

Reactivity and Resistance to Evaluation Devices

Ulla Forseth, Stewart Clegg, and Emil André Røyrvik

Abstract

This paper explores the trajectory of a novel evaluation device for customer satisfaction with service encounters and the performance of financial advisors. Drawing on literature on quantification and commensuration, boundary objects and bipartite collaboration, we explore the set-up and collaboration between employer and trade union in the design phase and the actual use and translation of the valuation system. The data stems from a multi-year, multi-site study of banking in two Nordic countries. The analysis illustrates how, when some operational managers started using the device to suit their own purposes, the process morphed from an initial agreement into a dispute. The paper shows how quantitative systems of evaluation easily diverge from their initially proposed purposes when in use, producing reactivity and resistance among organizational members. Some financial advisors were positive about the evaluation device and the opportunity it afforded to improve performance, whereas others regarded it as another surveillance attempt for enhancing management control. We also contribute to the literature by elaborating on the relationship between reactivity and resistance on individual and collective levels.

Key words: commensuration; evaluation; boundary object; bipartite collaboration; customer encounters; financial services

Introduction

The [evaluation] device represents a structured way of improving customer satisfaction [...]. How they [financial advisors] perform in encounters with customers is pivotal for the customers' decision to buy something or search somewhere else. (Manager/Vice-President, 2010)

*Ulla Forseth, Department of Sociology and Political Science, Norwegian University of Science and Technology (NTNU), 7491 Trondheim, Norway,
ulla.forseth@ntnu.no*

© 2019 Author

LiU Electronic Press, DOI 10.3384/VS. 2001-5992.196131

<http://valuationstudies.liu.se>



We emphasized that this is a tool for development and not for control and surveillance. It went totally wrong in one region [...] and it ended in full ‘baluba’ [dispute]. (Shop steward, 2011)

Systematic evaluation, the making of a judgement about the number, amount or value of some phenomena based upon quantification and measurement is becoming inherent in nearly all aspects of human activity, not least in the sphere of formal work life (Ball 2010). Increasingly, as ‘selves’ quantified by such evaluation devices, we live in an audit society and culture characterized by the widespread rise of indicators, standards and rankings (e.g. Roberts 1991; Power 1997; Sewell et al. 2011; Shore and Wright 2015a), characterized by panoptical devices and self-surveillance. Under conditions of the latter, evaluation shades into valuation: an estimation of one’s worth by one’s self reflexively using evaluation devices for this purpose.

New systems of performance measurement have been introduced to work places in a range of industries, often justified in positive ways by reference to the need to meet global competition, deal with market forces and increase performance and quality for the good of all parties involved. Since the 1980s, especially, front line workers in banking have experienced fundamental changes in job demands, summarized as a transition from being tellers to becoming competitive sellers (Regini et al. 1999; Forseth 2005; Bjørkeng and Clegg 2010; Forseth et al. 2015). The specific dimensions and impacts of these changes have been addressed in our previous research such as ambivalences in frontline work (Forseth 2010) and comparing transformation of work in retail and banking (Forseth and Dahl-Jørgensen 2002). Novel monitoring and new key performance indicators (KPIs) have been introduced, along with the customer, into the employer–employee relation. The terms auditing, monitoring or mapping are often used to denote impartiality; however, the everyday acceptability of these terms obscures how the systems that sustain them construct categories central to power relations in organizations (Osborne and Rose 1999; MacKenzie 2003; Law and Urry 2004; Larsen and Røyrvik 2017). A growing literature in the social sciences suggests that evaluation should be viewed in terms of processes that enact “how people, things and idea(l)s are ordered in relation to one another” (Kjellberg and Mallard 2013: 17), as the outcome of extensive institutional effort and social practices devoted to “rendering heterogeneous resources commensurable” (Styhre 2013: 52).

Our point of departure is the introduction of a novel evaluation device; a survey for measuring customer satisfaction with the service encounter between customers and financial advisors. The name of the valuation device is “Moment of Truth” or MOT. The term initially stemmed from bullfighting and was introduced into the service management literature in the 1980s (Normann 1984 [2001]) denoting the first contact or interaction between a customer and a service provider.

CEO Janne Carlzon, in his efforts to make the Scandinavian airlines more customer oriented, popularized the term (Carlzon 1985, 1988). Transferred to a service encounter the term is taken to represent a binary situation in which the salesperson is the responsible agent. Customer satisfaction is signified by the decision to purchase and customer dissatisfaction by the salesperson's failure to close a sale opportunistically for the merchant or service provider. In cases of satisfaction managerial discourse can talk about "magic moments" and in case of dissatisfaction it refers to "moments of misery". Binaries define reality and make salespersons responsible for the decisions that customers make. In the financial industry where products are very similar, advertising that firms excel in customer care and in creating "magic moments" for customers is an important element used to make one firm stand out from the others and is an essential tool in the competitive struggle (Forseth et al. 2015). Of course, as in other retail areas, the salesperson becomes the responsible agent such that management can hold them accountable for the ratio of moments of "magic" to "misery".

By looking at the trajectory of the use of the evaluation device in regard to front line sales in a Nordic bank, we ask a research question the answer to which contributes novel insights in at least three ways. The question is simple: Over time, how are evaluation devices' function as boundary objects sustained, justified and resisted between employers and unions? First, we analyse the evaluation device in relation to a literature of social studies of quantification and commensuration (e.g. Power 1997; Espeland and Sauder 2007). Second, we contribute to the literature on bipartite collaboration between employers and unions. Industrial relations are portrayed in adversarial terms in Anglo-American contexts, unlike the Nordic setting where management and labour are conceived of as having conflicting and common interests, as "counterparties and collaborators" (Bungum et al. 2015; Dølvik et al. 2014). Thus, in our case management enrolled the union to facilitate collaboration between the two parties. They did so by using the evaluation device as a boundary object (Star and Griesemer 1989; Star 2010; Lainer-Wos 2013). Nonetheless, the enrolment was not unproblematic as it created collaboration as well as resistance. Longitudinal data, drawn from a multi-year, multi-site study of banking in two Nordic countries from 2007–12 with a follow-up in 2017, enables us to follow the trajectory of the valuation device over time from consent to controversy and beyond. Third, we bring frontline sales and sales management into the realm of social studies of finance and evaluation in order to provide a more detailed discussion of finance as a social phenomenon separate from consideration of it as an economic object of "high finance".

In the following section we outline the theoretical resources we will deploy before we present the research design and methodology. We di-

vide the data analysis into three phases: first, the set-up and introductory phase of the valuation system; second, the phases of its utilization, translation and becoming a site of emerging conflict; third, the resolution and transformative phase. We contribute to the literature by illustrating how espoused purposes marking the introduction of an evaluation system can have unintended consequences when organizational actors start using the system for personal appraisal – for valuation of themselves and others.

Our results illustrate how the process of evaluation morphed from an initial agreement among the parties concerning its use to it becoming an object of controversy when some operational managers, who had not been part of the design negotiations, started using the device for their own purpose. We renewed contact with the field eight years after our first research encounter and include results from this engagement. In the wake of a new turn to customer orientation the evaluation system has been transformed and there is more emphasis on the qualitative aspects. Finally, we discuss the contributions and the implications of our research.

Theoretical resources

The first theoretical pillar we draw on is that of social studies of quantification and commensuration, fact making, accountability and auditing (e.g. Porter 1995; Poovey 1998; Merry et al. 2015). Building on the work of Power (1997), Espeland and Sauder (2007: 1) note that: “In the past two decades demands for accountability, transparency, and efficiency have prompted a flood of social measures designed to evaluate the performances of individuals and organizations”. Their interest is in how various rankings have emerged and the effects these have, a process that they refer to as commensuration – the way that qualitative phenomena are quantified and measured to facilitate comparison. Espeland and Sauder assert that actors are reflexive – meaning they are self-aware and thoughtful about the situation they find themselves in – and try to perform well in rankings. They refer to this as reactivity: “Although definitions of reactivity vary across approaches, the basic idea is the same: individuals alter their behaviour in reaction to being evaluated, observed, or measured” (Espeland and Sauder 2007: 6). Commensuration and reactivity are a powerful mix: they constitute employees who are well aware that they are subjects of surveillance; that this surveillance will be expressed in quantitative ranking of their performance and, when they are aware of the ranking purposes, will alter their behaviour in consequence. A large literature in accounting illustrates the effects of performance measures on individuals and organizations (see for instance Roberts 1991 on accountability and the sense of self and Jeacle and Carter 2012 on accounting as a mediator between creativity and control). Metrics invented for purposes of seemingly objective evaluation can become used for something else,

such as reflexive personal auditing of performance as well as personnel evaluation (Armstrong 2002).

In the Nordic model of working life, bipartite collaboration is a cornerstone regulated by law (Bungum et al. 2015; Dølvik et al. 2014). Indeed, collaboration between employer and union representatives can go beyond the basis regulated by law, such as strategic issues related to shutdowns and hiring of managers, particularly in major enterprises (Levin et al. 2012: 98–104). In our case, the employer and the trade union shared a history of “boxing and dancing” (Huzzard et al. 2004; Forseth and Torvatn 2015), of disagreement and cooperation. Management invited the union to be involved in discussion and commentary on the proposed evaluation device. Doing so, they established that the literature on boundary objects (Star and Griesemer 1989) was pertinent for our case, given the role that the evaluation device assumed.

The concept of boundary object was initially developed as a way of analysing how artefacts can facilitate collaboration in the absence of consensus within science and technology studies. Boundary objects facilitate the ways that actors from different social realms manage to cooperate while maintaining different viewpoints, interests, values, beliefs and different interpretations of reality. Any such object has to have a certain interpretative flexibility (Star 2010). A boundary object might be any object that is part of multiple social worlds that is “plastic enough to adapt to local needs and the constraints of the several parties employing them, yet robust enough to maintain a common identity across sites” (Star and Griesemer 1989: 393). The concept of boundary objects has been adopted beyond the realm of the science and technology literature. In organization studies, for instance, it has been used to discuss artefacts and engineering prototypes, patients and patient records, project management tools, as well as software specifications and spin-off organizations (Lindberg and Czarniawska 2006; Lainer-Vos 2013; Miele 2014).

Lainer-Vos (2013: 518) cautions against the simplification of the concept of boundary objects referring merely to cooperation between groups, a tendency to employ the term boundary objects in an anecdotal manner: not *all* objects can be boundary objects. He urges researchers to focus on the non-human object *and* the persons involved, the process itself and how the properties of the object may or may not foster cooperation. In addition, it is important to look at the broader context in which these objects are embedded, keeping in mind that boundary objects might not always be able to bridge gaps between different parties over time. Agreements can turn into disputes as processes evolve from situations of cooperation to occasions where the boundary object serves as a phenomenon through which discord between various groups is articulated (Miele 2014). For instance, Boltanski and Thevenot (1999) focus on those critical moments that occur when people coordinating their actions experience something going wrong.

In such cases, old stories, forgotten things and accomplished acts might return to haunt a previously unruffled process with new and contested ways of making sense. Responsibility is the key contested process in the case of such disputes: one needs to tease out the grounds on which “responsibility for errors is distributed and on which new agreement can be reached” (1999: 359). Since there can exist many modes of justification, disputes can be understood as disagreements either about “whether the accepted rule of justification has not been violated or about which mode of justification to apply at all” (359). By bringing the justification process centre stage, they highlight the legitimacy of the agreement, rather than forging an explanation solely styled in terms of contingency, deceit or force. In this way, they aspire to describe and discuss the actors’ sense of justice or injustice, through which people ground their stances on a notion of legitimate worth.

Responsibility is the key issue in power relations, notes Lukes (2005). To be able to identify a process as an exercise of power the assumption is that it was in the exerciser’s power to act differently: “an attribution of power is at the same time an attribution of (partial or total) responsibility for certain consequences” (Lukes 2005: 58). It is for this reason that the justification process is so central: where it is the case that actors act differently in regard to attributions of responsibility, resisting some while accepting others, power relations are clearly in play. As Juris and Sitrin (2016) note, the emphasis in the literature has been on individualized, unorganized and defensive forms of resistance as a power relation. The anthropologist James Scott’s writings on the “prosaic but constant struggle between subordinates and their overseers” (Courpasson and Vallas 2016: 3) has been particularly formative in its emphasis on “infrapolitics”, the “wide variety of low-profile forms of resistance” (Scott 1990: 19), which we also find in our case.

To sum up, we draw on several theoretical pillars and concepts to shed light on the reactions to and impacts of the trajectory of the novel evaluation device as well as providing insight into job demands in front line sales in banking and contributing to the conceptual development of the relationship between reactivity and resistance to evaluation devices.

Our case study affords insight into a significant research question concerning the infrapolitics of organization and how, over time, evaluation devices function as boundary objects sustaining, justifying and resisting the relations they inscribe between employers and unions. These infrapolitics are irredeemably processual rather than structural. Traditionally employers and unionists are antagonists in industrial relations arenas, but in a Nordic context also conceived of as collaborators, and their relations of power are not fixed: they may be more or less concordant or dissonant and diverse devices can facilitate or hinder this concordance or dissonance.

In the next section, we will outline the context, research design and methodology.

Context and methodology

During the global financial crisis of 2007–08, the first author managed a research project on sales strategies and practices in finance. The turmoil that the crisis unleashed in the financial services sector brought some aspects related to sales to the fore; nonetheless, several negotiations initiated on our part proposing research collaboration came to an abrupt halt. To gain access proved very difficult in the changed climate of post-crisis financial services. After several presentations to financial business circles to establish academic relevance and legitimacy, we were able to rely on network contacts among bankers and shop stewards from previous research in finance to enrol bankers from six Nordic banks into a research project. For this paper, we draw on data from one Nordic bank that offers a full range of financial services in several European countries. The extent of its activities as well as their range made it particularly vulnerable to the crisis, such that it suffered significant economic losses compared with many other Nordic banks. The case is exemplary because of the severe impact of the crisis on this bank, making some post-crisis processes more evident.

There had been many complaints about the bank having an “introvert culture” and an organizational situation of declining customer satisfaction and loss of customers, all before the financial crisis unfolded. The crisis aggravated the situation and contributed to a further decrease in customer trust, tarnishing the image of the bank as well as reducing returns for shareholders. The overall customer satisfaction with the bank marginally improved over 2009–10 but was still considerably lower than it had been before the global financial crisis. Partly in response to the crisis an important issue was formulated in the industry as “right selling” (matching sales products to customer requirements, income and investment profile) in contrast to what had become known as mis-selling (selling inappropriate products for the customer) (cf. Brannan 2017) and overselling (selling more product than the customer could afford), which generated a lack of customer trust (Forseth et al. 2015).

In an attempt to increase customer satisfaction and pursue the aim of becoming more customer-centric, management had initiated a project to develop a new evaluation system directed towards frontline operations and the performance of financial advisors, those responsible for sales work. Additionally, its baseline data was meant to serve as a point of departure for improving individual performance among the advisors. Management enrolled the presidency of the trade union in the project, with the union becoming a representative spokesperson for the evaluation device. The performance measure, which was geared

towards a non-financial dimension of performance and a longer-term goal (customer service and customer satisfaction), introduced threshold values that became forward-looking performance targets.

We interviewed a wide range of participants as Table 1 denotes. A detailed description of our interaction with the field and the different types of empirical material in the larger study are presented in a table in the Appendix. The data for this paper mostly stems from document and website analyses and interviews conducted in two countries with a strategic sample of people covering different levels and positions in a Nordic bank. The research approach adopted was interpretative and qualitative in its research methodology.

Table 1. Position, numbers and identification of interviewees.

Before the interviews, we went through available documents and websites to learn more about the history, organizational self-presentation, key facts and figures as well as the espoused core values and strategies of the focal banks. These were important inputs when creating open-ended interview guides for different categories of personnel. After collecting documentary material, we conducted face-to-face interviews with managers at different levels, financial advisors, shop stewards and union representatives. We strove to maintain an open dialogue during the interviews and sought to understand how it was possible for these informants to be able to reason as they did. The interviews lasted from 45 minutes to 2 hours.

Position of interviewees	Numbers	Identification
Manager strategic level (M_S)	4	#1- #4
Manager operative level (M_O)	2	#1- #2
Senior financial advisor	4	#1- #4
Financial advisor	1	#1
Financial advisor/union representative (local level)	4	#1- #4
Shop steward (presidency of union)	3	#1- #3
Customer service representative	1	#1
Manager strategic level (M_S), follow-up 2017	1	#5
Total	20	

From the outset interviewees talked to us about the evaluation device (called MOT) in both positive and negative terms. Its novelty was that it used customer evaluations of staff, making the staff politically responsible for how they were judged. Intrigued, at the end of 2011, we were able to familiarize ourselves with an article written by a journalist in the trade union magazine on the topic of this evaluation device. In this four-page article, a bank manager (vice-president) and the vice-president of the professional trade union expressed their views

and discussed the purpose, practice and consequences of the new evaluation system. On two occasions, during interviews with a manager and a senior financial advisor, we were shown the MOT score reports on their data screens. We took notes about the particular items on display and recorded how they interpreted the results and the nature of their comments about the pros and cons of the MOT device. Later, we received a printed copy of the questionnaire distributed to the customers. Re-entering the field some years later we sent emails to three former interviewees, to which they replied detailing current monitoring of performance and customer interaction. Finally, we accessed the revised survey via email and undertook an additional interview with a middle manager who elaborated on the new customer-centric strategies and the subsequent revision of customer mapping and evaluation devices. During the research process, we moved between theoretical frameworks, empirical data and interpretations (Eisenhardt and Graebner 2007). The process of data analysis and interpretation proceeded through several iterative steps (Alvesson and Kärreman 2011). We first read through the transcripts from the interviews and relevant documentary material several times. In the next step, we selected important quotes representing the different functions. We coded the quotes and put them in a table for comparison. The themes identified were recurrent in the material and were resonant with interviews that we conducted, serving as a description of the system, purpose, use and consequences (intended and unintended) of the evaluation system. We then consulted literature on the relevant topics and re-read the transcripts, looking for more details and nuances in interpretation as well as relevant interview quotes. After the initial review process, we expanded the scope of the article and reread the transcripts from the interviews. A selection of quotes is presented verbatim to make the storytelling of the actors come alive through raw data that we subsequently analyse and discuss.

Assessing service encounters in banking

Phase one: developing the device

Management emphasized that it was essential that financial advisors were “willing to go the extra mile for customers” in order to enhance customer satisfaction. The innovation termed MOT was regarded as an important tool aiding such development as well as a step towards increasing competitive edge and market share. A unique feature of this particular case was that the unions were involved from an early stage. Indeed, our story began when the first author was contacted by a shop steward inquiring about her research on service encounters in finance (Forseth 2001, 2005) because their financial institution had hired a consultancy firm and was planning to set up an evaluation device

named MOT for mapping customer satisfaction with the service encounter, a measure developed in addition to their global measure of customer satisfaction. A manager elaborated on the collaboration with the union:

There is one particular aspect I would like to add, and that is that I have been explicitly enrolling the trade union. So, when we decided to go from team level to individual level, I set up a project where we invited the union. They were actually allowed to define, very, very much of our new customer survey. (Manager_S #2, 2010)

Shop stewards confirmed that they were enrolled at an early phase, collaborating constructively with the employer: “We were involved from the start and could provide input and thoughts” (Shop steward #1, 2011). The trade union, however, emphasized that it was not “their” project from the start so much as one driven by management, with the union’s primary motive being to focus on the principles beyond the evaluation device.

Our point of departure has been to safeguard in the best possible way that this is a development tool, and that it is not the only thing that is being done to attain more customers. (Union vice-president, trade union magazine: 7)

The trade union also proposed that while all financial advisors could not reach the top scores, the scores could nevertheless serve as a stimulus for improving personal performance. Indeed, it is worth noting the emotional terms in which they talked about the valuation device: “As long as it [the device] functions as a caring push” (union vice-president, trade union magazine: 7). The term “caring push” is here used to denote a positive aspect of the evaluation tool (cf. Sewell and Barker 2006), which was a prerequisite for the trade union’s support for the new system. In their opinion, the most important aspect of the baseline data produced by the device was to help low performers to develop and discover those areas of the counselling (sales) process that they needed to improve. The device took on a “human” caring dimension by virtue of this emotional expression. Although each individual salesperson would be responsible for what they achieved as measured by the device the implicit agreement was that they would not be held responsible and be punished for low scores but be assisted to improve. Management stressed that the new evaluation tool had several positive advantages: it would serve as a stimulus and guidance for the individual financial advisor, afford systematic follow-up (coaching by operational sales managers and a tool kit) and be a benchmark for teams.

It’s not the measurements themselves which are good, but the knowledge that the measurements gives, which provides an opportunity to discuss what an advisor does well and what to do better. MOT is more objective than two peoples’ sub-

jective opinions, because it is the customer's honest opinion we get. (Manager/vice-president, trade union magazine: 6)

The quote illustrates several aspects of managerial discourse. Data from the monitoring of performance is presented as being more “objective” than the “subjective” views of a manager (or the individual financial advisor) because it is based on the “customer's honest opinion”: it is assumed that the customers do not express subjective views, something actually backed up by the vice-president of the trade union: “It is a more objective tool compared to what a single manager can do making studies within his own brain” (trade union magazine: 7). Neither union nor management problematize opinion from an uncontrollable entity, the customers, at this point. Instead they seem to take the results and the measures from this survey at face value. Doing so reflects a belief that has become ubiquitous under neoliberalism, that “the market” is objective and “always right” (Røyrvik 2018). Although the knowledge generated by the customers has a different position in relation to the employer–employee dyad, as a third-party view, it is definitely not “objective”.

There was an initial agreement, after the crisis, regarding the need to improve customer orientation and customer satisfaction with the service encounter shared by the union and management. They agreed that the MOT device could be one measure for achieving that goal. The two parties, however, underscored different aspects of the value of this performance evaluation device in their accounts.

We did not get a single complaint – or rather, the union did not get a single complaint. That is worth noting. It has to do with the way one can break the prejudices one can have about each other, and then work unbelievably constructively together. I think that I have a super collaboration with the vice-president and his people [in the union], and I also think that they experience that we are very responsive towards them. But from time to time we disagree, and then we discuss it and so on. (Manager_S, #2, 2010)

In the quote, he talks about “we” in the implementation process but later spells out that it was the union that did not get a single complaint. He also seems to attribute the lack of resistance partly to the process of bipartite collaboration, in which the two parties had shared their viewpoints and discussed details regarding the implementation and use of the new valuation device, reaching agreement. The manager told us that in the past there had been strong adversarial relations between the social partners, even conflicts. This project, he continued, had demonstrated that adversarial relationships did not have to be the norm, that one can share visions and viewpoints and come to an agreement if one establishes a respectful dialogue early. As such it was an example of “dancing” (Huzzard et al. 2004) together for enhanced customer orientation and improved performance, notwithstanding that

the parties could at times have different definitions of the situation (cf. Rosness and Forseth 2014). Thus, in the first phase the union became a representative spokesperson for the performance device. Two years later, a shop steward told us that systems for performance measurements were actually one of their main areas of concern. We will subsequently present data from union members and financial advisors expressing more nuanced views and critical comments regarding the use and translation of the evaluation device.

Phase two: using the device

The protocol for using the evaluation device was that after a service encounter with a financial advisor, every fifth customer would receive an email and be asked to answer a questionnaire about the counselling and the service they received, using a scale from one to ten for each parameter. The questions covered the following topics: first impressions (including physical premises); preparation and performance of the advisor; were they satisfied with the service encounter? did they gain more than they expected from the encounter? would they recommend this bank to friends and relatives based on this particular encounter? These questions were intended to tap into those dimensions of behaviour that were controlled by the financial advisors. Around twenty items in the survey dealt with the advisor.

At first the average scores were aggregated and presented only at the team level but from 2010 onwards summary reports were distributed to each individual financial advisor every sixth month. These reports were identical with MOT but quite symbolically were called “I-MOT”, literally referring to “Individual MOT”¹. These electronic reports gave the individual advisor feedback on performance and served as input for a conversation with their operational sales manager about the results.

We discuss the experiences of the customers, why they have experienced this, and what we can do in order to ensure that they get even better experiences. (Manager_vice-president, trade union magazine: 6)

In order to improve performance, individual advisors would have to move out of the “comfort zone”. During the conversation with their operational manager they would together formulate an action plan for what the advisor could do “significantly differently”. Management, however, underscored that this was only one element in a bigger plan for developing employees, enhancing customer orientation and introducing a programme for sales training. The baseline data was meant to serve as a good point of departure for improving individual perform-

¹ In Scandinavian languages (“imot” “imod” “emot”) can mean both “towards” or “against”, an indexical meaning in terms of the use to which these electronic reports would be put.

ance, but the results were also aggregated on a team level to enable comparison between teams, branches and brands.

In the following we will elaborate and discuss different reactions to this mapping as well as what different stakeholders summarized as the strong and weak points of the new evaluation device. The reactions to the new device were threefold; some welcomed it and were glad to learn how their customers rated them. One senior financial advisor (#3) was among those who welcomed the device, saying he was very happy about receiving individual feedback – “many of us like to be seen”, he added. He was among the top ten performers and usually had high scores on all parameters listed above, except for the selling of complex financial products. He added that experienced financial advisors tried to comfort newcomers who were nervously occupied with their presentation of self after poor scores on the I-MOT. By comforting their colleagues, they tried to counteract the negative impact. Other financial advisors regarded the MOT device as just an add-on to the other performance indicators, although they raised the question of why they as professionals had to be monitored on all aspects of their job. The third group of financial advisors and several union representatives were critical, and they complained that the MOT measure added to the enhanced sales pressure and the strain of the job.

It is yet another key performance indicator which contributes to more stress. The problem is that there is always something new – and it has to be measured, and there are no reductions in the number of previous measures. [---] I am completely disciplined, and if I get a poor I-MOT[result], I have to work hard. (Financial advisor/union representative #6, 2011)

He further commented that he thought most of the financial advisors would have liked to dismiss the MOT device (yet some advisors had opposing views and used the term “exciting” about the device). Besides, he underscored that it was very negative if operational managers did not make their own decisions but relied too heavily on these performance mappings. In his opinion, there was a general pattern that younger financial advisors accepted new evaluation devices more easily because they were used to them and enjoyed receiving feedback, whereas some of the more senior financial advisors disliked such feedback, something that touches on a wider discussion about monitoring and its impacts. A shop steward (#2, 2011) talked about what he saw as double communication within the financial industry in general: “competent professionals are told that they are free to perform but in reality they are controlled from A to Z.” He added that he was surprised that there was not more resistance towards this kind of control and surveillance. One manager (Manager #2, 2010) said that employees in banking had been “subject to 15 years of torture”, for which he used the term “KPI-bulimia” to denote the large number of KPIs in

use. Personally, he would have liked to reduce the numbers, but he was in favour of the MOT evaluation device. At an in-house workshop run by the financial union several employees talked about the increase in monitoring, referring to it as “kindergarten” mentality. They experienced such monitoring as infantile and as an insult to their professional identity.

In an answer to our question on what the union regarded as the strong and weak points of the MOT device, we received the following response:

The strongest part is to visualize if we are good at customer orientation. It is also a tool for development and improvement. There are tool kits for every part, and different measures such as coaching. The negative part is if MOT becomes related to pay, and there is a display of results without a focus on improvements. Someone might also take the opportunity of manipulating with MOT; it is very simple to manipulate. (Shop steward #1, 2011)

Not only this shop steward but also several other financial advisors emphasized that I-MOT should be a tool for the individual financial advisor and the operational manager, summarizing the positive aspects from their point of view, albeit that the union was also concerned about possible downsides of the mapping if it was related to pay and employment relations. When we asked financial advisors how one could manipulate the evaluation device, one of them gave us an example while showing us his I-MOT score report:

By not closing the case [making a registration] from a service encounter with a negative outcome, such as when customers are denied a loan. Such encounters are likely to get negative scores from the customers. I closed a meeting and told the customers: “Sorry, but you cannot realize your dream about buying a shop and starting a business.” Of course, they will not give you good scores. (Senior financial advisor #3, 2011)

He went on to tell us that he had heard that other colleagues who did not provide the customer with what they wanted, did not register such encounters to avoid receiving a negative outcome. “Because if two out of twenty give you bottom scores, you get a very poor result”, he added. One of his customers had also commented on the framing of the survey questions and written a letter to the advisor’s superior. The customer had objected to the question as to whether or not she gained more than expected out of the encounter with her financial advisor. She wrote that she always got what she expected and that she did not expect any more. But if she answered that she got what she expected, it would be interpreted as a negative result. He also said that his colleagues on site received low scores on the question about first impressions, with one common explanation for this being that customers were counselled in tiny cell offices without windows. There was one exception, which might have to do with how she greeted her custom-

ers, the colleagues reflected. Customers could experience service encounters differently even if the physical space was the same. Using the evaluation device at group levels made sense for identifying shortcomings with the service encounter but made less or no sense for assessing the individual performance of advisors. Seemingly, none of our informants questioned either the specific name of the evaluation device or the “truth” aspect in its title. One explanation might be that they were familiar with the term MOT from the service management literature, or that they just regarded it as a name with a fancy acronym. Neither of the representatives involved in establishing the device, from either the employer or the union side, questioned the legitimacy of a performance measure called the MOT being rated by a third party. In situ it was perceived as a different assessment from one that would obtain from the subjective views of a manager – as if customers did not offer subjective views.

Management underscored that customer satisfaction was pivotal for improving the competitive edge, that the MOT valuation device was an attempt to have a way of mapping performance in this regard as well as a structured way of following up the results. When asked about the validity of the MOT valuation tool, a manager answered the following:

I am not so concerned about people interpreting the questions differently as long as you have enough observations. [---] If one should be critical, there are two other questions: First, do the customers answer in an honest way? They don't; customers always answer by saying they are more satisfied than they really are. Because we find it a little difficult, at least in this country, to say something negative about each other. Then there is the other point: One has to be careful and not make Moment of Truth into a global measure for customer satisfaction. (Manager_S #2, 2010)

The above quote discusses some aspects of the validity of the measure and identifies some ambivalence on part of the employer. Previously, management underscored the objectivity of this measure because it was a judgement made by customers as a third party. The manager was not concerned about customers' interpretation of the questions, rather that customers tended to give higher scores in terms of their satisfaction because they liked to be nice and not criticize their advisor too harshly. This contradicts former claims by management and the union about customers not being biased but expressing their “honest views”, as illustrated in previous quotes. He also points to the fact that the device is only one way of tapping customer satisfaction with services in the bank. Next, we will look into what happened during and after the occurrence of a critical incident.

Phase three: performing the MOT

The union started receiving complaints from their members about the use of the evaluation device in 2011, after which the union and its members began to assume a more critical position towards the device, increasingly seeing it as a tool for surveillance and control. Some operational managers had started employing the evaluation device to satisfy their specific need for monitoring performance and ranking individuals. In one region of the company, we were told, some managers had ranked their sales personnel, the financial advisors, in categories related to performance and put it on public display within the organization. “Results have been posted on the board so that everyone can see it. It does not matter so much but it is a new turn of control and discipline” (financial advisor/union representative #3, 2011). A financial advisor/union representative (#4, 2011) added that “it was not official that it had an effect on wages, but we know it does”. More critical views emerged as illustrated in this quote: “When we experience that some managers can use objectives such as total of 8.5 scores [on the mapping] to ‘strike people in their heads’, there is something wrong going on.” (Union vice-president, trade union magazine: 6). Clearly, the capacity to close a given volume of sales depends not just on the salespersons powers of selling but the customers disposable income, something highly variable between different branches. The MOT device, in becoming an individual stratifying practice, diverged from the initial purposes that management and the trade union representatives had agreed upon. Indeed, both the union and management had lost control over how the measure was used in some parts of the focal organization.

Shop stewards disliked how employees were being construed as deviants because they failed to reach the aim of 8.5 formulated by management or to hit the top scale. Trade union members complained that the valuation device was expanding in importance and becoming all-embracing, encompassing aspects that had not previously been mentioned. The trade union coined the term “failure percent” to illustrate a worst-case scenario. They felt it was completely wrong for employment relations to be influenced because one failed to reach the goal, receiving 8 instead of 8.5 on the scale. Members also reported feeling threatened when I-MOT was used in relation to rewards systems and employment relations. Besides, trade union members claimed that operational managers were also using the device as a tool for dismissal. Some union representatives even talked about “management by fear” among employees, as it was thought that managers would start using the rankings to eliminate low performers in upcoming processes of layoffs caused by a round of “redundancies”. Using the metrics in this way was a dramatic shift from the expressed intentions in the design negotiations, not at all what the union had agreed to in the initial discussions. Union members also started blaming the union for its collab-

oration, accusing it of “only playing on the keyboard of the bank and being indifferent to members of the union”. Negative voices and resistance that had earlier been silent and/or silenced came to the fore, with the union becoming cast as advocates of increased surveillance by management. The union initially had been enrolled by management in an area (performance measurements) of management concern and through its support helped shape a frame for thinking about the evaluation system within the organization that was initially hegemonic in its translation of employee interests to the device. Eventually, however, the union assumed a more critical position and underscored the control aspect of the measure that they had supported their members using. Having initially interpreted the device in positive terms, and acquiesced in the development on these terms, they now came to translate what the device did as something antagonistic to its members’ interests.

One shop steward (#1, 2011) said that while bipartite cooperation with management was important it was sometimes double-edged, presenting challenges. He had played an active part in the initial phase and maintained that even if the union made some critical points, they became co-producers with management. He elaborated on what often happens in translation processes after the initial agreement on implementation and use.

We are presented with one “canvas” that we discuss. Then we get a different “canvas” from the [operational] managers. The pictures and thoughts do not come forward in a consistent way – there are several messages – and Human Resource partners in various regions interpret the message in different ways and the practice becomes different. There is a lack of follow-up that ensures that everyone gets the same understanding, the same practice and the same follow-up. There are too many images to keep track of. When the original “canvas” is presented the motive as well as the frame have to be included. (Shop steward #1, 2011)

The metaphor of a canvas is used in order to exemplify the translation process that occurred and how different stakeholders interpreted and used this process differently, especially those who had not participated in the initial discussions. In this case, a new metric invented for one purpose, as a development tool for more customer orientation, became translated into something completely different – a performance measurement for control and discipline purposes. In this case the initial agreement morphed into a disagreement regarding the use of the valuation tool by some operational managers. In order to discuss the use and manipulation of the device and solve the controversy a meeting was arranged between the employer’s human resources department and the union. The outcome was a document spelling out “facts or fictions” about the MOT device, what had been agreed (the initial intentions) and disagreed. Regarding the future use of the device, a shop steward (#1) made a prophecy that it would fade away after a period

of time if increased returns did not flow to the bank as a result of its deployment.

Discussion

In the wake of deregulation in the industry since the 1980s, the introduction of new technologies and a range of new financial products and services, employees in banking have become subject to additional KPIs and novel devices for monitoring performance. We have shed light on the trajectory of a novel evaluation device for customer satisfaction with service encounters and the performance of financial advisors; its use and translation, the reactions and impacts.

Acts of quantification and commensuration – turning qualitative phenomena into quantities for making comparisons – underpin contemporary work life. We are accustomed to hearing about the performance of individuals and organizations being expressed in numbers. Espeland and Sauder (2007) identified three characteristics of commensuration that produce reactivity. They are “its capacity to reduce, simplify, and integrate information; the new, precise, and all-encompassing relationships it creates”, and “its capacity to elicit reflection on the validity of quantitative evaluation” (Espeland and Sauder 2007: 16). The first point is that commensuration reduces large amounts of data to a single number in a ranking table: “Numbers circulate more easily and are more easily remembered than more complicated forms of information” (2007: 18). The second point identifies that commensuration builds a common relationship between separate individuals by comparing them using the same metric. Simultaneously, the metric differentiates between different people – it creates a set of relations between them in terms of their distribution on the metric’s scores. The third point concerns what the numbers produced in the act of commensuration actually mean. Espeland and Sauder (2007) illustrate that the producers of rankings see what they produce as a “real” account of relations between different individuals that tend to be accepted because of the legitimation that metrics afford.

Leonardi’s (2011) study of computer simulation popularized the notion of a boundary object (such as an information and communications technology (ICT) system) as affording certain possible actions. The evaluation device, as a means of producing a metric, provided certain affordances, most notably that employee ratings construct a social reality that is then taken to represent the reality of the organization. Such an affordance is not an objective property of the device being used so much as of that which, in a second order manner, the device makes possible. What the device makes possible is a means of formulating and doing differentiation in performance. Such quantitative measures and rankings do not simply register phenomena already existing in social reality or represent them in a neutral way. On the contrary, such systems are generative and performative technologies that

co-construct and help shape the phenomena they are meant to “record” (Osborne and Rose 1999; MacKenzie 2003; Law and Urry 2004; Larsen and Røyrvik 2017), something explicitly addressed in contemporary marketing and management research (Poon 2007; Azimont and Araujo 2010; Mason et al. 2015). Furthermore, such systems not only alter behaviour but also give rise to new subjectivities and forms of thinking that come fused with normative premises that facilitate regimes of (self-)control and systems of discipline (Shore and Wright 2015b). Reactivity is thus part of a cultural transformation, in this case within the organization, spurred by these measurements.

In our focal organization we identified very different patterns of reactivity and resistance to being evaluated; a wide range of low-profile forms of resistance (infrapolitics), as well as positive reactions, were observed among interviewees and their colleagues. The *high performers* rejoiced and expressed satisfaction at the possibility of keeping track of their sales figures and being/becoming visible. At the other end of the scale, *the critics* blamed the “system” and complained about increased sales pressure, stress and strain and enhanced surveillance. In between these categories, some employees tried to do their best and distance themselves from the performance monitoring, sales pressure and fierce competition.² We identified interesting patterns relating to reactivity and resistance illustrative of how the two concepts are patterned. Regarding reactivity, financial advisors expressed very different opinions towards the MOT device. Voicing criticism, however, did not necessarily lead to taking action. One senior financial advisor described himself as a resister refusing to sell complex financial products (one important KPI at that time). He underscored the ethics of fully understanding the products he offered to his customers. Others could be categorized as “passive resisters” because they had decided to do their best at sales without becoming obsessed with the MOT figures. Some, in the pursuit of top scores, manipulated the system by not registering customer encounters with negative outcomes on the part of the customer. The practice illustrated variance in individual actions but individual complaints to the union also spurred on what we identified as collective action. Indeed, a dispute arose between management and the trade union after a critical incident when some managers started using individual results from the evaluation device to suit their own purpose. Harsh criticism was made of management, but union members also blamed the trade union for teaming up with management instead of taking a critical position. Finally the union took on a more critical stance towards the evaluation device due to its translation and

² These categorical distinctions fit well with a typology of sellers identified in a previous case study in insurance; “super seller”, “service seller” and “system seller” (for further details see Berge et al. 2009).

use for surveillance in contrast to the development tool to which they had agreed.

The MOT rating system was a device that, in making individual advisors commensurably and publicly responsible for securing sales, exercised discriminatory power within the bank. The ratings led, wittingly or otherwise, to a fundamental reshaping of how financial advisors were encouraged to see themselves and how others saw them. The critical juncture happened when the I-MOT system³ partially transformed from a development tool for learning how to be more customer oriented to an individual accountability device geared towards surveillance, ranking and disciplining performance. For managers used to a culture of spreadsheets and comparisons, this was just another tool for ranking employees, one that they could use to produce numbers without necessarily reflecting on the consequences of doing so.

Publicly ranking individual's measures triggered employee resistance and the trade union shifted its position from being a spokesperson for the system to becoming critical. Reactivity partially changed to resistance. Such quantitative measures and rankings did not simply register phenomena already existing in social reality or represent them in a neutral way. On the contrary, such systems were generative and performative technologies that co-constructed and helped shape the phenomena they were meant to "record" (Larsen and Røyrvik 2017). Furthermore, such systems not only altered behaviour but also gave rise to new subjectivities and forms of thinking that fused with normative premises facilitating regimes of (self-)control and systems of discipline (ibid; Shore and Wright 2015b). Reactivity is thus part of a cultural transformation, in this case within the organization, spurred on by these measurements.

In a postscript to the research, some eight years after our first visit, we followed up on the research to see what use of the device was being made. We sent emails to some of our informants asking about the current state of play in terms of the use of the MOT device. According to an email response from the union "it used to be a big issue among the members", but the bank did not use this device anymore, and we were advised to contact the Human Resources department for further explanation. Besides, they wrote, "Mystery Shoppers were the thing", and in addition, the bank had hired a consultancy to investigate its customer base. A financial advisor replied that they still used an evaluation device but in a different form, which he later forwarded after

³ Interestingly, the acronym MOT has a double meaning when read as a word in Scandinavian languages. It can denote both "courage" and "against". With the turn to I-MOT, the individualized accountability and discipline-oriented version of the system, the acronym read as a word shifted meaning from "courage", while its denotation as "being against" strengthened. This shift in meaning signified the simultaneous turn to an attitude of resistance against the system in the organization on the part of the union.

checking with his superiors. This email survey about a specific service encounter is distributed to a representative sample of the customers except those customers who have just received the global customer satisfaction survey. The name MOT had disappeared, and at first glance the new opinion monitoring was a simplified version with only three questions to provide learning for the individual advisor. Core questions in the new device were recommendation of the bank to friends and relatives, and satisfaction with the counselling encounter, including the possibility for a qualitative comment. The final question had to do with what the “advisor did particularly well” (seven different options with ‘nothing’ as a final category) and what s/he could have done better (the same seven categories). We were told that the qualitative comment and the customers’ stories constituted the most important part of the responses.

Questions outside the control of the individual advisor, such as first impressions (physical premises), had apparently been removed. Management seemingly had been through a process of learning by doing and realized that in order to enhance customer satisfaction the bank had to change the strategy and shift focus from sales to the customer’s point of view. Besides, by listening to customers, the financial advisors and the union they had learned that some questions in the MOT device were difficult to interpret and answer. When advisors started telling customers about the importance of the mapping and how they should interpret and answer some of the questions (e. g. “did you get more than expected”) there was a problem of validity.

Sales targets for individual advisors had been abolished and they no longer received average sales scores. Instead management mapped customer satisfaction and customer development (retention of customers and influx of new customers). In addition, they monitored how the different branches were performing in economic terms. We were told that the officers were relieved but also somewhat frustrated as they no longer received computed average scores. Managers, used to a culture of spreadsheets and comparisons, also felt uneasy when they could not rely on individual sales figures but were told to keep track of more long-term customer relations. By shifting focus, and doing less sales monitoring, customer satisfaction had initially increased but then decreased after the introduction of more graduated fees and charges depending on the bank account and the number of services the customer was using.⁴

⁴ After media coverage of the “Panama papers” in 2016 and the extent of money laundering and tax evasion afforded by tropical tax havens, customer satisfaction decreased somewhat. It is not only the “moments” of the service encounter that are important: global issues also have a significant impact on the reputation of individual banks and other financial institutions.

Management listened to the criticisms that developed but a decisive point came when customers also voiced critical comments about the content of the MOT device. Management realized that they had to think twice about their assumption that this was a strategic tool for mapping customer satisfaction. In order to bring customers and their perspectives centre stage, management decided to monitor different things: customer satisfaction and customer development; retention and influx of new customers; dropping individual sales targets for the financial advisors. In so doing they redefined the power of advisors: they were no longer comparatively and differentially more or less responsible in terms of procuring more sales. After some adjustments a modified and simplified evaluation system continued its organizational existence but with emphasis on qualitative comments and customers' stories.

Analysing the evaluation device as a boundary object, a pattern emerges. Initially the device successfully enacted and in a sense "negotiated" a constructive collaboration for organizational development between the work-life partners – management and the trade union. Being involved in construing the purpose and design of the system, the trade union was enrolled as a developmental agent. The project of shaping the device organized consent and enacted a "dance" between the partners. However, when aspects of surveillance, ranking and control of the evaluation device I-MOT were accentuated, rather than being a boundary object that facilitated a dance of collaboration between partners the device became the occasion for staging a "boxing" match between them (Rosness and Forseth 2014). The evaluation system seemed not so much to provide a connecting bridge between adversarial opponents but rather became an occasion for articulating discord (Miele 2014).

The evaluation system continued its existence in a modified version such that, even after the main disagreements were seemingly resolved, the technology did not lose all of its boundary object qualities. Indeed, in spite of the strong resistance and criticism heaped upon management because of the use made of the measures by certain managers, the partners continued to collaborate and sought to fix the evaluation system and the situations it created and measured. With the renewed customer-centric strategy, less sales monitoring at the individual level and a more long-term perspective, collaboration between the parties returned to the "dance floor".

Conclusion

Recall the research question with which we guided our research: how do evaluation devices function as boundary objects sustaining, justifying and resisting relations inscribed between employers and unions? Our case study affords insight into a significant research question concerning the infrapolitics of organization and how, over time, evaluation

devices function as boundary objects sustaining, justifying and resisting the relations they inscribe between employers and unions. These infrapolitics are irredeemably processual rather than structural: they are not inscribed in employment relations per se but the processes whereby these are monitored, measured and managed.

The evaluation device, as a boundary object, did not function as a device offering unchanging affordances. Boundary objects have been, since the inception of the concept, celebrated for their role in facilitating collaboration and cooperation. Thinking of the evaluation device as a boundary object, it is evident that its effects were variable, with the variance being enacted by changing power relations that were themselves inscribed in the uses to which they were put. In the wake of the global financial crisis of 2007–08, when the whole area of financial services was being stigmatized, the evaluation device became an occasion for union/management cooperation to restore consumer confidence. The device, in the abstract, seemed to afford a way of enhancing collaboration between financial advisors and the management of the bank: it was a tool used to engender cooperation.

The shock of the global crisis receded and as it did the second order effects of the implementation of the device became more apparent. In practice, the device became a tool for management with which to differentiate the performance of financial advisors in terms of their powers as salespersons, judged and justified by the responsibility attributed to them for the sales they made, irrespective of the context and customer characteristics of the branches in which the sales were made. In consequence, the affordances of the device as a boundary object changed from being primarily productive of cooperation and collaboration to becoming an object of discord. The use of the device provided an occasion on which to project concerns about the extent to which the devices' possibilities were perceived differently and were being used in vastly different contexts. These contextual factors changed with the relative normalization of the industry after the crisis but the demographic differences in the customer base persisted.

In terms of infrapolitics it is evident that while the MOT system was introduced responsibly by agreement between both sides of the employment relation, union and management, as its use developed as a device joint responsibility eroded. The device was an instrument and it was not that the instrument changed but the uses to which it was put were extended without joint responsibility for doing so being sought. Once the union was challenged with responsibility for the changed uses to which the device was being applied, including personal and public rating and ranking, the infrapolitics of collaboration changed. Individual resistances followed by collective resistance by the union marked the beginning of the end of the device in an instrumentally intrusive manner, as we discovered when the field was revisited some years later.

Over time, initial agreements between labour and employer morphed into discord and resistance. Once the project was implemented the device functioned in a system of representations long established and routine. These representations were not just about the personal use of the feedback as a learning device occasioned by the metrics but also the personnel uses to which an overall knowledge of the distribution of the data allowed. While the individual organizational member might learn from their individual feedback, the organizational management learned from the distribution of the metrics across the population of financial advisors. The affordances shifted on a number of grounds. First, there was the amount of data available to the individual, with relative position in the management hierarchy of those using the data being the crucial variable. Second, there was the demographic distribution of familiarity with notions such as the quantified self, largely the preserve of the younger financial advisors and some high performing senior advisors, according to our interviewees. Third, the device favoured those that were demographically favoured in their client base and stigmatized those that were not. Hence, what were the affordances of the evaluation device varied with its use, not its design as an instrument. The affordances were an effect of hierarchical access to the overall data; to more youthful familiarity with the use of personal metrics of performance and to some extent were also a function of customer demography, as a major determinant of “magic moments”.

Some of the downsides of the MOT mapping, as we have described them, were acknowledged by management such that today there are no individual sales targets for financial advisors but there is monitoring of customer satisfaction and customer development. Hence, resistance to the use of the device was tempered and the current device has been renamed, simplified and qualified to remove some difficult questions for customers and elements beyond the control of the personnel concerned. Besides, there is more emphasis on the learning potential from the qualitative comments and customers’ stories, and new ways of mapping customers’ preferences for and use of different channels, with a shift away from commensuration and its justifications.

In conclusion, we may say that, theoretically, the case illustrates some valuable lessons when a qualitative phenomenon (customer satisfaction) is quantified and measured to facilitate comparisons, as described in the social studies of literature on quantification and commensuration (Power 1997; Espeland and Sauder 2007). Individuals alter their behaviour as a consequence of monitoring and surveillance, which changes the relations between service provider and customers, between employees and employers and the presentations of self made by service providers. In our case, in lay terms the figures were seen as reflecting instead of constructing reality, with the evaluation device contributing to enhanced sales pressure instead of more customer orientation. In addition, a metric invented for one purpose became used

by some managers for both personal auditing and personnel evaluation (Armstrong 2002).

Acknowledgement

We are grateful for support from the Research Council of Norway, 2008–11, [grant no. 185055/520] and the Department of Sociology and Political Science, Norwegian University of Science and Technology. Thanks to Tove Håpnes for her contribution in data collection, and feedback from our interviewees and stakeholders in finance. We are also grateful for inspiring comments from the anonymous reviewers and the editor.

Appendix: Overview of interaction with the field, empirical material and methods.

Interaction with the field 2007–17	Size	Topic	Empirical material	Analysis
Public documents	950 pages	Context data, self-presentation, sales strategies, performance, incentive schemes, financial instruments, financial crisis, reputation, legal court cases, regulations	Websites, media clippings, trade union magazines	Used as context data and input to formulation of interview guide and analysis
Meetings/seminars with management, employees, trade union in-house, 5 conferences in the industry	8 occasions attendees varying from 3 to 50 persons	Context data, sales rhetoric/incentives/practices, impacts of financial crisis, trust, reputation, regulation	Presentations, audio recording, field notes	Data gathering, context, check for resonance and relevance, reinterpretation of data material
Face-to-face interviews, follow-up via telephone: (the total number of interviews exceeds the number of persons as we had several interviews with some managers and shop stewards) Total: 41	Managers, financial advisors, customer representatives, shop stewards, union reps. (31). Board members union and trade organizations (5), advisor financial service authorities (1)	Sales and customer orientation post financial crisis	Verbatim transcriptions, audio files	Looking for patterns, anomalies, the empirical material has been coded and categorized, discussed with involved parties, recoded and reinterpreted. Comparisons across regions and branches, management and employees
Email interviews / correspondence on crisis, sales	Trade union seminar (30). Threads from 1 reply to 20	Sales goals vs practices: seller skills, customer orientation, performance, etc.	Email texts	
Quantitative mapping	Interviewees, trade union seminar	Organization, layout of premises, informal talk	Excel charts, numbers	
Observation	15 visits, 7 premises (each 1–2 days)	Photographs, field notes		
Internal documents from the banks	250 pages	Institutional self-presentations, sales, customer orientation, reputation and trust post crisis	Reports, presentations, customer surveys	Background material informing data gathering, checked for consistency with interpretations

References

- Alvesson, Mats, and Dan Kärreman. 2004. *Qualitative Research and Theory Development: Mystery as Method*. London: Sage.
- Armstrong, Peter. 2002. "The Costs of Activity-Based Management." *Accounting, Organizations and Society* 27(1-2): 99–120.
- Azimont, Frank and Luis Araujo. 2010. "The Making of a Petrol Station and the "On-the-move consumer"." Classification Devices and the Shaping of Markets. *Industrial Marketing Management* 39(6): 1010-1018.
- Ball, Kirstie. 2010. "Workplace Surveillance: An Overview." *Labor History* 51(1): 87–106.
- Berge, Liv, Ulla Forseth, and Tove Håpnes. 2009. "Salg, salg, salg – over hele linja! Mandag hele året? Om mestringsstrategier og arbeidsstiler i salg av forsikringer." *Sosiologisk tidsskrift* 17(3): 195–216.
- Bjørkeng, Kjersti, and Stewart R. Clegg. 2010. "Becoming DragonBankers: Constructing Practice through Processes of Socially Situated Learning." *Society and Business Review* 5(1): 48–65.
- Boltanski, Luc, and Laurent Thévenot. 1999. "The Sociology of Critical Capacity." *European Journal of Social Theory* 2(3): 359–377.
- Brannan, Matthew J. 2017. "Power, Corruption and Lies: Mis-selling and the Production of Culture in Financial Services." *Human Relations* 70(6): 641–667.
- Bungum, Brita, Ulla Forseth, and Elin Kvande (eds.). 2015. *Den norske modellen – Internasjonalisering som utfordring og vitalisering*. Bergen: Fagbokforlaget.
- Carlzon, Janne. 1985. *Riv pyramiderna: En bok om den nya människan, chefen och ledaren*. Stockholm: Bonniers.
- Carlzon, Janne. 1988. *Moments of Truth*. New York: Ballinger.
- Courpasson, David, and Steven Vallas (eds.). 2016. *The Sage Handbook of Resistance*. Los Angeles, CA and London: Sage.
- Dølvik, Jon E., Tone Fløtten, Jon M Hippe, and Bård Jordfald. 2014. "The Nordic Model towards 2030. A New Chapter?" Fafo-report No. 2015:07. Oslo: Fafo. www.fafo.no/images/pub/2015/20412.pdf, accessed August 15, 2016.
- Eisenhardt, Kathleen M., and Melissa E. Graebner. 2007. "Theory Building from Cases: Opportunities and Challenges." *Academy of Management Journal* 50(1): 25–32.
- Espeland, Wendy N., and Michael Sauder. 2007. "Rankings and Reactivity: How Public Measures Recreate Social Worlds." *American Journal of Sociology* 113(1): 1–40.
- Forseth, Ulla. 2001. "Boundless Work – Emotional Labour and Emotional Exhaustion in Interactive Service Work." Unpublished PhD thesis. Trondheim: Norwegian University of Science and Technology (NTNU).
- Forseth, Ulla. 2005. "Gender Matters? Exploring How Gender is Negotiated in Service Encounters." *Gender, Work & Organization* 12(5): 440–445.

- Forseth, Ulla. 2010. "Ambivalenser i frontlinjearbeid." *Tidsskrift for Arbejdsliv* 12(3): 9–24.
- Forseth, Ulla, and Carla Dahl-Jørgensen. 2002. "Sur la ligne du feu: transformation du travail relationnel dans une centre commercial et une banque en Norvège." *Travailler, Revue Internationale de Psychopathologie et de Psychodynamique du Travail* 9: 73–97.
- Forseth, Ulla, and Hans Torvatn. 2015. "Det norske verneombudet i et globalt arbeidsliv". In *Den norske modellen – Internasjonalisering som utfordring og vitalisering*, edited by Brita Bungum, Ulla Forseth, and Elin Kvande, 71–87. Bergen: Fagbokforlaget.
- Forseth, Ulla, Emil A. Røyrvik, and Stewart R. Clegg. 2015. "Brave New World? The Global Financial Crisis' Impact on Scandinavian Banking's Sales Rhetoric and Practices." *Scandinavian Journal of Management* 31(4): 471–479.
- Huzzard, Tony, Dennis Gregory, and Regan Scott (eds.). 2004. *Strategic Unionism and Partnership: Boxing and Dancing?* Basingstoke: Palgrave Macmillan.
- Jeacle, Ingrid, and Chris Carter. 2012. "Fashioning the Popular Masses: Accounting as Mediator between Creativity and Control." *Accounting, Auditing & Accountability Journal* 25(4): 719–751.
- Juris, Jeffrey, and Marina Sitrin. 2016. "Globalization, and Social Transformation." In *The Sage Handbook of Resistance*, edited by David Courpasson and Steven Vallas, 31–50. Los Angeles, CA and London: Sage.
- Kjellberg, Hans, and Alexandre Mallard et al. 2013. "Valuation Studies? Our Collective Two Cents". *Valuation Studies* 1(1): 11–30.
- Lainer-Vos, Dan. 2013. "Boundary Objects, Ones of Indeterminacy, and the Foundation of Irish and Jewish Transnational Socio-Financial Networks." *Organization Studies* 34(4): 515–532.
- Larsen, Tord, and Emil A. Røyrvik (eds.). 2017. *Trangen til å telle. Objektivisering, måling og standardisering som samfunnspraksis*. Oslo: Scandinavian Academic Press.
- Law, Johan, and John Urry. 2004. "Enacting the Social." *Economy and Society* 33(3): 390–410.
- Leonardi, Paul L. 2011. "When Flexible Routines Meet Flexible Technologies: Affordance, Constraint, and the Imbrication of Human and Material Agencies." *Management Information Systems Quarterly* 35(1): 147–167.
- Levin, Morten, Tore Nilssen, Johan E. Ravn, and Lisbeth Øyum (eds.). 2012. *Demokrati i arbeidslivet. Den norske samarbeidsmodellen som konkurransefortrinn*. Bergen: Fagbokforlaget.
- Lindberg, Kajsa, and Barbara Czarniawska. 2006. "Knotting the Action Net, or Organizing between Organizations." *Scandinavian Journal of Management* 22(4): 292–306.
- Lukes, Steven. 2005. *Power: A Radical View*. London: Palgrave.
- MacKenzie, Scott B. 2003. "The Dangers of Poor Construct Conceptualization." *Journal of Marketing Science* 31(3): 323–326.

- Mason, Katy, Hans Kjellberg and Johan Hagberg. 2015. "Exploring the Performativity of Marketing: Theories, Practices and Devices." *Journal of Marketing Management* 31(1-2): 1-15.
- Merry, Sally E., Kevin E. Davis, and Benedict Kingsbury. 2015. *The Quiet Power of Indicators: Measuring Governance, Corruption, and the Rule of Law*. New York, NY: Cambridge University Press.
- Miele, Francesco. 2014. "Organizations as Boundary Objects: Organizing Business between Research and the Market." In *Information Infrastructure(s): Boundaries, Ecologies, Multiplicity*, edited by Alessandro Mongili and Giuseppina Pellegrino, 238–257. Newcastle-upon-Tyne: Cambridge Scholars Publishing.
- Normann, Richard. 1984 [2001]. *Service Management. Strategy and Leadership in Service Businesses*. New York, NY: Wiley.
- Osborne, Thomas and Nikolas Rose (1999) "Do the Social Sciences Create Phenomena?: the Example of Public Opinion Research." *The British Journal of Sociology*, 50(3): 367-396.
- Poon, Martha (2007). "Scorecards as Devices for Consumer Credit: the Case of Fair, Isaac and Company Incorporated." *The Sociological Review*, 55(2): 284-306.
- Poovey, Mary. 1998. *A History of the Modern Fact: Problems of Knowledge in the Sciences*. Chicago, IL: University of Chicago Press.
- Porter, Theodore. 1995. *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton, NJ: Princeton University Press.
- Power, Michael. 1997. "Expertise and the Construction of Relevance: Accountants and Environmental Audits." *Accounting, Organizations and Society* 22(2): 123–146.
- Regini, Marino, Jim Kitay, and Martin Baethge (eds.). 1999. *From Tellers to Sellers. Changing Employment Relations in Banking*. Cambridge, MA: The MIT Press.
- Roberts, John. 1991. "The Possibilities of Accountability." *Accounting, Organizations and Society* 16(4): 355–368.
- Rosness, Ragnar, and Ulla Forseth. 2014. "Boxing and Dancing – Tripartite Collaboration as an Integral Part of a Regulatory Regime." In *Risk Governance of Offshore Oil and Gas Operations*, edited by Preben H. Lindøe, Michael Baram, and Ortwin Renn, 309–339. Cambridge: Cambridge University Press.
- Røyrvik, Emil A. 2018. "The Measure of Sociality: Quantification, Control and Economic Deviance." In *Economy, Crime and Wrong in a Neoliberal Era*, edited by James G. Carrier, 66–92. New York: Berghahn Books.
- Scott, James. 1990. *Domination and the Arts of Resistance*. Hidden Transcripts. New Haven, CT: Yale University Press.
- Sewell, Graham, and James R. Barker. 2006. "Coercion Versus Care: Using Irony to Make Sense of Organizational Surveillance." *Academy of Management Review* 31(4): 934–961.

- Sewell, Graham, James R. Barker, and Daniel Nyberg. 2011. "Working under Intensive Surveillance: When Does 'Measuring Everything That Moves' Become Intolerable?" *Human Relations* 65(2): 189–215.
- Shore, Cris, and Susan Wright. 2015a. "Audit Culture Revisited: Rankings, Ratings, and the Reassembling of Society." *Current Anthropology* 56(3): 421–444.
- Shore, Cris, and Susan Wright. 2015b. "Governing by Numbers: Audit Culture, Rankings and the New World Order." *Social Anthropology* (23)1: 22–28.
- Star, Susan Leigh. 2010. "This is not a Boundary Object: Reflections on the Origin of a Concept." *Science, Technology & Human Values* 35(5): 601–617.
- Star, Susan Leigh, and James R. Griesemer. 1989. "Institutional Ecology, 'Translations' and Boundary Objects; Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39." *Social Studies of Science* 19(3): 387–420.
- Styhre, Alexander. 2013. "The Economic Valuation and Commensuration of Cultural Resources: Financing and Monitoring the Swedish Culture Sector." *Valuation Studies* 1(1): 51–81.

Ulla Forseth is Professor at the Department of Sociology and Political Science at the Norwegian University of Science and Technology (NTNU) with extensive experience from research in private and public enterprises. She has published on transformation of work and organization in finance, emotions and emotional labour, gender relations, power and resistance, risk, regulation and safety. After her position as visiting professor at the University of Trento, Italy in 2017–18, she is currently working on comparisons between the Norwegian model of working life and welfare and the "Italian model" and affective ethnography. Her publications include "Boxing and Dancing: Tripartite Collaboration as an Integral Part of a Regulation Regime" (with Ragnar Rosness, 2014), and she is co-editor of *Den norske modellen: Internasjonalisering som utfordring og vitalisering* (with Brita Bungum and Elin Kvande, 2015).

Stewart Clegg is Distinguished Professor of Management and Organization Studies at the University of Technology, Sydney. He has published widely in management, organizations and politics literatures in many of the leading journals. He is a Visiting Professor at EM-Lyon, France and at Nova School of Business and Economics in Lisbon, Portugal. Widely acknowledged as one of the most significant contemporary theorists of power relations he is also one of the most influential contributors to organization studies, recognized by his being chosen both

as an EGOS (European Group for Organizational Studies) Honorary Member and as a Fellow of the Academy of Management.

Emil André Røyrvik is a social anthropologist and Professor at the Department of Sociology and Political Science at the Norwegian University of Science and Technology (NTNU). His research focuses on ethnography and theorizing in the context of financialization, organizational and managerial cultures, and economic and cultural globalization. Some of his recent publications include “The Measure of Sociality: Quantification, Control and Economic Deviance” (in *Economy, Crime and Wrong in a Neoliberal Era*, edited by J. Carrier, 2018), and the co-edited collection *Trangen til å telle. Objektivisering, måling og standardisering som samfunnspraksis* (with Tord Larsen, 2017). His monograph *The Allure of Capitalism: An Ethnography of Management and the Global Economy in Crisis* (published with Berghahn Books) won SINTEF’s (independent, Norwegian research organization that conducts contract research and development) award for excellence in research 2011.