA)”) coupled to a performance management database (“digiBAN”). Both were introduced to the LMO work routine in the first quarter of 2013. Here it is noteworthy that the BAN system in use by the Inspectorate is founded on the assumption that hospitals awarded a high BAN score for their inquiry report have effectively learned from the AE at hand.

In short, the historical contextual setting outlined above reveals the influence of external drivers and factors on regulatory processes, whereby the changing social and political climate played a key role to switch gears (Jordana & Levi-Faur 2004). Other factors, such as those defined by Baldwin & Black (2016), including a regulator's risk and problem definition, operational constraints, and reputational factors, also played a role. The new work routine, standards, and performance measurement system were introduced to regulate the field of healthcare organizations and make it easier to monitor their performance. But the “sharpened” routine was also introduced as a trust device (Porter 1996; Halfman 1998); the standardizations needed to discipline the Inspectorate's work by: objectifying the evaluation practices, limiting inspectors' regulatory discretion, preventing regulatory

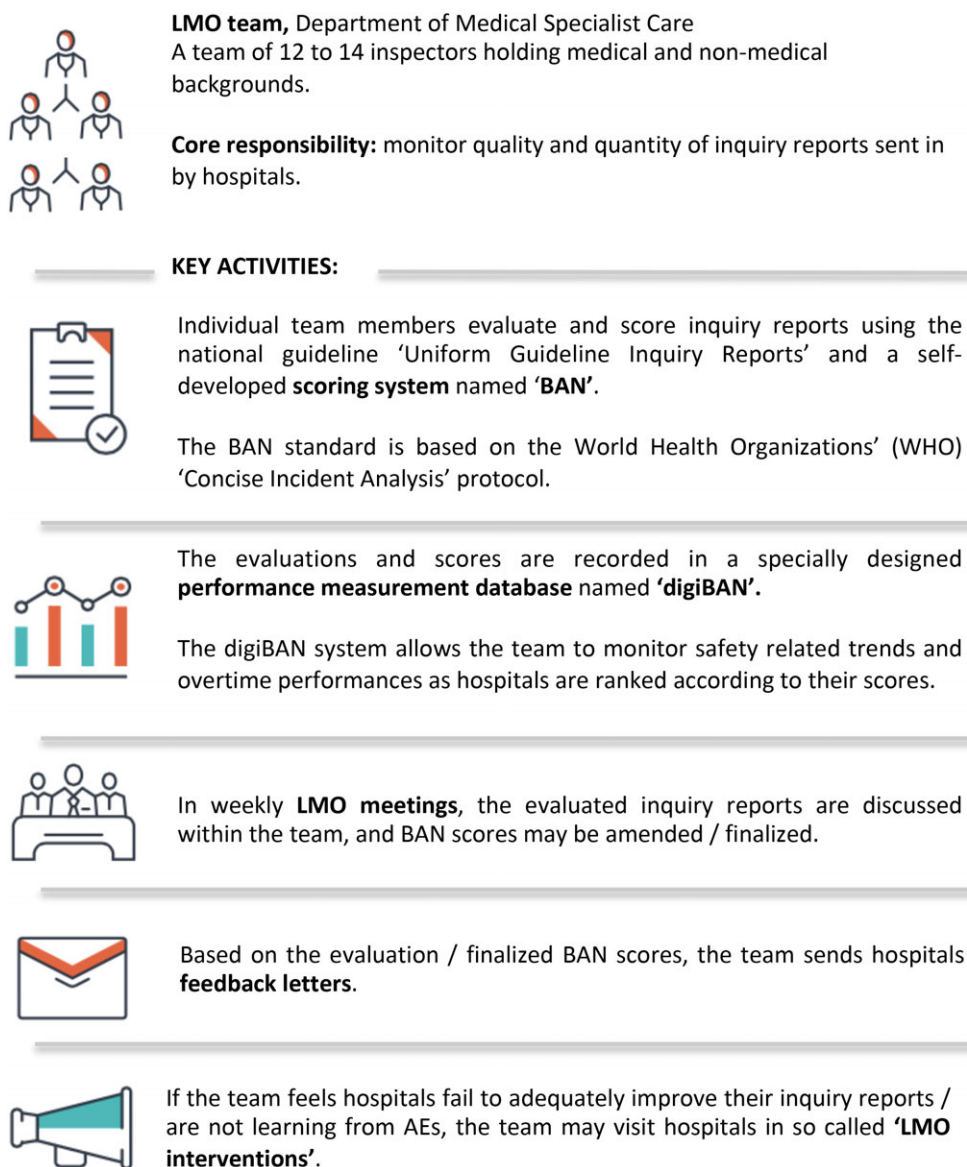


Figure 1 National Incident Reporting Consultation (NIRC) team, responsibilities and activities. scoring analysis (SA), ***

capture, creating traceable outputs, and speeding up work processes. It is this objective of the sharpened work routine that we analyze in this paper. The next section lays out the theoretical perspective applied in our analysis.

1.2. Theoretical framework

1.2.1. Learning from mistakes and the "theory of learning"

The concept of learning is widely used in relation to patient safety and quality-improvement initiatives (Rowley & Waring 2011). Since the turn of the century there has been a general consensus within the safety movement that mistakes and human errors made within the complex realm of healthcare – or any other high-risk industry for that matter – are inevitable (Kohn et al. 1999; Rowley & Waring 2011; Mitchell et al. 2015). Scientists, healthcare professionals, and policymakers alike have campaigned for the importance of using mistakes as a catalyst for learning, advocating that inquiries into AEs should stimulate the continuous improvement of patient safety (Kohn et al. 1999; Tucker & Edmondson 2003; Iedema et al. 2006). The Inspectorate shares this notion, declaring that drawing lessons in the wake of an AE is an imperative (IGZ 2016).

However, *what* must be learned, *who* must learn and *how* one – or an organization – can learn to improve the safety of patients are debated. With regard to *what* and *who* must learn, Jensen (2008) illustrated that health-care is often viewed as a system (see e.g. the 1999 well-known and celebrated Institute of Medicine report “To Err is Human”), but in the wake of an AE it is often the medical professional who receives the blame: to err is *human*, not system-based. This, Jensen argued, is a contradictory state of affairs, as he explained that dysfunction(s) in a system can be attributed to a wide range of actors and processes included in that system: individual end-users (i.e. doctors and nurses), but also the engineers or designers of products and procedures used in the system. The focus and/or where the line of inquiry is drawn ultimately dictates what and who can or needs to learn from the incident at hand. Illustrative of this is a study by Behr et al. (2015) revealing that the manner in which a medical error or incident is framed by the actors involved shapes the way an inquiry is carried out and presented in a report. This ultimately influences what is – or can be – learned from the event and by whom (Behr et al. 2015).

How organizations or individuals in an organization can learn is also much debated. Tucker and Edmondson (2003), for example, show the difficulties hospital employees face using problems and mistakes as opportunities for improvement because of existing system and cultural barriers. Anderson et al. (2013) also highlight the complexities of trying to learn from inquiry reports, as the reports do not provide unambiguous data on how to improve safety. Some scholars (Rowley & Waring 2011; Hollnagel et al. 2013) have even argued that aside from the practical difficulties of learning from mistakes it is more effective to learn from the things that go right rather than concentrating on the things that go wrong, thereby questioning the whole idea of learning from error.

In light of these debates it is clear that learning is a controversial matter and therefore it is interesting to explore how the standards made and used by the Inspectorate constitute learning and dictate what the process of “learning from mistakes” must entail.¹ To explore this “theory of learning” we use the concept of pedagogy as originally introduced by Johann Herbart (Kenklies 2012). Herbart suggested that the concept refers to the assumptions of an educator who acts upon these assumptions using a specific set of skills, with a deliberate end goal in mind (Kenklies 2012). An educator can be anyone or any organization who intends to “teach,” either implicitly or explicitly, by explaining, demonstrating, correcting and the like, and in so doing to be a felt presence, as well as having a reverberating influence on the knowledge or behavior of another party (Hansen & Laverty 2010). Thus, for our analysis we identify the Inspectorate as an educator of healthcare organizations, stimulating those organizations to learn from (the analysis of) AEs.

Denoting a regulator as an educator is not entirely new. Moving away from the traditional dichotomy of rule-orientated/legalistic versus cooperative/conciliatory enforcement approaches (Scholtz 1984) – regulatory scholars increasingly describe enforcement approaches as holding a variety of different interactional styles (Ayres & Braithwaite 1992; May & Winter 2000; Lo et al. 2009), an educational approach being one of these possible dimensions. But although the educative approach of a regulator has been recognized, empirically little is known about what such a role entails in practice.

1.2.2. Standards and performance measurement systems as “instruments of control”

As “learning” is envisioned as an important way to stimulate and improve patient safety, the Inspectorate has made the concept into something that needs to be governed and monitored. To do so, the Inspectorate devised standards in the form of guidelines dictating what healthcare organizations need to do and investigate in the wake of an AE. The BAN scoring system was developed to aid inspectors in the process of evaluating and ranking the inquiry reports sent in (Leistikow et al. 2017).

Numerous social scientists have pointed to the way in which the creation and use of standards shapes new realities and impacts relationships and subsequent behavior (Bowker & Star 1999; Lampland & Star 2009; Dahler-Larsen 2012, 2014). As Bowker and Star (1999) explain, each standard valorizes some point of view and silences another. In other words, standards are not only helpful to control or govern a messy reality (Lampland & Star 2009) but also define this reality as well, revealing what is deemed to be important and what is not. Moreover, as Porter (1996) explains, standards are often introduced to replace human judgment but are simultaneously created and used by humans who carry with them ideas and interpretations of what those standards mean and represent. Thus, as we try to underpin the Inspectorate’s pedagogy, it will be fruitful for us to closely examine their (use of) standards.

The standards formulated by the Inspectorate feed into a performance measurement system ranking health-care organizations' learning abilities. We know that rankings commensurate practices. That is, rankings transform qualities into quantities so as to reduce and simplify information into easily comparable figures (Espeland & Stevens 1998). However, we have very little understanding of how these help organizations govern activities (Wallenburg et al. 2016). Thus, the way rankings in the Inspectorate's performance measurement system are performed in a social process also deserves our attention.

This "performance" or practical use of rankings and standards by the actors involved is another matter of interest as we wish to examine the regulatory power (Slager et al. 2012) of these instruments. Standardization and ranking practices are increasingly recognized and acknowledged as a form of regulation (Slager et al. 2012) and standards are regarded as "instruments of control" that carry the capacity to generate order and facilitate coordination and cooperation, thereby creating similarity and homogeneity (Brunsson & Jacobsson 2000). As Slager et al. (2012) state, standards facilitate coordination by defining the appropriate attributes of the standardized subject – in our case what hospitals need to do and investigate in the wake of an AE – rendering these aspects visible to external inspection and opening up the possibility of sanctioning noncompliance. This gives the standard makers power: external regulatory power. But standard makers should then also be regulated by those same standards, given that standards limit discretion and influence the relationship between regulators and the regulated.

To examine the regulatory power exerted internally, Dahler-Larsen's (2012, 2014) concept of "constitutive effects" may help us. The concept acknowledges the profound influence standardizing and evaluating activities may have on what learning is and/or should entail (Hulst & Segerholm 2012). The concept can help us understand how the infrastructure in use by the Inspectorate has formative effects on the standard makers themselves – in this case the Inspectorate.

In sum, in our data analysis we set out to explore if and how standards affect decisionmaking by the Inspectorate and how the Inspectorate's pedagogy is constituted through its work routine and instruments.

2. Methods

Fitting our interest to examine the effects of standards and a performance measurement system *in use*, we employed an ethnographic fieldwork approach. The merit of this approach is that it allows for insight into the internal processes of a group and organization and to recover the distinct meaning given to these processes by the actors involved (Emerson et al. 2015; Rhodes 2015). Moreover, "being there," closely observing and listening to the inspectors at work allowed us to get below the surface of official accounts by providing texture, depth, and nuance (Rhodes 2015).

Over the course of seven months (February–August 2015), the first author studied the work performed by the Inspectorate's LMO team, which was mandated the task of monitoring the quality and quantity of inquiry reports sent in by hospitals (see Fig. 1 for a brief LMO team description). To gather comprehensive insight into the team's daily work practices and increase the validity of our data, we triangulated diverse methods: a quantitative analysis of the data in the database, observations, informal interviews, document analysis, and a focus group discussion (see Table 1 for a schematic overview).

During the observations we upheld an overt researcher role (Green & Thorogood 2006), openly jotting field notes (Emerson et al. 2015) and, when appropriate, asking questions. To respect the confidential nature of the discussions observed, no audio or video recordings were made. Instead, the field notes were quickly drawn up into observation reports to safeguard as many details as possible (Green & Thorogood 2006; Emerson et al. 2015) and generate detailed "thick" descriptions (Geertz 1973). To enrich the quality of the observation reports, internal documents, such as meeting agendas and PowerPoint presentations, were also obtained.

For our quantitative analysis, we asked the team of inspectors to send us details on the LMO interventions they had executed in 2014. These records were matched with data from the digiBAN database, allowing us to document which hospitals were visited, what their average BAN scores were (the quality of the inquiry reports as assessed by the Inspectorate), and how many AEs were reported.

Lastly, a focus group discussion was organized to interview the team about their daily work. Rather than using a tightly structured topic guide (Green & Thorogood 2006), the first author presented observation excerpts and quantitative data from our database analysis, asking the team to reflect on these findings. Realizing that this

Table 1 Overview of fieldwork activities conducted at the Inspectorate, February–August 2015

Fieldwork activity	Description	Raw data collected	Output for analysis
Qualitative data analysis	Analysis of BAN scores in digiBAN database and interventions performed	BAN figures and intervention	Graphs and tables
Document analysis	Collection of internal documents	Diverse protocols, guidelines, and internal digital communication	N/A
Observations including informal interviews [†]	LMO meetings (9 meetings)	20.75 hours of field notes/meeting minutes and agenda	Observation reports
	LMO interventions (2 interventions)	2.5 hours of field notes/PowerPoint presentation, meeting minutes, and agenda	Observation reports
	Inspectors prepare meetings/score inquiry reports (5 different inspectors)	10 hours of field notes	Observation reports
Focus group discussion	Inspectors' LMO meetings	1.5 hours of audio recording/field notes	Interview transcript

[†]See Green & Thorogood (2006, p. 80)

approach could possibly influence the inspectors' daily work thereafter, it was deemed necessary to capture the inspectors' reactions to reveal their underlying assumptions, which were of interest to us for this paper. Besides being a data collection tool, we thus also embraced the focus group as a “member check” or reflexive instrument (Alvesson 2003) to validate our interpretations and collectively reflect on them. With permission, the group interview was audio recorded and transcribed verbatim.

In all documents, the names of hospitals and inspectors were anonymized to protect their privacy. Also, all (confidential) internal documents were obtained with the permission of the LMO team members. The documents were systematically labeled; labels we will refer to when citing these nonpublic data sources.

The observation reports, transcript, and collected internal documents were inductively coded in the form of a thematic content analysis (Green & Thorogood 2006). The thematic exploration – a joint effort by the first and third authors – was followed up by a deductive analysis, recoding and assembling the data using the concepts from our theoretical framework.

It is appropriate to note that during the fieldwork study, the second author was a member of the LMO team. The second author allowed the first author easy access to data sources and provided helpful insider-knowledge of processes and procedures (Rutz et al. 2017). This “easy access,” as well as the close proximity of the first author to the studied setting – often the case in ethnographic studies – increased the risk of being unable to liberate oneself from some taken-for-granted ideas (Alvesson 2009). To minimize this risk, regular meetings were held with all three authors to reflect on the research findings.

2.1. Findings

In the following passages we discuss our findings on the three interlinked activities that we examined at the Inspectorate. First, the scoring and evaluation of reports using the BAN standard; second, the LMO meetings; and third, the process of providing hospitals with feedback and executing interventions.

2.1.1. Scoring and evaluating reports: Standards creating meaning

The evaluation of an inquiry report is an intensive, time-consuming process. It is repetitive – “essentially it's production work, done for hours on end” – but the very nature of the reports makes the evaluation process “far from easy or routine.”²

While reading the inquiry report the inspector mumbles, her hand covering her mouth: “Jeez.” I don't react. We continue to read. She sighs, frowning: “Huh? Oh, no.” She shifts on her seat, wiggling her legs restlessly, still reading. Hissing at her computer screen, she shrugs her shoulders: “Ow, people, this is such terrible care. This is truly unbelievable!” (...) After 15 minutes we reach the last page of the report. She turns to me, smiling. I am confused. As if the inspector could read my mind, she explains: “Well, I do think they [the hospital]

The chart on the right is an overview of the quality and quantity of the inquiry reports sent to the Inspectorate by hospital X, during 2014. The digiBAN has a scoring table for each Dutch hospital.

On the x-axis, the bars in the chart represent the individual inquiry reports. The y-axis reveals the quality of these inquiry reports (BAN score of 0–100).

The dark purple line denotes the moving-average of hospital X (BAN score average of the last 5 reports), revealing its overtime progress. The three gray lines indicate the mean moving-average score of all other hospitals at that point in time: lowest ranking hospitals, average and highest scoring. This overview allows the Inspectorate to compare hospital X to all other Dutch hospitals.

Due to the limited scope of quantitative data we were not able to perform a significant statistical analysis to determine if the LMO team's interventions produced a significant effect on the quality of the BAN scores. The figure does demonstrate that the quality of hospital X's inquiry reports yo-yos greatly. The moving-average of all Dutch hospitals appears to be relatively steady over the year 2014.

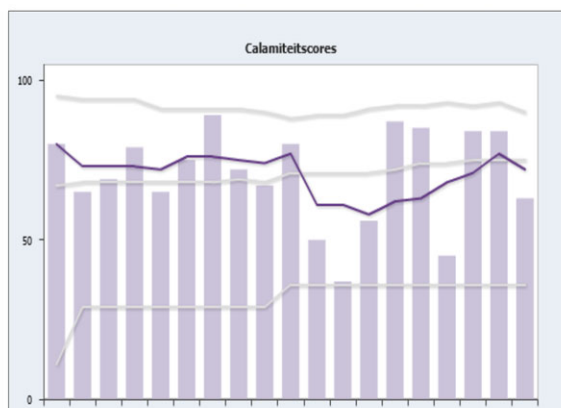


Figure 2 Screenshot from the digiBAN database, 27 March 2015: BAN scores for hospital X.

have investigated the incident well. Let's have a look how my BAN score will turn out." And she opens the digiBAN database to start scoring.³

As this observation illustrates, each inquiry report is evaluated by a "general" inspector who scores the document and drafts a letter to the hospital commenting on the inquiry process, that is, what the hospital has done right and what can – or should still – be improved. However awful an AE may be in terms of content, inspectors are searching for the potential lessons that are drawn from the event and can thus be pleased if the event has been examined thoroughly. If the quality of the inquiry report is deemed adequate (a BAN score of 80 or more on a scale of 100, see Fig. 2) and/or the reported AE is not overly complex in nature, the case is closed and a letter is sent to the hospital. All other reports make the LMO meeting agenda, requiring the initial assessment to be followed up by other team members – usually two or three "specialized" inspectors – who read the report and proposed return letter and provide feedback on the text. The inquiry report and the team's feedback are discussed during the LMO meeting. The outcome of this discussion determines how the Inspectorate ultimately judges and responds to the inquiry report.

Determining the complexity of an AE and its accompanying inquiry report is only partly formally defined. Certainly, internal guidelines dictate when inspectors need to collectively discuss a specific report, for example, if the inquiry reveals that the AE was caused by the culpable behavior of a medical professional. There are, however, informal dynamics at play. For instance, if an inquiry report is difficult to understand because it has been "poorly written,"⁴ is "very technical in nature,"⁵ or if the incident is "outrageously shocking!"⁶ We also observed instances where reports were set on the LMO agenda because the evaluating inspector had a "hunch" about a safety trend specific to the hospital or the healthcare field in general.

Although informal dynamics are at play, when it comes to determining the quality of an inquiry report, the BAN standard plays a key role. As presented in Figure 1, the BAN scoring system is based on the World Health Organization's (WHO) "Concise Incident Analysis Protocol" dictating that an incident inquiry must be timely (start as soon as possible); interdisciplinary and impartial; involve the actors familiar with the event; continuously dig deeper by asking "why, why, and why" at each level of cause and effect; and, finally, identify the changes that should be made to prevent reoccurrence (IGZ, internal document #2013C1). To be thorough, the inquiry must include a description of the incident, the underlying causes and effects must be analyzed, all contributing factors must be identified, and the findings and recommendations must be documented and formalized. Furthermore, to be credible, an inquiry should include the participation of leadership (managers and/or board of directors) and those closely involved; it should address conclusions with recommendations for reducing risk, include

consideration of relevant literature and other sources of information, and include an evaluation plan to determine if recommendations are implemented (IGZ, internal document #2013C1). In addition to the WHO inspired criteria, the Inspectorate added extra criteria to their BAN standard – principles they “find particularly important.”⁷ These criteria include involving the patient or their next of kin in the reconstruction of the event, as well as a description of the aftercare process for both patients and involved professionals (i.e. “second victims”) (IGZ, internal document #2013B1).

All of these criteria coincide with the uniform guideline for inquiry reports (IGZ 2014), which is sent to a hospital once it has reported an AE and starts an inquiry. In theory, hospitals thus know what is expected of them – and how the Inspectorate judges their inquiries. In practice, the quality of the inquiry reports – as determined by the Inspectorate – varies greatly, not only across but also within hospitals (see Fig. 2). Likewise, we observed that not all hospitals cover and/or interpret the criteria of the guideline in the same way. Some hospitals, for instance, refuse to involve patients in their inquiry report, which is one of the Inspectorate’s requirements. Or, in another example, when the LMO team concludes that a hospital has not identified all of the causes and effects in its analysis, frustrating a team eager for their “pupils” to do better and pick up on pointers provided in earlier feedback letters:

Chairperson: “This is a bad analysis! We helped them last time but they still haven’t improved!”⁸

For the Inspectorate, a good inquiry into an AE ticks all of the boxes of the BAN scoring form; the more ticked boxes, the higher the BAN score. The BAN thus molds how an inquiry process is (to be) executed: it constitutes norms to which hospitals must adhere and (to a certain extent) regulates behavior (Lampland & Star 2009), thereby exerting regulatory power on the hospitals (Brunsson & Jacobsson 2000), especially because hospitals that consistently score poorly on specific criteria of the BAN can be reprimanded.

The BAN standard and subsequent scoring system has practical advantages for the Inspectorate. The standard helps to organize a sometimes “messy reality” (Lampland & Star 2009) and aids inspectors to read through the inquiry reports – sometimes exceeding 30 pages in length – in a systematized manner; while reading sometimes physically checking-off the various criteria before they give the inquiry report a formal score:

As we read, Inspector 4 scribbles in her notebook, mumbling to herself: “Date is correct.” She writes down the date.

“Reaction board of directors? Yes. Ok.” Check.

“Aftercare patient? Yes.” Check.

“Do I understand what has happened here?” Scrolls back and forth through the pages, rereading, and concludes “Um yes. I understand.” Check.⁹

The standard ultimately defines and communicates to the hospitals what the Inspectorate believes is important. This is a good example of assisted sense-making whereby a constructed mechanism for creating meaning organizes a complex reality (Dahler-Larsen 2012). This is practical but it has consequences and potential downsides. One pitfall concerns the procedural focus of the BAN, neglecting content-related details. By concentrating on ticking off the BAN criteria boxes, the Inspectorate risks losing sight of other elements that may have been important. This became clear when a dismissed inquiry report was reopened and put on the LMO meeting agenda. The quality of the inquiry report had been sufficient; the hospital had met all the criteria of the BAN standard and thus the case was closed. Later, however, the Inspectorate received news that the hospital had failed to mention essential medical details. If the evaluating inspector had focused on these medical specifics in the inquiry report, the inconsistency in the reconstruction of the event may have come to the fore earlier. Inspector 9 explains:

This is the pitfall of this type of regulation. The problem is caused by the way we judge the inquiry reports. Sometimes you miss things by not looking at the medical details.¹⁰

This example also illustrates the relationship of dependence between the Inspectorate and hospitals; it is a continuous balancing act between regulator and regulated (Hawkins 1983). The LMO team needs to trust the reporting hospitals to convey all relevant details and produce a truthful representation of what went wrong. Equally, the team needs to trust hospitals to follow through on their proposed improvements, as the Inspectorate

does not monitor them.¹¹ This recurring state of tension fuels arguments at the weekly LMO meetings, which we will discuss next.

2.1.2. LMO meetings: Negotiations and collective sense-making

The weekly LMO meetings serve as a platform to jointly discuss inquiry reports, reflect on, and (re-)negotiate the awarded BAN scores and subsequent feedback letters:

I sit in the corner of the room, notebook on my lap, watching the inspectors settle down at the oval-shaped conference table. Coffee cups in hands, friendly chat[ter], laughter here and there. A homemade pie is cut into generous pieces and passed around on paper plates; someone's birthday treat. I find the ambiance amicable, these colleagues seem to know each other well.

At 10:40, slightly delayed, the meeting starts. Thirteen inspectors, women in the majority, gaze at a case file projected on a flat television screen at the far end of the room. "We have a busy schedule today!" the chairperson notes. I glance at the agenda; 22 inquiry reports need to be discussed in the coming two and a half hours [...]

At 11:06 the third inquiry report is presented; a young patient unexpectedly died shortly after she was hospitalized. In its report, the hospital concludes that the patient was misdiagnosed: a fatal mistake. Inspector 7, who has evaluated the inquiry report, suggests the case can be closed. "Although the quality of their [the hospital's] investigation can be improved, I think they have devised appropriate measures to minimize the risk of reoccurrence." I observe head[s] nodding quietly around the table but then Inspector 6 sits up and raises her arms: "I don't agree [with closing the case]. It [the incident] is so serious! You all just assume that the hospitals are going to carry out the measures they suggest in their reports. Let them report back to us, to prove that they have followed up on their promises!" Inspector 14 is quick with his response: "Everything we discuss here is serious. The point of our LMO meetings is to jointly establish that hospitals demonstrate that they are learning from their mistakes. We [the Inspectorate] need to trust hospitals to execute the improvements they formulate in their reports. Legally they have a responsibility to do so. So all adverse events reported to us, no matter how serious, are good. If we [inspectors] ask this hospital to re-do their inquiry or report back to us again, we may discourage them from reporting their mistakes in future. I mean, look at the BAN score, this inquiry report is better than the last one" [...]

As the discussion wraps up, in line with Inspector 7's suggestion the case is closed (...). The meeting continues, the inspectors press on; 19 reports to go.¹²

For the team members it is tiresome and at times emotionally straining work. The subject matter of the meetings emotionally affects inspectors, powerfully exemplified in this excerpt:

These meetings and the preparation [assessing reports and drafting feedback letters] cost so much time. The work pressure is tremendous. It's a never-ending flood of misery. [...] You know their [the hospital's] intentions are good but sometimes you lose sight of that when you read and discuss inquiry reports all day. Quite frankly I feel disappointed in my former colleagues [practicing physicians]. I feel like: Jesus man, please stop doing this! Stop making these mistakes!¹³

During the meetings there is room for inspectors to vent their frustration and concern – exclamations such as "Oh, this poor patient!" are commonplace – but most of the discussions concentrate on the quality of the inquiry reports. The team openly deliberates on whether the hospital has learned from the event under examination. Sometimes the inspectors see eye-to-eye, sometimes they don't. This in itself is interesting because the BAN standard was introduced to assess the quality of the reports uniformly and to objectify the inspectors' evaluation. Certainly individual discretion is limited but the emotions of the inspectors are never far away.

When closely examining the discussions it becomes clear that it is difficult for the team to purely use a set standard, the BAN, to evaluate the reports. The occasionally shocking details of an event sometimes cloud one's judgment or cause conflicting positions. Emotions regularly overflow the boundaries of the standard. For example:

Inspector 16: *"The doctor made the wrong decision!"*

Inspector 2: *"Yes, but the reconstruction of the event is correct [according to the BAN standard], so I am going to give them a point for that."*¹⁴

When confronted with this and other similar observations in the focus group discussion, Inspector 3 explained:

At one point we decided that we would focus on the process instead of judging the care, [w]hen we formed the LMO. But you can see how hard it is to do.

The team considers the quality of the inquiry report and the medical care to be two values that are indeed different but closely interlinked. Inspector 6 explains:

You can have a situation where a doctor has made an error of judgment. One of many things that can go wrong in the care process. If they [the hospital performing the inquiry] don't take that aspect into consideration in their inquiry, they may have performed a fair analysis, according to our BAN criteria. But still, their assessment will not be complete because I feel like they have left something out. Because they forgot to ask one 'why' question. Why has the doctor made a judgmental error? So, they are two different values [process versus care content] but they are linked to each other.¹⁵

Throughout the meetings these different, yet intertwined, viewpoints are negotiated. The team uses each other's expertise – some have a surgical background, others have experience in the field of nursing etc. – to muddle through these queries and form a joint assessment. For every unique case they literally make sense of the situation, sometimes having to “read between the lines” of the BAN standard, determining what is important and what should be learned to improve the quality of care. These findings suggest that the BAN standard cannot just be “applied” and/or does not just “work.” Rather, the inspectors interpret the inquiry, use their individual discretion to make choices about the key learning lessons, and eventually confer on these choices in meetings with the rest of the team. While the standard surely influences the discussions, the internal regulatory power of the standard is influenced – in this case limited – by the context in which it is used. To be precise, in this setting the context is shaped by the relationship of dependence between Inspectorate and hospitals, the (informal) knowledge individual inspectors carry about hospital practices, and inspectors' affective response to the AE at hand.

2.1.3. *Providing feedback: Pedagogic reasoning*

Once the BAN score is agreed, the content of the proposed feedback letter is discussed, settled, and sent to the hospital. This return letter is a preformatted document that is amended and personalized to suit the feedback the LMO team decides to communicate. In general there are three options: (i) a case can be closed, in which event the Inspectorate merely provides recommendations for possible improvements to future inquiries; (ii) the Inspectorate may request the hospital to “dig deeper” by asking the hospital to answer imminent questions or re-do the inquiry entirely; or (iii) when there are serious care-related quality concerns and the inquiry report is not up to the Inspectorate's standards, the LMO team will send the hospital a letter announcing that the Inspectorate will start its own inquiry, performed by a specialized team of inspectors.

Whichever course is decided on, the common point is that the communicated message is carefully constructed: content as well as tone matter. The team's goal is to stimulate hospitals to learn from AEs; the inspectors feel it is best – in most cases – not to be too “harsh,” even if the inquiry has not been performed according to the Inspectorate's standards. The team is aware that hospitals invest a lot of time in investigating AEs and do not want to discourage them, even if there is room for improvement:

We shouldn't send back too many points [of improvements] and remarks because we may overwhelm them [the hospital] with a flood of information.¹⁶

Or

You shouldn't give an unsatisfactory score. That will only scare them off. It's risky because they might stop reporting their adverse events.¹⁷

Like teachers, the inspectors carefully deliberate on what to address and how their feedback should be constructed. The team recognizes that their feedback letters to the hospitals have a disciplinary effect. Aside from illustrating the pedagogic reasoning behind the Inspectorate's feedback to hospitals, these quotes once more stress the relationship of dependence between the Inspectorate and the hospitals. The Inspectorate reasons that

hospitals must report and investigate AEs, otherwise they miss learning opportunities (IGZ 2016) and therefore the LMO team is challenged to uphold a delicate balance between evaluating, stimulating, and/or reprimanding the hospitals. The need to uphold a continuing relationship of trust shapes the team's compliance strategy (Hawkins 1983).

We observed that the LMO team is eagerly trying to find (new) ways of maintaining this balance. For instance, inspectors recently started to telephone hospitals as well as sending return letters. When asked about this development, Inspector 3 explains:

Yes, we do that more often. (...) [We call] to ask them a question but also to announce that a return letter is coming. It's like decorating our unpopular message with a small red ribbon. By calling we give them [the hospital] some context, so they understand where we [the Inspectorate] are coming from when we ask them to re-do the inquiry. We explain why we think it will help them.

Inspector 4, adds:

[I]f you send a critical letter then on paper it comes across different than if you do it verbally. Over the phone you can be more refined. So when you call in advance [before the feedback letter is sent] and clarify your critical message, it comes across better.¹⁸

The BAN standard and coupled digiBAN database are also a source for the team to fine tune their pedagogic approach, as the scores allow inspectors to track overtime progress for individual hospitals but also for the field as a whole. Knowing which elements a hospital is struggling with and/or what the overall quality of inquiries are provides inspectors with the opportunity to formulate their feedback accordingly. When asked what role the BAN plays in monitoring and steering the hospitals, inspectors explained that the standard and database do indeed "provide context" (Inspector 4) and are therefore "looked at systematically" (Inspector 14). Moreover:

When we evaluate the inquiry reports, we often look at the BAN score in relation to earlier inquiries. Then you determine: have they [the hospital] improved? Have they worsened? And this codetermines what you [the inspector] write in your letter.¹⁹

Our data, however, do not convincingly support these claims. While observing inspectors ($n = 5$) score reports and draft return letters, earlier BAN scores were consulted in only two of the 12 cases. Likewise, at LMO meetings we did occasionally hear inspectors refer to BAN scores; for example "Look at the BAN, they're doing better now than last time" or "They have an average score of 80!," but this was not a consistent occurrence.

With regard to using the BAN score as a ranking tool to become "sharper" in monitoring learning development and acting accordingly – that is, deciding when and where an intervention is necessary (IGZ internal document #2013C1) – we were surprised to discover that the LMO team apparently does not use the data systematically in that way (see Fig. 2, illustrating what the scoring tables look like).

In our fieldwork we discovered that the LMO team performed 11 interventions in 2014 that addressed the quality of inquiry reports. Unexpectedly, these interventions were not targeted at the poorest scoring hospitals. Informed of our findings, the team was surprised:

"Why did we go there?" (Inspector 14)

"I didn't expect [name hospital] to score so high." (Inspector 11)

"One case I find interesting is [name hospital]. They have a low ranking. But I feel like they have the capacity to learn and we could help them. But we haven't gone there." (Inspector 8)²⁰

These statements once more illustrate that the inspectors use informal knowledge and/or personal conviction to assess the (learning) needs of a hospital, ultimately shaping their course of action. The next excerpt confirms this:

Inspector 11: *My first thought when I read [name poorly scoring hospital], well, you know. Sometimes it's just not worth flogging a dead horse. (...)*

Interviewer: *What do you mean by that?*

Inspector 11: *Well, it is a hospital that I, personally, feel is not strong at learning. Let me leave it at that.*

Interviewer: *Ok. So is that a reason not to go there?*

Inspector 11: *Well, it wouldn't be my first pick.*

Inspector 3: *But we are worried about them* [the hospital].

Inspector 11: *Yes, we are worried. (...) But if you go around* [performing interventions to support hospitals in their quest to learn from mistakes] *I can imagine you would first visit the hospitals where you have a good hope that your visit will have an effect.*

Inspector 8: *But this is not something we formally decide or discuss.*

Inspector 11 & 3, simultaneously: *No.*

Inspector 8: *So, it's something that just happens.*

Interviewer: *There is no protocol?*

Inspector 3: *Nope.*²¹

The “willingness” and “ability” of a hospital to learn thus plays a large role in what regulatory action, if any, is taken by the Inspectorate. However, there is no formal standard to measure “willingness” and when the team is asked what the learning “ability” of a hospital entails exactly, there is no concrete answer.

We found that the informal knowledge used to determine a hospital’s “willingness” and “ability” to learn or to generally establish a pedagogic approach comes in various forms. We can provide countless examples of when inspectors not only used the BAN standard to formulate their feedback or decide on their course of action, but also relied on their own experiences, personal networks, professional expertise, and/or awareness of other ongoing regulatory programs:

Who's going to call? It's [name responsible board member], she hates us!²²

With their [the hospital's] ongoing development project and renovations, they are really in over their heads. This is really one of our high-alert merger hospitals.

Ok. Then we'll make our [feedback] letter stricter and we'll send it to the account holder,²³ to alert her as well.²⁴

Such statements reveal the informal dynamics at play at the LMO conference table, coloring the debates, stipulating enforcement tactics, and influencing the ultimate regulatory actions taken. On occasion, personal sentiments and experiences even dominate the discussions and the team members need to remind each other to uphold their own standards and protocols. For example, after a heated discussion, the chairperson pleaded:

Come on guys, please. Next time I will remove the [hospital's] name from the report and then we'll see how you judge it [the inquiry report].²⁵

As a concluding note it is interesting to mention that during the focus group discussion we confronted the LMO team with our observations on how they use informal knowledge and (sometimes) struggle to adhere to their own BAN standard. During the observed LMO internal governance meetings, which the first author attended after the focus group, the team again discussed these “struggles” and the inspectors expressed their willingness to “use” the BAN standard and digiBAN database the way they were (originally) intended to be used. However, their work routines have not changed, as became apparent in follow-up interviews.

3. Discussion

In this paper we explored how standards and a performance management system influence decisionmaking by the Dutch Health and Youth Care Inspectorate and how the Inspectorate's pedagogy is constituted through its work routine. Theoretically, these findings are relevant because they draw attention to the understudied effects of standards and performance measurement systems on standard makers themselves, furthering our understanding of the workings of regulatory power (Slager et al. 2012). Also, empirically, our findings are interesting because they shed light on the multifaceted nature of an educative regulatory style.

The Inspectorate's BAN standard and subsequent guidelines aim to constitute the learning processes of hospitals. The standard dictates what hospitals must do in the wake of an adverse event (AE), how they must investigate it, who to involve etc. The standard ultimately requests hospitals to identify specific root causes and use

these to formulate suitable measures. If a hospital does not comply with the standard it risks receiving a reprimand and/or regulatory sanctions. In this way the BAN standard may have considerable external regulatory power as it molds hospital behavior; it is in the hospitals' interest to comply with the Inspectorate's checklists. From a pedagogic notion, one may say that the Inspectorate's BAN standard, just like an educator's classroom rules, has a reverberating influence on the behavior of another party (Hansen & Lavery 2010).

Internally for the Inspectorate, the introduction of the BAN standard and digiBAN database has had a profound influence on the way in which the LMO team executes its work in practical terms. For instance, the standard dictates the way inspectors read through the inquiry reports and has standardized work practices; reports are judged using the BAN, feedback letters are drafted, LMO meetings take place to discuss the BAN scores etc. The introduced standard thus manages to standardize daily work practices, possibly even improving efficiency and traceability; which were important goals of the newly introduced "sharpened" work routine. But there are limits to the standards' internal regulatory power.

Stemming from a changed definition of risk, the BAN standard reflects the Inspectorate's intention to evaluate inquiry reports based on process instead of content; learning from what went wrong is deemed more important than what actually went wrong. In practice, however, our findings show that splitting these two values is difficult and sometimes even problematic, fueling negotiations at the LMO conference table. This leads us to conclude that in terms of content, the standards used by the Inspectorate have limited internal regulatory power as the assessment of inquiry reports and the subsequent feedback provided to hospitals is colored by the content of the incident as well as other forms of informal knowledge about hospital practice. That is, the standard in itself does not have the power to fully discipline inspector's behavior and objectify their assessments; additional "work" needs to be performed to make the standard an "instrument of control" (Brunsson & Jacobsson 2000), as exemplified by the many discussions in the LMO team. Expert judgment, ideas and interpretations are needed (Porter 1996). Continuously, informal knowledge, emotions, and the medical expertise of the inspectors slips through the Inspectorate's actions – precisely those elements that the standard seeks to keep out in light of the public debate about the Inspectorates' presumed capture by the hospital sector and the discretionary space of individual inspectors. One may argue that the "sharpened" work routine has indeed limited the discretion of individual inspectors but our study demonstrates that regulatory discretion has not been eliminated. Rather, regulatory discretion has acquired a collective nature (Rutz *et al.* 2017) as it has become embedded within the LMO team during their meetings, as well as through the build-up of common routines. This collective discretion becomes apparent in the overflow of substantive and affective reasoning in relation to the otherwise processual standard, and is equally visible in the pedagogic reasoning underlying decisionmaking.

Because of the focus on "learning," the existing relationship of dependence between the regulator and the regulated (Hawkins 1983; Legemaate *et al.* 2013) is cultivated further, thereby not necessarily limiting the politically criticized capture-style relationships. In order to live up to its regulatory promise, the Inspectorate depends on hospitals to report AEs, provide truthful inquiry reports and follow through on proposed improvement measures. As an internal constitutive effect, this dependency forces the Inspectorate into a teacher or mentor role and may explain the cautious behavior and substantial pedagogic reasoning we observed throughout the LMO team's regulatory activities. It appears that balancing this relationship at times weighs heavier than upholding the logic of the BAN standard; feeding the pedagogic approach wherein the LMO team wraps its reprimanding feedback with "red ribbons" to soften or tone down their critical feedback letters. Pupils must remain motivated to keep learning, thus enforcement tactics are carefully weighed in order to preserve, even nurture, the continuing relationship (Hawkins 1983). An educative regulatory style is therefore multifaceted, holding both deterrence and cooperative style elements and is influenced by the historical and institutional setting, both of the regulator and the regulated. In part, it entails a communicative strategy to educate the regulated of responsible behavior (Lo *et al.* 2009), in our case stimulating healthcare organizations to learn from adverse events. Based on the underlying theory of learning and depending on the relationship between regulator and regulated, as well as the emotions at play, deterrent or more cooperative strategies can be played out to attain results.

The BAN standard used by the Inspectorate has allowed for the creation and subsequent utilization of a performance measurement system. Based on our findings we can identify three types of rankings in this system. First, an absolute ranking, whereby a BAN score allows the quality of an individual report, with all of its different elements, to be compared to an absolute maximum score. Second, a longitudinal ranking revealing the progress

of scores over time for an individual hospital. Third, a ranking to compare the performance of an individual hospital to all other hospitals, in general or on specific elements. In theory, all of these rankings provide the Inspectorate with a situational context on which to base its (pedagogic) feedback and/or other regulatory activities. But surprisingly, the rankings are not systematically used that way. Instead, other sources of information, such as informal knowledge about a hospital, previous experiences, professional knowhow, and emotions play a mediating role in the assessment of inquiry reports. Thus, both the BAN standard and performance measurement database – although introduced to benchmark learning and support learning processes over time – shape the process as a whole but do not manage to standardize it. Instead, there are many overflows that influence the ways in which the standard is put to use.

Clearly our findings demonstrate that standards and performance measurement systems are not by definition “instruments of control” (Brunsson & Jacobsson 2000). They do in part constitute the practices that they measure but these constitutive effects are (under)determined by the relationships and affects in which they are enacted. Although the standard has changed the relationship between the regulators and the regulated, the Inspectorate remains dependent on hospitals as they need hospitals to report AEs and be honest about what they report. The LMO team also negotiates and is influenced by personal contextual knowledge about hospitals – both in terms of their knowledge about hospital practices in general and by specific hospital organizations. That is, information from other sources, personal emotions, earlier experiences, and subjective ideas about the willingness of a hospital to work on improving quality and safety and learn from their mistakes play a role alongside the standard and rankings. For policymakers and researchers this is a relevant point to take away from this study, as the functioning of standards is often taken for granted. Our study illustrates that one should be mindful of not only the script in use (what is the purpose of the standard) but also the context in which the script unfolds, closely examining how it is used in regulatory practices, how users make sense of the standard, and which relationships influence this use and sense-making. In this case, the constitutive effects of the standards and rankings in the performance measurement system seem to hinge on the working practices of the regulator who propagates these very standards and rankings. Standards and rankings do not “just” have the power to control, objectify, limit regulatory discretion, capture relationships, or regulate. We stress that when one determines that they do carry constitutive effects, one needs to closely examine what the exact source of the effect is.

This conclusion is underlined by the limited impact of our focus group discussion, wherein the LMO team was confronted with their unintentional use of their own standards and performance measurement system. The inspectors were surprised by the findings, yet despite the expressed intention, their daily work routines have not changed. We might hypothesize that this is not because the inspectors are unwilling but because the nature of the dependent relationship between the LMO team and hospitals is stronger than the BAN standard and digiBAN system. Thus the tendency to focus on “good students” and ignore “bad apples” (Bardach & Kagan 1982) is not just a functional hiccup. Rather, it may be a direct product of the chosen regulatory approach. That is, the focus on “learning from an AE” rather than the specifics and associated risks of an AE, executed inside a web of dependency relationships.

With regard to the Inspectorate’s “theory of learning,” we conclude that their assumptions about what hospitals should learn in the wake of an AE, are not just formally embedded in the standards and performance measurement system used. Formally, the individual elements of the BAN standard and coupled system carry with them an underlying assumption that they reflect effective learning. In practice, the definition of effective learning does not end there, as our study shows that the Inspectorate’s pedagogy is collectively negotiated. Using a pool of informal knowledge, the LMO team decides where they need – or want – to turn a watchful eye. This is clearly demonstrated by the decision to visit hospitals that do not necessarily score the lowest (the “bad apples”) but do hold the greatest potential to benefit from an intervention. These grounds for “learning potential” are not formally embedded in the LMO team’s protocols and are not standardized; they are informally enacted. In this case revealing the assumptions of an educator who feels that “*sometimes it’s just not worth flogging a dead horse.*”

4. Conclusion

This paper sought to cast light on the function of standards and performance measurement systems in regulatory contexts. For scholars, policymakers and regulators, our findings are interesting as they lend weight to the notion

that evaluation, monitoring, and ranking practices – increasingly introduced and used in a regulatory context confronted with a complex mix of political, societal, and reputational demands, as well as limited resources – do not operate neutrally and affect both sides of the regulatory equation: regulators and regulatees. Although we show that standards and performance measurement systems indeed carry the power to constitute the very practices that they measure, we stress that they are not by definition “instruments of control.” Their constitutive effects are (under)determined by the relationships in which they are enacted. As an implication we thus recommend that policymakers and regulators actively monitor and critically reflect on their own work practices and use of standards to educate themselves about the limitations and implications of these evaluation mechanisms. Scholars of regulation evaluating the effects of standards and performance measurements would do well to examine not only how these affect the regulated, but also how the consequences of such regulatory instruments are mediated by the practices in which they are put to use. As Power (2000) argued for auditing practices, the consequences of each standard, monitoring, or ranking system should be appraised on its own merits.

To conclude, we are hopeful that our ethnographic study contributes to the gap in knowledge surrounding the practical execution of regulatory work by providing a glimpse into the everyday practices and struggles of a particular regulatory body, as well as the multifaceted nature of an educational enforcement approach, holding both rule-orientated and cooperative style elements. Now that we know how the Inspectorate uses standards and have determined how standards influence the internal regulatory context, we can extend our exploration inside healthcare organizations. Is the Inspectorate’s pedagogic approach a fruitful way to support learning from AE’s and patient safety more generally? Does the chosen regulatory methodology actually make healthcare safer? For future research, the insights provided in this study allow us to better understand the effect of regulation practices and standards on the ways in which hospitals organize their internal inquiry processes and learn from mistakes, as well as how regulatory activities have an impact on patient safety in general.

Acknowledgments

The authors wish to thank the Dutch Health and Youth Care Inspectorate for opening its doors to facilitate data collection. The inspectors of the LMO team in particular have our gratitude for generously sharing their stories and work routines.

Additionally, we owe kind words of appreciation to the HCG group at the Erasmus School of Health Policy and Management and the anonymous reviewers for their critical comments and valuable recommendations to earlier drafts of this paper.

Notes

- 1 Here we must stress that although the questions “how, what and who learns in the wake of an adverse event” are relevant, deserving thorough scientific research, they lie beyond the scope of this study. We argue that in order to answer such questions it is first necessary to gain insight into the Inspectorate’s attempts to steer this learning process. Thus, our focus is on the Inspectorate’s “theory of learning,” or pedagogy.
- 2 Observation, field notes, Inspector 7, 11 August 2015.
- 3 Observation, field notes, Inspector 8 scoring, 20 July 2015.
- 4 Field notes, 16 July 2015.
- 5 Field notes, 21 August 2015, 16 July 2015, 24 February 2015.
- 6 Field notes, 24 February 2015.
- 7 Inspector 7 and 14, field notes.
- 8 Observation, field notes, LMO meeting, 7 April 2015.
- 9 Observation, field notes, Inspector 4 scoring, 21 August 2015.
- 10 Observation, field notes, LMO meeting, 7 October 2015.
- 11 At the time of the fieldwork project the Inspectorate had no programs in place to monitor these suggested improvements. We have since learned, however, that the Inspectorate plans to start monitoring this in the (near) future.
- 12 Observation, field notes, LMO meeting, 10 February 2015.

- 13 Inspector 16, Observation, field notes, LMO meeting, Inspector 16, 10 August 2015.
- 14 Observation, field notes, LMO meeting, 17 February 2015.
- 15 Observation, field notes, 16 July 2015.
- 16 Observation, field notes, LMO meeting, Inspector 4, 10 February 2015.
- 17 Observation, field notes, LMO meeting, Inspector 8, 17 February 2015.
- 18 Quotes, focus group discussion, 2 June 2015.
- 19 Focus group discussion, Inspector 11, 2 June 2015.
- 20 Quotes, focus group discussion, 2 June 2015.
- 21 Excerpt, focus group discussion, 2 June 2015.
- 22 Observation, field notes, LMO meeting, Inspector 16, 3 March 2015.
- 23 For its supervision of hospitals, the Inspectorate has assigned an “account holder” at each hospital. These (senior) inspectors are generally not on an LMO team, but may be consulted on specific decisions.
- 24 Observation, field notes, LMO meeting, dialogue Inspector 3 and chairperson, 3 March 2015.
- 25 Observation, field notes, LMO meeting, 17 February 2015.

References

- Alvesson M (2003) Beyond Neopositivists, Romantics, and Localists: A Reflexive Approach to Interviews in Organizational Research. *Academy of Management Review* 28, 13–33.
- Alvesson M (2009) At-Home Ethnography: Struggling with Closeness and Closure Chapter 8. In: Ybema S, Yanow D, Wels H, Kamsteeg F (eds) *Organizational Ethnography: Studying the Complexities of Everyday Life*. SAGE, London. pp. 156–174.
- Anderson JE, Kodate N, Walters R, Dodds A (2013) Can Incident Reporting Improve Safety? Healthcare Practitioners’ Views of the Effectiveness of Incident Reporting. *International Journal for Quality in Health Care* 25, 141–150.
- Ayres I, Braithwaite J (1992) *Responsive Regulation: Transcending the Deregulation Debate*. OUP, Oxford.
- Baldwin R, Black J (2016) Driving Priorities in Risk-Based Regulation: What’s the Problem? *Journal of Law and Society* 43, 565–595.
- Bardach E, Kagan RA (1982) *Going by the Book: The Problem of Regulatory Unreasonableness*. Temple University Press, Philadelphia, PA.
- Behr L, Grit K, Bal R, Robben P (2015) Framing and Reframing Critical Incidents in Hospitals. *Health, Risk & Society* 17, 81–97.
- Boot JM (2013) *De Nederlandse gezondheidszorg. Gezondheid en eigen initiatief. Voorzieningen. Overheid en beleid*. Bohn Stafleu van Loghum, Houten.
- Bowker GC, Star SL (1999) *Sorting Things out. Classification and its Consequences*. The MIT Press, Cambridge, MA.
- Brunsson N, Jacobsson B (2000) *A World of Standards*. OUP, New York.
- Buijsen AJM (2014) *Wetgeving gezondheidszorg. 3de herziene uitgave*. Maklu, Apeldoorn-Antwerpen.
- Dahler-Larsen P (2012) *The Evaluation Society*, Stanford CA, Stanford University Press.
- Dahler-Larsen P (2014) Constitutive Effects of Performance Indicators: Getting beyond Unintended Consequences. *Public Management Review* 16, 969–986.
- Dute J (2015) Toezicht op de kwaliteit van de gezondheidszorg: de IGZ als kop van Jut Chapter 6. In: Havinga T, Verbruggen P, Waele HD (eds) *Toezicht tegen het licht. Kernwaarden, kansen en knelpunten*. Wolters Kluwer, Deventer, The Netherlands. pp. 75–92.
- Emerson RM, Fretz RL, Shaw LL (2015) *Writing Ethnographic Fieldnotes*. University of Chicago Press, Chicago.
- Espeland WN, Stevens ML (1998) Commensuration as a Social Process. *Annual Review of Sociology* 24, 313–343.
- Garfinkel H (1967) *Studies in Ethnomethodology*. Polity Press, Cambridge, UK.
- Geertz C (1973) *The Interpretation of Cultures. Selected Essays*. Basic Books Inc, New York.
- Green J, Thorogood N (2006) *Qualitative Methods for Health Research*. SAGE, London.
- Halfman W (1998) *Standardization as a Trust Device. The Politics of Chemical Risk: Scenarios for a Regulatory Future*. Springer, Dordrecht, The Netherlands.
- Hansen DT, Lavery MJ (2010) Teaching and Pedagogy Chapter 15. In: Bailey R, Barrow R, Carr D, McCarthy C (eds) *The SAGE Handbook of Philosophy of Education*. SAGE, London pp. 223–236.
- Hawkins K (1983) Bargain and Bluff: Compliance Strategy and Deterrence in the Enforcement of Regulation. *Law & Policy* 5, 35–73.
- Hollnagel E, Braithwaite J, Wears RL (2013) *Resilient Health Care*. Ashgate Publishing Limited, Farnham.
- Hulst A, Segerholm C (2012) Inspection effects through the eyes of the inspectors: Swedish notions. ECER Conference, Cadiz. ECER Organization, Cadiz, Spain.
- Iedema R, Jorm C, Long D, Braithwaite J, Travaglia J, Westbrook M (2006) Turning the Medical Gaze in upon Itself: Root Cause Analysis and the Investigation of Clinical Error. *Social Science & Medicine* 62, 1605–1615.
- IGZ (Inspectie voor de Gezondheidszorg)(2013a) *Leidraad meldingen IGZ 2013*. IGZ, Utrecht.
- IGZ (2013b) *Werkplan 2013*. IGZ, Utrecht.

- IGZ (2014) *Uniforme richtlijn calamiteitenrapportages voor zorgaanbieders*. IGZ, Utrecht.
- IGZ (2016) *Meerjarenbeleidsplan 2016-2019*. IGZ, Utrecht.
- Jensen CB (2008) Sociology, Systems and (Patient) Safety: Knowledge Translations in Healthcare Policy. *Sociology of Health & Illness* 30, 309–324.
- Jordana J, Levi-Faur D (2004) *The Politics of Regulation: Institutions and Regulatory Reforms for the Age of Governance*. Edward Elgar, Cheltenham, UK.
- Kenkies K (2012) Educational Theory as Topological Rhetoric: The Concepts of Pedagogy of Johann Friedrich Herbart and Friedrich Schleiermacher. *Studies in Philosophy and Education* 31, 265–273.
- Kohn LT, Corrigan JM, Donaldson MS (1999) *To Err is Human: Building a Safer Health System*. National Academies Press, Washington DC.
- Lampland M, Star SL (2009) *Standards and their Stories. How Quantifying, Classifying and Formalizing Practices Shape Everyday Life*. Cornell University Press, London.
- Lascoumes P, Le Gales P (2007) Introduction: Understanding Public Policy through its Instruments—From the Nature of Instruments to the Sociology of Public Policy Instrumentation. *Governance* 20, 1–21.
- Legemaate J, Grit K, Plomp E, Die MD, Bovenkamp HVD, Broer TEA (2013) *Thematische wetsevaluatie bestuursrechtelijk toezicht op de kwaliteit van zorg*. ZonMw, Den Haag.
- Leistikow I, Mulder S, Vesseur J, Robben P (2017) Learning from Incidents in Healthcare: The Journey, Not the Arrival, Matters. *BMJ Quality and Safety* 26, 252–256.
- Lipsky M (2010) *Street-level Bureaucracy, 30th Ann. Ed.: Dilemmas of the Individual in Public Service*. Russell Sage Foundation, New York, NY.
- Lo CW-H, Fryxell GE, Van Rooij B (2009) Changes in Enforcement Styles among Environmental Enforcement Officials in China. *Environment and Planning A* 41, 2706–2723.
- May PJ, Winter S (2000) Reconsidering Styles of Regulatory Enforcement: Patterns in Danish Argo-Environmental Inspection. *Law & Policy* 22, 143–173.
- Mitchell I, Schuster A, Smith K, Pronovost P, Wu A (2015) Patient Safety Incident Reporting: A Qualitative Study of Thoughts and Perceptions of Experts 15 Years after ‘To Err Is Human’. *BMJ Quality and Safety* 25, 92–99.
- Ombudsman N (2011) *Rapport over rol UMCG en IGZ in zaak ernstige handicap jongen na behandeling UMCG*. Nationale Ombudsman, Den Haag.
- Ombudsman N (2012) *Geen gehoor bij de IGZ. Signalen over de Inspectie voor de Gezondheidszorg*. Nationale Ombudsman, Den Haag.
- Porter TM (1996) *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton University Press, Princeton, New Jersey.
- Power M (2000) The Audit Society—Second Thoughts. *International Journal of Auditing* 4, 111–119.
- Power M (2007) *Organized Uncertainty. Designing a World of Risk Management*. OUP, Oxford.
- Rhodes R (2015) Ethnography. In: Bevir M, Rhodes R (eds) *Routledge Handbook of Interpretive Political Science*. Routledge, London.
- Robben P, Grit K, Bal R (2015) Inspectie voor de gezondheidszorg Chapter 14. In: Mertens FJH, Muller ER, Winter HB (eds) *Toezicht. Inspecties en autoriteiten in Nederland*. Wolters Kluwer, Deventer pp.373–392.
- Rowley E, Waring J (2011) *A Socio-Cultural Perspective on Patient Safety*. Ashgate Publishing Limited, Farnham.
- Rutz S, Mathew D, Robben P, De Bont A (2017) Enhancing Responsiveness and Consistency: Comparing the Collective Use of Discretion and Discretionary Room at Inspectorates in England and The Netherlands. *Regulation & Governance* 11, 81–94.
- Schippers EL (2013) Kamerbrief. Kabinetsreactie onderzoeksrapporten IGZ. Tweede Kamer der Staten-Generaal, The Hague, The Netherlands.
- Scholtz JT (1984) Cooperation, Deterrence, and the Ecology of Regulatory Enforcement. *Law and Society Review* 18, 179–224.
- Slager R, Gond J-P, Moon J (2012) Standardization as Institutional Work: The Regulatory Power of a Responsible Investment Standard. *Organization Studies* 33, 763–790.
- Sorgdrager W (2012) *Van incident naar effectief toezicht. Onderzoek naar de afhandeling van dossiers over meldingen door de Inspectie voor de Gezondheidszorg*. VWS, Den Haag.
- Sparrow MK (2000) *The Regulatory Craft. Controlling Risks, Solving Problems and Managing Compliance*. Brookings Institution Press, Washington DC.
- Tamuz M, Franchois KE, Thomas EJ (2011) What’s Past is Prologue: Organizational Learning from a Serious Patient Injury. *Safety Science* 49, 75–82.
- Tucker AL, Edmondson AC (2003) Why Hospitals Don’t Learn from Failures: Organizational and Psychological Dynamics that Inhibit System Change. *California Management Review* 45, 55–72.
- Vries PWD (2011) *Onderzoekscommissie inzake Jelmer. Rapport naar aanleiding van het onderzoek naar de werkwijze van de Inspectie voor de Gezondheidszorg in de casus Jelmer*. IGZ, Utrecht.
- VWS(Volksgezondheid, Welzijn en Sport)(2012) *Nieuwsbericht 02-04-2012: Schippers wil herstel vertrouwen IGZ*. VWS, The Hague, The Netherlands.
- Wallenburg I, Quartz J, Bal R (2016) Making Hospitals Governable: Performativity and Institutional Work in Ranking Practices. *Administration & Society* 1–27. doi.org/10.1177/0095399716680054.
- Weick KE (2001) *Making Sense of the Organisation*. Blackwell Publishing Ltd, Oxford.