Safeguarding the unknown: performance measurement, academic agency and the meaning of research quality in practice

David S. Bedford
Accounting Discipline Group, University of Technology Sydney, Sydney, Australia

Markus Granlund
Department of Accounting and Finance, University of Turku, Turku, Finland, and

Kari Lukka
Department of Accounting and Finance, University of Turku, Turku, Finland and Department of Accounting and Finance, Turku School of Economics, Turku, Finland

Abstract

Purpose – The authors examine how performance measurement systems (PMSs) and academic agency influence the meaning of research quality in practice. The worries are that the notion of research quality is becoming too simplistically and narrowly determined by research quality's measurable proxies and that academics, especially manager-academics, do not sufficiently realise this risk. Whilst prior literature has covered the effects of performance measurement in the university sector broadly and how PMSs are mobilised locally, there is only little understanding of whether and how PMSs affect the meaning of research quality in practice.

Design/methodology/approach – The study is designed as a comparative case study of two university faculties in Finland. The role of conceptual analysis plays a notable role in the study, too.

Findings – The authors find that manager-academics of the two examined faculties have rather similar conceptual understandings of research quality. However, there were differences in the degree of slippage between the "espoused-meaning" of research quality and "meaning-in-practice" of research quality. The authors traced these differences to how the local PMS and manager-academics' agency relate to one another within the context of increasing global and national performance pressures. The authors developed a tentative framework for the various "styles of agency". This suggests how the relationship between the local PMS and manager-academics' exerted agency shapes the "degrees of freedom" of the meaning of research quality in practice.

Originality/value – Given that research quality lies at the heart of academic work, the authors' paper indicates that exploring the three matters – performance measurement, the agency of manager-academics and the meaning of research quality in practice – in combination is crucial for the sustainability of the academe. The authors contribute to the literature by detailing the way in which local PMS and manager-academics' agency have material impacts on what research quality means in practice. The authors conclude by highlighting the

pressing need for manager-academics to exercise the agency in efforts to safeguard a broad and pluralistic understanding of research quality in practice.

1. Introduction

"Quality" is a term frequently heralded and celebrated in the academe. It is viewed as a positive aspect of academic work overall, not least with regard to scholarly research, where seeking high quality is a taken-for-granted phenomenon, even a constituting feature of it. It is therefore surprising how little attention the definitional difficulty that surrounds the notion of quality has received. As Pirsig (1974) demonstrates in *Zen and the Art of Motorcycle Maintenance*, defining quality is far from unproblematic. Through an insightful philosophical exploration into the meaning of quality, Pirsig concludes that "true quality" is inherently undefinable. Rather than trying to resolve the inherent challenge of defining the notion, Pirsig's analysis provides a helpful "way of thinking" about quality. It forms a starting point for our study examining how the *meaning of research quality in practice* is understood in the current circumstances of the academe.

A significant amount of academic inquiry has revealed the influential role of performance measurement systems (PMSs) in how universities are managed, especially as a mechanism to determine and evaluate academic output (e.g. Modell, 2003; Sousa *et al.*, 2010; ter Bogt and Scapens, 2012; Kallio and Kallio, 2014; Kallio *et al.*, 2016, 2017). Whilst it is often suspected that increasing the quantity of output reduces the overall quality of research produced, and especially its originality and creativity (e.g. Guthrie and Parker, 2014; Alvesson *et al.*, 2017), this is notoriously difficult to demonstrate, for two main reasons. First, what becomes directly and "easily" visible are only the quantity of research outputs, citations and the rankings and impact factors of their outlets. This has led to increasing reliance on, and trust in, the stories that these figures tell (Guthrie and Parker, 2014; O'Connell *et al.*, 2020). Second, even if we wished, the issue embedded in the first point is difficult to counter-argue, given that the notion of research quality arguably escapes any attempts to pre-define and measure it (Pirsig, 1974; Harvey and Green, 1993; Wittek and Kvernbekk, 2011). These two observations lay the motivational ground for our study and link to our key worry: The difficult-to-define notion of research quality has increasingly become determined by its measurable proxies (e.g. Willmott, 2011; Parker and Guthrie, 2013; Humphrey and Gendron, 2015; Martin-Sardesai and Guthrie, 2018; Picard *et al.*, 2019).

Scholarly research is an example of what Karpik (2010) refers to as "singularities" – unique goods or services whose value cannot be determined by observable technical properties (see also Picard *et al.*, 2019). Whilst assessments of singularities are always characterised by uncertainty, various kinds of "judgment devices" can provide some guidance as to their value. However, as Karpik (2010) points out, devices that attempt to standardise the basis of comparison will always pose a risk of failing to capture the essential characteristics that make each singularity unique. Consequently, judgement devices, such as journal impact factors and rankings, pose a risk of not capturing the virtue of quality in the way that would reflect the underlying pursuit of "good scholarship" in its full meaning (Hopwood, 2007; Czarniawska, 2011; Alvesson *et al.*, 2017). A fundamental question is whether academics sufficiently realise the risk of the "tail starting to wag the dog" and how they deal with this risk. Letting the (incomplete) performance measures (the proxies) become the central target for activities rather than the relevant aspects of scholarly research itself, threatens the practice of good scholarship (Alvesson *et al.*, 2017), necessary for its long-term sustainability (Johnston and Riemer, 2014).

The dysfunctional effects of the "performance measurement regime" on the academe (Becker and Lukka, in press) have not gone unnoticed in accounting and management

literature (for an overview, see Welpe *et al.*, 2015). This research typically depicts PMSs, rooted in "New Public Management" (NPM) (Hood, 1995), as hostile to academic freedom and detrimental to the overall creativity and quality of research. Even though resistance is often mentioned, academics are typically viewed as helpless and innocent victims (Thomas and Davies, 2005; Manes-Rossi *et al.*, in press). This antagonistic view is arguably too structuralist and pessimistic, which has been noted by a small, yet burgeoning area of research paying serious attention to the agency of academics themselves (Barry *et al.*, 2001; Anderson, 2008; Sousa *et al.*, 2010; Clarke and Knights, 2015; Englund and Gerdin, 2020; Soin and Huber, 2021; Manes-Rossi *et al.*, in press). Our exploration not only continues along these lines but importantly links it with the fundamental issue of how the meaning of research quality in practice is in danger in these circumstances.

Our objective is to examine the *meaning of research quality in practice* in the contemporary "performance measurement era" in the academe and to analyse the role of the *agency of academics* in this regard. We specifically focus on "manager-academics" – heads of departments as well as top managers of the faculty and the university [1]. Manager-academics must navigate expectations from (at least) two directions simultaneously – the administrative and the scholarly. This can easily lead to conflicts and tensions in manager-academics' work, not least regarding the trade-off between quantity and quality in research [2]. We explore how they respond to these tensions arising from global and local performance pressures, and what kinds of resistance, if any, they employ. Our objective drives the following research question: *How do the local PMS and manager-academics' agency influence the meaning of research quality in practice?*

To examine this question empirically, we conducted a comparative case study between two largely kindred faculties (labelled Case A and Case B) from two universities in Finland, a European country with a reputation of high educational quality (Schatz, 2015). We employed not only the interview method and analysis of documentary materials, but also reliance on introspection (Alvesson, 2003). Our case findings show that manager-academics of the two faculties have similar conceptual understandings of research quality. However, there were notable differences regarding the degree of slippage between the "espoused-meaning" of research quality and its 'meaning-in-practice". We trace these differences to how variations in the local PMS and manager-academics' agency relate to one another. Based on our analysis, we develop a tentative framework that theorises varying *styles of agency* amongst the manager-academics (i.e. conformers, passive mediators, active mediators, resistors and generators), that tend to arise dependent on the relationship between the local PMS, in terms of how narrow or broad it is in defining academic performance and the manager-academic's perceived opportunity for, and willingness to exert, their agency.

We connect these styles of agency to the "degrees of freedom" available for how research quality is understood in practice. We argue that research quality in practice (in the Pirsigian spirit) needs sufficient "degrees of freedom" for the ideal of "good scholarship" to be maintained. It is leaving sufficiently open-ended the boundaries of what quality *can* mean which is of critical importance. Yet our study suggests that the "degrees of freedom" in practice are easily becoming constrained. We find that most manager-academics were reluctant to exercise their agency, or often did not perceive the meaning of research quality to be a concern that they should actively address. Instead, the determination of research quality is often left to be "automatically" taken care of, either through the self-governance of academics under global performance pressures and surrounded by seemingly objective judgement devices (like journal quality classifications), or more typically through narrowly designed local PMSs.

Our study shows that the meaning of research quality in practice is shaped by the PMS, with a broad and inclusive local PMS leading to higher "degrees of freedom" than a narrow and exclusive one, but in conjunction with how manager-academics perceive and engage their

agency. Consequently, the emancipatory argument of the paper is that we need to take seriously the fact that the nature of the local PMS matters for research quality and the need for manager-academics to be conscious of their agentic potential. Manager-academics have the capacity to contribute to "safeguarding the unknown" – the many-sided, unpredictable and incommensurable nature of research quality in practice – by exercising their agency actively, especially with regard to enacting or resisting changes to local PMSs. Although most manager-academics in our study were relatively passive, a few of them, especially from one of the two faculties, indicated a significant inclination to employ determined and long-term agentic efforts. We see such efforts as essential for safeguarding a broad and pluralistic meaning of research quality in practice.

The remainder of this study is structured as follows. In the following section, we review the existing knowledge on performance measurement and its changes in the academe, the agency of individual academics in this context and the meaning of quality in research. In section three, we detail our method and in section four we present the empirical findings. In section five, we discuss the implications of our results, before providing a conclusion to the study in section six.

2. Literature review

2.1 Performance measurement in academia

The general theoretical domain of our study is performance management in knowledge-intensive public organisations, such as universities. Central to performance management in universities is not only the increasingly pervasive employment of PMSs to monitor and evaluate performance, but also a close relationship to the publish-or-perish culture, which nowadays is a pervasive part of academic life at universities globally (Alvesson *et al.*, 2017; Becker and Lukka, in press). Whilst universities have long used some measurements to follow and control their performance, the rapid intensification of their use was kicked off by the NPM movement (Hood, 1995). The use of PMSs has spread rapidly, reaching a point which justifies the use of the expression "performance measurement era" in universities (Grossi *et al.*, 2020).

As Enders (2004) notes, performance measurement tends to transform qualities into measurable quantities, making it easier to access and process information. Devices used to measure research quality have come to be seen as trustworthy technologies that facilitate transparency and enable more objective performance evaluation of universities, faculties, as well as individual academics, at least seemingly facilitating commensurability (Karpik, 2010; Bialecki *et al.*, 2017). Indicators of research output are often connected to globally applied journal lists and rankings, or ones determined by national evaluation panels. They have arguably introduced perverse incentives into the academe (Edwards and Roy, 2017). For example, incentives rewarding the number of publications may lead to massive amounts of substandard, incremental papers, reduced quality of peer review and, ultimately, "bad science" (Alvesson *et al.*, 2017).

Regarding quality and judgement devices, Karpik (2010) introduces a novel way of understanding the crux of the concern [3]. His concept of "singularities" refers to goods and services with three main features: (1) they are multidimensional in that they are made up of various overlapping dimensions; (2) they are incommensurable, because there is no common and obvious measure that enables an indisputable evaluation of their quality; and (3) they are characterised by uncertainty, as evaluating quality of singularities is complex and often cannot occur until the good is bought or the service is provided. Karpik's analysis of singularities is relevant for us since only with slight adaptation we can transform these thoughts to a piece of research as the "good" in focus.

Karpik (2010) suggests two analytical distinctions ("types of involvement") for quality assessment: active vs passive and autonomous vs heteronomous judgement. The former

distinction deals with how much of one's own time and effort is devoted to the assessment, whilst the latter is about the degree of independence of assessment. In the context of evaluating research quality, the combination of passive and heteronomous refers, broadly speaking, to "outsourcing' the judgemental work to some outside party (e.g. forming assessments based just on journal rankings or impact factors). In contrast, active and autonomous assessment of quality means that the evaluator focusses directly on the substantive contents of the piece, and whilst rankings might be used as a broad reference, essentially an effort of independent overall judgement is made. Karpik analyses the various kinds of techniques and technologies – which he calls "judgment devices" – that offer short-cuts for assessors. Journal and university rankings are prominent examples of such devices.

Performance measurement within universities has been widely studied (Grossi *et al.*, 2020). Most of the existing research focusses on the consequences of regulatory changes and the encroaching managerialism in universities and has been conducted at the institutional or organisational levels of analysis (e.g. Parker and Jary, 1995; Modell, 2003; Guthrie and Neumann, 2007). This literature establishes the potential negative effects of increasing managerialism and performance measurement on academic life and the general opposition to them. However, it provides little insight into how the micro-practices of individual academics, especially manager-academics, are implicated. Survey studies in the area often generalise observations on managerialist practices, such as PMSs, in the university sector without considering that there are likely many differences in their design and use at the local level. In that regard, the agency of academics may play a significant role. The next section reviews the literature on the interaction between PMS and the agency of academics.

2.2 Agency of academics
The question of agency is closely linked to the philosophical question of whether humans are subject to deterministic mechanisms that govern their behaviour or have the capacity to exert free will. Whilst this issue remains unresolved, it is typical in social science to lean towards the latter view, treating humans at least *as if* they have discretion in their actions (Lukka, 2014). Structuration theory, for example, describes how knowledgeable agents are often able to conceptualise and reflect upon the practices they entertain (Giddens, 1984). Accordingly, agents can, for instance, trigger or hinder the actualisation of presumed causal mechanisms to implement their agenda. However, they typically cannot, at least in the short term, change those causal mechanisms themselves (Durand and Vaara, 2009). That said, since knowledgeable agents often have the possibility of "acting otherwise", their intentions and actions are potentially changeable, leading also to the possibility of new causal mechanisms emerging.

In this study, we are interested in the agency of manager-academics in shaping their organisation's members' thoughts, understandings and actions (both their own and those of others) regarding research quality. Recently, Alvesson *et al.* (2017) remarked that academics seem to be largely unaware of, or not using to its full potential, the agency that they have at their disposal: "Many academics downplay the freedom they have in constructing their identity" (Alvesson *et al.*, 2017, p. 100) and "most researchers have more power than they realize" (Alvesson *et al.*, 2017, p. 139) as it is precisely academics themselves who make evaluative judgements about research and researchers as well as taking part in academic management, funding bodies and advisory boards. Our empirical study will shed light on these types of matters, especially how manager-academics either use and reproduce or in various ways challenge, those structures (i.e. PMSs) that lead to research quality being understood in particular ways.

There are only a few studies that shed light on how PMSs at the local level are intertwined with academic practice and the agency of academics. Sousa *et al.* (2010) studied how researchers working as managers dealt with the pressures imposed by new research

performance management systems, such as PMSs. They argue that the agency of academics is typically underestimated, observing that manager-academics adapt their negotiation and leadership styles towards performance norms. Englund and Gerdin (2020) reported on a research group that exhibited high levels of non-compliance regarding what was institutionally deemed important and legitimate in performance evaluation. They identified four qualities of the research group that inhibited the encroachment of the "academic performer ideal." These relate to the pre-existing ideal, in that it is 1) central and long-established, 2) orthogonal to the academic performer ideal as materialised by the PMS, 3) largely shared within the research group and 4) externally legitimate. These qualities tended to form an important basis for criticising and contesting the "performer' ideal created by the prevailing PMS.

There are also studies that look at the PMSs in the academe by drawing on the distinction between complying and resisting forms of agency. Barry et al. (2001) explored how far managerialism had spread in the UK in a comparative study of two universities and concluded that whilst managerialism did not (yet) feature very deeply, features of resistance were still already found. In interview-based studies, Anderson (2008) reported many subtle ways of resistance, not just "rebellious and uprising" ones (p. 254). Clarke and Knights (2015) revealed how managerialism led academics to chase "the illusive sense of a secure self through 'careering'; a frantic and frenetic individualistic strategy designed to moderate the pressures of excessive managerial competitive demands" (p. 1). However, such compliance was not the only reaction; forms of resistance were found, too. Soin and Huber (2021) suggest compliance is not just non-agency but a social practice that goes beyond mere consent and find three variations of compliance, whilst Manes-Rossi et al. (in press) argue that resistance need not be only antagonistic, observing seeds of "generative resistance" in an Italian university.

Despite this awakening interest in the varying forms of academic agency, the question remains as to how the impact of performance measurement and the various potential agentic possibilities of academics (especially manager-academics) play out regarding how the meaning of research quality is shaped and understood in practice. We explore this uncharted territory by employing Karpik's (2010) ideas on the two analytical dimensions for quality assessment and the foundational idea of Pirsig (1974) according to which quality is hard to even define, not to mention measure, in practice.

2.3 On the notion of quality and its meaning in research practice

Quality is often discussed as if everyone understands what it means. This is, however, a far from unproblematic assumption, as the very concept of quality is challenging to even define. In one major inspiration for our study, Robert Pirsig's novel *Zen and the Art of Motorcycle Maintenance* (1974), the key character Phaedrus takes seriously the issue of defining quality. He distinguishes between "romantic" and "classic" ideas of quality. The former posits that quality is an immediately present and directly perceived excellence, yet difficult to explicate and measure. The latter views quality as being based on rational analysis, whereby many relevant aspects can be separately evaluated, often through quantitative measurement. After a meticulous analysis and taxing introspection, Phaedrus concludes that quality is a combination of the two perspectives, whilst still maintaining that a true, "absolute quality" cannot be clearly defined. For Pirsig, quality is a transcendental notion that pre-exists any definable and observable categories.

A few higher education researchers have linked their analysis closely to Pirsig's (1974) thoughts, as well as distinguishing between the notion of quality and its determination in practice. Shields (1999) draws on Pirsig's idea about connecting the romantic and classic ideas of quality, yet observes how easily performance measurement in practice tends to squeeze out

the romantic aspect. Harvey and Green (1993) also argue that quality should be viewed as a relative concept, but remind us that being content with the philosophical position that quality cannot be defined is not feasible in practice where judgements have to be made. On the other hand, it is also perilous to take a purely pragmatic "this-just-has-to-be-done' approach, essentially pointing to a problematic aspect of quality indicators in use: Rather than being viewed as only proxies of the underlying fundament, they tend to become *the* fundaments themselves.

Wittek and Kvernbekk (2011) are especially helpful for clarifying the challenges relating to the notion of quality. They argue that quality cannot be clearly defined, because the notion is ambiguous (it has many meanings) and vague (it is difficult to decipher at its boundaries). Situational stipulations help in the former regard, but regarding the latter, we just have to accept the uncertainty around the notion. Wittek and Kvernbekk borrow from Wittgenstein (1953) the idea of "family resemblance" suggesting that there is still something that the various stipulative expressions of quality refer to, even if having potentially multiple meanings and being vague at their very boundaries. The spirit of Wittek and Kvernbekk (2011) is brought to the fore in the following quote:

> [. . .] even in the absence of an agreed-upon, unified definition of quality, we all (think we) recognize quality when we see it. [. . .] We can tell the difference between good and poor student papers when we see them, even if we cannot pinpoint exactly the basis of our judgment. Art experts agree that one painting is better than another, even if they can point to no objective criteria. This is interesting, given the lack of a clear definition of quality. We still (think we) know what it is. (p. 675)

Hence, the real issue is not so much recognising quality as we face it – despite the often tacit nature of these understandings – but how we can unproblematically explicate and communicate our evaluations. Whilst quality cannot be clearly defined, there still is some common thread keeping the various formulations of the idea together. For our analysis in the context of scholarly research, we will take the Pirsigian not-strictly-definable excellence as the cornerstone of what research quality is supposed to be all about. From this notional standpoint, we will explore the ways in which research quality gets explicated, communicated and used in the everyday practice of academics, embedded in the current performance measurement regime and conditioned by their agentic possibilities.

3. Methods, research materials and analysis

Our research is empirically designed as a comparative case study. The target organisations were two faculties, labelled as Case A and Case B, situated within separate universities in Finland [4]. These two faculties were selected, firstly, because they represent the same major disciplinary area with quite similar focus and scope and, secondly, due to the expectation, based on anecdotal expectations, that there would be differences in how manager-academics are acting at them. For instance, the different historical trajectories of the two faculties and personnel selection might be indicative of variations in the way manager-academics act regarding the questions of interest for us.

The Finnish academic environment, in general, is not an exception to what we can find elsewhere in the Western world (Mathies *et al.*, 2020; Kivistö *et al.*, 2017). New public management initiatives have resulted in a performance-based state funding model, which has further diffused into the internal funding models of universities. Career structures in Finland are also similar to those elsewhere, being increasingly tenure track based. So, Finland serves here only as an example: Academic activities, PMSs, as well as managerial practices in Finland, are generally what can be observed elsewhere. Additionally, the issues we examine are not dependent on disciplinary fields. The same results-based state funding devices, mechanisms of gaining external funding and career structures apply to all disciplinary fields.

Therefore, the fact that we cannot reveal the disciplinary field for securing anonymity, is not an issue considering our research question.

We interviewed eight manager-academics in similar positions from both organisations. Those interviewed included top managers of the universities, top managers of each faculty and many heads of department [5]. We employed triangulation of research methods. Accordingly, in addition to conducting interviews, we also collected formal documents from publicly available sources. Given the research topic, introspection was one significant research method, too (see Alvesson, 2003). The fact that one of the authors is currently a manager-academic and another has significant prior experience of being a manager-academic greatly facilitated this. Nevertheless, the most important empirical method of the study was interviews, and especially through them, our analysis was able to go beyond the formal procedures towards the routines in the informal domain.

We conducted the interviews using a carefully prepared semi-structured research guide in line with the idea of theme-based interviews (see, e.g. Qu and Dumay, 2011). This approach allowed room also for free association by the interviewees as well as for asking further questions on topics that happened to surface [6]. At the start of the interview, we informed the interviewees only about our general research interest, to avoid any potential leading effects. The interview lengths ranged from 70 to 132 min, the average being 96 min 14 of the 16 interviews were conducted face-to-face and two through Skype (further information on the interviews is contained in Table A1). The interviews were conducted in most cases by two of the researchers of the team, a few times by all three and a few times by just one of us. All interviews were recorded and transcribed verbatim.

Over the course of our research, we kept a detailed, chronological research diary. This included brief notes after each interview to capture immediate impressions and observations. Then, once the interviews were transcribed, we prepared memos on each of them, which focussed on the perceived highlights of each interview. In a close dialogue of going abductively back and forth between our original theoretical ideas, concepts, data and interpretation (Dubois and Gadde, 2002; Ahrens and Chapman, 2006; Lukka and Modell, 2010), we gradually developed a broad understanding of the major structural drivers that shape the meaning of research quality in practice. Guided by this understanding, our analysis continued by focussing on the interplay between the local PMSs and manager-academics' agency. The theoretical storyline of our paper matured in the research diary, which greatly facilitated the write-up stage. The overall spirit of our study is interpretive – capturing the meanings of the people interviewed plays a notable role in this study. Despite the research process being generally abductive, we chose to write the paper up in a rather deductively reconstructed manner to improve its readability.

4. Findings
In this section, we present our central empirical findings regarding the PMSs, manager-academic agency and the meaning of research quality in practice – and their interplay – from our comparative case study.

4.1 Performance measurement systems
Both faculties closely monitor academic performance, with significant emphasis placed on research outputs. Table 1 summarises the main features of the PMS in each case.

The PMS in Case A was implemented less than 5 years before the interviews, prior to which there was no systematic measurement of academic performance. The PMS was implemented, along with wider structural reforms, with the aim to achieve "world-class" performance. The PMS focussed almost singularly on publications in highly ranked journals, especially on publications in one exclusive global journal ranking system (which we refer to with GJRS1).

	Case A	Case B
Scope	Primarily publications, GJRS1 publications emphasised. Other aspects of academic performance are discussed informally (e.g. in salary negotiations)	"Research points" from publications, professional academic tasks, citations, editorships, journal reviews and external funding. Publication numbers vis-à-vis the JUFO[1] classification are followed separately as well
Publication categories used	Primarily GJRS1[2], but also GJRS2[3]; to some extent JUFO	JUFO
Connection to basic (state) funding	Largely indirect, but strengthened by the fact that the GJRSs correlate with the JUFO[4]	Strong link because the state funding model is also based on the JUFO
Career paths/ development	Tenure track is the foundation of the HRM system, with GJRS1 publications used as the primary criterion	Multiple paths, wider consideration of the publication and other record
Other rewarding	Monetary reward for publications in the GJRSs (for GJRS1 eight times more and for the highest level in GJRS2 four times more, than the reward received in Case B)	Monetary reward for publication in the highest level of the JUFO (before certain budget cuts, used to be also for the second highest rank publications)
Time orientation	Figures calculated annually, in relation to budget funding 3-year averages applied	Figures calculated annually, in relation to budget funding 3-year averages applied

Note(s): [1] On the Finnish JUFO classification, see https://julkaisufoorumi.fi/en/publication-forum (retrieved 30 May 2022)
[2] We call GJRS1 one particular global journal ranking system. To get listed in GJRS1, a journal needs to reach high values in bibliometric analyses, but there is also a voting mechanism where selected university faculties have a vote. The journals included are geographically highly concentrated as comes to their editorial teams and publishers
[3] We call GJRS2 another globally applied journal ranking system, alongside GJRS1
[4] The correlation regarding GJRS1 is though only partial, as it covers only a subset of the journals listed in the JUFO and GJRS2
Source(s): Own creation/work

> It's relatively straightforward. It's the number of publications, especially the [GJRS1] [. . .] lower quality journal articles are not counted. So, basically one key KPI. (M-A 4A)

The focus on highly ranked journals is the basis for performance evaluation, incentive and promotion decisions within the faculty. This is especially the case with the tenure-track system, where GJRS1 plays an important, even decisive, role.

> This tenure track model, it is tough in that if you don't get the publications [in GJRS1 journals], it's up or out, and it's bye bye. [. . .] I may think that they are doing interesting stuff but there's no way they are getting tenure. (M-A 8A)

In Case B, a comprehensive PMS had been in place for two decades. It was crafted to signal desired expectations regarding research and other academic outcomes, which were based on consultations between top management and senior academics. The basic mechanism of the PMS is the allocation of "research points" across a wide range of academic activities, such as publications, journal editorships, journal reviews, external funding and citations.

> The main idea is that it tries to capture all kinds of activities that a good researcher could do, very liberally, very inclusively, going from publications to normal things considering nowadays the JUFO, your review and evaluation tasks, your editor board memberships, your citations . . . I think everything you can imagine. (M-A 4B)

The local PMS of Case B aims to signal that it is not only publications at the highest level of the ranking that are valued, but also those at lower levels. In relation to publications, the PMS is tied to the national journal ranking system, the JUFO, a classification system intended for the evaluation of research across all scholarly disciplines. A top-ranked JUFO publication ("JUFO three") is awarded one third more points than what is awarded to publications in the second highest rank ("JUFO two") and four times more than what is awarded to the third highest rank ("JUFO one").

Overall, the most important differences between the PMSs of the two faculties, relevant to our exploration, is that the PMS of Case A is substantially narrower in focus as compared to the broader and multidimensional PMS at Case B.

4.2 Manager-academic agency

The general perception in both cases was that although their position as a manager-academic enabled them to exert at least *some* degree of influence upon the academic agenda of their area of responsibility, the potential to influence practices locally as well as in the university more broadly was diminishing. However, the extent to which agency has diminished, and importantly, the willingness to actively intervene in university and departmental processes varied significantly between the two cases. Findings for each case are reported in the following subsections.

4.2.1 Case A. The introduction of the tenure track and PMS were part of recent wider structural changes at Case A, which encompassed a significant centralisation of administrative and strategic decision-making. One of the top manager-academics of the faculty noted that, given the current environment in which universities operate, there is some degree of inevitability about the changes:

> This process, I don't pretend to be very important here. I'm just a vehicle for this organization, the board, the top management. If I wouldn't be here, somebody else would be. (M-A 2A)

Even though senior academics are nominally part of the strategic decision-making process, they feel they have had little capacity, at the individual level, to influence decision outcomes, including the design of the PMS. Much of the decision-making is highly centralised, with the purpose of participative dialogue, in the words of the top manager-academic, to "seduce" senior academics to buy into the changes:

> Just because I'm a [top manager], I can't make decisions that the professors don't like [. . .] I need to persuade them, I need to seduce them to, kind of, my side. Otherwise, you can't accomplish these kinds of changes in a university. (M-A 1A)

Department heads were generally of the opinion that their capacity to influence strategic decisions is minimal and even intentionally bounded through increased centralisation — participation is largely viewed as merely an illusion. One manager-academic was particularly scathing in assessing the administrative changes at the university:

> They call it leaders' dialogue, but we call it leaders' monologue. They go around, they show a couple of slides, allow a couple of questions, and then go home and continue with the same set of slides. Feels like you are talking to a wall. (M-A 8A)

It is perhaps not surprising, then, that most interviewees at Case A noted that they took their managerial positions only reluctantly.

> I was not applying for this job. I was dragged into this job. Not dragged, but I mean, on the third effort I agreed to an interview and okay, I took the job. (M-A 2A)

> I'm sort of hoping that at the end of this year [another professor] will come back to the department and will take this so I can get rid of this duty. (M-A 3A)

Manager-academics saw their capacity to directly affect the quality and quantity of research output as being quite limited. One stated that "there's nothing in my power that I can do to squeeze GJRS1 publications [from staff] except for saying 'don't submit to [discipline specific] journals" [outside of the GJRS1] which I'm not going to do" (M-A 7A). As such, efforts were focussed on influencing the "necessary micro-cosmos" (M-A 5A) of the department, to provide the conditions conducive to academic success and the necessary resources, such as administrative support and funding, to allow researchers to preserve time for research. Practices such as formal performance evaluations, promoting research seminar attendance, developing intra- and inter-departmental research efforts, mentoring tenure track staff and engaging with doctoral students, were also mentioned. Yet several interviewees questioned whether their positions were even necessary for achieving desired performance outcomes.

> I think most of us are self-driven in the sense that they don't need much of a leadership. (M-A 3A)

Another noted that there is little need for any formal management processes if academics "have the right values" (M-A 7A) – the "right" values being those consistent with pursuing research amenable to publication in top-ranked journals, which the PMS serves to reinforce. Whilst a few of the manager-academics indicated awareness of the potential downsides of a strict adherence to how performance is measured and evaluated through the PMS, there was little indication of an appetite for change – manager-academics either did not perceive these downsides as being particularly problematic or did not feel like they had the agentic capacity to address them. One example relates to the potential temptation to use short-cuts in conducting research under notable time pressure.

> I was sort of in the opinion that if we are striving for quality, those [GJRS1] publications don't come in two years, they normally take more time [. . .] I can see an example from my department right now when [the academic's] tenure came up, they rushed to send some stuff to a journal. Which I wouldn't have advised them to do. But then again, what can I do? (M-A 4A)

Manager-academics are subject to the same performance evaluation regime as their subordinates. Most perceived the need to make trade-offs between what they should or could be doing in their managerial roles, and what they actually did, to preserve time for their own research-related activities. Only one manager-academic at Case A explicitly noted any real need to engage their agency to intervene in the potential tension between the increasing pressure to publish in top-ranked journals and research quality:

> It has become more and more sort of this KPI-driven type of thing, so people need to publish, publish, publish, no matter whether the articles actually make sense or provide any contribution in the sense that they actually deal with interesting issues [. . .] I try to sort of fight against that, but it seems a pretty hard battle. (M-A 4A)

The sense of agentic potential in Case A is largely limited to the upper echelon of managers. Through a shift to a centralised structure and the institutionalisation of narrow systems of evaluation, the "room to manoeuvre" for lower-level manager-academics is restricted. Yet, interviewees generally accepted the aim to become "world-class" as an important thing to pursue and understood that GJRS1 publications are imperative to reputation-building efforts. The spectre of "global pressures" is perceived as a largely immutable force.

4.2.2 Case B. When asked about the reasons why they had taken upon managerial duties, only one interviewee at Case B indicated that they had done so reluctantly – upon joining the university, this manager-academic was promised to be able to do "professor's work only, but that lasted about two months" (M-A 6B). Others, some of whom had occupied the department head position for long periods of time, responded that they hold the position to actively shape the development of their unit or department as well as the university more broadly. One

interviewee mentioned that it was in their nature to want to influence the department's trajectory.

> I have a feeling of responsibility, a kind of too big feeling, that's a personal characteristic, maybe. That's why I have been involved so much in the development work over the years [. . .] I don't want to give the academic leadership to just [anyone], because I see it as a long-term building up of an academic culture here [. . .] It's this atmosphere of striving for high quality in everything. (M-A 3B)

Another noted that they were "too conscious of the need to take care of some bigger issues" as they felt the need for "working for some greater good, to make things kind of work" within the changing environment in which the university now operates. But this is at great personal strain as they "seldom work the kind of normal office hours" (M-A 8B).

The general perception of manager-academics was that whilst they could exert some degree of influence upon the academic agenda of their area of responsibility, the potential to influence administrative practices locally as well as in the university more broadly was diminishing.

> In terms of what we see as important, developing the school and the research here, agency for that is broad and I could probably do much more and be free to do that. But then in the everyday management, there the room for agency has diminished somewhat. (M-A 3B)

Other responsibilities noted by department heads were to "buffer" staff from administrative tasks to preserve time for academic work and to actively fight against the increasing pressures placed on academics to publish in the short term. One department head saw a strong tension between the pressures to publish and the "compromising [of] research quality", especially amongst young academics:

> This tendency to be rushing and instrumentalist, playing the publish or perish game. So that's the major tension, that's something I have to fight every day and it's close to every day, it's really all the time. (M-A 4B)

Others shared similar sentiments:

> One of our researchers that has been in sort of an insecure position for some time, I've been all the time trying to mentor him to get rid of this idea that 'I have to publish, I have to publish alone, I have to publish fast,' but you should publish together, slow and good quality. (M-A 6B)

Other efforts of manager-academics at Case B were focussed on informal practices geared towards developing a culture that prioritises research quality – practices such as creating a culture around research seminar participation and informal discussion groups, encouraging and facilitating collaboration within and across departments, providing feedback and guidance to colleagues and mentoring junior staff members, were frequently mentioned. One department head felt that leading by example as well as developing decisively the research culture is of vital importance, as comes to the fore in the following quotes:

> It is that I always talk about this: let's remember to think, let's invest a lot in the early process, [. . .] let's be active, work together. And I always try to show this model myself and talk about that and act like this in the seminars and discussion group meetings. (M-A 4B)

> I like [the local PMS], I think it's good. But it's only a starting point. I think that the rest is academic leadership and it has to be there, you have to start organizing these things, you have to make sure that people understand the seminar is the big day. You have to be there, having read the paper, ready to say something. (M-A 4B)

Most manager-academics regarded research quality as a core concern in taking on the role of an academic leader. One interviewee, when asked what they wanted to achieve as an academic leader, responded:

[Long silence] That's a tricky question [. . .] I think I'm doing the greatest impact through my doctoral candidates. It's kind of trying to make them see things, to do innovative research, to see science as a broad and open-minded exercise. I hate narrow-mindedness, I hate only single-theory, only single-method, only single rules; this kind of dogmatism is horror for me. (M-A 3B).

4.3 Research quality

Interviewees in both cases demonstrated largely consistent *conceptual* understandings of research quality, aligning closely with Karpik's (2010) notion of singularities. First, research quality was understood as being multidimensional, with manager-academics tending to demarcate quality between two dimensions that roughly correspond to Pirsig's (1974) classic and romantic ways of thinking about quality. One dimension was that research must conform to basic research standards and principles ("quality as conformance"). A department head from Case A referred to this as "basic quality" which comprised "must-haves":

> I think one aspect of quality are those must-haves. I mean [. . .] craftsmanship in terms of rigor. So conceptual rigor, methodological rigor, that concepts are properly defined or that the research question is properly defined, the concepts, the data collection or the empirical setting is properly described, and data is reported. I get angry when I have the feeling that this basic quality is not achieved, because this is what we should be able to produce as good researchers. (M-A 5A)

Other manager-academics from Case A emphasised that research should follow the appropriate "scientific method" (M-A 3A) and that there is a strong methodological foundation so that "you can trust the results" (M-A 4A). In Case B, one department head referred to quality as conformance as the "narrower take on quality" (M-A 4B) that can be considered a checklist of items that should be ticked off for the research to meet an acceptable standard. Others referred to this kind of quality as "engineering quality" (M-A 5B) or "hygiene factors" such as "good dataset, good analysis, good writing, good flow" that should be present in any piece of research (M-A 2B).

The second dimension of research quality, "quality as excellence", was perceived as important, but more difficult to establish. As one manager-academic noted succinctly – "all that matters in the end is the contribution" (M-A 2B). Contribution can, however, take many forms. Most manager-academics commented that high-quality research must contain something that is "really new", such as a novel theory that opens a new stream of research, counterintuitive or surprising findings, exploiting a unique dataset or context, developing a new concept, problematising the status-quo, or in general, making you think differently on a topic. Manager-academics also commented that high-quality research was rarely of the type that generates just incremental knowledge within an established field. Some noted that such research had its merits (it is "good research", but not of high quality), whilst others expressed concern about the increasing production of research that has high quality in terms of technical execution but only marginal contribution to the knowledge of a discipline.

> It's not just kind of testing a hypothesis in a bigger model or it's not just adding bits and pieces to existing knowledge but it's somehow questioning are we doing the right things. (M-A 3B)

The second aspect consistent with singularities is that research is incommensurable regarding quality. Nearly all interviewees alluded to the difficulty in specifying the boundary concerning whether a piece of research is of high quality or not. One responded that there is "no way of defining it [but] you know it when you see it" (M-A 7A). Several others indicated that high-quality research was not something definable but more a "feeling" (M-A 4B) that they had – the research is exciting, inspirational and evokes "wow-feelings" from the reader.

[A]n exceptional piece is something that you still are . . . you almost can't get sleep in the evening, there is something so striking, so new, so inspiring you just know it when you see it. So maybe the degree of novelty or the originality of the contribution, but sometimes it's quite difficult to draw the line. (M-A 2B)

This alludes to quality, at least in terms of the second dimension (i.e. "quality as excellence"), as being inherently based on tacit understandings. Nevertheless, certain "conformance" criteria are relatively objective and definable in nature; a checklist of standards that must be fulfilled to reach a basic level of quality. These are often well established within a field, enabling consensus between evaluators concerning the presence or absence of conformance criteria that make it relatively easier to be achieved. Yet these criteria can vary by discipline, resulting in incommensurability for even this more basic form of quality when attempting to make interdisciplinary judgements.

The third defining feature of singularities is that evaluation is complex and often marked by uncertainty. A few manager-academics commented that there can be a significant evaluation lag in forming a judgement about the quality of a piece of research, and furthermore, the perception of the quality of a piece of research can change over time, for instance, depending on the evaluating party. Several interviewees also pointed to the notion of quality as a socially constructed understanding within a community and that you could only understand the meaning of quality if you are embedded within that community. This points to the difficulty for an outsider to assess quality, with little tacit understanding of what constitutes quality within a specific community.

> There's an interesting book by Michèle Lamont [. . .] about how difficult it is really to pinpoint the particular issues that you have to have present in making the distinction between good and excellent; it relies on your own expertise within the field. (M-A 8B)

Whilst conceptual understandings were closely aligned at both organisations, this was not the case when it came to the meaning of research quality *in practice*. We consider this issue for each case organisation in the following sections.

4.3.1 Case A. In Case A, there were substantive difference between the conceptual understanding of research quality and its meaning in practice. In a telling example, one manager-academic held the following view of journal rankings, *in principle*:

> I think this is good for communicating some kind of intent that we really take seriously the idea that you should publish in top outlets. It should never be used as a substitute for actually sitting down and reading the papers and making a judgment on the quality of the people that you are evaluating. (M-A 7A)

Yet when asked about how they evaluate *in practice* the performance of staff in their unit:

> So [sigh] . . . I have to rely on this performance in terms of the journals in which you get it published as a signal of that because I can't judge every field of [the discipline] myself as being interesting or not so interesting. (M-A 7A)

The PMS of Case A equates quality very directly with publications, and especially with what are classified as top journals, to the exclusion of other measures:

> [I]f you think about our reward system, for example, we don't reward on citations. We reward on publications (M-A 3A).

In describing the research culture of the faculty, one top manager emphasised how "quality research" has become ingrained into the mindset of academics at the university.

> I think that because our choice to try to get world class, the emphasis is all the time on quality [. . .] people understand that we have to produce quality, we have to aim higher, we have to improve. I think it is built into this university now. (M-A 3A)

However, the idea to have GJRS1 as the major benchmark for tenure-track academics, and further promotion decisions, was largely imposed by faculty top management. Department heads acknowledged that there are many discipline-specific journals outside of the GJRS1, where high-quality research can be, and is, published, but admitted that the reliance on judgement devices like GJRS1 does serve the purpose in signalling the strategic intentions of the university. Nevertheless, the importance of publications in top-ranked journals to the aims of the entire university has trickled into the practices of department heads, even for those who expressed criticism to how academic performance is defined.

> When we have these department meetings, I'm showing who's been publishing what. I'm trying to put a little bit of peer pressure as well so that everybody is actually knows where they stand in terms of their research within the faculty. So again, I've done that so far, but I'll try to do that mainly to push a little bit more pressure on some individuals. (M-A 4A)

The position that publishing in "top-tier" journals is what ultimately matters has become the dominant ethos, being central to tenure, promotion and compensation decisions. Most manager-academics at Case A indicated that there is a degree of convergence between the measurement of research outcomes through their PMS with the quality of research. However, two interviewees expressed serious reservations about the calculability of academic performance through the kind of PMS they have, especially research quality. One lamented the increasing emphasis on quantitative measures as the basis for performance evaluation as academic work is "also about the nontangible things and you cannot measure everything" (M-A 5A). They understood, though, how it facilitates the bureaucratic management of the university:

> I think it makes the lives of the decision makers easier by referring to KPIs [. . .] But the truth is not as simple, it cannot be expressed in KPIs. It's more complex than that. (M-A 5A)

From previous experience at universities abroad, this department head considered that placing such substantial emphasis on a single measure of performance, the GJRS1, can have detrimental consequences on the type of research that is pursued. Instead, the department head prefers to let "the market" (the academic community of their discipline) determine the performance expectations and incentives for researchers.

> The market defines the targets [. . .] if you want to have a salary increase or a job, look at the market requirements (M-A 5A).

Another manager-academic at Case A (M-A 8A) gave a quite critical, even cynical account of the "unnecessary bureaucracy" in how the research output of the department is evaluated.

> [The faculty head] is pretending to evaluate me and I'm pretending to be evaluated. [The faculty head] is looking at the numbers of the department [. . .] one number went up or down, and [they] said that's good or that's bad, and I say actually that's not good or bad, that is just random chance that anybody with a basic understanding of statistics will understand, and they said OK [. . .] it's just for show. It's like airport security check; they have security theatre, and we have performance theatre. (M-A 8A)

This interviewee continued by explaining that the administrative performance expectations were "destructive" to the culture they had fostered, which privileges academic freedom and conducting quality research that has an impact. What truly matters to academics is how they are perceived by others in their field. Like M-A 5A, this interviewee considered the academic community to be the primary evaluator of an academic's contribution:

> Our University pushes much more towards quantity [. . .] This is about one of these examples of double-speak, that the university and the [faculty] say that, of course we don't really think this, but you are supposed to get the numbers because that's where the money is from [. . .] To advance in your academic career you are much more dependent on what people in your field think of you abroad than

you are of what some administrators think here in [the country], that's very rarely relevant. So, the culture of your academic field always is the stronger one — that's what rewards the quality in the end. (M-A 8A)

Interestingly, whilst this manager-academic strongly supported the ideal of academic freedom, they argued that the top journals in the discipline field were becoming increasingly narrow in theoretical and methodological scope. When asked whether this narrowing in the field created a conflict for academics to pursue their own agenda, they responded:

Well, yes and no. No in the sense that we are absolutely free to do whatever research we please [. . .] But of course there are incentives, right. People tend to follow the incentives, at least in our field. So, if you get rewarded, maybe it's Pavlovian, you get rewarded for publications, you go for the publications. Maybe it's not even the calculated decision, it just makes you feel good, makes you salivate. (M-A 8A)

4.3.2 Case B. In contrast to Case A, the coupling between the PMS and the evaluation of academic performance was relatively loose. One department head, when asked about whether they use journal rankings to evaluate performance and research quality, replied:

No, quality comes from the people and their orientation and understanding of the academic quality and what we try to achieve here. It doesn't come from the managerial control systems. These are just checking that we are performing okay, fine in terms of numbers. (M-A 6B)

This department head commented further:

They all know that we should publish according to the JUFO classification, but I do actually just the opposite, that I tell them that "yes it's important but it's not that important; it's more important you bring about a credible research track as an individual and then you are able to take the others with you and work together and build a research community of your own". I'm not sure if they are buying into this, but that's what I'm trying to tell them. (M-A 6B)

One manager-academic saw the PMS at Case B as an integral mechanism to signal that academic performance beyond journal publications is valued. It was noted that the initial PMS was not designed primarily for formal performance evaluations, but as a tool for self-reflection. Nevertheless, journal rankings have become a more important component of the system:

I think [the PMS] has been instrumental, or somehow a tool, a signalling device, for paying attention to both quality and quantity. And, of course, over the years there might have sometimes been some misunderstandings that it's more for quantity [. . .] But now the recent discussions and the recent revisions to our system really emphasize more and more [research quality]. For instance, you get relatively more points now for JUFO three and JUFO two publications which is a clear signal in my view. (M-A 2B)

Despite the intended purpose of the PMS as a multidimensional signalling device, some manager-academics perceived it as being akin to a "results-table" without any real systematic effect on academic behaviours. However, most interviewees regarded the PMS as a valuable tool for annual development discussions. The following excerpt is indicative of perceptions more generally at Case B, noting both the positive intent of the PMS design and the fact that in everyday practice its use may be limited:

Honestly, it doesn't play that big role at our department [. . .] sometimes we discuss them but we don't, I don't use these points systematically [. . .] When you collect this information, it has an impact on people. They know that these types of things are measured, so they know that it counts. (M-A 7B)

Interviewees at Case B repeatedly mentioned that there is, and should be, an active judgemental element of one's own in determining the quality of research. Interviewees commented that "quality is not the JUFO classification" (M-A 7B), that quality was not

determined "by the journal where it is published" (M-A 2B) and that "quality comes first, and not the number of publications" (M-A 8B). Another lamented the increasing reliance by the wider research community on journal rankings for determining the quality of research:

> As a community, we have simplified it a lot. If it's published in a good journal, it must be good. But that's not necessarily a good thing. Because not everything that is in good journals is very great, and not every great piece of science is published in the best journals. (M-A 5B)

There was also a general recognition amongst manager-academics of Case B that too strict reliance on rankings and objective measures of research performance can be detrimental to the research culture within the faculty as well as to individual academics.

> Rankings are bad when the rankings trickle down to the valuation of people. They become values as themselves [. . .] If they become kind of a marker for a value of somebody who's in academia, that can be very deteriorating. (M-A 8B)

This faculty head commented that with the PMS in place, department heads have a significant amount of autonomy in terms of how they improve academic performance. But some of them also suggested that the pressures from the field are today such that there may not be any longer a need for the current research points system in place, at least in terms of incentivising academics to pursue "top-tier" journal publications, given the pervading external pressures to publish. The view of the top manager-academic at Case B was that journal rankings served a useful purpose in gauging research quality. Metrics such as journal rankings inevitably are, at least to some extent, relied upon because many manager-academics are "acting at a distance' and therefore may have less discipline-specific knowledge. They noted, though, that "objective measures" such as journal rankings should be seen as just a proxy, and not equivalent to, the notion of quality. Several manager-academics of Case B commented that the understanding that quality was separate from journal rankings was engrained within faculty and departmental cultures.

4.4 Summary of findings
Our empirical findings are briefly summarised in Table 2. We will theorise on these observed differences regarding the meanings of research quality in practice and the reasons for their variation more thoroughly in the ensuing section.

5. Discussion
5.1 Research quality: conceptual notion vs practical meaning
In this study we examine the meaning of research quality in practice from the perspective of manager-academics, using comparative case study materials from two university faculties. Our findings reveal a relatively consistent interpretation of the *conceptual* meaning of quality by manager-academics at the two case sites. The conceptual understanding of quality closely aligns with Karpik's (2010) notion of singularities: Research quality is multidimensional, roughly encapsulating Pirsig's (1974) classic and romantic aspects of quality and since especially the latter aspect is largely based on tacit understandings and assessments of research quality are typically marked by incommensurability and uncertainty. This is in line with the notion that any true attempt to objectify quality in research will inevitably capture "something less than Quality itself" (Pirsig, 1974, p. 236).

Despite little variation within and between the two faculties regarding the conceptual meaning of research quality, there were substantive differences in the *meaning of research quality in practice*. These differences reflect the degree of slippage between the conceptual or "espoused-meaning" of quality and its "meaning-in-practice" (see Argyris and Schon, 1974). Our study focusses on demonstrating how differences in such slippage can be traced to, first,

	Case A	Case B
Performance measurement		
Local PMS design	Narrow. Primarily focussed on a single performance measure – the number of publications, especially in GJRS1 listed journals	Broad. A points-based system that captures a wide variety of academic activities (e.g. publications, professional academic tasks, citations, editorships, journal reviews and grants)
Local PMS use	Tightly connected to the evaluation of academic performance. Basis of tenure track, performance evaluation, incentive compensation and promotion decisions	Loosely connected to the evaluation of academic performance. Primarily a tool for self-reflection and as a basis for discussions concerning academic development
Research quality		
Conceptual meaning	Conceptual understanding in both case organisations is consistent with Karpik's (2010) notion of singularities. Research quality is seen to be multidimensional and incommensurable and its evaluation complex and marked by uncertainty. The conceptual dimensions of research quality are representative of Pirsig's (1974) classic (quality as conformance) and romantic (quality as excellence) perspectives on quality	
Meaning in practice	Meaning strongly influenced by measurable proxies. Judgement devices (i.e. local PMS and global journal rankings) used to infer research quality and are constitutive of the "academic performer ideal" (Englund and Gerdin, 2020)	Meaning partly influenced by measurable proxies. Judgement devices (i.e. local PMS and global journal rankings) used to inform assessments of research quality rather than to determine them. Separation of quantifiable performance outcomes and what it means to be a good academic
Agency		
Managerial role participation	Reluctant. Little desire to assume managerial responsibilities, especially as department heads. Short tenures, typically designated on a rotation basis between senior academics	Committed. Often strong desire to take position. Some hold the position for long periods of time
Perception of agentic potential (research quality and academic performance)	Passive. Typically, "hands off" approach to managing academics. Main concerns about buffering academics from administrative burdens and preserving time for research. Academic performance expectations and understanding of research quality driven by the PMS or external performance expectations	Active and passive. A few manager-academics active and conscious about the need to buffer external performance pressures. Most of them concerned about shaping how academics understood research quality and what it means to be a good academic, even if their actual use of agency would have been limited
Perception of agentic potential (administrative and institutional practices)	Limited. Little inclination to seek structural or administrative change. Institutional changes viewed as inevitable consequences of changing global imperatives. Strategic direction centralised at the university level, no real capacity, or visible initiative to try, to shape institutional decisions	Towards more limited. Desire to influence wider changes in the university, but actual capacity to do so has diminished. Yet some indicated significant motivation and efforts to shape the long-term trajectory of their departments

Source(s): Own creation/work

Table 2.
Summary of empirical findings in the two case organisations

the differing ways in which judgement devices (Karpik, 2010) of performance are mobilised, both those that originate within and outside the university, and second, the degree to which, and how, individual agency is exercised by manager-academics. We consider these two factors in turn below, before outlining a framework that builds abductively on our empirical findings.

5.2 Performance measurement, judgement devices and research quality
The way the PMS was designed and mobilised in making judgements was notably different between the two faculties. In Case B, the PMS was deliberately designed to be broad and inclusive, stressing the many sides of academic work and that not only high-ranked publications matter. The PMS of Case B was perceived to be primarily a mechanism for self-reflection and a basis for development discussions, with looser connections to tenure-track and promotion decisions. In Case A, the PMS was designed to be narrow and exclusive and perceived as inevitable in the pursuit of being regarded as a "world-class" institution. The PMS was tightly linked to more significant financial incentives, the ability to secure tenure and the likelihood of promotion. This made it clear that publishing in GJRS1 journals is not only expected but a requirement for long-term employment at Case A (van Dalen and Henkens, 2012). These reinforcing mechanisms created far fewer "degrees of freedom" in Case A than in Case B for autonomous interpretations in making judgements about research quality and what it means to be a "good scholar" (Hopwood, 2007; Czarniawska, 2011; Alvesson *et al.*, 2017).

It must also be noted that there were several similarities between the two institutions. Despite the different approaches between the PMSs and their uses, publications in high-ranked journals were an important marker of research quality with academics generally seeking and many achieving, publication in the most prestigious research outlets. Manager-academics at both institutions, albeit at times reluctantly, also took "shortcuts" in making judgemental evaluations by relying on measures captured by the local PMS. Yet there was still a greater tendency for manager-academics at Case A to defer to the PMS – together with its "outsourcing agents' like GJRS1 – as the primary judgement device. Judgements about the quality of research and academic performance more generally, were also broader at Case B than at Case A and in the former faculty, there was also more insistence on the need to use autonomous scholarly judgement. Acknowledgement of the need for active judgement, rather than just leaning on journal rankings, surfaced also at Case A, but this seemed to be more a matter of principle than that of actual practice. Ultimately, the narrow focus of the PMS and its more rigid application in Case A appeared to increase the divergence between the espoused meaning of quality and its translation in practice.

The critical question of course is whether this narrowing of the practical meaning of quality is problematic. After all, academics generally aspire to publish in the best journals, and a narrow and transparent PMS would appear to be an effective mechanism for focussing efforts more sharply towards this. Furthermore, as Karpik (2010) argues, commensuration enabled by judgement devices is not a threat to incommensurability if it is "pluralist and reversible" (p. 12) and "safeguards the diversity of personal interpretations" (p. 18). However, in practice, securing incommensurability is, in the current performance measurement regime, under severe threat due to performance pressures sourcing from all levels (global, national and local), the combined effect of which peaks at Case A. As several of our interviews indicated, even academics who seek to employ active and autonomous modes of evaluation are likely to find it difficult to entirely escape from pressures towards commensuration. As the romantic element of quality is inherently subjective and difficult to determine, it is more likely to be marginalised in attempts to capture quality through measurable proxies. In the

longer term especially, such marginalisation poses a risk regarding the sustainability of good scholarship and the uniqueness and diversity of research.

5.3 Manager-academic agency and research quality

Prior literature suggests that, at least in principle, manager-academics have the capacity to "safeguard the unknown" – the many-sided, unpredictable and incommensurable nature of research quality in practice – by exercising their agency actively. As Sousa *et al.* (2010) argue, academics' power to manage potentially harmful situations caused by managerialism is often underestimated. Yet we observed, somewhat surprisingly, that many manager-academics do not seem particularly conscious about their agentic opportunities or the need to exercise them (see Alvesson *et al.*, 2017). Manager-academics generally perceived their capacity to influence research quality as limited and regarded their agentic capacity as diminished – pressures from global competition and changes in national funding arrangements often induced a feeling of inevitability regarding changes to university management and the evaluation of academics and their research.

Nevertheless, it was evident that manager-academics at Case B were more willing and perceived a greater potential, to exercise agency, whilst those at Case A were more apt to lend their agency away, trusting either the self-governing capacities of individual academics or the academic field to define and make judgements on the quality of research. Even though manager-academics at Case A perceived the local PMS as overly narrow, most either did not see the need or were reluctant, to stage any acts of more active resistance (see Anderson, 2008; Kalfa *et al.*, 2018) in the form of, for instance, attempting to change the local PMS design. They also felt little incentive to take the role of a manager-academic, especially as a department head and were willing to dedicate only so much time to their managerial duties. After all, they are subject to the same pressures as those without managerial responsibilities. Hence it is little surprise that manager-academics at Case A placed greater reliance upon various kinds of judgement devices to make evaluations about research quality. This is consistent with the argument of Espeland and Stevens (1998) that as commensuration is cemented within the practices and processes of organisations, "it becomes more taken for granted and more constitutive of what it measures" (p. 329). Whilst certain indications of resistance, tending to be subtle and covert rather than open, were observed, the overall mode of agency at Case A was compliance with the local PMS, in the form of very active and competitive, even frantic "careering" (Clarke and Knights, 2015).

Manager-academics at Case B demonstrated more of a belief that quality can be actively influenced. Many manager-academics at Case B saw the main threat to research quality arising from the global and national context of academia, rather than the local PMS. They perceived that the PMS of Case B could in fact be mobilised to support an alignment between the espoused meaning of quality and its meaning in practice, as it encapsulated multiple performance measures to reflect the quality and looser ties to formal performance and promotion decisions. This provided greater "degrees of freedom" for manager-academics to exert at least some amount of agency in defending the incommensurability of research quality.

Even though we observed only a few manager-academics at Case B making intensive and determined long-term efforts to alleviate the increasing performance pressures from beyond the local context that were affecting research quality in a detrimental manner, this observation still marks a notable difference from what we observed at Case A. This indicates that there are significant agentic possibilities to safeguard the many-sided, unpredictable and incommensurable nature of research quality in practice – if manager-academics are inclined to seize these opportunities in a decisive, long-term manner. These findings also suggest the potential for variation in academics' dealings with performance pressures – not only going

from compliance to active resistance, but also including strategies where performance pressures are used as an activating discussion partner in pursuing generative and innovative forms for safeguarding the romantic aspects of research quality in practice (Soin and Huber, 2021; Manes-Rossi *et al.*, in press). These resisting and generative agentic approaches were found notably more in Case B than in Case A. It seems fair to argue that this difference arises from the differences in the design and use of the local PMS, given that the global and national performance pressures are identical in Cases A and B. It is worth noting that several of the interviewees from Case B more or less actively participated in the design of the local PMS of Case B, from its initial form and development over the preceding two decades.

5.4 A tentative framework

Building abductively on our empirical findings, we present a tentative framework in Table 3 that captures how variation in PMS design (narrow and exclusive vs broad and inclusive) and whether manager-academics are active or passive in buffering internal and external pressures, are related to *styles of agency* and the meaning of research quality in practice. The framework encapsulates our core argument that there need to be sufficient "degrees of freedom" for the meaning of research quality in practice to be consistent with the commonly

| | | Active or passive manager-academic agency | |
		Active	Passive
Local performance measurement system design	Broad and inclusive	*Resistors and generators* The broad and inclusive local PMS design is actively used in resisting the external pressures for commensuration by pluralistic internal evaluations. Active manager-academic intervention may even reach over and above resistance to innovating generative possibilities Sufficient degrees of freedom available to allow for a broad, pluralistic understanding of research quality in practice	*Passive mediators* The potential of the broad and inclusive local PMS design is only passively used through somewhat inclusive internal evaluations, but with only limited effectiveness given the role allowed for the increasing forces of external pressures. Restricted degrees of freedom, increasing the likelihood that the meaning of research quality in practice is more narrowly defined by external judgement devices
	Narrow and exclusive	*Active mediators* Despite the narrow and exclusive local PMS, the multiple pressures towards commensuration are actively attempted to get mediated by manager-academic intervention, but with only limited effectiveness given homogenous internal and external pressures. Restricted degrees of freedom, increasing the likelihood that the meaning of research quality in practice is narrowly defined by internal and external judgement devices	*Conformers* Manager-academics just comply with the narrow and exclusive local PMS, which does not mitigate the multiple external pressures towards commensuration. Hence, external pressures are not locally buffered Very restricted degrees of freedom, resulting in a high likelihood that the meaning of research quality in practice is defined by internal judgement devices, reinforced by the external ones

Source(s): Own creation/work

shared understanding of its conceptual meaning, hence allowing for its full open-ended plurality. We argue that both the PMS design and the role of manager-academics are important factors in determining the degrees of freedom available.

The willingness and ability of manager-academics to exert their agency towards safeguarding the meaning of research quality in practice were more observable in Case B than in Case A. This is at least partly a function of the broader PMS design, which is less coupled to tenure and promotion decisions, making it easier for manager-academics to resist leaning too heavily on the PMS as a judgement device. Yet, even at Case B, we observed only a few manager-academics that made substantive, persistent and long-term efforts to mitigate the global pressures towards commensuration. Referred to as *resistors*, these manager-academics enable the conditions for more pluralistic understandings of research quality to surface. Importantly, enacting this pattern of agency leaves more room for the Pirsigian romantic notion of quality, resulting in a closer alignment between the conceptual understanding of research quality and its meaning in practice. At least in one case encountered in Case B, a manager-academic developed an innovative strategy in dealing with the performance pressures, which reached over and above mere resistance. This manager-academic could be viewed as a *generator* by their style of agency, as they intervene actively to open up alternative ways for understanding the meaning of research quality in practice. Although manager-academics at Case B were generally concerned about influencing academic behaviours and what research quality means in practice, most would fall closer to the *passive mediators* type. Whilst internal pressures towards commensuration are lessened by the presence of an open and inclusive PMS, there is an increased risk of global pressures encroaching on the meaning of research quality in practice.

As for Case A, most manager-academics are consistent with the style of agency that we term *conformers*. Manager-academics feel little need to exert their agency towards influencing the meaning of research quality in practice and the passive and heteronomous outsourcing of research quality assessments to judgement devices dominates. Whilst the narrowly focussed PMS and the associated increase in managerialism are not widely popular, there is little active resistance. As such, the meaning of research quality in practice is extensively defined by the local PMS and further reinforced by global pressures. However, two manager-academics at Case A did express strong reservations regarding the calculability of research quality through the local PMS and actively intervened to reduce its influence as the primary evaluative device. As such, these manager-academics are closer to what we call *active mediators*. However, evaluations externalised to "the market" effectively result in the same pressures towards commensuration, since the judgement devices at the global level are essentially the same as those embedded in the local PMS (i.e. widely adopted rankings like GJRS1).

The core insight from this framework is that academics need to be situated in a context where they are actively buffered from external pressures and where local measurement practices are broad and inclusive with regard to academic performance, for the practical meaning of research quality to have a chance of aligning with commonly shared conceptual understandings. We fear, however, that this situation is far from the norm for most institutions. The more narrowly the PMS captures academic performance, and the more tightly connected it is to decisions regarding evaluation, promotion and tenure, the greater the likelihood that the meaning of research quality in practice succumbs to pressures towards commensuration. Likewise, a context in which manager-academics are relatively passive regarding pressures on academics to conform to narrowly defined ideals of performance the more likely for slippage between the practical and conceptual meanings of research quality. In instances where both internal and external pressures are reinforcing, we see a very high likelihood that the meaning of research quality in practice becomes synonymous with how it is captured in judgement devices. The framework, therefore, illustrates the power of both

academic agency and the construction of the PMS in providing sufficient degrees of freedom for a pluralistic understanding of research quality.

It is important to emphasise that global and national pressures do not automatically lead to changes in the PMS, nor the ability or willingness to exert agency. In Case B there was sufficient resistance to preserve the underlying role of the PMS as a tool for relatively independent self-reflection, whilst in Case A, the introduction of the PMS was considered an inevitable response by top manager-academics to perceived global and national pressures and the explicit desire to be a "world-class" institution. Furthermore, as Anderson (2008) and Englund and Gerdin (2020) reveal, even within contexts where the PMS is highly constraining, academics can effectively contest the "performer ideal' being imposed. Our major concern is, however, that the conditions specified by Englund and Gerdin (2020) that decrease the probability of yielding to the pressures of a narrowly focussed PMS (alternative ideals that are: long-established, strongly shared by group members and externally legitimate) are far more likely to emerge *prior* to such PMSs being implemented. Attempting to "carve out spaces for thinking and acting otherwise" (p. 932) is inevitably easier when the PMS does not yet act to rigidly reinforce a restrictive understanding of the meaning of research quality.

6. Conclusions

We argue that research quality in practice (in the Pirsigian spirit) needs "degrees of freedom" for the academic work to have a chance to carry out its full potential. This means that leaving the boundaries of what quality *can* mean sufficiently open-ended is of critical importance. We contend that the meaning of research quality in practice is conditioned by the local PMS, yet in interaction with the agency of manager-academics. From our empirical observations, we derive a tentative framework to explain the conditions which enable or constrain the meaning of research quality in practice. We suggest that both the PMS and active intervention by manager-academics are critical for engendering a pluralistic understanding of research quality in practice.

Our study has two main implications. First, manager-academic agency is important, but rather infrequently mobilised, for safeguarding the romantic meaning of quality in research. Certainly, superfluous leaning on the romantic notion of quality can lead to an overemphasis on subjectivity in evaluations. However, without including this aspect in the practical meaning of quality in the academe, there is a high risk of excessive commensuration, likely in the long run leading to less heterogeneity in research topics and approaches (Lukka and Mouritsen, 2002; Chapman, 2012). In such a case, the "biodiversity" necessary for the long-term sustainability of the academe (Johnston and Riemer, 2014), would become endangered. Consequently, quality assessment should not be viewed as only a technical operation that takes place straightforwardly through "handy" judgement devices.

Second, the manager-academics' agency requires a sufficiently broad local PMS to offer the necessary room for manoeuvre. Hence, we argue that the way in which the local PMS of academic units are designed and used really matters. If it simply reinforces the pressures of the global publish-or-perish culture, an instrumental understanding of quality becomes embedded or 'built in" to the institution. This is shown very explicitly in Case A, where a straightforward commensuration of research quality has become *de facto* the default. Even in Case B, where the PMS is notably broader and less coupled to career decisions, a commensuration tendency regarding manager-academics' practices related to research quality appears, too. This suggests that under the current performance measurement regime in the academe at large, we are, worryingly, more likely to see movement from the situation of Case B towards Case A than the other way around.

Whilst a broad practical understanding of the notion of research quality – fundamentally at least partly an "unknown" thing – is under threat, there is still hope. It is important to realise that it

is not only an exogenous systemic danger but rather notably an endogenous one as well: It is up to the academics themselves. Academics should emancipate themselves from positioning too easily as mere passive "victims" in the iron cage of the various performance pressures imposed on us to realise our own potential for intervention. Indeed, the needed resisting or generative agency is grounded on manager-academics more clearly realising the agentic potential they have. As our study suggests, in the short term, manager-academics' agency concerning the meaning of research quality in practice is framed by the current design and the routinised usage patterns of the local PMS, albeit with at least some potential leeway. But in the longer term, there is the opportunity for manager-academics to actively shape the local PMS towards being less narrow and rigid, thereby allowing for different kinds of realisations of research quality to emerge.

On a larger scale, initiatives such as the San Francisco Declaration on Research Assessment (DORA) are encouraging. The academe should also keep in mind that some of the bigger leaps in the progress of our scholarly knowledge are often embedded in radically new takes on research. In a nutshell, quoting Elster (2000, p. 267): "When conventions accumulate to reduce the scope of artistic choice, a revolution may be required to liberate creativity. But even before that point is reached, rebellions against convention can serve as a vehicle for creativity." Very similar ideas – once the word "artistic" is replaced by "scholarly" – are worth paying serious attention to in supporting sustainable academic scholarship, which must be based, we argue, on a notable openness as for how quality in research is understood and practised. This will require active resistance as well as openness to finding new generative forms of intervention from all academics, especially from manager-academics.

Notes

1. In Finland, the source of our empirics, most of the heads of department and deans of faculties are full professors. Rectors/Vice-Chancellors/Presidents of universities tend to have backgrounds as full professors too, whilst according to Finnish legislation, they at least need to have earned a doctoral degree.

2. This study does not seek to empirically measure the quality of research, or their difference. Indeed, using Pirsig's (1974) conceptualisation implies that "true quality" is inherently not measurable and precisely from that point starts our central storyline. Hence, we focus on the *meaning* of research quality in practice, not on the measurement of quality.

3. Karpik (2010, p. 10) contemplated between using "quality" or "singularities" as his anchor term. The reason for choosing "singularities" ("the unique") is relevant for our analysis: "For a long time, I used the term *quality* (or *qualities*) and its derivatives *quality goods* and *economics of quality* or *qualities* [footnote]. But the word's affinity with a unidimensional reality, its increasingly frequent use, the growing diversity of its meanings and the misunderstandings it prompted led me to replace it with the notion of *singular products* (goods and services) or, more simply, *singularities* [. . .]."

4. During the time of the empirical work, it remained typical that top university and faculty managers are academics having assumed for a certain, fixed period of time a managerial role. Hence, the term that we use, manager-academic, should be viewed as illustrative.

5. To ensure anonymity, our write-up of this study is especially careful concerning what we disclose about our interviewees, including the case organization they represent. We use the general term "manager-academic" for all interviewees, and the abbreviation "M-A" (e.g. M-A 1A, referring to manager-academic coded as number 1 representing Case A) when referencing interview quotes.

6. The interview guide is available upon request.

References

Ahrens, T. and Chapman, C.S. (2006), "Doing qualitative field research in management accounting: positioning data to contribute to theory", *Accounting, Organizations and Society*, Vol. 31 No. 8, pp. 819-841.

Alvesson, M. (2003), "Methodology for close up studies – struggling with closeness and closure", *Higher Education*, Vol. 46 No. 2, pp. 167-193.

Alvesson, M., Gabriel, Y. and Paulsen, R. (2017), *Return to Meaning: A Social Science with Something to Say*, Oxford University Press, Oxford.

Anderson, G. (2008), "Mapping academic resistance in the managerial university", *Organization*, Vol. 15 No. 2, pp. 251-270.

Argyris, C. and Schon, D. (1974), *Theory in Practice: Increasing Professional Effectiveness*, Jossey Bass, San Francisco.

Barry, J., Chandler, J. and Clark, H. (2001), "Between the ivory tower and the academic assembly line", *Journal of Management Studies*, Vol. 38 No. 1, pp. 87-101.

Becker, A. and Lukka, K. (in press), "Instrumentalism and the publish-or-perish regime", *Critical Perspectives on Accounting*. doi: 10.1016/j.cpa.2022.102436.

Bialecki, M., O'Leary, S. and Smith, D. (2017), "Judgement devices and the evaluation of singularities: the use of performance ratings and narrative information to guide film viewer choice", *Management Accounting Research*, Vol. 35, pp. 56-65.

Chapman, C.S. (2012), "Framing the issue of research quality in a context of research diversity", *Accounting Horizons*, Vol. 26 No. 4, pp. 821-831.

Clarke, C.A. and Knights, D. (2015), "Careering through academia: securing identities or engaging ethical subjectivities?", *Human Relations*, Vol. 68 No. 12, pp. 1865-1888.

Czarniawska, B. (2011), "Successful research: in whose eyes?", *European Accounting Review*, Vol. 20 No. 1, pp. 53-55.

Dubois, A. and Gadde, L.-E. (2002), "Systematic combining: an abductive approach to case research", *Journal of Business Research*, Vol. 55 No. 7, pp. 553-560.

Durand, R. and Vaara, E. (2009), "Causation, counterfactuals and competitive advantage", *Strategic Management Journal*, Vol. 30 No. 12, pp. 1245-1264.

Edwards, M.A. and Roy, S. (2017), "Academic research in the 21st century: maintaining scientific integrity in a climate of perverse incentives and hypercompetition", *Environmental Engineering Science*, Vol. 34 No. 1, pp. 51-61.

Elster, J. (2000), *Ulysses Unbound: Studies in Rationality, Precommitments and Constraints*, Cambridge University Press, Cambridge.

Enders, J. (2004), "Higher Education, internationalisation, and the nation-state: recent developments and challenges for governance theory", *Higher Education*, Vol. 47 No. 3, pp. 361-382.

Englund, H. and Gerdin, J. (2020), "Contesting conformity: how and why academics may oppose the conforming influences of intra-organizational performance evaluations", *Accounting Auditing and Accountability Journal*, Vol. 33 No. 5, pp. 913-938.

Espeland, W.N. and Stevens, M.L. (1998), "Commensuration as a social process", *Annual Review of Sociology*, Vol. 24 No. 1, pp. 313-343.

Giddens, A. (1984), *The Constitution of Society. Outline of the Theory of Structuration*, Polity, Cambridge.

Grossi, G., Kallio, K., Sargiacomo, M. and Skoog, M. (2020), "Accounting, performance management systems and accountability changes in knowledge-intensive public organizations: a literature review and research agenda", *Accounting, Auditing and Accountability Journal*, Vol. 33 No. 1, pp. 256-280.

Guthrie, J. and Neumann, R. (2007), "Economic and non-financial performance indicators in universities: the establishment of a performance-driven system for Australian higher education", *Public Management Review*, Vol. 9 No. 2, pp. 231-252.

Guthrie, J. and Parker, L. (2014), "The global accounting academic: what counts", *Accounting, Auditing and Accountability Journal*, Vol. 27 No. 1, pp. 2-14, 2014.

Harvey, L. and Green, D. (1993), "Defining quality", *Assessment and Evaluation in Higher Education*, Vol. 18 No. 1, pp. 9-34.

Hood, C. (1995), "The "new public management" in the 1980s: variations on a theme", *Accounting, Organizations and Society*, Vol. 20 Nos 2-3, pp. 93-109.

Hopwood, A.G. (2007), "Whither accounting research?", *The Accounting Review*, Vol. 82 No. 5, pp. 1365-1374.

Humphrey, C. and Gendron, Y. (2015), "What is going on? The sustainability of accounting academia", *Critical Perspectives on Accounting*, Vol. 26, pp. 47-66.

Johnston, R.B. and Riemer, K. (2014), "On putting the score ahead of the game", *Communications of the Association for Information Systems*, Vol. 34 No. 1, pp. 849-856, available at: http://aisel.aisnet.org/cais/vol34/iss1/47.

Kalfa, S., Wilkinson, A. and Gollan, P.J. (2018), "The academic game: compliance and resistance in universities", *Work, Employment and Society*, Vol. 32 No. 2, pp. 274-291.

Kallio, K.M. and Kallio, T.J. (2014), "Management-by-results and performance measurement in universities–implications for work motivation", *Studies in Higher Education*, Vol. 39 No. 4, pp. 574-589.

Kallio, K.M., Kallio, T.J., Tienari, J. and Hyvönen, T. (2016), "Ethos at stake: performance management and academic work in universities", *Human Relations*, Vol. 69 No. 3, pp. 685-709.

Kallio, K.M., Kallio, T.J. and Grossi, G. (2017), "Performance measurement in universities: ambiguities in the use of quality versus quantity in performance indicators", *Public Money and Management*, Vol. 37 No. 4, pp. 293-300.

Karpik, L. (2010), *Valuing the Unique. The Economics of Singularities*, Princeton University Press, New Jersey.

Kivistö, J., Pekkola, E. and Lyytinen, A. (2017), "The influence of performance-based management on teaching and research performance of Finnish senior academics", *Tertiary Education and Management*, Vol. 23, pp. 260-275.

Lukka, K. (2014), "Exploring the possibilities for causal explanation in interpretive research", *Accounting, Organizations and Society*, Vol. 39 No. 7, pp. 559-566.

Lukka, K. and Modell, S. (2010), "Validation in interpretive management accounting research", *Accounting, Organizations and Society*, Vol. 35 No. 4, pp. 462-477.

Lukka, K. and Mouritsen, J. (2002), "Homogeneity or heterogeneity of research in management accounting?", *European Accounting Review*, Vol. 11 No. 4, pp. 805-811.

Manes-Rossi, F., Spanò, R., Martin-Sardesai, A. and Guthrie, J. (in press), "Resistance to changing the way performance is measured: discourses at an Italian university", *Journal of Public Budgeting, Accounting and Financial Management*.

Martin-Sardesai, A. and Guthrie, J. (2018), "Accounting for the construction of research quality in Australia's research assessment exercise", in Widener, S., Epstein, M. and Verbeeten, F. (Eds), *Performance Measurement and Management Control: the Relevance of Performance Measurement and Management Control Research*, Emerald, pp. 221-241.

Mathies, C., Kivistö, J. and Birnbaum, M. (2020), "Following the money? Performance-based funding and the changing publication patterns of Finnish academics", *Higher Education*, Vol. 79 No. 1, pp. 21-37.

Modell, S. (2003), "Goals versus institutions: the development of performance measurement in the Swedish university sector", *Management Accounting Research*, Vol. 14 No. 4, pp. 333-359.

O'Connell, B., De Lange, P., Stoner, G. and Sangster, A. (2020), "Impact of research assessment exercises on research approaches and foci of accounting disciplines in Australia", *Accounting, Auditing and Accountability Journal*, Vol. 33 No. 6, pp. 1277-1302.

Parker, L. and Gurthrie, J. (2013), "Accounting scholars and journal rating and benchmarking", *Accounting, Auditing and Accountability Journal*, Vol. 25 No. 1, pp. 4-15.

Parker, M. and Jary, D. (1995), "The McUniversity: organization, management and academic subjectivity", *Organization*, Vol. 2 No. 2, pp. 319-338.

Picard, C.F., Durocher, S. and Gendron, Y. (2019), "Desingularization and dequalification: a foray into
ranking production and utilization processes", *European Accounting Review*, Vol. 28 No. 4,
pp. 737-765.

Pirsig, R.M. (1974), *Zen and the Art of Motorcycle Maintenance. An Inquiry into Values*, Bodley Head,
London.

Qu, S.Q. and Dumay, J. (2011), "The qualitative research interview", *Qualitative Research in Accounting
and Management*, Vol. 8 No. 3, pp. 238-264.

Schatz, M. (2015), "Toward one of the leading education-based economies? Investigating aims,
strategies, and practices of Finland's education export landscape", *Journal of Studies in
International Education*, Vol. 19 No. 4, pp. 327-340.

Shields, P.M. (1999), "Zen and the art of higher education maintenance: bridging classic and romantic
notions of quality", *Journal of Higher Education Policy and Management*, Vol. 21 No. 2,
pp. 165-172.

Soin, K. and Huber, M. (2021), "Compliance and resistance: how performance measures make and
unmake universities", *Organization*, pp. 1-22.

Sousa, C.A., de Nijs, W.F. and Hendriks, P.H. (2010), "Secrets of the beehive: performance management
in university research organizations", *Human Relations*, Vol. 63 No. 9, pp. 1439-1460.

ter Bogt, H.J. and Scapens, R.W. (2012), "Performance management in universities: effects of the
transition to more quantitative measurement systems", *European Accounting Review*, Vol. 21
No. 3, pp. 451-497.

Thomas, R. and Davies, A. (2005), "Theorising the micro-politics of resistance: new public
management and managerial identities in the UK public services", *Organization Studies*,
Vol. 26 No. 5, pp. 683-706.

Van Dalen, H.P. and Henkens, K. (2012), "Intended and unintended consequences of a publish-or-
perish culture: a worldwide survey", *Journal of the Association for Information Science and
Technology*, Vol. 63 No. 7, pp. 1282-1293.

Welpe, I.M., Wollersheim, J., Ringelhan, S. and Osterloh, M. (Eds) (2015), *Incentives and Performance:
Governance of Research Organizations*, Springer, Cham.

Willmott, H. (2011), "Journal list fetishism and the perversion of scholarship: reactivity and the ABS
list", *Organization*, Vol. 18 No. 4, pp. 429-442.

Wittek, L. and Kvernbekk, T. (2011), "On the problems of asking for a definition of quality in
education", *Scandinavian Journal of Educational Research*, Vol. 55 No. 6, pp. 671-684.

Wittgenstein, L. (1953), *Philosophical Investigations*, Blackwell, Oxford.

(The Appendix follows overleaf)

Appendix

Code	Faculty	Position	Date	Length	Number of interviewers
M-A 4B	Case B	Head of Department, Professor	15.3.2016	1 h 47 min	One
M-A 8B	Case B	Director of a research centre, Professor	15.3.2016	1 h 29 min	Two
M-A 5B	Case B	Head of Department, Professor	16.3.2016	1 h 45 min	Two
M-A 7A	Case A	Head of Department, Professor	17.3.2016	1 h 27 min	Three
M-A 2B	Case B	A top manager of the faculty, Professor	22.3.2016	2 h 12 min	One
M-A 2A	Case A	A top manager of the faculty, Professor	23.3.2016	1 h 30 min	Two
M-A 6B	Case B	Head of Department, Professor	24.3.2016	1 h 31 min	Three
M-A 8A	Case A	Head of Department, Associate Professor	29.3.2016	I h 22 min	Two
M-A 1B	Case B	A top manager of the university, Professor	30.3.2016	1 h 30 min	One
M-A 3A	Case A	A top manager of the faculty, Professor	31.3.2016	1 h 34 min	Three
M-A 4A	Case A	Head of Department, Professor	31.3.2016	1 h 26 min	Three
M-A 7B	Case B	Head of Department, Professor	1.4.2016	1 h 51 min	Two
M-A 3B	Case B	A top manager of the faculty, Professor	5.4.2016	1 h 26 min	Two
M-A 6A	Case A	Head of Department, Professor	13.6.2016	1 h 47 min	Two
M-A 5A	Case A	Head of Department, Associate Professor	13.6.2016	I h 40 min	Two
M-A 1A	Case A	A top manager of the university, Professor	14.6.2016	1 h 10 min	Two

Table A1.
Interview schedule

Source(s): Own creation/work

Corresponding author
Kari Lukka can be contacted at: kari.lukka@utu.fi