


RESEARCH ARTICLE

University research by the numbers: Epistemic methods of using digitized performance measures and their implications for research practices

Sara Giovanna Mauro¹  | Lino Cinquini²  | Margit Malmrose³  | Hanne Nørreklit^{2,3} 

¹Università degli Studi di Modena e Reggio Emilia – Dipartimento di Economia Marco Biagi (*University of Modena and Reggio Emilia – Department of Economics "Marco Biagi"*), Modena, Italy

²Scuola Superiore Sant'Anna (*Sant'Anna School of Advanced Studies*) – Institute of Management, Pisa, Italy

³Aarhus University, Aarhus, Denmark

Correspondence

Sara Giovanna Mauro, Università degli Studi di Modena e Reggio Emilia – Dipartimento di Economia Marco Biagi (*University of Modena and Reggio Emilia – Department of Economics "Marco Biagi"*), Viale Jacopo Berengario, 51, Modena 41121, Italy.
Email: saragiovanna.mauro@unimore.it

Funding information

European Academy of Management (EURAM) - EURAM 2020 Grant, Grant/Award Numbers: Application ID, 48

Abstract

This paper explores the way by which universities create meaning of digitized performance measures on research quality and their effects on university scholars' actions. Drawing on pragmatic constructivism, we scrutinize the epistemic methods by which the digitized performance measures of research quality are handled and used in the governance of research activities in two disciplinary fields in two university settings (Denmark and Italy) and their implications for constructing scholarly research practices. The analysis elucidates exemplars of two epistemic methods of building meaning of and using digitized performance measures: one reflective and interactive, and one authoritative and mechanical. The latter constrains the researchers' scholarly reasoning and communication and, hence, infringes upon the scholarly fundamentals of university practices. The paper concludes that if the issues of misconceptions of research quality related to the transitions from analog to digital language are neglected, the digital transformation results in dysfunctional human and social practices.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 The Authors. *Financial Accountability & Management* published by John Wiley & Sons Ltd.

KEYWORDS

digitalization, digitized performance measures, language game, pragmatic constructivism, university research

1 | INTRODUCTION

Intended to improve efficiency and effectiveness, new public management (NPM) reforms have infiltrated higher education systems worldwide, transforming universities into enterprises and scholars into entrepreneurs and performers (Gendron, 2008; Parker, 2011). In this context, universities have extensively implemented performance management systems dominated by quantitative and procedural forms of assessment (e.g., Craig et al., 2014; Kallio et al., 2017; Northcott & Linacre, 2010; Parker, 2022; Pianezzi et al., 2020). Accounting research reveals that in response to this, universities and academics have become increasingly oriented to achieving measurable results (Chua, 2019; Gerdin & Englund, 2019; Espeland & Sauder, 2016), with consequent dysfunctional effects that range from the growing simplicity and standardization of higher education activities to the lack of innovation and stagnation of scientific production, as well as the redefinition of academic values and roles (e.g., Gebreiter, 2021; Gerdin & Englund, 2019; Humphrey & Gendron, 2015; Tandilashvili & Tandilashvili, 2022; ter Bogt & Scapens, 2012; Vakkuri & Johanson, 2020). The traditional and normative ideal of a university as a democratic, autonomous, collegial, and innovative space has been replaced with a new ideal that values centralization, hierarchical accountability, and efficiency (Kallio et al., 2020; Lucas, 2006; Schmidt & Langberg, 2007). An “authoritative management” with a “judgmental evaluation” (ter Bogt & Scapens, 2012) approach to directing and governing university activities might jeopardize the fundamental scholarly values of academia (Du & Lapsley, 2019, p. 475). Yet, the lack of improvements in effectiveness might be enabled by the use of performance measurement systems that are not capable of comprehending the complex work of academics (Kure et al., 2021) and the multiple social functions performed (Craig et al., 2014).

However, this stream of studies generally puts little or no emphasis on digital technologies and their specific effects on academic practices, or the quality of measurement (Agostino & Arnaboldi, 2017; Agostino, Bracci, et al., 2022; Agostino, Saliterer, et al., 2022; Lavertu, 2016). Yet digital technologies strongly support such performance management systems (Agostino & Arnaboldi, 2017; Andrews, 2019) because they are an active element in synthesizing, accumulating, and transmitting information for the ratings and rankings that are applied in universities’ decision-making and performance evaluations. The production of data on system-interoperable platforms has brought the promises of greater empowerment for users and thereby more effectiveness in public sector activities (Agostino & Arnaboldi, 2017). However, the digital technologies enabling “multicentric, real-time production of data” pose important challenges in terms of the quality of the digitally produced data (Agostino, Bracci, et al., 2022; Agostino, Saliterer, et al., 2022). Specifically, the quality of the digitally produced data is influenced by both a “digitization” procedure, which assumes that a 1:1 transition from analog to digital format occurs for objects (processes, forms, products, services, etc.), and a “digitalization” procedure that changes “the processes beyond [the] mere digitising of existing processes and forms” (Mergel et al., 2019). The digital format is simple as it comprises a binary code based on a two-symbol system (typically zero and one), which, for instance, is used to represent the actors’ truth positions of false or true for a statement on service output, implying that truth positions can be counted and ranked (Agostino, Bracci, et al., 2022; Agostino, Saliterer, et al., 2022; Nørreklit et al., 2019)—the so-called digital language. However, in relation to phenomena such as research quality, it is questionable whether the actor’s truth positions on some basic statements can grasp the complexity of research quality, and hence, the 1:1 translation is questionable. If the digitalization procedure by which the meaning of digitized measures is created is not able to validate their quality, it might facilitate a digital transformation with profound and potentially dysfunctional “cultural, organizational, and relational changes” of university practices (Mergel et al., 2019, p. 12). In view of this, it is relevant to explore the quality of digitized measures and the data “translation” process with the aim of developing an understanding of epistemological conditions

of digital performance measurement (Agostino & Arnaboldi, 2017; Agostino, Bracci, et al., 2022; Agostino, Saliterer, et al., 2022; Lavertu, 2016; Micheli & Mari, 2014).

This study contributes to such research by providing insights into the epistemological soundness of the digitalization procedures by which university governance practices create meaning for digitized performance measures of research quality, and their effect on the fundamental values of scholarly practices (see the call for papers on these issues—Agostino, Bracci, et al., 2022). We apply the concept of epistemic method (Nørreklit & Trenca, 2021) to emphasize a concern about a practice epistemology of knowledge creation involving organized processes and techniques of measurement (Nørreklit & Trenca, 2021). Thereby, we distinguish the practice processes of knowledge creation from scientific research methods. We suggest that different epistemic methods might not be equally suited to creating meaning for digitized measures of research quality, and this might affect scholars' actions in relation to their research activities in a way that may infringe on some fundamental scholarly values of university practices, where power and domination might determine research quality rather than sound scholarly reasoning.

Against this background, this paper addresses the following research questions:

1. What is the soundness of the epistemic methods applied by universities' actors to create meaning(s) for the digitized measures of research quality?
2. What are the implications for researchers' actions of the soundness of the epistemic methods (with a view to evaluating whether they jeopardize the value of researchers' sound scholarly reasoning)?

To identify the epistemic methods applied and evaluate their soundness, this study draws on pragmatic constructivism (PC), which offers a useful ontology and epistemology for understanding and investigating practices (Nørreklit, 2017a, 2017b; Nørreklit et al., 2016). As the core of scholarly research practices, we consider that sound reasoning should set the criteria for what governs universities' knowledge creation, rather than money, power, or domination (Nørreklit et al., 2019). Analyzing the epistemic methods and their soundness empirically, the research scrutinizes the experiences and insights of 18 academics employed at five publicly funded university units within the sciences and business in Denmark and Italy.

The results of this research suggest that some academic fields are governed by a dialogical and reflective epistemic method when engaging with digitized performance measures, whereas other fields use an authoritative and mechanical epistemic method. The choice of epistemic method has implications for the values driving scholars' attention, the factual possibilities for action, and communicative interactions. In particular, an authoritative and mechanical application may strongly constrain academics' factual possibilities of pursuing their intrinsic motivations to undertake research projects that are valuable for society. As the academics become disempowered, the scholarly fundamentals of university practice are infringed upon.

Providing insights into how digital technologies are used in the performance measurement of complex public services (Craig et al., 2014) such as research activities, and their effect on academic actions, this study addresses a neglected field in accounting research (Agostino, Bracci, et al., 2022; Agostino, Saliterer, et al., 2022). Although studies have discussed the dysfunctional implications of performance measurement in the governance of university activities (see, e.g., Humphrey & Gendron, 2015; Vakkuri & Johanson, 2020), the epistemic methods associated with digitized performance measures and their scholarly implications remain largely under-investigated. Particularly, accounting research has investigated control styles and approaches with a focus on the characteristics of organizational processes of producing and using accounting information and, hence, with little attention to the epistemic quality of the measures (Ahrens & Chapman, 2004; Davila & Ditillo, 2017; Liboriussen et al., 2021; Simons, 1995). Moreover, the accounting literature mainly involves itself within the academic field of accountancy (see, e.g., Chua, 2019; Gebreiter, 2021; Humphrey & Gendron, 2015), and few studies conduct in-depth interviews into academics' reasoning and actions around both the technical and social aspect of digitized measures, which are necessary to unfold larger nuances, divergences, and effects of measurement (Nørreklit et al., 2016; ter Bogt & Scapens, 2012). Engaging with different academic fields, the present study reveals such nuances and novel insights on how different fields create, use, and

respond to digitized performance measures that are otherwise taken for granted. Digitized measures can be created and used in different ways depending on the epistemic method applied, which can result in different (non-)intended effects. These insights regarding effects and the connection to epistemic methods of producing and using digitized performance measures are made possible by conducting interactive interviews and analyzing the language games in use.

The structure of this paper is as follows. The next section illustrates the theoretical approach. The third section describes the research setting and the fourth section the methodological approach. The fifth section illustrates the findings of the empirical analysis, and the last two sections discuss the findings and draws conclusions.

2 | EPISTEMIC METHODS OF PERFORMANCE MEASUREMENT

In order to identify and evaluate the effects of the epistemic methods of research performance measurement, a paradigmatic foundation is needed that acknowledges that human actors draw on epistemic methods of performance measurement to construct organizational realities, while also recognizing that not all epistemic methods are equally well functioning in facilitating organizational effectiveness. In this section, we first describe the shortcomings and effects of the epistemologies of representational realism and social constructivism as the basis for an epistemic method for performance measurement. Furthermore, we reveal that the paradigmatic basis of digitalization might encourage an epistemic method of digitized performance measurement that accelerates such shortcomings. Finally, we argue that PC offers a useful ontology for understanding how functioning practices can be created and an epistemology that helps us formulate a set of meaningful criteria and processes for creating sound digitized performance measures.

2.1 | Representational realism and social constructivism

In the field of accounting, performance measurement is reasoned to emanate from an epistemology of representational realism (Chua, 1986). Representational realism considers measures to be the results of an epistemological process that creates “a bridge between reality, to which the object under measurement belongs, and the linguistic/symbolic realm to which the measurement results belong” (Mari, 2007, p. 42). Measurement involves “a process of empirical, objective assignment of symbols to attributes of objects and events of the real world, in such a way as to represent them or to describe them” (Micheli & Mari, 2014, p. 150). To this purpose, the epistemology of measurement should include an operation to uphold that the measurement results are based on empirical observation, that is, “empiricity,” and created independently of the observers, that is, “objectivity” (Micheli & Mari, 2014, p. 150). The production of measurement independently of the observer requires meaningful measurement concepts and scales outlined by qualities of content, reference, criteria, and consistency.

Accounting aims to make measurement “objective” by making rules that must be followed in the measurement processes (Porter, 1996). In general, it is assumed that accurate performance measures and fair standards can be obtained through a scientific process. This is rooted in the thoughts of Taylor, who aimed to take the control of the machine shop out of the hands of the workmen and place it completely in the hands of the management, by scientifically searching for the objective work procedure (Taylor, 1916). Following a cybernetic view of control, the objective and rule formulation, and the linkage of negative/positive feedbacks to deviations, are assumed to work coercively on human actors, enforcing them to conform to objectives and rules (Ahrens & Chapman, 2004), as is visible in the current performance management of universities.

Intensively, the epistemology of realism in relation to performance measurement has been challenged by scholars supporting a position of social constructivism. For instance, Tinker (1991) argued that there is no correspondence between the scientific level of accounting theory and practice, and the “events” in economic reality. There is no representational faithfulness because there is no objective economic reality “out there.” Measurement relates to a more

or less complex social phenomenon, socially constructed. Furthermore, the interpretation of rules and standards may change with the speaker, situation, and “traces” of previous articulations (Tinker, 1991). Reality is a construction reified by human thoughts, interactions, and agreements. Replacing human intentionality and reflection with enforcement to follow the rules, that is, mechanical objectivity, might destroy good management practices and hence facilitate dysfunctional results (Micheli & Mari, 2014).

Despite the extensive scholarly criticism of representational realism as the basis for performance measurement, little attention has been given to providing insights fruitful for developing a more robust epistemology of performance measurement (Micheli & Mari, 2014). One might argue that an interactive and enabling approach to performance management is an improvement on the cybernetic approach (Davila & Ditillo, 2017). Such an approach reveals design characteristics of how management should organize the involvement of employees in the production and use of accounting information; however, it does not give much detailed insight into how managers can create sound performance measures. Moreover, one might argue that since the turn of the century, organizations have invested increasing resources in implementing various performance measurement models (e.g., the balanced scorecard and integrated reporting). However, in such models, the measurements of properties of complex phenomena are often “equalled to the simplest and most easily measurable aspects of the activity” (Micheli & Mari, 2014, p. 153). Thus, the capability of the measurement system determines what is treated as important (Micheli & Mari, 2014), and reality is reduced to what is easy to measure. The measures operate together with mechanical determinism as the mode of explanation of organizational performance. The idea that quantitative performance measures together with mechanical determinism make up the business world is supported by clichés such as “what gets measured gets done” (Kaplan & Norton, 1992). As the relationships between actual and observed phenomena are speculative, the models constructed by numbers become a simulacrum of reality, that is, self-referential models without much resemblance to any profound reality (Baudrillard, 1981).

Thus, we can conclude that the epistemological bases of performance measurement models are fragile, and therefore, their effects are uncertain. We seem to witness the acceleration of a reductive epistemology of quantification that inflates itself to grasp the whole ontology (Micheli & Mari, 2014, p. 152). Next, we explain the role of the paradigm behind digitalization in relation to such development.

2.2 | Paradigm of digitalization

Systems of information technology (IT) build on a particular form of symbolic language rooted in the philosophy of language of Wittgenstein’s (1921) early work, *Tractatus Logico-Philosophicus*. Wittgenstein argued that the meaning of an everyday expression should be clarified by translating it into basic atomic observational statements called *elementary sentences* (Nørreklit et al., 2019). The elementary sentences can be mirrored in the world or not, thus establishing whether the relation between a proposition and the world is true or false. Thus, elementary sentences concern only observational matters, and a statement’s truth proposition has a binary status—a digital language—true or false. This binary status implies that truth propositions can be counted. Concerned about true representation, Wittgenstein’s early philosophy of language conforms to ideals of representational realism. Later, however, Wittgenstein (1953) in *Philosophical Investigations* considered the digital language as inadequate for understanding living human practices and replaced it with the philosophy of live language (described in the next section) (Nørreklit et al., 2019). Nevertheless, his philosophy forms the basis of IT.

Not only is IT machine language based on such a digital language, consisting of binary codes with instructions used to control computers, but also live thinking and understanding are translated into truth positions (i.e., true or false) through binary code (i.e., zero or one) that is the digital logic that permits IT to make measures (Nørreklit et al., 2019). When the digital language of the IT system is used to represent numerical variables and mathematical operations of calculative accounting models that are symbolic in nature, the 1:1 transition might be feasible. However, in relation to complex phenomena of the human life-world, the assumption of digitization as a 1:1 transformation between the

analog and digital format of objects is questionable, because the nature of the two languages, analog and digital, differs. In human practices, everyday language is replete with concepts with vague borders and many layers of understanding and meaning. These nuances cannot be captured by the clear-cut truth positions of elementary statements (Nørreklit et al., 2019). Thus, symbolic sentences might not be formed as elementary sentences that mirror the world. We label such sentences that do not represent phenomena *pseudo-elementary sentences*.

Pseudo-elementary sentences are an issue, for instance, when the complexity of research quality is transformed into a binary code through the ratings of scholarly work based merely on some people's assessments on whether simple statements are true or false. Universities can then evaluate the performance of each academic based on measures such as citation indexes, impact factors, articles published in certain journal categories, and so on, which are heterarchically generated at a distance through digital technologies (Agostino, Saliterer, et al., 2022). However, due to its complex, idiosyncratic, and discretionary nature, the quality of university research cannot be mirrored in an elementary sentence; hence, statements on which the digitized measures are based are pseudo-elementary sentences (Nørreklit et al., 2019). For instance, impact factor systems are binary in nature because a citation/non-citation is perceived as a true/false answer to an elementary sentence on whether the article is "good." However, behind, the citations might lie a variety of arguments and interests. Similarly, Financial Times journal ranking is based on digital language and pseudo-elementary sentences (Nørreklit et al., 2019). Thus, voting for journals is based on asking business schools' officials whether a publication belongs to category one or zero, representing whether it should be counted. Overall, embedded in the digital production and communication of performance measures such as the *h*-index and journal rankings, there are hidden views processed by concealed algorithms. Such digitized measures of research quality are not outlined by a criteria-based conceptual structure but created through a heterarchical network of producers' subjective views, expressed through truth propositions of simple statements. When measurement concepts are not clearly and meaningfully identified, the measures' meanings are blurred and open to interpretation.

The discrete and heterarchical production of digitized performance measures of research quality is challenging for users who need to validate their quality and understand their underlying meanings to make sound performance evaluations (Agostino, Bracci, et al., 2022; Agostino, Saliterer, et al., 2022; Nørreklit et al., 2019). Research reveals that despite their opaque meaning, such digitized measures might be used through a "digitalization" procedure where they "function autonomously as a data assemblage able not just to mimic markets but, increasingly, to enact them" (Burrows, 2012, p. 355). In a context of heterarchical generated digitized performance measures of research quality, there is an inherent risk of users losing sight of the attributes of the measures, and hence, it is uncertain on what basis decisions are made and what elements people are made accountable for (Agostino, Bracci, et al., 2022; Agostino, Saliterer, et al., 2022; Lavertu, 2016).

On this basis, we think it is fair to conclude that the epistemology by which digitized measures of research quality are produced and used does not seem to diminish the problem that measuring complex phenomena is "equaled to the simplest and most easily measurable aspects of the activity" (Micheli & Mari, 2014, p. 153). Digital technologies offer an open, interactive, and system-interoperable information platform where answers to complex questions can easily be transformed into a binary code and computed, and so they might accelerate the problem of reliance on what is measurable so that such measures make up the whole ontology. If we want to improve the quality of how performance assessment of university research is carried out, there is a need for a more robust epistemological process. This should enable the underlying meaning of the digitized measure to be made known, by developing definitions and criteria through proximity to the phenomenon behind the numbers and developing capabilities to integrate them in actor thinking and understanding. In the next subsection, we explain how the paradigm of PC can be a suitable basis for doing this.

2.3 | Pragmatic constructivism

To understand and develop the epistemic method by which digitized performance measures are used and their implications for university practices, we draw on the PC paradigm, which offers a useful ontology for understanding how

functioning practices are created (Nørreklit, 2017a, 2017b; Nørreklit et al., 2016). Additionally, it includes an epistemology that helps us formulate a set of meaningful criteria and processes for developing and evaluating epistemic methods of performance measurement of complex phenomena such as university practices that are effective at facilitating the creation of intentional outcomes. Later, we explain the core concepts of PC using the idea of the conventional university as an exemplar.

2.3.1 | Language games and the conventional university

PC considers human beings as creative, reflective actors who, with a view toward creating intentional results, construct and organize their practices around the use of a relatively complex set of language games. In these language games, actors' communications are interwoven with other action forms (Wittgenstein, 1953, §7). The specific type of language used in the actors' communications speaks to the specific field of practice and how it is organized and controlled (Wittgenstein, 1953). Although language conventions dictate the use of words, the specific meaning of language is associated with the local practices within which it is learned, used, and developed. People in local practices produce experiences and understandings through which knowledge and skills accumulate around the creation of functioning language games that might become more advanced over time (Nørreklit et al., 2019). Such complex human cognition is expressed in the notion of conceptual habitus (i.e., the embodied dispositions of a person or group that organize the ways in which they perceive and understand the situation and act upon it¹—Nørreklit et al., 2019).

The language games of capable scholars are based on scholarly habitus that are highly complex, have multiple layers, and are constantly developing. In their cognition, scholars draw on symbolic languages, such as those of mathematics and logic, to develop models with special grammars. Symbolic language is the foundation of scholarly language games; however, it does not express meaning by itself. Symbolic language acquires meaning in relation to its use in specific language games; it thereby functions through its interplay with conceptual formations of the scholarly, habitus-based language games of the academics, in which the scholarly habitus refers to the questioning, reasoning, and interacting of scholars (Nørreklit et al., 2019).

To construct reality successfully, however, language games of university governance must develop and integrate the following four dimensions in their actor–world relationships: values, facts, possibilities, and communication (Nørreklit, 2017b). First, the scholarly actors' values motivate them to guide their choices among action possibilities. If the scholarly actors experience their values as fulfilled, they are intrinsically motivated to be creative and responsible problem solvers. In addition, institutions are driven by values. Following the ideals of the Humboldtian tradition, the fundamental values of universities are the cultivation of human reasoning and character (Du & Lapsley, 2019; Nørreklit et al., 2019). Sound reasoning rather than money and power is considered a basic value that should govern knowledge creation and social interaction at universities.

Furthermore, for actors to achieve their intentional values, factual possibilities within the actors' value ranges are required. Facts are not the thing. The term fact is used to signal that proposed informational claims are considered trustworthy because they are based on evidence. There can be facts about both objective and subjective ontological phenomena existing in the world (Nørreklit, 2017b). Facts are a necessary basis for successful actions, but alone they are insufficient because actions cannot occur if actors have no possibilities among which they can choose. For universities to create high-quality research outputs, facts and possibilities must be integrated in the form of scholarly habitus, resources, and institutional environments that can create innovative results (factual possibilities). To procure the best possible conditions for the cultivation of science and humanity, the scholarly, habitus-based, language games must be *free*.

This means that research and teaching should develop based on scholarly principles and argumentation that should unfold in the open space of communicative interaction among peers independent of any political authority or economic interest (Magna Charta Universitatum, Bologna, 1988). Academics have a strong role in (co-)authoring their language

games through an interactive dialogical process, where they suggest arguments challenging one another (Nørreklit et al., 2019). Scientific knowledge is assumed to develop in the open space of communicative interaction among peers.

On this basis, we conclude that the fundamentals of universities are a community of free academics, who are self-motivated to do innovative research, have an advanced scholarly habitus within a specific field of research, and are engaged in communicative interaction with others through sound reasoning. If these characteristics are jeopardized through institutional power and domination, it might influence the academics' actions in such a way that universities' fundamental values are at stake.

2.3.2 | Language games of performance measurement

For the construction of epistemic methods of performance measurement that are effective at facilitating the creation of intentional outcomes, PC outlines some principles for the conceptual qualities of measures and the process of producing and using them (Kure et al., 2021; Nørreklit et al., 2016). In Table 1, column 2 provides an overview of features of an interactive and reflective epistemic method based on a language game of performance measurement that includes such conceptual qualities of measures and processes, and column 3 reports the features of an authoritative and mechanical epistemic method.

More specifically, columns 1 and 2 show that the establishment of conceptual qualities is crucial to the presence of meaningful concepts for the observation of research performance. They imply that the measurement concepts must be outlined through clarification of content and exemplary references. The *content* of a concept outlines the abstract idea of the property of the measurement object, which enables people to determine the exemplary references of the concept. The *exemplary references* establish a shared horizon of understanding of what the abstract idea implies in its practical use.

Furthermore, to increase precision, a concept might be given a supplementary *criterion-based meaning*. Criteria have the ability to overcome issues of subjectivity by transforming the qualitative basis of the conceptual content into numbers. Formulating and weighting different categories of performance of an activity can be turned into binary symbolic language (i.e., digitized) and quantified. Such a quantification of performance is not about measuring specific countable properties of the research, but about assigning a property to the phenomenon under measurement with the intention of controlling it. Thus, it is a measure of a non-apparent "property intended to measure" (Micheli & Mari, 2014, p. 151). To adequately address a measurement task in relation to purpose, such measures involve the reason-based construction of a conceptual model that can be pragmatically contested. There is incomplete available knowledge about the "true value" (i.e., facts) that should be assigned to describe the property of the object under measurement, but an approximation of the facts that is reasonable in relation to the goals of the measurement system can be justified based on its effects on actual outcome, that is, its pragmatics (Micheli & Mari, 2014, p. 152; Nørreklit et al., 2016).

However, the conceptualized measures function only if they become meaningfully integrated into the language game of the specific practice, considering all four dimensions of reality, through the actors' reflective reasoning and judgment. Due to the complexity of advanced habitus-based practices, a language game of performance measurement of such activities must involve the habitus of the organizational actors in the narration and interpretation of the performance measures (Nørreklit et al., 2019). Through communication, such actors need to be engaged in a reflective and dialogical process to create a coauthored meaning of the measurement concept in relation to the specific reality construction. To facilitate suitable intentional actions, actors in local organizational practices must agree on the function of the measurement concept to facilitate function practices. This means that a good concept must provide possibilities that are based on facts. It implies that the organizational actors should ensure that the measurement concept is identifiable by properties of appearance (facts) and is related to the possibilities embedded in the properties. Furthermore, the factual possibilities outlined in the concept must be within the value range embedded in the actors' habitus and the organizational goals. An institutionalized heterarchical digitized measurement system might be introduced in the process by managers as a point of departure to categorize research publication outlets. However, creating a meaningful

TABLE 1 Features of an interactive and reflective epistemic method versus an authoritative and mechanical epistemic method of performance measures.

	<i>Interactive and reflective epistemic method (column 2)</i>	<i>Authoritative and mechanical epistemic method (column 3)</i>
Concepts (column 1)		
Features of definitions		
Content	Abstract idea of "what is research quality" is formulated	Abstract idea of "what is research quality" is unclear
Reference	Examples of aspects of specific research outcomes signifying research quality are formulated	It is unclear what aspects of the research outcomes signify research quality
Criteria	Measures are based on meaningful reasoning and tested pragmatically	Measures are based on what is measurable
Features of a narrative scheme		
Communication	Reflective and dialogical interactions; focus on the integration of the four dimensions of reality	Monological voice, hierarchical top-down implementation of measures
Facts	Evidence of research quality behind measures of research	Postulation of facts
Possibilities	Reflection on the possibilities embedded in the factual properties	Deterministic, rigid, rule-built suggestions for actions
Values	Linking to values embedded in the actors' habitus and organizational goals	Linking to the values embedded in the institutionalized measurement system
Pragmatic truth	It facilitates intentional results	The system is right

conceptual understanding of such measures of academics' performance requires that managers engage in a dialogical and reflective interaction with the researchers to understand the facts, possibilities, and values behind the measures (Nørreklit, 2017b; Nørreklit & Trenca, 2021). Thus, the organizational actors will be able to identify the underlying meaning of the digitized performance measures and hence be able to integrate them reasonably in the language games of specific practices. The actors' reflective and dialogical interaction around the establishment of measures should be governed by meaningful reasoning in relation to the integration of the four dimensions of reality.

When a language game of performance measurement of research quality is governed by such conceptual features of an interactive and reflective epistemic method, it should facilitate the creation of intentional outcomes. Actors should proactively consider such features; however, the ultimate issue is whether the conceptual model facilitates the creation of intentional results, that is, *pragmatic truth*. Past experiences of dysfunctional effects give reason for a reflective learning process to develop conceptual models with a view to creating better pragmatic truth.

In contrast, we consider that the dysfunctional measurement problem is accelerated when managers without specific field insights use an authoritative and mechanical language game in their interpretation and evaluation of scholars' performance based on digitally produced measures. As depicted in column 3, an authoritative and mechanical language game of performance management does not carefully consider the establishment of conceptual qualities of content, reference, and criteria for measurement, but rather it is the capability of the measurement system that defines what is treated as important (Micheli & Mari, 2014). It involves a monologic voice that tends, in accordance with its own values, to make postulations about the interpretation of the numbers; and outlines rigid, rule-built suggestions for actions based on the results (Nørreklit & Trenca, 2021). An authoritative and mechanical management language game does not question statements presented as facts and possibilities and does not integrate the values and knowledge of others, but it might integrate factual possibilities that strongly constrain the other's actions. Accordingly, the language games of performance measurement might either support or constrain scholars' factual possibilities for creating intentional values, and in so doing, they might sustain or transform the scholars' language games.

3 | RESEARCH SETTING

This research analyses cases of publicly funded universities in Denmark and Italy. In both countries, university management is increasingly digitized (OECD, 2019), but this study particularly focuses on their digital research evaluation systems. Such systems are typically developed as part of NPM-inspired reforms, which have been widely implemented in both Denmark and Italy. Among Scandinavian countries, Denmark has had the most radical NPM-oriented reforms (Schmidt & Langberg, 2007), whereas Italy shows an incremental pattern of adopting NPM-oriented reforms (Hynman et al., 2014). Additionally, Denmark's higher education sector has a decentralized decision-making system, whereas Italy has a strong centralized system of university management. Thus, the university systems in Denmark and Italy are based on two different sociopolitical systems, which brings relevant nuances to our study.

3.1 | The research performance measurement system in Denmark

In Denmark, the 2003 University Act marked the beginning of a new era of NPM in Danish universities (Kristensen, Nørreklit & Raffnsøe-Møller, 2011). The governance structure was radically transformed by establishing new "boards," with mostly external members and a minority of academic representatives. Members are elected by the Minister of Science. The board appoints the university's rector, deans, and department heads. A later reform further extended the control exercised by the board and the rector to include decisions about the organization of the academic units and their managerial authority (Degn & Sorensen, 2015). Due to the "autonomy" gained, universities are expected "to demonstrate value for money by regularly documenting and evaluating institutional efforts" (Nielsen, 2017). The

board is made accountable to the Ministry of Science, Technology, and Innovation through development contracts, which are agreements made on future performance goals, including quantitative targets and output indicators.

Regarding research, Danish universities receive a fixed basic research grant (85%) as well as a basic grant distributed according to a results-based allocation model (15%) (Danish Ministry of Higher Education and Science, 2021). The results-based allocation model for funding universities is weighted as follows: 45% is allocated based on educational activities, 20% is allocated based on research activities financed from external funds, 25% is allocated based on research rankings, and 10% is allocated according to the number of completed PhDs. The research rankings are based on the Danish BFI (Bibliometric Research Indicator) list, which includes three levels: normal, high, and highest. Points are earned based on university employees' publications in particular channels (i.e., journals, series, and particular publishers) (Danish Ministry of Higher Education and Science, 2021). The listed publication channels rankings are revised regularly by members of 67 field-specific committees appointed by the Rector College's Executive Committee (Nielsen, 2017). The committee's reasoning for the rankings is not publicly available, and they do not outline the conceptual content or criteria for what make up research quality. All the research outputs of each university are registered by individual faculties in a database called PURE (Publication Research) that is supported and validated by Danish libraries.

The BFI and PURE systems are used to allocate funding to universities at the organizational level, but they are not intended to be applied for performance evaluation on individual levels (Danish Ministry of Higher Education and Science, 2021). Within universities, departments and their managers have the autonomy to decide how and to what extent such systems should be applied as part of the decision-making and evaluation process. The university's management is given autonomy to develop local ranking lists, which demonstrates a decentralized system.

3.2 | The research performance measurement system in Italy

In Italy, publicly funded universities have been under growing scrutiny since the 1990s (Bruno & Dal Molin, 2022; Cinquini, 2001; Reborá & Turri, 2010). On the one hand, universities have been given greater financial autonomy; on the other hand, several tools and processes have been introduced to control and manage performance in ways that pursue efficiency and focus on results (Bracci et al., 2020). At the university level, the key institutional actors in the governance system are the rector, academic senate and board of directors, general manager, board of auditors, and evaluation body (Corte dei Conti, 2021). The rector and senate, both elected, are internal university members, whereas the board of directors includes both internal and external experts and professionals, some of whom are elected and some nominated by the senate, who then appoint the general manager following the rector's proposal.

The Italian Ministry of University and Research (MUR) provides universities with ordinary funding (e.g., for salaries, maintenance, and research) and specific funding for large investments and projects. A portion of ordinary funding is allocated based on the university's results in the ANVUR (National Agency for the Evaluation of Universities and Research Institutes) evaluation; according to the latest available information on the MUR website, around 30% of available resources are allocated based on a rewarding logic, and 80% of this part of the funding is determined by the VQR (*Valutazione della Qualità della Ricerca*—Assessment of Research Quality) results. ANVUR is the national agency established to manage the external evaluation of public higher education institutions, under the control of the MUR, determining the “quality rankings” of universities and departments. Regarding research, Italian universities and departments are evaluated based on their research outputs (e.g., number and type of publications—classified as papers in international/national journals, book chapters, books, etc.—and quality of research based on citations, journal impact factors, and/or journal rankings) through the VQR procedure managed by ANVUR. The evaluation of research outputs is performed by a national group of experts, drawn by the MUR from a list of candidates prepared by ANVUR, selected from full professors of various disciplines. The result is a ranking of universities and departments based on their research outputs and third mission activities (social and economic impact). National journal ranking lists are developed for each academic field and used as a reference point to assess research quality for the VQR.

These are developed by ANVUR through a working group of experts, using bibliometric indexes and international database information. The processes of monitoring results and guaranteeing quality have increasingly influenced resource allocation. For instance, so-called departments of excellence receive additional financial resources based on their (i) research quality, as determined by a departmental performance indicator set by central authorities, and (ii) development projects.

To facilitate evaluation procedures, each university's research outputs are registered in an online database called IRIS. The use of digital tools allows information to be collected periodically, providing the content reference to assess the research results of departments and individuals by central governments and universities themselves. The information is also considered by MUR's national group of experts to assess whether individuals are qualified to become full or associate professors, mainly based on the number and type of their publications, citations for bibliometric disciplines, and international funded projects and collaborations.

Accordingly, in Italy, we find a centralized use of digitized research measures that assumes the measures to be reliable. However, each university can elaborate its own regulations and guidelines concerning performance measurement and develop its own system to allocate funds among departments, and manage careers. In doing so, the universities are influenced by the parameters used in national evaluation by ANVUR, and they often use journal rankings as a reference point for assessing research quality. Thus, the Italian case demonstrates greater central control of what makes up research quality compared to the Danish case.

4 | RESEARCH METHODOLOGY

We adopted an explorative approach to address our research questions empirically, examining various universities. By exploring different epistemic methods of performance measurement and their effects on scholarly practices, we seek to go beyond an interpretive paradigm (Liboriussen et al, 2021). Thus, instead of purely explaining, we seek to investigate how academics cognitively link their perceptions of performance systems to the construction of actions to produce intentional outcomes. Therefore, we designed a data collection method that would ensure sufficient insights into university scholars' reasoning and actions around the use of digitized performance measurement of university activities. To do this, we engaged in a dialog with the informants and discussed their worldviews by using pre-produced texts that illustrated examples of performance measurement-related situations. This supported a reflection on the actors' different perspectives regarding such systems. On this basis, the empirical findings informed the development of our conceptual framework on epistemic methods, which shaped our case analysis. Thereby, the descriptions of the empirical epistemic methods of measurement in the next subsections make up exemplar cases of the theoretical development based on the PC paradigm.

4.1 | Data collection

This study takes a qualitative approach with the aim of exploring different specific contexts in-depth. The analysis focuses on "paradigmatic cases" (Flyvbjerg, 2006) of "positive," good, performance according to national performance rankings. In the Danish case, good scholarly reputation and good positioning in international rankings were considered in the case selection process. In the Italian case, rankings based on the VQR were used to identify the departments and universities considered good performers.

Three university departments (units) in Italy and four in Denmark were selected to represent diverse disciplinary fields with different features and historical traditions. The chosen departmental subject areas were (i) natural sciences, (ii) arts, (iii) political science, and (iv) business. Additionally, secondary data were collected and analyzed to establish the research background, including internal guidelines and protocols, information available on the universities' institutional websites, and legislative documents produced by national governments or agencies.

TABLE 2 Interviewees

	Italy	Denmark
Natural science	3	3
Full professor	1 (male)	1 (male)
Associate professor	1 (female)	1 (male)
Assistant professor	1 (male)	1 (male)
Business	3	9
Full professor	1 (male)	3 (male)
Associate professor	1 (male)	4 (male and female)
Assistant professor	1 (female)	2 (male and female)

Semi-structured interviews were employed for primary data collection to elicit interviewees' opinions, experiences, and beliefs, while gathering detailed information regarding the empirical context. Based on the understanding obtained by reading the literature on performance measurement and digitalization, we prepared an interview guide with open-ended questions about the production and use of digitized performance measures. To capture employees' ways of perceiving their use (epistemic method), the interviews were arranged as interactive dialogs (Nørreklit et al., 1986). For this purpose, the interview script included real-life scenarios simulating potential situations occurring within universities to identify the interviewees' perspectives and experiences. Real-life constructs, by being realistic, should pave the way for the discussion of alternative, even diverging, positions to explore interviewees' reactions and reflections on such situations (Argento & van Helden, 2022; Lapsley & Llewelyn, 1995). The real-life examples were shared with the interviewees by asking them to read a slide illustrating the scenario. Then, the interviewees were asked about the existence of such a context in their real work-life setting and were asked for their reactions and opinions.

The academics interviewed within the departments were systematically selected to provide a spectrum of views by involving actors with different types of experiences (full and associate professors and researchers, both male and female). In Italy, in two cases, the full professors were the heads of their departments, and in Denmark, four department managers were interviewed. The interviewees cannot be considered the representative of the opinions, beliefs, values, and experiences of their respective departments. Instead, they illustrate different examples of individual perspectives, experiences, and understandings (Nørreklit et al., 1986), which ultimately pinpoint in-depth reasoning. Thus, they were identified and interviewed to share their views as exemplars of their universities' actions. We do not aim to represent the whole spectrum of actors' perceptions and reasoning about performance measurement, nor to generalize our results to the entire disciplinary field/university department. Because the features of an epistemic method are contextually specific and we investigate individual perspectives, experiences, and understandings, we aim to identify some of the epistemic methods in place and their effects. To do so, we collected multiple views from multiple departments, and then, after coding our transcripts, we have chosen to illustrate only five units representing two of the subject areas of the sample: natural sciences (1 department per each country) and business (one department in Italy and two departments in Denmark). Table 2 reports the list of interviewees. We chose these because our analysis showed that these two types of departments exhibit insights in different approaches to and effects of the performance measurement systems in place.

Regarding the validity of the data on the epistemic methods of digitized performance measurement, we considered consistency in interviewees' perceptions of how the measurements are used, and we checked their perceptions against the secondary data. We did not identify any inconsistencies. Regarding actions, we are concerned about conceptualization of the various ways of acting on the measurements. We have identified some ways according to the interviewees' experiences, and we have focused on these as exemplary of the context, but we recognize that there might be also other ways of acting.

In total, 861 min of interviews were recorded, transcribed, and analyzed. The interview transcripts were read by the researchers and analyzed to identify key issues concerning the types of digitized performance measures produced, how they have been used, and the primary perceived effects and implications of digitalization according to the interviewees. These key issues were then analyzed in light of the PC approach described above, codifying the recurrent and relevant issues by using the key concepts of PC, as explained in the next section.

4.2 | Analyzing the cases

To analyze the language games driving the epistemic methods of digitized performance measures and their effects, one can examine the interviewees' talk in relation to the features of the epistemic method by which measures are used and their action on such measures (e.g., Nørreklit & Trenca, 2021). We conceptualize the language games in use based on their resemblances with the features of the epistemic methods outlined in Table 1 and the scholarly actions they induce.

Specifically, we analyzed the first research question on the epistemic method used to create meaning for digitized measures by investigating the use of digitized performance measurement of research activities in relation to the language games described by interviewees around recruitment, promotion, and funding. First, we focused on whether the interviews revealed a language game in which actions are taken to ensure that the digitized measures are trustworthy expressions of research quality because they are outlined through the clarification of conceptual content, exemplary reference, and criteria, or whether they are rigidly taken for granted. Thus, we paid attention to whether the digitized performance measures were communicated hierarchically in a top down manner as facts on valuable research without considering the evidence, or whether they were questioned through a dialogical interaction with organizational scholars governed by a reflective reasoning process to investigate the evidence of research quality according to shared set criteria and develop digitized measures suitable for establishing a more sound factual basis for evaluating the level of research quality. Moreover, we considered how the digitized measures were linked to choices among action possibilities. We are interested in whether choices of possibilities are linked in a deterministic way, that is, mechanically, to digitized performance measures or whether such measures together with other observational points are reflected upon to make considered decisions about which action possibilities are likely to lead to valuable research results. On this basis, we can conclude whether the epistemic method is authoritative and mechanical, creating a poor factual basis for action facilitating research quality (i.e., it poorly integrates the four dimensions of reality), or whether the epistemic method is dialogical and reflective for making reason-based judgments that integrate the four dimensions of reality and hence form a sound basis for actions fulfilling intentional values (see Table 1).

Similarly, we addressed the second research question on the implications of the epistemic method of using digitized performance measures on researchers' actions by looking at how doing so shapes the researchers' practices in relation to their factual possibilities for undertaking action that is aligned with the value set of the scholarly habitus, and being engaged in communicative interactions with others through sound reasoning. Based on this evaluation, we discuss whether, in relation to research, the scholars draw on a scholarly, habitus-based language game, or if we are witnessing a transformation toward a new form of language game that infringes upon the scholarly fundamentals of university practices.

In view of our conceptual framework, our findings provide exemplary cases of epistemic methods of digitized performance measures and their implications for scholarly practices (Liboriusen et al., 2021). Whether a case is exemplary depends on whether its content extends the border outlined by the conceptual content of the theoretical framework. Deviations of the actors' case descriptions from the theoretical descriptions do not necessarily imply that the case does not match the conceptual content, as the lines drawn by the theoretical framework are not absolute (Nørreklit et al., 2016). Absolute evidence can never be obtained from such cases, but we sought to enhance the robustness of our case analysis by applying critical methodological standards of checking the evidence of the depictions.

5 | EMPIRICAL ANALYSIS

This section analyses the language games embedded in the selected university departments' use of digitized performance measures and their implications for scholars' actions, first in Denmark and then in Italy. Table 3 summarizes the key findings discussed in the following subsections.

5.1 | Natural sciences department in Denmark

The natural sciences department is a large unit that includes approximately 10 section leaders. In some fields, the department is among the best in the world. Scholars are charged with publishing top-tier papers within their fields, which means that only a small number of high-quality journals are relevant for them to publish in. The interviews reveal an awareness of the conceptual qualities of such digitized measures of research quality. A very small number of particular journals within a field are considered to publish top-quality innovative research, (e.g., *Nature* and *Science*). There is a shared understanding among peers within a field that these particular journals publish research of high scientific merit, that is, important, original, surprising, and well-performed research. Historically, these journals have published articles with pragmatically documented scientific significance. The narrow identification of such top journals implies that these can be justified as measures for "high quality, innovative research." Thereby, the binary logic is supported in which the value I means that the sentence "high quality, innovative research" is true, and O means it is false despite the fact that it is not shaped as an elementary atomic sentence that mirrors the world but involves many layers of understanding and meaning.

Moreover, we witness an awareness of author-level impact factor implications. For example, the department manager raised specific concerns about the quality of the scientific habitus that might underlie the measurement of the number of citations an author has. To support this statement, he referred to the British mathematician Andrew Wiles, "who solved one of the most significant mathematics problems of the last 400 years [Fermat's Last Theorem]. Andrew Wiles received 856 citations for his work, which is not much compared to a technical administrative person that might have 28,000 citations". However, the technical administrative person does not have a scientific habitus level equivalent to that of Andrew Wiles. Accordingly, citations and impact factors are not considered trustworthy measures of research quality; that is, they are not sufficient solid facts.

Publication in a top journal is a necessary requirement when recruiting and promoting people, but it is not a sufficient condition. The department manager states that one should be careful with these discussions because they tend to focus on one factor, but scientific qualifications are a multidimensional space. Therefore, it is also important to examine the scientific habitus, which says something about the factual possibilities for engaging in top-tier, innovative research. When evaluating a scholar's scientific habitus, the managers and employees involved discuss and reflect upon information about collaborations, start-up abilities, and the attraction of external funding (demonstrating communication). The intention is to evaluate an applicant's factual possibility of creating advanced scientific insights. Thus, digitized measures are not the only information sources used in decision-making. The following quote by an academic also shows that the department does not use journal ranking mechanically: "We look more at the human side of things. I know at some departments they count number of publications and normalise the impact factor and all that sort of mechanical way of looking at things. That doesn't exist at our department."

Overall, the interviewees emphasize that the department uses an interactive and reflective epistemic method to establish a confined journal ranking that is outlined by features of conceptual qualities (see Table 3). This process implies that the digitally produced measures meaningfully represent the highest standard of innovative research within the field. Furthermore, when evaluating a researcher, the department evaluates the scholar's factual possibilities for contributing to the advancement of scientific insights, that is, scientific habitus, based broadly on factual information, dialogical interaction, and reflective reasoning. This epistemic method implies a solid factual basis for

TABLE 3 Case study findings

	Science (DK)	Business (DK)	Science (I)	Business (I)
Epistemic methods of using digitized performance measures (RQ1)				
<i>Epistemic methods</i>	Reflective and interactive	Authoritative and mechanical	Authoritative and mechanical	Authoritative and mechanical
<i>Values</i>	Quality in research output	Quality in research output	Quality in research output	Quality in research output
<i>Facts</i>	Journal ranking developed conceptually and scholarly habitus scrutinized	Digitized measures are postulated to be representation of research quality	Digitized measures are used (with criticism) as representation of research quality	Digitized measures are postulated to be representation of research quality
<i>Possibilities</i>	Reflective	Mechanical institutionalized	Mechanical ambiguity	Mechanical institutionalized
<i>Communication</i>	Scholarly dialog	Strongly, hierarchical	Hierarchical	Strongly, hierarchical
<i>Integration</i>	Good	Poor	Poor	Poor
Effects of epistemic methods on the academics' actions (RQ2)				
<i>Values</i>	Create highest standard of innovative research	Produce scores for the measurement system	Produce scores for the measurement system	Produce scores for the measurement system
<i>Factual possibilities</i>	Intrinsic motivated researchers Scholarly habitus engaging in top-tier innovative research	Disempowered academics Research as a matter of publication in high-ranked journals Crowding out of relevant research fields	Disempowered academics Research largely as a matter of publication stars (citations, h-index) Crowding out of relevant research topics and areas	Disempowered academics Research as a standardized process Crowding out of relevant research topics and areas
<i>Communication</i>	Scholarly	Digital metrics	Digital metrics	Digital metrics
<i>Pragmatic truth</i>	YES	NO	NO	NO
	Intentional values facilitated	Intentional values inhibited	Intentional results inhibited	Intentional results inhibited

recruiting and promoting researchers with top-level scientific habitus, who can conduct innovative research to be published in the top journals within the specific fields.

The department's conceptualization of research quality is aligned with the values and factual capabilities of the university's scholars. For instance, one researcher expressed that he "would never even consider publishing in a non-top-rated journal. I mean, why bother? You get good results in good journals that people read, and that's it." The scholars are intrinsically motivated to conduct research and produce scientific results (values) rather than being extrinsically motivated through reward systems. In accordance with the focus on intrinsic motivation, advancements in scientific insights and not journal rankings are the focal point of attention in the researchers' daily communications. Because the researchers are highly capable and intrinsically motivated to publish in top-tier journals, management has no need to pressure researchers to increase their quantitative output measures (e.g., ranking, impact, and number of publications). The use of an interactive and reflective epistemic method to establish solid facts on the quality of the academics' scientific habitus helps to explain why the department's researchers are extremely successful in publishing in these journals. Overall, the epistemic method provides space (factual possibilities) for such top-level researchers to use their intrinsic motivation to create innovative research in interaction with each other and hence to realize both their and the organization's values (see Table 3).

5.2 | Business department in Denmark

We examine two Danish universities' research units, both of which are situated within business faculties that encompass a broad range of disciplinary fields, including, governance, accounting, and finance. The faculty units use digitized measurement systems that are strongly influenced by institutionalized journal ranking systems such as the Academic Journal Guide (CABS, 2021) to evaluate individual scholars' research performance. To ensure consistency across disciplines, the institutionalized journal ranking systems are enacted at the top management level outside the specific fields in which the quantitative output norms are established. Both selected research units have no outline of the conceptual content of research quality or specific exemplary references demonstrating significant research quality. This implies that it is uncertain what the digitized measures signify. One academic illustrated the sometimes-skewed representation of the measures: "... the ranking of journals within a given list is, to a large extent, arbitrary. It is biased or corrupted or whatever you call it, because sometimes people tend to have a few journals in the top group. And then I can hear people say, we know what the four top journals are, but we better include other journals in that top list because otherwise, we have too few public publications in that top list. And that is, of course, a sign of corruption in the system." Accordingly, the interviewees consider the journal rankings to be a questionable representation of research quality because they are driven by emotional self-interests of those who constructed the system. The journal rankings are considered elementary sentences depicting the quality of the research contributions, although they are, in fact, pseudo-elementary sentences. Similarly, one researcher noted that the journal rankings in accounting tend to be dominated by what some British or American scholars with particular types of research interests want to read. Consequently, one of the academically considered world's best journals on accounting information systems is ranked at level two.

Although the distorted nature of the measures was noted by some scholars, management strongly emphasize such digitized measures and rely on them as factual representation of research quality that should be used in a hierarchical, top-down, mechanical fashion in deriving possibilities for action (see Table 3). Poorly conceptualized measures of research quality are presented as the whole reality that should be used mechanically to steer the decision-making process on recruitment, promotion, and rewards. As expressed by one researcher, "everything that is signalled to us is that this is the one matrix that matters." In particular, when evaluating the research qualifications of applicants and employees, the focus is predominantly on higher level journals. Such an authoritative and mechanical epistemic method of using digitized performance measures implies that action possibilities are not derived on the factual basis of research quality (values).

Our findings reveal that the authoritative and mechanical use of a digitized measurement system that is—according to the academics—poorly based on facts about research quality has serious dysfunctional consequences for scholars' factual possibilities for doing research. One negative consequence of such highlighted corrupted journal rankings is the crowding out of research focusing in certain important areas for business and society. In one department, for instance, important disciplines were nearly eliminated because all of them were required to meet identical norms of digitized performance measures. In another department, an academic noted that the low journal ranking of accounting information systems implied that it was not factually possible to give promotions to scholars doing research in the area of accounting and IT, "which does not make sense in a stage of entering the digital age" (Academic). Biased to some research areas, the measures affect some scholars' factual possibilities of building a scholarly career. For example, one academic explained that he changed from doing deep research in management accounting systems to more broadly focusing on the effects of management styles on decision-making because he could not build a career upon his first focus. Accordingly, research areas of high importance to practice, and hence to teaching, are not considered factually possible for academics to conduct research.

Consequently, in some fields it can be difficult to find research that is relevant to teaching: "I must admit, it's rare that I see an article in a journal where I think this is exactly what I need for my teaching" (Academic). Additionally, the choice of scholars based on digitized performance measures does not obviously lead to good, innovative research. For instance, one scholar stated that "everyone was quite bored and quite unimpressed by the job presentation performance of the last associate professor they hired; but quite openly he was hired for the reason that he had a really good CV, papers in the pipeline for top journals and a good network."

Finally, the digitized measures are the focal point of conversation not only at department meetings and among employees but "also at lunch talk or something like that" (Academic). Academics explain how the poorly constructed, digitized measures create a hierarchical status of who has value for the organization and who does not. Accordingly, it is described that scholarly content is crowded out in communicative interactions; hence, researchers become visible through digitized measures that express their value abstractly.

Overall, the authoritative and mechanical use of poorly conceptualized digitized performance measures seems to facilitate a research practice that focus on producing scores for the measurement system, and, thereby, crowds out academic fields and research projects important for business and society (see Table 3). To pursue an academic career path, researchers suppress their intrinsic motivations and values to develop relevant scholarly insights within their academic field, because publishing in highly ranked journals within that field is not considered factually possible. The researchers are disempowered to be governed by their scholarly habitus. Furthermore, in the communicative interaction, researchers become visible through digitized measures that provide an abstract idea of their value rather than as academics with scholarly habitus to engage in advanced research around the core problems of the field.

5.3 | Natural science department in Italy

Contrary to Denmark, at the Natural Science Department in Italy, digital platforms are increasingly used to make information easily available. In response to the national evaluation system, the department has increasingly adopted and used digitized performance measures of research quality such as *h*-index to inform its decision-making and performance evaluation procedures. However, the interviewed academics questioned the validity of the digitized performance measures by identifying the problem that the indices and rankings on which they are based are ambiguous and "can reflect things that may not be real." This implies that the system is perceived as an aggregated system of factors that is not outlined through conceptual content and exemplary references. For instance, because the meaning of the *h*-index is ambiguous, it becomes impossible to conclude that a scholar with a high index is the best: "If one has an *h*-index equal to 100, and one has an *h*-index equal to 120, this does not mean that the one with the *h*-index of 120 is better than the one with the *h*-index equal to 100" (Academic). Moreover, the evaluation of an academic's research contribution was emphasized as "not an easy task and [one that] requires tools that go beyond the banal digitalisation"

(Academic). Although the measures of research quality are poorly constructed, the digitized measures of research quality are communicated as factual reference points for research quality and hence value. They are considered elementary sentences, although they are, in fact, pseudo-elementary sentences. Overall, we witnessed in the department an authoritative and mechanical epistemic method of using digitized performance measures (see Table 3).

However, the aforementioned ambiguity of the measures implies that the national evaluation procedure used to evaluate the department and university also appears ambiguous and challenging for the applicants to navigate. This is because, for the academics' evaluation procedures, they "had to understand which papers to select, but to select on the basis of what? On the number of citations?" (Academic). The ambiguity in criteria reveals that the mechanical linking of performance to funding and promotion might take place only at the surface level without a full recognition of the meaning of the digitized measures. Nevertheless, the mechanical use of digitized performance measures seems to be prevalent in the evaluation of individuals, because it is communicated that the academics have to reach certain metrics to be promoted.

As the epistemic method provides a poor factual basis for choices among possibilities, it produces dysfunctional effects. In fact, the digitized performance measure "has created serious distortions in the behaviour of researchers" (Academic) because the factual possibility for an academic to have a career depends on their subscribing to such measures. Academics might react to measures and rankings following a mechanical stimulus–response pattern of action: "People say 'I have to reach such an h-index, so I need to publish this number of papers'" (Academic). In this way, they become driven by an extrinsic motivation to produce results for the digitized measurement system. Consequently, academics cannot choose how to act according to their scholarly values and intrinsic motivations but must succumb to the misleading measures of research quality imposed by the digitized performance measurement system. As mentioned by one academic, the conventional thinking that perceives measurement systems as serving the purpose of providing information inputs for decision-making has been convoluted so the system has become the overall purpose to which the academics must provide the input. Such compliance with the digitized measurement system promotes certain research topics within the field and crowds out others. Thus, some research areas with high popularity "grow, while the rest lie desolate. And this will cause a serious cultural impoverishment" (Academic). Moreover, academics are influenced to highlight their publications on digital platforms making the publications visible to the community and to institutions. On top of this, digitized measures of research quality have implemented a sort of mechanical thinking "used for disempowering" the "scholarly voice of academics," but when "a machine decides who is good and who is not, we will pay for the consequences" of that decision for "the reason that the machine does not think."

Overall, the authoritative and mechanical epistemic method of digitized measurement system provides a poor factual basis for promoting researchers and allocating funding, which incentivize deformed research practices (see Table 3). This epistemic method does not encourage factual possibilities for researchers to undertake actions driven by their intrinsic motivations and scholarly reflective reasoning. Instead, the system induces an environment of disempowered researchers, which implies that some research topics and areas within the field are crowded out. The researchers are disempowered and not governed by their scholarly habitus and express weighty critic of the poorly conceptual outlined digitized measures.

5.4 | Business department in Italy

In the business department, the performance measurement system is based on the number and type of publications, relying on the national rankings to classify journal publications within their specific field. Thus, contrary to the Danish business units and the Italian natural science unit, the digitized performance measures are perceived by some scholars to be trustworthy and acceptable because they are "the result of digitalisation" together with "the willingness to find an objective measurement based on IT tools and rankings that are assignments of scores through the use of internet platforms" (Academic). Moreover, the digitized performance measures appear to be strongly institutionalized. For instance, the allocation of additional funding based on the department's "excellence," as measured by the nationally

outlined digitized performance measurement system, has prompted the department to establish an employment committee of full professors and to adopt regulations indicating the precise targets and objectives for scholar recruitment and career advancement. Furthermore, there are “technicians providing visibility of the knowledge” (Academic) in charge of making data on research production available and communicating them to the academic and institutional environment. Overall digitalization in itself and the institutional willingness to use it seem to be the criteria for solid measures of research quality. However, the factual basis and meaning of such measures are not debated, and the logic beyond their production is not shared. We do not witness any consideration on the conceptualization of research quality with respect to abstract ideas, exemplars, or criteria of quantification. They are considered elementary sentences, but they are, in fact, pseudo-elementary sentences.

We find that the department has strongly institutionalized an authoritative epistemic method of using digitized performance measures (see Table 3). The information produced by the digitized performance measures is communicated and perceived by scholars as factual reference points for research quality and hence value, and the measures are used mechanically to make choices among the action possibilities of funding and promotion.

The digitized performance measures have accelerated the prospect of making researchers’ knowledge creation efforts auditable, holding researchers accountable for their “performance” and use of public resources. However, this use of digitized performance measures seems to influence the academics to produce an increasingly high quantity of outputs without reflecting on their relevance or quality: “Digitalisation pushes us to do more and more, because you have a number as a target and obviously you are incentivised to overcome that number since you have a clear objective. The fact that there is the possibility of measuring [your performance] digitally allows the definition of a clear threshold: either you are above or you are below it” (Academic). The statement shows that academics are motivated to achieve numerical targets to benefit their careers, and to some extent, they appreciate knowing what to achieve. However, when accountability is translated into compliance with digitized performance measures, our interviews indicate that it supports the standardization of research outputs rather than creativity and effectiveness. The digitized measurement system limits the scholars’ factual possibilities of being driven by an intrinsic motivation to make innovative research. Overall, academics feel that they are under increasing surveillance due to the increased level of accountability produced by digitalization.

Furthermore, it is felt that digitization of the performance evaluation of research quality might make “the tool more relevant than” human reasoning and judgments (Academic). For instance, an interviewee mentioned that digital technologies have created an “obsession” for communicating and making results visible by counting them, and this counting “has been strengthened, expanded and gotten out of hand” (Academic). An interviewee perceived the extensive use of a “very calculative form of IT-tools” as an attempt by heads of department to free themselves from a “sense of guilt with respect to mechanisms of choice and evaluation that could be of a subjective type,” so becoming disempowered. In view of this, “delegating to the algorithm is a form of moral comfort, of leaving these processes to an objective mechanism” (Academic).

Overall, the authoritative and mechanical epistemic method of digitized performance measures provides a poor representation of what comprises relevant and good research, which distorts scholarly decision-making. The authoritative and mechanical epistemic method might not only tightly constrain the factual possibilities of doing research within certain areas and topics in the academic field but also create a milieu of mechanical surveillance-style evaluations in which human interactions, reasoning, and ethical judgments are crowded out. The strongly institutionalized system induces an environment of increasingly disempowered researchers.

5.5 | Conclusions on the empirical analysis

The empirical analysis elucidates exemplars of two epistemic methods of using digitized performance measures: one reflective and interactive, and one authoritative and mechanical (see Table 3). At the Department of Natural Sciences in Denmark, we found a reflective and interactive epistemic method for producing and using digitized performance

measures. This department has developed digitized performance measures that are considered to conceptualize the content, exemplars, and criteria of research quality due to a limited ranking of journals that counts only those top journals that are considered to accept the highest quality and most important scientific research in their field. By providing meaning to digitized measures through conceptualization, the department handles the problem that research quality cannot be expressed as an elementary atomic sentence that mirrors the world. Nevertheless, the department does not use the digitized performance measures mechanically but evaluates the researchers' scientific habitus based on many facets of information and reflective reasoning. In effect, the researchers at the department possess advanced scientific habitus and intrinsic motivations to undertake innovative research, and the performance management approach makes it factually possible for them to take action to realize their scholarly values. The reflective and interactive language game fosters the factual possibilities for university language games driven by free, cultivated, scholarly habitus aiming to produce scientifically important results.

In contrast, the empirical analysis displays three examples of authoritative and mechanical language games in the use of digitized performance measures: one at Danish universities' business departments and two at Italian university units (see Table 3). We notice a difference between the digitized measurement system in Danish business units, where the same publication norms are applied across research fields, and the Italian system, where the publication norms differ across the disciplines. However, in all the three examples of authoritative and mechanical language games, the measures are taken as facts and are used authoritatively and mechanically by management in decision-making on recruitment and promotion in order to obtain and maintain the status of excellence and to obtain resources. The digitized performance measures are not conceptually outlined by content, exemplars, and criteria and were developed disconnected from the complex thinking and understanding of university scholars' research work; that is, no one-to-one relation exists between these questionable measures of research quality and scholarly perceptions of research quality. Embedded in the digitized performance measures may be a variety of hidden arguments and interests, the deformation or corruption of which are widely recognized among researchers. The digitized measures are considered to be based on elementary sentences depicting the quality of the research contributions, although they are, in fact, pseudo-elementary sentences.

The authoritative and mechanical epistemic method strongly constrains researchers' factual possibilities to pursue their intrinsic motivations to engage in research activities within their value ranges. An implication of this epistemic method is that researchers must often suppress their intrinsic motivations and thus conflict with their values to develop relevant scholarly insights within their academic field, because acting on these intrinsic motivations and values at the same time as pursuing careers is factually impossible. This is most damaging when a common standard is developed across disciplines. Thus in the Danish business units, where the same publication norms are applied across disciplines, the outcome is a competitive environment between academic fields governed by a distorted valuation measure, which ends up crowding out academic fields that are important to business and society. In this context, the researchers seem to be concerned about how to create their career paths by shifting their research focus. In Italy, where the publication norms are developed by scholars within a discipline, the crowding-out effect facilitated by the distorted performance measures happens in relation to research topics and areas within the discipline. In this context, the researchers seem to be concerned about the digit world accelerating a milieu of robotic evaluation in which human beings are disempowered. Overall, when the values and action possibilities embedded in the scholars' language games in relation to research are pressured by the authoritative and mechanical use of poorly constructed digitized measures of research quality, we witness a suppression of scholarly values and reasoning.

6 | DISCUSSION

Our findings reveal that university managers have factual possibilities to handle digitized measures of research quality based on different epistemic methods, and thus, they are not victims of performance management tools per se, as the literature sometimes suggests (e.g., Chua, 2019). Moreover, the study displays that the choice of epistemic method has

implications for the soundness of digitized measures of research quality. In particular, the authoritative and mechanical epistemic method of using digitized measures facilitates the poor construction of concepts for performance evaluation of universities' activities. Poor data quality of performance measures is in-line with previous findings (e.g., Craig et al., 2014; Kallio et al., 2017; Kure et al., 2021). However, the analysis in this paper reveals that some of the misconceptions of research quality are related to the transitions from analog to digital language. This is visible, for instance, in ranking systems that assign the property of research quality to certain publications, but where we do not know the content of the research quality they measure. Such systems do not provide a conceptual understanding of what their measures aim to signify; in other words, the models do not have a factual basis. In particular, embedded in the multicentric digital production of data, there is an "inherent risks of losing control over their quality" (Agostino, Saliterer, et al., 2022, p. 161). Unless the meaning of the digitized measures of research quality is developed through interactive and reflective reasoning to outline their conceptual content, reference, and criteria of quantification, no shared understanding can emerge of what the measures signify. The authoritative and mechanical epistemic method lack of sound factual basis of digitized measures poses challenges in terms of sound decision-making and performance evaluation. This is in-line with the findings of Bracci (2022), who argues that there is a risk of biased outputs related to how digitized measures are designed, and the reliance on such measures in decision-making and control without considering this risk increases the difficulty in representing the performance achieved and ensuring legitimacy and trust.

However, institutional actors seem not to be always very concerned about whether performance measures have a questionable factual basis and dysfunctional effects. This behavior can be explained by the institutional value system that may assign questionable measures of research quality to the researchers' work with the intention of controlling the researchers. Such an authoritative view is supported by ter Bogt and Scapens (2012), who highlight an increasing application of managers' subjective judgmental forms of performance evaluation of researchers, rather than a dialogical interaction. If the measurement concept is poorly constructed, the meaning of the digitized measures of research quality is open to become flexible and exposed to corruption, power, and domination (Kure et al., 2021). An authoritative use of poorly constructed digitized measures implies that the possibilities linked to the measures become imposed and difficult or impossible to escape; hence, they play a performative role in research quality (Soin & Huber, 2021). The digitized measures facilitate a competitive environment—a pseudo market—that influences academics to modify their actions to fit the digitized measurement system, instead of focusing on what is more important in terms of impact for the community (see e.g., Agostino, Bracci, et al., 2022; Agostino, Saliterer, et al., 2022; Burrows, 2012, p. 355). The managers' interpretations of the digitized measures define what is treated as important research (Micheli & Mari, 2014). Accordingly, a major implication of digitalization is the space for a power switch in who gets to translate the meaning of such numbers (Agostino, Bracci et al., 2022 & Agostino, Saliterer et al., 2022). For instance, accountability originally belonged to accountants, due to their technical skills, but digitalization has brought a power shift away from the professional field to IT, social media, and, in our case, university and department managers, which has implied that the traditional criteria of sound performance evaluation is out of use. Thereby, "managing by the digitized numbers" become functional for implementing new managerialism into the university research system to increase not only efficiency but also for accelerating power and domination (e.g., Tandilashvili & Tandilashvili, 2022). Overall, we witness that the authoritative and mechanical use of digitized measures inhibit the university researchers' practices by taking control of the concepts of what makes research quality (Nørreklit et al., 2019). There is no interactive and reflective reasoning to establish general conditions of truth to govern the social interaction (Bracci, 2022; Nørreklit et al., 2016).

In addition, the study suggests that the roles of values and measures may be reciprocal, meaning that measures may shape scholarly values, but scholarly values also affect how digitized performance measures are developed and used. In the four cases where the authoritative and mechanical epistemic method is used, researchers become scholarly-disempowered. However, similarly to Gerdin and Englund (2019), we find that academics have different tactics for acting on the enforcement of digitized measures of research, and hence, the disempowerment of academics unfolds differently. For instance, academics in the Danish business departments find a way to have a career by steering their research focus away from the core of the scholarly field, whereas academics in Italy to have a career must subscribe to rules developed by scholars within the discipline, where the enforcement by which the rules are implemented seems

stronger in the department of business than in the department of natural sciences. Overall, academics in the business departments tended to take the performance measures of research quality for granted, despite concerns about their effects, whereas academics of the natural sciences are confident in their own critical assessment and reflections on the numerical measures and have their own professional and academic ways of meaningfully conceptualizing measures of research quality. This suggests that different academic traditions may influence understanding of the epistemic methods applied, which is further justified in the following by elaborating on the epistemological basis of the two fields.

The realist and constructivist views associated with the natural sciences fields of physics provide strong insight into the powers and limits of measurability and measurement processes and have founded robust theoretical bases that can be used to improve the theories and practices of performance measurement. Although the realist view of representational truth has a strong basis in natural sciences, the constructivist view is also present where interpretive models are a turning point for acquiring knowledge about reality (Micheli & Mari, 2014, p. 150). Within the constructivist view, a crucial epistemological concern is the construction of a conceptual model to adequately address a measurement task in relation to purpose. Uncertainty regarding measures is related to incomplete available knowledge about the true value that “should be assigned to describe the object under measurement,” but only information that is meaningful in relation to the goals of the measurement system is required. The dependence of the measurement model's foundational ability to create measures of sufficiently high quality can be justified based on its ability to create pragmatic results. Thus, the measurement model works together with an epistemological measurement process in which the model is tested pragmatically against a referential point of observation, that is, “everything does not go.” With such an epistemological basis at the fundamental level, it seems not surprising that the scholars of natural science exhibit an insightful understanding of performance measures and hence express weighty discomfort with poorly conceptualized digitized measures.

Although the accounting literature includes an extensive discussion between the paradigms of, for instance, realism and social constructivism (Chua, 1986), it has given little attention to insights that might be fruitful for developing a more robust epistemology of performance measurement (including digitized systems) in relation to issues around the notion of performance measurability and sound measurement processes (Micheli & Mari, 2014). Promoters of realism tend to envy physics and take its quantitative epistemologies as an ideal for social science, but they are not very concerned about conceptualization of measures or pragmatic testing of measurement models. Although the validity of quantitative estimates may not be established with real conviction, they are often given considerable weight (Porter, 1996, p. 8). The replacement of human intentionality and reflection with enforcement of rules, that is, mechanical objectivity, has been supported by social science research rooted in representational realism, “and it has a powerful appeal to the wider public” (Porter, 1996, p. 4). Taking the position that facts cannot be established and that all measures are shaped through power and domination, researchers take a radical social constructivist stand and look at the problem from a political and ideological perspective rather than engaging in the development of measures that can improve the performance measurement of the bodies in their power. These features might partly explain why the theoretical bases for performance measurement within social science are fragile, and hence their effect uncertain. This fragility of performance measurement within the social science may relate to business scholar's familiarity and hence comfort with poorly conceptualized digitized measures.

7 | CONCLUSION AND FURTHER RESEARCH

Prior studies have highlighted a shortage of accounting studies on the use of digital technologies in the governance of universities and their effect on academic work and values (Agostino & Arnaboldi, 2017; Agostino, Saliterer, et al., 2022; Lavertu, 2016). This paper addresses this research gap by focusing on the nature of digitized performance measures of research quality and the soundness of the epistemic method by which university governance practices create meaning for such measures and their effects on the fundamental values of scholarly practices. We first engaged with the problem theoretically, building on PC to outline two epistemic methods of using digitized performance measures—one

reflective and interactive and the other authoritative and mechanical—and their effects on the functioning of scholarly practices. Next, we explored the problem empirically by presenting the findings of how epistemic methods unfold at two different types of academic departments in two different national settings. Relating the global phenomenon of measures of research quality in two different national settings gives nuances of the conceptual framework and also enables more distinct conclusions (Parker, 2022).

The empirical findings reveal that both epistemic methods were in use at the different departments. The analysis indicates that the two epistemic methods deal differently with the misconception of research quality related to its transition from analog to digital language. We find that the authoritative and mechanical epistemic method of digitized performance measure does not facilitate the construction of sound concepts for the digitized measures of research quality but limits scholars' freedom to take responsible, intrinsically motivated actions to advance scholarly insights within their academic fields. In such a process, researchers become scholarly-disempowered, and important research topics, areas, and fields are crowded out. The authoritative and mechanical use of digitized measures inhibit the university researchers' practices by the taking control of the concepts of what makes research quality, whereas the numbers become functional for implementing new managerialism into the university research system to increase not only efficiency but also for accelerating power and domination. However, different academic traditions may influence the scholars' comfort with adapting to and action on poorly conceptualized digitized measures. Overall, the authoritative and mechanical epistemic method seems to facilitate a digital transformation that has a dysfunctional effect on the language game of scholarly practices at universities to the extent that it might infringe upon the free language and reasoning of the conventional university rooted in the Humboldtian tradition and subscribed to by many European universities in the Magna Charta Universitatum, Bologna (1988) (Agostino & Arnaboldi, 2017; Du & Lapsley, 2019; Nørreklit et al., 2019). In contrast, we show that to grasp the quality of the university's research performance, academics' scholarly habitus must be involved in the interpretation and conceptualization of the research underlying the digitized performance measures. No shared understanding can emerge of what the measures of research quality signify unless the meaning of such measures is outlined by conceptual criteria for their content, reference, and quantification involving human interaction and reflective reasoning (Agostino, Bracci, et al., 2022; Agostino, Saliterer, et al., 2022). In other words, an interactive and reflective epistemic method is the most suitable approach to sustain a scholarly open and sound academic field.

The study contributes to the existing research on the implications of digitalization and performance measurements in higher education (e.g., Agostino & Arnaboldi, 2017; Du & Lapsley, 2019; Kallio et al., 2020; Nørreklit et al., 2019; Pianezzi et al., 2020; Tandilashvili & Tandilashvili, 2022). Specifically, we contribute to this literature by revealing that some of the misconceptions of research quality may be related to the transitions from analog to digital language. The analysis of the authoritative and mechanical epistemic method contributes to the accounting research on the performance measurement of universities revealing that performance measures of research do not necessarily lead to increased efficiency and effectiveness but may produce techniques that are destructive of academic practices and virtues (Bracci, 2022; Dunleavy et al., 2005). Furthermore, we show that digitized measures of research quality are not necessarily taken for granted but can be handled in different ways applying different epistemic methods with different effects. Moreover, by engaging with different academic fields, the study adds novel insights into different ways of responding to otherwise taken-for-granted digital performance systems.

On top of that, we have developed a more robust conceptual framework for the performance management of research. We have argued that PC provides a theoretical basis for bridging the two views of realism and constructivism and that it offers an epistemology that helps us formulate a set of meaningful criteria and processes for creating sound digitized performance measures for university research. We also contribute by documenting the case of a natural sciences department that demonstrates how the interactive and reflective epistemic method can be applied with the effect of enabling and sustaining academic practices. Emphasizing the importance of the conceptual quality of measures, the paper adds to the stream of literature on the interactive and enabling control approaches of performance management where the technical side tends to be neglected (Ahrens & Chapman, 2004; Davila & Dittillo, 2017; Simons, 1995).

We focus on the nature of the digital language in relation to the production and use of digitized research performance measures. It is not known how widely these epistemic methods are used. However, because concepts are by their nature inherently concerned with general perspectives, it is reasonable to assume that the revealed epistemic methods based on a small number of cases can be applied to other situations (Nørreklit et al., 2016). Thus, the inquiry into what epistemic methods exist for using digitized measures and their effects on academic actions provides relevant insights that could inspire actors in the charge of designing and implementing such systems within universities. Following Payne and Williams (2005), however, we suggest that individual actors are responsible for establishing whether our findings are transferable to their specific practical context (see, e.g., Nørreklit et al., 2016). Our contextual descriptions of the cases in which these epistemic methods are applied should enable actors to assess where generalization to other contexts is or is not justified (Payne & Williams, 2005). Future studies may further investigate the conditions in which reflective and interactive language games of digitized performance measurement systems can be used successfully. A limitation of our study is that we do not have knowledge on the values, interests, and actions driving the design of the centralized digitized performance measures. Further research is needed on this. Additionally, future research can analyze the epistemic methods of the digitized performance measurement of research and their implications in other university contexts, academic disciplines, and in teaching and administrative activities. Finally, it seems pivotal to create further insights into the implications of digitized measures of research quality for human interaction, reasoning, and ethical judgments.

ACKNOWLEDGMENTS

The authors are grateful to the editors and the anonymous reviewers for their constructive feedbacks and comments on previous versions of this paper. Any errors remain of the authors.

DATA AVAILABILITY STATEMENT

Research data are not shared due to privacy reasons.

ORCID

Sara Giovanna Mauro  <https://orcid.org/0000-0003-4309-3625>

Lino Cinquini  <https://orcid.org/0000-0002-0899-5351>

Margit Malmrose  <https://orcid.org/0000-0002-8507-0731>

Hanne Nørreklit  <https://orcid.org/0000-0002-6636-7863>

NOTE

¹<https://en.wikipedia.org/wiki/Habitus>.

REFERENCES

- Agostino, D., & Arnaboldi, M. (2017). Social media data used in the measurement of public services effectiveness: Empirical evidence from Twitter in higher education institutions. *Public Policy and Administration*, 32(4), 296–322.
- Agostino, D., Bracci, E., & Steccolini, I. (2022). Accounting and accountability for the digital transformation of public services. *Financial Accountability & Management*, 38(2), 145–151.
- Agostino, D., Saliterer, I., & Steccolini, I. (2022). Digitalisation, accounting and accountability: A literature review and reflections on future research in public services. *Financial Accountability & Management*, 38(2), 152–176.
- Ahrens, T., & Chapman, C. S. (2004). Accounting for flexibility and efficiency: A field study of management control systems in a restaurant chain. *Contemporary Accounting Research*, 21(2), 271–301.
- Andrews, L. (2019). Public administration, public leadership and the construction of public value in the age of the algorithm and 'big data'. *Public Administration*, 97(2), 296–310.
- Argento, D., & Van Helden, J. (2022). Are public sector accounting researchers going through an identity shift due to the increasing importance of journal rankings? *Critical Perspectives on Accounting*, 102537.
- Baudrillard, J. (1981). Simulacra and simulations. In *Crime and media* (pp. 69–85). Routledge.

- Bracci, E. (2022). The loopholes of algorithmic public services: An "intelligent" accountability research agenda. *Accounting, Auditing & Accountability*, 36, 739–763.
- Bracci, E., Deidda Gagliardo, E., Paolo, E., & Gobbo, G. (2020). Coerenza e qualità dell'architettura di programmazione e controllo delle Università. Riflessioni sulle consonanze e dissonanze strategico-gestionali. In M. Paoloni, & E. Deidda Gagliardo (Eds.), *Il Performance Management nelle Università Statali: Stato dell'arte e nuove sfide*. Franco Angeli.
- Bruno, A., & Dal Molin, M. (2022). The implementation of PMS in Italian universities: A longitudinal analysis. *Management Control*, 2022(1), 59–78.
- Burrows, R. (2012). Living with the h-index? Metric assemblages in the contemporary academy. *The Sociological Review*, 60(2), 355–372.
- CABS. (2021). <https://chartere.dabs.org/academic-journal-guide-2021/>
- Chua, W. F. (2019). Radical developments in accounting thought? Reflections on positivism, the impact of rankings and research diversity. *Behavioral Research in Accounting*, 31(1), 3–20.
- Chua, W. F. (1986). *Radical developments in accounting thought: The accounting review* (pp. 601–632). American Accounting Association.
- Cinquini, L. (2001). Quale ruolo per la contabilità direzionale nelle Università. In P. Miolo Vitali (Ed.), *I sistemi di misurazione economico-finanziaria nelle Università italiane: problemi e prospettive*. CEDAM.
- Corte dei Conti. (2021). *Referto Sul Sistema Universitario*. May, 2021, 1–248.
- Craig, R., Amernic, J., & Tourish, D. (2014). Perverse audit culture and accountability of the modern public university. *Financial Accountability & Management*, 30(1), 1–24.
- Danish Ministry of Higher Education and Science. (2021). <https://ufm.dk/en/research-and-innovation/statistics-and-analyses/bibliometric-research-indicator/bfi-rules-and-regulations>
- Davila, A., & Ditillo, A. (2017). Management control systems for creative teams: Managing stylistic creativity in fashion companies. *Journal of Management Accounting Research*, 29(3), 27–47.
- Degn, L., & Sørensen, M P. (2015). From collegial governance to conduct of conduct: Danish universities set free in the service of the state. *Higher Education*, 69(6), 931–946.
- Du, J., & Lapsley, I. (2019). The reform of UK universities: A management dream, an academic nightmare? *Abacus*, 55(3), 452–482.
- Dunleavy, P. (2005). New public management is dead: Long live digital-era governance. *Journal of Public Administration Research and Theory*, 16, 467–494.
- Espeland, W. N., & Sauder, M. (2016). *Engines of anxiety: Academic rankings, reputation, and accountability*. Russell Sage Foundation.
- Flyvbjerg, B. (2006). Five misunderstandings about case-study research. *Qualitative Inquiry*, 12(2), 219–245.
- Gebreiter, F. (2021). A profession in peril? University corporatization, performance measurement and the sustainability of accounting academia. *Critical Perspectives on Accounting*, 87, 102292.
- Gendron, Y. (2008). Constituting the academic performer: The spectre of superficiality and stagnation in academia. *European Accounting Review*, 17(1), 97–127.
- Gerdin, J., & Englund, H. (2019). Contesting commensuration: Public response tactics to performance evaluation of academia. *Accounting, Auditing & Accountability Journal*, 32(4), 1098–1116.
- Humphrey, C., & Gendron, Y. (2015). What is going on? The sustainability of accounting academia. *Critical Perspectives on Accounting*, 26, 47–66.
- Hyndman, N., Liguori, M., Meyer, R. E., Polzer, T., Rota, S., & Seiwald, J. (2014). The translation and sedimentation of accounting reforms. A comparison of the UK, Austrian and Italian experiences. *Critical Perspectives on Accounting*, 25(4–5), 388–408.
- Kallio, T. J., Kallio, K.-M., & Blomberg, A. (2020). From professional bureaucracy to competitive bureaucracy – Redefining universities' organization principles, performance measurement criteria, and reason for being. *Qualitative Research in Accounting & Management*, 17(1), 82–108.
- Kallio, K. M., Kallio, T. J., & Grossi, G. (2017). Performance measurement in universities: Ambiguities in the use of quality versus quantity in performance indicators. *Public Money & Management*, 37(4), 293–300.
- Kaplan, R. S., & Norton, D. P. (1992). The balanced scorecard – Measures that drive performance. *Harvard Business Review*, 70(1), 71–79.
- Kristensen, J., Nørreklit, H., & Raffnsøe-Møller, M. (Eds.). (2011). *University performance management: The silent managerial revolution at Danish universities*. DJØF.
- Kure, N., Nørreklit, H., & Røge, K. M. (2021). Objective and results-based management of universities: Constructing reality or illusions? *Financial Accountability & Management*, 37(2), 204–230.
- Lapsley, I., & Llewellyn, S. (1995). Real life constructs: The exploration of organisational processes in case studies. *Management Accounting Research*, 6(3), 223–235.
- Lavertu, S. (2016). We all need help: "Big data" and the mismeasure of public administration. *Public Administration Review*, 76(6), 864–872.

- Liboriussen, J. M., Nørreklit, H., & Trenca, M. (2021). A learning method of trust building: Beyond the performance management of artistic events. *Qualitative Research in Accounting & Management*, 18, 516–544.
- Lucas, L. (2006). *The research game in academic life*. McGraw–Hill Education.
- Magna Charta Universitatum, Bologna. 18 September 1988. <http://www.magna-charta.org/magna-charta-universitatum/the-magna-charta-1/the-magna-charta>
- Mari, L. (2007). Measurability. In M. Boumans (Ed.). *Measurement in economics*. Elsevier.
- Mergel, I., Edelman, N., & Haug, N. (2019). Defining digital transformation: Results from expert interviews. *Government Information Quarterly*, 36(4), 101385.
- Micheli, P., & Mari, L. (2014). The theory and practice of performance measurement. *Management Accounting Research*, 25(2), 147–156.
- Nielsen, M. W. (2017). Gender consequences of a national performance-based funding model: New pieces in an old puzzle. *Studies in Higher Education*, 42(6), 1033–1055.
- Northcott, D., & Linacre, S. (2010). Producing spaces for academic discourse: The impact of research assessment exercises and journal quality rankings. *Australian Accounting Review*, 20(1), 38–54.
- Nørreklit, H. (2017a). *A philosophy of management accounting: A pragmatic constructivist approach*. Taylor & Francis.
- Nørreklit, H., Nørreklit, L., & Mitchell, F. (2016). Understanding practice generalisation—Opening the research/practice gap. *Qualitative Research in Accounting & Management*, 13(3), 278–302.
- Nørreklit, H., & Trenca, M. (2021). Performance management in a milieu of customer participatory measurement: Beyond the ratings and rankings of strictly come dancing. *The British Accounting Review*, 53(6), 100873.
- Nørreklit, L. (2017b). Actor–reality construction. In *A philosophy of management accounting* (pp. 23–71). Routledge.
- Nørreklit, L., Pedersen, S. L., Tuft, K., & Prangsgaard, B. (1986). *Aktørsmetoden—en indføring i erhvervsøkonomisk projektarbejde [The actor approach – An introduction to business project work]*. Poseidon, Aalborg.
- Nørreklit, L., Jack, L., & Nørreklit, H. (2019). Moving towards digital governance of university scholars: Instigating a post-truth university culture. *Journal of Management and Governance*, 23, 869–899.
- OECD. (2019). *Education at a Glance*. OECD Publishing.
- Payne, G., & Williams, M. (2005). Generalization in qualitative research. *Sociology*, 39(2), 295–314.
- Parker, L. D. (2022). Public university research engagement contradictions in a commercialisation higher education world. *Financial Accountability & Management*, 1–18. <https://doi.org/10.1111/faam.12341c>
- Parker, L. (2011). University corporatisation: Driving redefinition. *Critical Perspectives on Accounting*, 22(4), 434–450.
- Pianezzi, D., Nørreklit, H., & Cinquini, L. (2020). Academia after virtue? An inquiry into the moral character(s) of academics. *Journal of Business Ethics*, 167(3), 571–588.
- Porter, T. M. (1996). *Trust in numbers*. Princeton University Press.
- Rebora, G., & Turri, M. (2010). Critical factors in the use of evaluation in Italian universities. *Higher Education*, 61(5), 531–544.
- Schmidt, E. K., & Langberg, K. (2007). Academic autonomy in a rapidly changing higher education framework: Academia on the Procrustean bed? *European Education*, 39(4), 80–94.
- Simons, R. (1995). *Levers of control*. Harvard Business School Press.
- Soin, K., & Huber, C. (2021). Compliance and resistance: How performance measures make and unmake universities. *Organization*. <https://doi.org/10.1177/13505084211066810>
- Tandilashvili, N., & Tandilashvili, A. (2022). Academics' perception of identity (re) construction: A value conflict created by performance orientation. *Journal of Management and Governance*, 26(2), 389–416.
- Taylor, F. W. 1916. *The principles of scientific management*. Harper & Brothers.
- Ter Bogt, H. J., & Scapens, R. W. (2012). Performance management in universities: Effects of the transition to more quantitative measurement systems. *European Accounting Review*, 21(3), 451–497.
- Tinker, T. (1991). The accountant as partisan. *Accounting, Organizations and Society*, 16(3), 297–310.
- Vakkuri, J., & Johanson, J. E. (2020). Failed promises—Performance measurement ambiguities in hybrid universities. *Qualitative Research in Accounting & Management*, 17(1), 33–50.
- Wittgenstein, L. (1953). *Philosophical investigations*. Basil Blackwell.
- Wittgenstein, L. (1921). Logisch-philosophische abhandlung. *Annalen der Naturphilosophie*, 14, 185–262.

How to cite this article: Mauro, S. G., Cinquini, L., Malmlose, M., & Nørreklit, H. (2023). University research by the numbers: Epistemic methods of using digitized performance measures and their implications for research practices. *Financial Accountability & Management*, 1–27. <https://doi.org/10.1111/faam.12367>