

The theory and practice of performance measurement

Pietro Micheli, Luca Mari,

ABSTRACT

This paper builds on principles and techniques developed in measurement science, as currently understood in physical sciences and engineering, to improve the theory and practice of performance measurement. To do so, it firstly discusses three fundamental positions on measurement, characterized as metaphysical, representational and relativistic. Subsequently, it lays the foundations of a pragmatic epistemology of measurement in both physical and social sciences. Finally, these insights are integrated through the examination of possible advances in both the theory and practice of performance measurement in organizations.

Keywords: Performance measurement, Measurement science, Epistemology, Performance management.

“Our interest in measurement for driving performance improvements arose from a belief articulated more than a century [ago] by a prominent British scientist, Lord Kelvin (1883): “I often say that when you can measure what you can speak about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind. If you cannot measure it, you cannot improve it.” Norton and I believed that measurement was as fundamental to managers as it was for scientists” (Kaplan, 2009; p. 1253).

1. Introduction

Performance is a notion that permeates contemporary societies, as a means to assess the quality of individual and collective efforts (Corvellec, 1997). In management research, performance is often perceived as encapsulating the unitary purpose of organizations (March and Sutton, 1997). Indeed, organizations are required to ‘perform’ and to communicate their achievements to key stakeholders. As a consequence, organizational functions and processes are increasingly demanded to demonstrate their contribution to performance.

The need to establish links between planning, decisions, actions, and results has generated substantial interest in the measurement of organizational performance. Scholars from management accounting and other areas of management research have examined a wide range of issues related to the design, implementation, use, and review of performance measurement systems (see, for example, Chenhall and Langfield-Smith, 1998; Goold and Quinn, 1990; Hall, 2008; Henri, 2006; Ittner et al., 2003; Neely, 1999). In management practice, organizations have invested considerable resources to measure and demonstrate their performance (Hood et al., 2000; Micheli and Manzoni, 2010). However, there is still no conclusive evidence supporting either the benefits or the shortcomings of introducing performance measurement systems in either private or public sector organizations (Griffith and Neely, 2009; Malina and Selto, 2001; Power, 2004; Townley et al., 2003).

This paper argues that both research and practice in performance measurement (PM) suffer from an underdeveloped conceptualization of the notion of performance measurability, and of the derived measurement processes. While the study of PM has often led to the critique or support of specific frameworks, such as the Balanced Scorecard (Kaplan and Norton, 1992, 2008), in physical sciences and engineering the very concept of measurement has been extensively debated

also at a foundational level (Campbell, 1920; Carnap, 1966; Ellis, 1968; Hempel, 1952). Learning from the insights generated by these debates, we argue that a characterization of measurement in purely functional terms (Joint Committee for Guides in Metrology, 2008a) allows its application to non-physical properties without any reductionist or physicalist implications and, as such, it could inform studies in management research.

By examining PM epistemology, we aim at advancing both the theory and the practice of management in two major ways. First, we draw on fundamental debates on measurement science made in physical sciences and engineering to provide more robust theoretical bases to the study of the measurement of performance in organizations. Through an epistemological analysis of measurement, we re-examine key properties of measurement (e.g., objectivity, accuracy and precision), and on this basis we discuss and compare the implications of adopting a criterion of adequacy, as opposed to a criterion of truth, and of a model-based view, as opposed to a truth-based view, for characterizing measurement and its results. Second, we explain and help address several issues that have emerged in PM related studies, for example on the links between PM and strategy; on the roles of PM in organizations; and on the importance of reviewing PM systems.

We start by providing an overview of current debates on the theory and practice of PM. We then introduce some epistemological analyses on measurement in the physical and social sciences. Subsequently, we present an outline of a conceptual history of measurement, as mainly developed in physical sciences and engineering, by comparing three main paradigms: metaphysical, representational, and relativistic. This discussion leads to the examination of PM as a fundamentally epistemic and pragmatic act, rather than as the determination of the ‘true value’ of organizational performance. We conclude by discussing several implications of this standpoint for measurement of performance in organizations.

2. Performance measurement: benefits, limitations and shortcomings

Studies in performance measurement have often focused on procedures and tools that could improve the efficiency and the effectiveness of organizations (Franco-Santos et al., 2007; Kaplan and Norton, 1992). Several researchers have demonstrated that appropriate measurement and management of performance can facilitate:

- formulation, execution and review of organizational strategy (Ahn, 2001; Chenhall, 2005; Euske et al., 1993; Gimbert et al., 2010; Govindarajan and Gupta, 1985);
- communication of the results achieved to stakeholders, and strengthening of brand and reputation (Atkinson et al., 1997; McKevitt and Lawton, 1996; Neely et al., 2002; Smith, 1995a);
- strategic alignment, motivation of employees at all levels, creation of a performance improvement culture, and organizational learning (Gittell, 2000; Henri, 2006; Kolehmainen, 2010; Malina and Selto, 2001; Roos and Roos, 1997).

Several empirical studies have shown how PM can be generally productive and help improve organizational performance (Davis and Albright, 2004; Franco et al., 2012; Hoque and James, 2000; Ittner et al., 2003; Speckbacher et al., 2003). However, despite considerable resources invested (Neely et al., 2006), PM related initiatives can often fail to deliver on their promises (Neely and Bourne, 2000). Furthermore, if done poorly, they can be not only ineffective, but harmful and indeed destructive (Perera et al., 1997; Royal Statistical Society, 2005). Therefore, it is crucial to understand under which specific conditions performance measurement and management practices can actually enable performance improvements.

So-called 'alternative approaches' have looked at PM, considering it more as a social practice

rather than as a technical process (Covaleski et al., 1996). In this context, the need for deeper reflections on the conceptual and operative conditions required for the measurement of performance has been advocated (Chua and Degeling, 1993). Indeed, a number of scholars have remarked that PM is often regarded as the objective evaluation of reality by academics and practitioners (Morgan, 1988; Power, 1997). The customary use of adages such as ‘if you cannot measure it, you cannot manage it’¹ and ‘what gets measured gets done’ (Garvin, 1993; Johnson and Kaplan, 1987; Kaplan and Norton, 1992; Osborne and Gabler, 2002; Peters and Waterman, 1982) sometimes expresses not only the acknowledged importance of PM in organizations, but also, and far more generally and critically, a (usually unjustified) belief about the epistemic role of measurement, i.e., its capacity to acquire and communicate knowledge. These praising adages have two major implications: first, they suggest that measurement is a necessary condition for management and a sufficient condition for aligned attitudes and behaviors, whereas this is not necessarily the case in organizations (Kennerley and Mason, 2008; Pollitt, 2006; Smith, 1995b). Second, they are silent regarding all the key properties of measurement (e.g., objectivity, accuracy, and precision), almost as if they were unproblematic and could be taken for granted. A basic aim of this paper is to argue against the underpinnings of such a position and to challenge the widespread, albeit often implicit, view that PM could enable organizations to

¹

This cliché is easily falsifiable by acknowledging that, instead, ‘you can control (and, therefore, manage) even what you do not measure’. As a simple example, let us consider a feedback control system, such as a thermostat or a car cruise control, in which the set point is fixed according to the strategy ‘maintain this state’ (i.e., the current temperature, the current speed), whatever such state is and even without knowing its representation as a quantity value. Since feedback operates so as to minimize error, i.e., the difference between desired and current state, all is required in this case is an acquisition device stable enough to make such comparison reliable, even independently of the availability of any formalized information on the controlled quantity, and therefore of any measurement. The most explicit indication that this is ‘control without measurement’ is that a control system exploited in this way does not require any calibration for its proper usage, nor does it produce any information on the controlled system.

determine the ‘true value’² of their performance, and, as a consequence, that ‘perfect measurement’ leads to improved performance. To this goal, we analyze the epistemological bases of PM, a topic on which, interestingly, little has been written about. To do so, we build on debates on measurement in physics and engineering, where key aspects of the measurement process have been extensively discussed. These lead us to propose a position of pragmatic moderate relativism, which supports the view that the trade-off between acceptable quality and available resources - rather than ‘perfect’ objectivity, precision and accuracy - is the most general criterion for ‘good’ measurement.

3. Differences between sciences: epistemological preliminaries

While this paper is not aimed at investigating the complex topic of the (possible) methodological characterization of different scientific disciplines, it is important to consider the implications that different modes of explanation and theorization have on the theory and practice of PM. Indeed, important differences between sciences can be recognized in the process of formation of theories, in the way they are applied, and how such a use is deemed to affect the behavior of theorized systems (Ghoshal and Moran, 1996). The application of management theories, unlike theories in sciences of inorganic or organic matter, has generally a considerable impact on the object they relate to, e.g., the application of theories such as agency theory significantly affects the ways in which organizations operate. As organizations are adaptive systems, in the social sciences

² The term ‘value’ is unfortunately polysemic: it can be used with an axiological connotation, thus referring to the desirability of an entity in a context (“this object is of (high / low) value”), but also with a purely formal meaning, particularly to designate the result of the application of a function to an argument (if the function f is applied to the argument x , then $y = f(x)$ is the value of f for x), where indeed such an application is also called an ‘evaluation’. In this paper we will only deal with this latter meaning, according to the common modeling of properties as functions, which map objects (“having” such properties) to property values. Consequently, for example 1.23 m might be the value of the property length as applied / considered for a given object. Similarly, we could say that 98% is the value of a property, i.e., performance, as calculated for the logistics function of an organization (measured, for example, as percentage of on-time delivery of products).

theories tend to be self-fulfilling, whereas in physical sciences they clearly do not (Gergen, 1973). For example, the more a theory is based on strong assumptions about human self-interest and opportunism, as is the case of agency theory, the more its application will induce exactly those types of behaviors expected from the theory itself (Ghoshal and Moran, 1996). Indeed, attitudes and behaviors are both affected by “individual dispositions as well as by the situation that shapes the individual’s perceptions and instrumentalities” (Pfeffer and Sutton, 2006; p. 50). Therefore, in a principal – agent situation, there is a high risk that “opportunistic behavior will increase with sanctions and incentives imposed to curtail it, thus creating the need for even stronger and more elaborate sanctions and incentives” (Ghoshal and Moran, 1996; p. 14). In other words, the assumption that agents are self-interested and behave opportunistically would not only lead to taking measures to tackle such behavior, but it would also reinforce agents’ self-interest and opportunism. Also, considering organizations as bundles of employment contracts can make managers focus on controlling opportunistic behavior so much that they are distracted from the main task of running a business (see, e.g., Beer and Cannon, 2004; Pfeffer, 1998).

On the other hand, despite the existence of significantly different explanation modes, processes of theorization and epistemic roles of the human agent, many management scholars have adopted the “‘scientific’ approach of trying to discover patterns and laws, and have replaced all notions of human intentionality with a firm belief in causal determinism for explaining all aspects of corporate performance” (Ghoshal, 2005; p. 77). Such a concept of ‘scientific’ approach clearly refers to the ‘hard’ sciences, as conceptualized and progressively developed from the Galilean-Newtonian interpretation of the role of experiments in knowledge acquisition and conceptualization. In fact, despite being a social science, management often adopts modes of explanation and theorization that are derived from physics, in which the object of the study is

supposed not to be affected by the theory being adopted. As Ghoshal (2005; p. 79) notes, “management theories at present are overwhelmingly causal or functional in their modes of explanation”. In this regard, in his Nobel Memorial Lecture, Friedrich von Hayek stated that economics, like other social sciences, is subject to the so-called ‘physics envy’, which leads researchers to draw inappropriate conclusions and to forcefully adopt methodologies and methods drawn from physical sciences (von Hayek, 1989), mostly because of physics’ predictive abilities. Von Hayek referred to this issue as ‘scientistic error’.

Reflections on the differences between sciences have relevant implications on the epistemology of PM. A realist view assumes that the external world exists independently of us and can be accessed objectively, and that observations can be reported by means of a subject-neutral and theory-neutral language. It also supports the formulation of general laws for ‘closed systems’, where environmental influences can be controlled, or even completely eliminated (Tsoukas, 1989). A subjectivist, or, more generally, a non-objectivist, view denies the possibility of such an epistemological foundation, emphasizing the critical, constructive role of human agents, possibly with reference to the interactions with an independent external reality, which can constrain or facilitate human action (Johnson and Duberley, 2000).³

Social sciences entail operative interventions in social life and value judgments by investigators (Bhaskar, 1978): hence, instead of being primarily aimed at providing ‘true descriptions’ of the external world, their theories can be considered “a form of insight, i.e., a way of looking at the world, and not a form of knowledge of how the world is” (Bohm, 1980; p. 4). The polysemy of the fundamental concept of ‘explanation’ is a meaningful indicator in this regard: while in

³ On this point, see also Goodman’s (1978; p. 7) comment: “With false hope of a firm foundation gone, with the world displaced by worlds that are but versions, with substance dissolved into function, and with the given acknowledged as taken, we face the questions how worlds are made, tested, and known.”

physical sciences to explain mainly means to exhibit a law from which the explanandum can be derived by deduction – the so-called *nomological* explanation –, in social sciences to explain mainly means to identify a goal able to justify the explanandum as a consequence – the so-called *teleological* explanation (Nagel, 1961).^{4,5}

The adoption of an epistemological stance according to which theories can provide insight, rather than objective knowledge, has relevant implications for the measurement of organizational performance. First of all, PM will have to be regarded as a means to gather and analyze data in relation to organizational performance, rather than as a way to access (actual or potential) ‘true knowledge’ of how organizations perform. Second, our understanding of organizational performance will inevitably be limited. These implications, in turn, could have substantial repercussions on the design, implementation and use of PM systems. Since these claims call into question the very concept and possibility of measuring performance, the following section discusses the epistemology, and in some degree the ontology, of measurement in more depth. To do so, recent developments in measurement science, as currently understood in physical sciences and engineering, are analyzed and subsequently implications for PM are discussed.

4. The epistemology of measurement

Management scholars have emphasized how, according to the positivist paradigm, which

⁴ The historical development of ‘hard’ sciences such as physics and biology can be regarded as a progressive shift in explanation style, from teleological to nomological (e.g., why do heavy bodies fall when they are free to move above the earth surface? from: because they aim at reaching their natural place; to: because they are attracted by the gravitational field generated by the earth). Noteworthy is the fact that nomological explanations are predictive (their applicability conditions are generally well identified, so that explanations can be deductive), whereas teleological explanations are not (they refer to ‘intentions’, thus preventing repeatability and the possibility of generalization).

⁵ Although the primary focus of this paper is not on the long-standing debate over the nature of social sciences, it should be noted how the paradigm of the ‘design sciences’, inspired by Simon (1996), has received considerable attention in recent years (Bate, 2007; Denyer et al., 2008; Huff et al., 2006; Romme, 2003).

maintains an objectivist epistemological perspective, many studies in management are based on the following assumptions:

- “Variables other than the ones whose variation we would like to observe are perfectly controlled for.”⁶
- The empirical scales measure the constructs completely.
- There is no cognitive disagreement among social agents about the definition of the situation.
- Social agents have an ability to reflect, that is, they can think about their own thinking and that of others” (Numagami, 1998; p. 4-5).

Interestingly, these assumptions are aligned with traditional principles in physical science, for which measurement is “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” (Finkelstein, 2003; p. 41). On the other hand, we argue that, not only are the assumptions outlined by Numagami examples of ‘physics envy’ and ‘scientific error’, but also current debates in measurement science are questioning the validity of those very principles in physical sciences and engineering. With the aim of clarifying these issues and providing an alternative epistemological perspective on PM, an overview of the conceptual history of measurement of physical quantities is now provided.

4.1 Measurement: history and characterizations

The etymology of the Greek word for ‘measure’ provides an interesting insight into the concept

⁶ This assumption is often referred to as *ceteris paribus* hypothesis. In physics it has traditionally played a critical role to justify the superposition principle, for which, e.g., the problem of computing the state of a system of n interacting bodies, that in general is mathematically intractable for even low values of n , is reduced to the much simpler problem of computing the state of $n(n-1)/2$ systems of 2 bodies. Systems for which such a reduction leads to inadequate results are paradigmatically considered complex.

and history of measurement science (Mari, 2003). Indeed, two general meanings can be retraced for it: ‘subjective evaluation’, as in the well-known quotation attributed to Protagoras (Plato, *Theaetetus*, 152a), “man is the measure of all things, of the existence of things that are, and of the non-existence of things that are not”, and ‘objective description’, in particular because of the Euclidean systematization of geometry, according to which, “a magnitude is a part of a magnitude, the less of the greater, when it measures the greater; the greater is a multiple of the less when it is measured by the less; a ratio is a sort of relation in respect of size between two magnitudes of the same kind” (*Elements*, Book V, definitions 1–3). The common ground of these conceptions is that a measure has something to do with a comparison with a pre-identified entity, man according to Protagoras, a magnitude according to Euclid. On the other hand, consistent with the Greek philosophical attitude against empirical activities, neither of them emphasizes the experimental side of measurement. The path towards the current understanding(s) of the concept has been complex and everything but linear. Three main paradigmatic standpoints, also interpretable as approximately subsequent phases of a conceptual history of measurement, could be defined (Mari, 2003):

1. metaphysical
2. representational
3. relativistic.

While the authors of this paper do not endorse any of these standpoints, for the aims of this paper, it is important to review their main tenets and conclusions. The first position was expressed by the founders of physical sciences, such as Pythagoras and Aristotle on the one hand, and Galileo and Kepler on the other, for whom numbers were part of the empirical world, so that measures were inherent properties of the objects being measured (“numbers are in the

world”, as Kepler wrote (Letter to Micheal Maestlin, 1595)). According to them, phenomena were intrinsically quantitative, and measurement operated to determine pre-existing properties, i.e., as a discovery process of ‘true values’ of quantities (Mari, 1997). The classical distinction between ‘primary’, i.e., ‘intrinsic’, and ‘secondary’, i.e., subject-related, properties also clearly emphasizes the ontological assumptions underlying this position. As late as 1899, reliance on the fundamental possibility and power of measurement led A.A. Michelson⁷ to state that since “the more important fundamental laws and facts of physical science have all been discovered [...] our future discoveries must be looked for in the sixth place of decimals” (quoted in Richtmyer et al., 1969; p. 43), thus emphasizing the role of precision, as obtained by measurement, in the search of ‘truer and truer’ values.

During the 19th century and the first decades of the 20th century the effectiveness of measurement in physical sciences and engineering acted as a trigger to export that paradigmatic knowledge acquisition process to different disciplines, with the aim of increasing the objectivity in the evaluation of properties of psychophysical, behavioral, and social phenomena (Rossi, 2007). In such a broader context, the very concept of true value, usually assumed a real number (with a measurement unit), was generally deemed to be unmaintainable, and the focus shifted towards the measurability of ‘secondary’ properties. This led to the endorsement of the representational standpoint, an a-metaphysical – and sometimes anti-metaphysical, in some neo-positivistic interpretations – perspective, according to which measurement results cease to be true expressions of reality, and they are thought of instead as representation means, i.e., results of operations that preserve the relations observed among entities. Through the first representation theorems (Helmholtz, 1887; Hölder, 1901) and the theory of scale types (Stevens, 1946),

⁷ Michelson was awarded the Nobel Prize in physics in 1907.

measurement was then formalized as a (homo)morphism (Krantz et al., 1971, 1989, 1990), i.e., a function preserving the relations defined on its domain, thus emphasizing the internal consistency of representation, sometimes also characterized as ‘empirical meaningfulness’ (Roberts, 1979), instead of its truth.⁸ It is indeed this consistency constraint that guarantees measurement to be adequate for its goal of representing information on measured entities. Furthermore, this stresses the importance, if not the unavoidability, of interpretive models in acquiring knowledge about reality.

The move from the metaphysical to the representational position is crucial, as it implies a sharp shift of emphasis, from ontology (measurement as a means to know how reality is) to epistemology (measurement as a means to acquire and consistently express information on reality). As this position maintains, measurement only claims to act as a representational tool, a bridge between reality, to which the object under measurement belongs, and the linguistic/symbolic realm, to which the measurement result belongs: measurement relates to the available knowledge on the state of an object. While still requiring the information produced through measurement to be non-ambiguous and inter-subjective (i.e., to be interpretable in the same way by different subjects), and measurement results to convey information solely about the measured property (Cecconi et al., 2006), from a representational point of view the very constraints for measurement to be objective and empirical assume a new connotation. By objectivity is meant here that “the numbers assigned to a property by measurement must, within

⁸ A simple example of such consistency constraints is as follows. Let us suppose that: (i) an empirical procedure is available to assess, for any two organizations x and y , which of them has achieved a better performance, in symbols $x <_{\text{perf}} y$, or that they have the same performance, $x =_{\text{perf}} y$; (ii) such comparisons fulfill some usual conditions ($=_{\text{perf}}$ is reflexive and symmetric, $<_{\text{perf}}$ is anti-symmetric, both $=_{\text{perf}}$ and $<_{\text{perf}}$ are transitive, ...). Then according to the representational standpoint any mapping f from x, y, \dots to, e.g., natural numbers, which is a homomorphism (i.e., if $x <_{\text{perf}} y$ then $f(x) < f(y)$ and if $x =_{\text{perf}} y$ then $f(x) = f(y)$) formalizes a measurement. It should be noted that nothing prevents here the empirical assessments $x <_{\text{perf}} y$, or $x =_{\text{perf}} y$ to be subjective.

the limits of error, be independent of the observer” (Finkelstein, 2003; p. 41). Empiricity means that measurement “must be the result of observation and not, for example, of a thought experiment”; this is because “the concept of the property measured must be based on empirically determinable relations and not, say, on convention” (Finkelstein, 2003; p. 41). This emphasis can be understood by remarking that nothing in the representational formalization implies (homo)morphisms to describe objective and empirical processes (the results of a ‘thought experiment’, or just a subjective guess, can be in fact represented in (homo)morphic terms), so that being objective and empirical are independent constraints and, as such, they must be explicitly introduced (and indeed if what prescribed by the representational standpoint is assumed as a sufficient, and not only necessary, condition of measurability, then completely subjective estimations are measurements as long as they are consistent in the representational sense).

The further move towards the third, relativistic, position is also significant, as it entails an even more explicit emphasis on a model-based view, as symbolized in the shift from error to uncertainty (Joint Committee of Guides in Metrology, 2008b). The concept of error, as the expression of the discrepancy from a given reference, has been traditionally associated to the empirical inability in the measurement process of determining the true value of the property of an object. Rather, uncertainty is a “non-negative parameter characterizing the dispersion of the quantity values being attributed to a measurand, based on the information used” (Joint Committee of Guides in Metrology, 2008a). In this context, ‘measurand’ is defined as the “quantity intended to be measured” (Joint Committee of Guides in Metrology, 2008a), where this ‘intention’ further stresses the dependence on models. This characterization emphasizes that measurement is a knowledge-based assignment process, and not the empirical determination of a

pre-existing property value. Accordingly, quantification of phenomena is considered to be possible only within a given, although sometimes implicit, model, so that the notions themselves of ‘true value of a measurand’ and of its determination become meaningless. In order to explain the substantial implications of this claim, the next section discusses the relativistic paradigm in more detail, also pointing out some ongoing further developments.

4.2 The relativistic paradigm and beyond

From a relativistic perspective the concept of model becomes of primary importance. In particular, a model-based view takes into account pragmatic components, which are not considered in the two previous paradigms, and regards measurement as a process, which aims at attaining ‘adequate-to-purpose’, rather than ‘true’, results.⁹ Replacing the criterion of truth with a criterion of adequacy has profound consequences, since, according to the latter position, decisions over cost and quality of measurement become essential components of the measurement process itself. As a paradigmatic case, in this context the very concept of ‘absolute precision’ for measurement results is acknowledged to be meaningless, and the search for their increasing precision is considered relevant only as a trade-off with measurement resources, and therefore, in particular, costs. Similarly, the concept of accuracy could be re-conceptualized as the degree of trust that can be attributed to a value, as obtained by a suitable characterization of the measurement process, rather than its correspondence (its ‘closeness of agreement’, according to the Joint Committee for Guides in Metrology (2008a)) to the ‘true value’. If measurement is regarded as an assignment, rather than a determination, any measurement result could only

⁹ The concept of truth is a complex and controversial one, as the coexistence of several distinct “theories of truth” witnesses (see, e.g., Glanzberg, 2009). Even in the specific context of measurement science a significant transition is in progress, where ‘true value’, once referring to the value produced by an “ideal” measurement system operating in “ideal” conditions, is now defined as “quantity value consistent with the definition of a quantity” (Joint Committee for Guides in Metrology, 2008a).

generate information that is meaningful in the context of the model within which measurement is being carried out. Therefore, “measurement results must be assigned (and not determined) according to the goals for which the measurement is performed, with the consequence that they are adequate if they meet such goals” (Mari, 2007; p. 76).

From a relativistic standpoint, objectivity and inter-subjectivity, features already taken into account by the representational position (Finkelstein, 2003), are still considered to be target points for measurement. However, the traditional acknowledgment that experimental issues (e.g., environmental noise and instrumental instabilities) make it impossible to fully reach these conditions is complemented with the understanding that the unavoidable presence of models are also issues that make the measurement process and its results not ‘ideal’. This particularly applies to the problem of measurand definition. Once a neglected issue, under the implicit assumption that properties are given and as such do not need to be defined, it was particularly thanks to Bridgman (1927) that the importance of defining even classical physical quantities was affirmed in order to prevent ambiguities and misinterpretations. This model dependence is what in the social sciences (and psychometrics in particular) is customarily acknowledged in terms of the problem of construct validity, and it is the basis of current research activities aimed at giving a formal ground to the concept of definitional uncertainty, the “component of measurement uncertainty resulting from the finite amount of detail in the definition of a measurand” (Joint Committee for Guides in Metrology, 2008a). As a consequence, such experimental and conceptual issues are interpreted as critical elements in the general trade-off with resources devoted to perform measurement. According to the relativistic view, measurement itself can be characterized as a process able to acquire and formally express information that is ‘objective and inter-subjective enough’ for the given purposes.

On the basis of the previous discussion, the main features of the relativistic paradigm can be summarized as follows:

- Measurement is not a determination, but an assignment. Measurement results only generate information that is meaningful in the context of the model within which measurement is undertaken.
- Measurement results are informational, and not empirical, entities.
- The measurability of a property conceptually depends on the current state of knowledge of the property; therefore, it is not an ‘intrinsic characteristic’ of the property itself.
- The measurability of a property in practice depends on the availability of experimental conditions; therefore it cannot be derived from formal requirements.
- In a measurement process there is a general qualitative trade-off between specificity and trust, i.e., between precision and accuracy; both aspects depend on the available experimental knowledge of the measurand.
- Objectivity and inter-subjectivity are desirable, but not necessary, features of measurement.

While it is highly plausible that the relativistic standpoint described so far has been triggered by the dominant epistemological relativism of the last fifty years (see, e.g., Feyerabend, 1975; Hanson, 1958; Kuhn, 1970), it is questionable whether radical relativism (‘everything goes’, as in Feyerabend’s paradigmatic statement) could be applied to measurement. If the neo-positivistic concept of ‘protocol of truth’ can hardly be maintained today either in physical science or in accounting (Bierman, 1963), measurement has however critical ties to truth that any oversimplification such as ‘truth does not exist’ does not adequately express. On one hand, the dependence of measurement results on models and goals, and therefore the operative

meaninglessness of the very concept of an ‘absolute’ true value for measurands, can be taken for granted¹⁰. Furthermore, the fact that measurement results include some conventional information can be easily ascertained, given their structure of references to (conventional, indeed) measurement units or scales. On the other hand, such references have an empirical grounding, expressed by the requirement for measurement systems to be calibrated with respect to a measurement standard, i.e., be metrologically traceable to it. Hence, if measurement is defined as the process performed by suitably calibrated measurement systems, and, thus, if such systems and their calibration are in their turn properly characterized, then both the dependence of measurement results on models, i.e., their ‘intrinsic uncertainty’, and their foundational role of ‘truth approximations’ can be justified. According to this position of pragmatic moderate relativism, the acknowledged pivotal role of measurement as information acquisition process can be justified once again, although on new bases: set aside any metaphysical assumption, it is precisely on the functional structure of the process that an epistemology of measurement can be developed today. Measurement systems in both physical and social sciences should be designed taking into account a principle common to all production processes: while quality of products – in this case, measurement results – can be at least comparatively assessed and is the justifying condition to perform the process, the trade-off between acceptable quality and available resources remains the most general criterion for ‘good’ measurement.

¹⁰ An excellent example in this regard is given by the introduction, in the third edition of the International Vocabulary of Metrology (Joint Committee for Guides in Metrology, 2008a), of the concept of target measurement uncertainty, defined as “measurement uncertainty specified as an upper limit and decided on the basis of the intended use of measurement results”. These explicit references to ‘decisions’ and ‘intentions of use’ unambiguously highlight that target uncertainty is, in fact, a pragmatic specification, namely, the epistemic counterpart of tolerance, i.e., a parameter widely used in many industrial contexts to specify the ‘acceptable technical quality’ of products.

5. Implications for performance measurement

In the wake of the modern conception of science, Galileo argued that the world is written in mathematical terms. As a consequence, the task of scientists was deemed to count what is countable, measure what is measurable, and what is not measurable, make measurable. Over the last twenty years, organizations have invested increasing amounts of resources in the design, implementation, and use of performance measurement (PM) systems such as the Balanced Scorecard following a similar, albeit implicit, principle that measurement is a necessary activity for organizations to be successful. Also, the assumption has been that the more objective, accurate and precise the measurement, the higher its impact would be on organizational performance. This paper argues against this standpoint, and proposes an alternative position, according to which measurement should aim to produce ‘adequate-to-purpose’, rather than ‘true’ results. Also, the endorsement of such position could help us improve both the theory and the practice of performance measurement, thus strengthening the link between PM and organizational performance, which several empirical studies are increasingly questioning (see Franco et al., 2012 for a review).

Since the beginning of the 20th century, the metaphysical position expressed by Kepler and Galileo was challenged in a number of ways. As we have argued above, this has led to the establishment of two other perspectives on measurement – representational and relativistic –, which reflect our greater understanding of the powers and limitations of measurement itself. However, these advancements seem to have had little impact on PM research and practice. Indeed, PM practices are often driven by the (either implicit or explicit) belief expressed in the initial quote by Robert Kaplan that measurement is a necessary condition for both decision-making and improvement (Garvin, 1993) – an argument which is aligned to the outdated

metaphysical paradigm. This is supported by evidence that, in recent years, organizations in both private and public sectors have been investing increasing amounts of resources to obtain ‘true descriptions’ of their performance through the use of PM systems (Micheli and Manzoni, 2010). As several empirical studies have shown, however, the introduction of several PM initiatives in both private and public sector organisations has often had little or no impact on either decision-making or performance (see, e.g., Cavalluzzo and Ittner, 2004; Griffith and Neely, 2009; Poister, 2010).

From the analysis presented above we argue that PM suffers from the lack of solid theoretical foundations: a more rigorous examination of the epistemological issues of measurement processes could contribute not only to the development of PM from a conceptual standpoint, but also to its practice. First of all, even though social scientists have referred to ‘physics envy’ (von Hayek, 1989) to indicate the introduction and use of approaches developed in the physical science to study social phenomena, in the case of measurement the authors believe that, while debates in physics and engineering could inform management theory and practice, they seem not to have been sufficiently taken into account. More specifically, from an epistemological perspective, the main tenets of the pragmatic perspective, which is being developed within the relativistic standpoint, should inform current debates in PM; also, more clarity is required regarding the philosophical underpinnings of measurement in social sciences.

Secondly, it is important to remark that the process of measurement can relate, in principle, to three types of objects: physical, which exist in space and time; ideal, which exist outside space and time; and social, which have a beginning in time and are socially constructed (Ferraris, 2005; Husserl, 1907; Meinong, 1904). Measurement science in physics and engineering has traditionally focused on physical entities, such as phenomena, bodies, and substances (Joint

Committee for Guides in Metrology, 2008a). PM is most often concerned with properties of social objects (from stakeholders' satisfaction to brand management, from intellectual capital to patients' experience), which are often complex and difficult to define and measure. Therefore, it could be argued that in PM, even more than in physical sciences, complete empiricity and objectivity must be considered as goals and not as preliminary, i.e., necessary, conditions. PM should be seen as an assignment, rather than an empirical determination; the relevant influence of implicit and explicit models acknowledged; and the related presence of errors and uncertainties admitted and properly dealt with. Although tools and techniques could become more sophisticated, thus reducing errors and uncertainties in measurement, it will never be possible to reach absolute certainty on the value(s) to assign to what is being measured.

Thirdly, scholars have observed that performance in organizations is often reduced in its meaning and scope to what is technically easier to measure of performance itself (Smith, 1995b), thus following an apparently operational standpoint (Bridgman, 1927). The main implication is that a performance indicator is often equaled to the simplest, and most easily measurable, aspect of the activity or process being performed, e.g., number of complaints expressed by employees = employees' satisfaction; impact factor of publications = academic standing. This issue is particularly relevant and problematic, if we consider the great difficulties encountered in – and, frequently, the impossibility of carrying out – the measurement of properties of very complex activities and processes. Unfortunately, the consequence is that in the social sciences what is treated as important is what happens to be accessible to measurement (von Hayek, 1989). Furthermore, this claim confirms that the ontological and the epistemological dimensions are often conflated, thus reducing reality to what it is easy to access of reality itself.

Fourthly, to measure does not necessarily mean to associate empirical objects with numbers, but,

more generally, with information entities (Mari, 2007). This distinction is particularly relevant from a PM point of view, as the measurement of performance – both in practice and in theory – is most frequently equated to the mere association of numbers to empirical objects. This position is problematic for two reasons. First of all, measurement is an evaluation performed by means of a measurement system, and it is only such a system and its proper setup and usage that guarantee, up to a degree adequate to the given purposes, the objectivity and the intersubjectivity of measurement results. Therefore, it is not the assignment of numbers, i.e., quantification, but the adequacy of the measurement system that makes an evaluation a measurement. Secondly, the presumed ‘inherent objectivity’ of numbers often leads to little consideration given to supposedly ‘subjective’ qualitative indicators, in favor of substantial reliance on quantitative ‘measures’ of performance, which, in fact, may be equally problematic and even less informative (Dossi and Patelli, 2010).

6. Towards a pragmatic view of performance measurement

In physical sciences and engineering the metaphysical paradigm has been largely superseded; while the representational paradigm has not found a wide support plausibly because it is deemed too an abstract approach (Mari, 2000), the role of models in measurement, and therefore the relativistic paradigm moderated by the pragmatic instances analyzed above, is nowadays explicitly emphasized. The adoption of a pragmatic perspective in PM research and practice requires a re-conceptualization. First of all, the adoption of a model-based view, as opposed to a truth-based view, has substantial implications on the measurement process and on the interpretation of its results. From a model-based point of view, measurement is regarded as a

knowledge-based process, rather than a purely empirical determination. The object whose property is measured is assumed to exist in the empirical world, but it is acknowledged that the data collected about that object results from an interpretation process, i.e., an explicit or implicit model is involved, which belongs to the informational realm. As a consequence, the measurement procedure must be defined, and the object under measurement designed and set up by considering the context and the goals for which the measurement itself is being undertaken. This means that, for example, the reasons and purposes of introducing a performance indicator should be made explicit and agreed before the indicator is used. Furthermore, particular care has to be exerted when utilizing any data or information in a context other than that in which it was meant, as this has substantial implications on the drivers, purposes and uses of PM (Behn, 2003; Hatry, 1999). This is particularly significant in benchmarking exercises or compilation of league tables in both private and public sector contexts (Ammons, 1999; Goldstein and Spiegelhalter, 1996). Equally, from a model-based point of view, it would be nonsensical to state that a performance indicator is either 'good' or 'bad' in absolute terms. Rather, on the bases of its goals and other relevant factors (e.g., cost, quality), an indicator could be deemed adequate or inadequate-to-purpose. For example, an indicator that considers the average time an organization spends to produce a quote might be appropriate to monitor the organization's responsiveness to customers. This does not equate to saying that the average difference between date of verbal confirmation of receipt of quote by customers and the date of first contact by customers is the 'right measure' for responsiveness to customers. In fact, the calculation of variance in performance could also be informative. However, a simple average might be appropriate for monitoring purposes, whereas richer information would be necessary to support process improvement.

Aims such as those of accuracy, repeatability and objectivity in measurement emphasize the necessity to define the properties being measured, and to question and consider the influence of the measurer, and more generally the environment, on the measurement results. Although most organizations utilize measures of ‘absenteeism’, for example, the ways in which such property is effectively defined and measured can substantially vary (Hausknecht et al., 2008). Moreover, information on absenteeism would be influenced by the person undertaking the measurement, i.e., differences in responses are likely to emerge if measurement is performed either by management or by an independent agency. Reasons for measurement (e.g., understanding reasons for absence, launching a new program to motivate employees, or simply reducing absenteeism) would also influence how measurement is undertaken, how results are analyzed and communicated, and which subsequent management actions might follow. For example, if management mainly wants to understand reasons for absence, then it could start by gathering and analyzing data at both individual and unit-level; subsequently it could collect good quality data not only on employee satisfaction and commitment, but also over potential job alternatives in the areas in which it operates;¹¹ such datasets would then have to be reviewed in conjunction at board level, at least initially. If, instead, management wants to reduce absenteeism, it could target specific units where absenteeism has traditionally been high; link data over absence to current initiatives over job enrichment and rewards, for example; and review the data with specific individuals during their performance appraisals (see also, Hausknecht et al., 2008).

The relevance and unavoidability of interpretive models in the measurement process also highlight the importance of surfacing mental models in the design and use of PM systems. If

¹¹ In areas with low unemployment workers tend to be less concerned about absence from work, especially when they do not feel committed and/or are dissatisfied with their organizations (Hausknecht et al., 2008; Larson and Fukami, 1985; Markham, 1985).

mental models are considered as “deeply ingrained assumptions, generalizations, or even pictures or images that influence how we understand the world and how we take action” (Senge, 2006; p. 8), it is clear how PM could be both influenced by, and used as a way to challenge, mental models (see, e.g., Hall, 2011). This is particularly relevant, since long-term success often depends on the process through which management teams modify and improve the shared mental models of their organizations, their markets, and their competitors (Argyris, 1977; De Geus, 1999; Senge, 2006).

The consideration of the epistemic role of the measurer also has significant consequences in relation to organizational learning and review of PM systems. As Bohm (1980; p. 23) argued, “the attempt to suppose that measures exist prior to man and independently of him leads [...] to the ‘objectification’ of man’s insight, so that it becomes rigidified and unable to change, eventually bringing about fragmentation and general confusion”. If confusion exists between measured properties and measurement results, in such a way that measurement is seen as capturing the essence of objects, PM systems will inevitably be static. Research has demonstrated how the use of excessively rigid PM systems could lead to organizational paralysis, or ‘ossification’ (Smith, 1995b; Townley et al., 2003). Indeed, after investing substantial resources to design and implement large sets of performance indicators, organizations often decide not to modify them, also because they are perceived as ‘perfect representations’ of performance. On the contrary, only by analyzing the data gathered through the system and, in particular, reformulating the PM system itself, can organizations improve their performance.

Such an improvement could only be achieved if single and double loops of learning are generated. Indeed, “double-loop changes cannot occur without unfreezing the models of organizational structures and processes now in good currency” (Argyris, 1992; p. 11). Therefore,

reviews of PM systems must happen through an in-depth comparison between what is measured of the activities performed, and which activities really occur, since measurement is related to the knowledge about the state of an object, rather than the knowledge about the object 'in itself'. Consequently, reviews could be used to challenge not only the current PM system, but also the organization's strategy and its implementation (on the roles of PM systems in strategy formulation and implementation see, for example, Chenhall, 2005; Gimbert et al., 2010; Simons, 1990).

Replacing, or at least complementing, the criterion of truth with a criterion of adequacy implies that cost and quality of measurement should be considered relevant components of the measurement process, and, therefore, assessed both before and after measurement takes place. As a consequence, the introduction of PM systems must be considered as an investment from which to expect a certain return, rather than either an inherently fruitful use of resources (Kaplan and Norton, 1992) or a 'necessary evil' (Brignall and Modell, 2000). Error / uncertainty should be considered in relation to the empirical ability of obtaining appropriate information on the intended property: hence, specificity and trust also become essential features of PM, as performance could be measured with great accuracy, but precision can be misleading, as indicators can be precisely wrong (Mari, 2007).

As the etymology of "measurement" suggests, PM systems should be proportionate, i.e., they should consist of an adequate number of indicators, which can inform decision-making processes, rather than aim at providing 'complete and true representations' of performance.

Finally, the focus of measurement should shift from what is measurable, which is a prevalently epistemological act, to the objects that we want to measure, i.e., the actual organizational processes and activities being performed. Moreover, indicators should not be considered as exact

pictures of reality or as unveiling presumed truths. On the contrary, they ought to be used as ways to gather information about organizational performance that is as adequate as possible. Several authors have suggested ways to tackle the inherent incompleteness of PM systems (Chapman, 1997; Lillis, 2002; Wouters and Wilderom, 2008). Although more appropriate designs and implementations could certainly contribute to this aim, it is also important to fully acknowledge the limitations of measurement. In particular, when measuring performance the last word has to go to the object being measured, and not to the subject; as subjects we have to continuously confront ourselves with the object and not vice versa (Ferraris, 2005). This is especially relevant in the context of performance reviews, in which clarity must exist between the content of the review and the process of reviewing performance (Martinez et al., 2010), and data should be interpreted without excessive trust in numbers (Lenz and Lyles, 1985; Williams et al., 1994).

This paper has built on principles and techniques developed in measurement science to improve the theory and practice of performance measurement in organizations. The implications of the proposed pragmatic standpoint are many, ranging from the definition of measurement processes and the attributes of ‘good’ measurement, to the use of measurement results. Although it is acknowledged that great advancements in both theory and practice are certainly possible, they would have to go beyond ‘more and more precise measurement’.

References

- Ahn, H., 2001. Applying the Balanced Scorecard concept: an experience report. *Long Range Planning* 34(4), 441-461.
- Ammons, D.N., 1999. A proper mentality for benchmarking. *Public Administration Review* 59(2), 105-109.
- Argyris, C., 1977. Organizational learning and management information systems. *Accounting, Organizations and Society*, 2(2), 113-123.
- Argyris, C., 1992. *On Organizational Learning*. Blackwell Publishing, Cambridge, MA.
- Atkinson, A.A., Waterhouse, J.H., Wells, R.B., 1997. A stakeholder approach to strategic performance measurement. *Sloan Management Review* 38(3), 25-37.
- Bate, P., 2007. Editorial to the Special Issue 'Bringing the design sciences to organization development and change management'. *Journal of Applied Behavioral Science* 43(1), 8-11.
- Behn, R.D., 2003. Why measure performance? Different purposes require different measures. *Public Administration Review* 63(5), 586-606.
- Beer, M., Cannon, M. D., 2004. Promise and peril in implementing pay-for-performance. *Human Resource Management* 43(1), 3-48.
- Bhaskar, R., 1978. *Realist Theory of Science*. Longman Higher Education Division, New York.
- Bierman, H. Jr., 1963. Measurement and accounting. *The Accounting Review* 38(3), 501-507.
- Bohm, D., 1980. *Wholeness and the Implicate Order*. Routledge and Kegan Paul Ltd., London.
- Bridgman, P.W., 1927. *The Logic of Modern Physics*. Macmillan, London.
- Brignall, S., Modell, S., 2000. An institutional perspective on performance measurement and management in the 'New Public Sector'. *Management Accounting Research* 11(3), 281-306.

- Campbell, N. R., 1920. *Physics: the Elements*. Cambridge University Press, Cambridge.
- Carnap, R., 1966. *Philosophical Foundations of Physics*. Basic Books, New York.
- Cavalluzzo, K.S., Ittner, C.D., 2004. Implementing performance measurement innovations: evidence from government. *Accounting, Organizations and Society* 29(3-4), 243-267.
- Cecconi, P., Franceschini, F., Galetto, M., 2006. Measurements, evaluations and preferences: a scheme of classification according to the representational theory. *Measurement* 39(1), 1-11.
- Chapman, C.S., 1997. Reflections on a contingent view of accounting. *Accounting, Organizations and Society* 22(2), 189–205.
- Chenhall, R., 2005. Integrative strategic performance measurement systems, strategic alignment of manufacturing, learning and strategic outcomes: an exploratory study. *Accounting, Organizations and Society* 30(5), 395-422.
- Chenhall, R., Langfield-Smith, K., 1998. The relationship between strategic priorities, management techniques and management accounting: an empirical investigation using a systems approach. *Accounting, Organizations and Society* 23(3), 243-264.
- Chua, W., Degeling, P., 1993. Interrogating an accounting-based intervention on three axes: instrumental, moral and aesthetic. *Accounting, Organizations and Society* 18(4), 291-318.
- Corvellec, H., 1997. *Stories of Achievements: Narrative Features of Organizational Performance*. Transaction Publishers, New Brunswick, NJ.
- Covaleski, M.A., Dirsmith, M.W., Samuel S., 1996. Managerial accounting research: The contributions of organizational and sociological theories. *Journal of Management Accounting Research* 8(1), 1-35.
- Davis, S., Albright, T., 2004. An investigation of the effect of Balanced Scorecard implementation on financial performance. *Management Accounting Research* 15(2), 135-

153.

- De Geus, A., 1999. *The Living Company: Growth, Learning and Longevity in Business*. Nicholas Brealey Publishing: London.
- Denyer, D., Tranfield, D., van Aken, J. E., 2008 Developing design propositions through research synthesis. *Organization Studies* 29(3), 393–413.
- Dossi, A., Patelli, L., 2010. The inclusion of non-financial indicators in performance measurement systems used in relationships between headquarters and subsidiaries. *Long Range Planning* 43(4), 498-526.
- Ellis, B., 1968. *Basic Concepts of Measurement*. Cambridge University Press: Cambridge.
- Euske, K.J., Lebas, M., McNair, C.J., 1993. Performance management in an international setting. *Management Accounting Research* 4(4), 275-300.
- Ferraris, M., 2005. *Dove Sei? Ontologia del Telefonino, second ed.* Bompiani, Milan.
- Feyerabend, P., 1975. *Against Method - Outline of an Anarchistic Theory of Knowledge*. Verso, London.
- Finkelstein, L., 2003. Widely, strongly and weakly defined measurement. *Measurement* 34(1), 39-48.
- Franco-Santos, M., Kennerley, M., Micheli, P., Martinez, V., Mason, S., Marr, B., Gray, D., Neely, A., 2007. Towards a definition of a business performance measurement system. *International Journal of Operations and Production Management* 27(8), 784-801.
- Franco-Santos, M., Lucianetti, L., Bourne M., 2012. Contemporary performance measurement systems: a review of their consequences and a framework for research. *Management Accounting Research* 23(2): 79–119.

- Garvin, D.A., 1993. Building a learning organization. *Harvard Business Review* 71(4), 78–91.
- Gergen, K., 1973. Social psychology as history. *Journal of Personality and Social Psychology* 26(2), 309-320.
- Ghoshal, S., 2005. Bad management theories are destroying good management practices. *Academy of Management Learning & Education* 4(1), 75-91.
- Ghoshal, S., Moran, P., 1996. Bad for practice: a critique of the transaction cost theory. *Academy of Management Review* 21(1), 13-47.
- Gimbert, X., Bisbe, J., Mendoza, X., 2010. The role of performance measurement systems in strategy formulation processes. *Long Range Planning* 43(4), 477-497.
- Gittell, J.H., 2000. Paradox of coordination and control. *California Management Review* 42(3), 101-117.
- Glanzberg, M. (2009). Truth, in: Zalta, E. N. (ed.), *The Stanford Encyclopedia of Philosophy*. <http://plato.stanford.edu/archives/spr2009/entries/truth/>.
- Goldstein, H., Spiegelhalter, D., 1996. League tables and their limitations: statistical issues in comparisons of institutional performance. *Journal of the Royal Statistical Society* 159(3), 385-443.
- Goodman, N., 1978. *Ways of Worldmaking*. Hackett, Indianapolis.
- Goold, M., Quinn, J.J., 1990. The paradox of strategic controls. *Strategic Management Journal* 11(1), 43-57.
- Govindarajan, V., Gupta, A.K., 1985. Linking control-systems to business unit strategy - impact on performance. *Accounting, Organizations and Society* 10(1), 51-66.
- Griffith, R., Neely, A., 2009. Performance pay and managerial experience in multi-task teams: evidence from within a firm. *Journal of Labor Economics* 27(1), 49-82.

- Hall, M., 2008. The effect of comprehensive performance measurement systems on role clarity, psychological empowerment and managerial performance. *Accounting Organizations and Society* 33(2-3), 141-163.
- Hall, M., 2011. Do comprehensive performance measurement systems help or hinder managers' mental model development? *Management Accounting Research* 22(2), 68–83.
- Hanson, N., 1958. *Patterns of Discovery*. Cambridge University Press, Cambridge.
- Hatry, H., 1999. *Performance measurement: getting results*. Urban Institute, Washington DC.
- Hausknecht, J. P., Hiller, N. J., Vance, R. J., 2008. Work-unit absenteeism: effects on satisfaction, commitment, labor market conditions, and time. *Academy of Management Journal* 51(6), 1223–1245.
- Helmholtz, H., 1887. An Epistemological Analysis of Counting and Measurement, in: Kahl, R. (ed. and trans.), *Selected writings of Hermann von Helmholtz*. Middletown (CT): Wesleyan University Press, 1971.
- Hempel, C. G., 1952. *Fundamentals of Concept Formation in Empirical Science*. University of Chicago Press, Chicago.
- Henri, J.-F., 2006. Management control systems and strategy: A resource-based perspective. *Accounting, Organization and Society* 31(6), 529–558.
- Hölder, O., 1901. Die Axiome der Quantität und die Lehre vom Mass. *Berichte über die Verhandlungen der Königlich Sächsischen Gesellschaft der Wissenschaften zu Leipzig, Mathematisch-Physikalische Klasse*, 53, 1–46. (Part 1 translated by J. Michell and C. Ernst (1996). The axioms of quantity and the theory of measurement. *Journal of Mathematical Psychology* 40(3), 235–252.
- Hood, C., James, O., Scott, C., 2000. Regulation of government: has it increased, is it increasing,

- should it be diminished? *Public Administration* 78(2), 283-304.
- Hoque, Z., James, W., 2000. Linking Balanced Scorecard measures to size and market factors: impact on organizational performance. *Journal of Management Accounting Research* 12, 1-17.
- Huff, A., Tranfield, D., van Aken, J.E., 2006. Management as a design science mindful of art and surprise: a conversation between Anne Huff, David Tranfield, and Joan Ernst van Aken. *Journal of Management Inquiry* 15(4), 413–424.
- Husserl, E., 1907. *Die idee der Phaenomenologie*; translated by W.P. Alston and G. Nakhnikian. Nijhoff, The Hague, 1964.
- Ittner, C.D., Larcker, D.F., Randall, T., 2003. Performance implications of strategic performance measurement in financial service firms. *Accounting, Organizations and Society* 28(7-8), 715-741.
- Johnson, P., Duberley, J., 2000. *Understanding management research*. Sage: London.
- Johnson, H.T., Kaplan, R.S., 1987. *Relevance lost – the rise and fall of management accounting*. Harvard Business School Press: Boston, MA.
- Joint Committee of Guides in Metrology, 2008a. International vocabulary of metrology — Basic and general concepts and associated terms (VIM, 3rd edition). JCGM 200:2008 (also published in hard copy format as ISO/IEC Guide 99:2007, Geneva (Switzerland):ISO/IEC, 2007). <http://www.bipm.org/en/publications/guides/vim.html>.
- Joint Committee of Guides in Metrology, 2008b. Guide to the expression of uncertainty in measurement (GUM 1995 with minor corrections). JCGM 100:2008. <http://www.bipm.org/en/publications/guides/gum.html>.
- Kaplan, R.S., Norton, D.P., 1992. *The Balanced Scorecard - measures that drive performance*.

- Harvard Business Review 70(1), 71-79.
- Kaplan, R.S., Norton, D.P., 2008. Execution Premium. Linking Strategy to Operations for Competitive Advantage. Harvard Business School Press: Boston, MA.
- Kaplan, R. S., 2009. Conceptual Foundations of the Balanced Scorecard, in: Chapman, C. S., Hopwood, A. G. and Shields, M. D. (Eds.), Handbook of Management Accounting Research. Vol. 3, Elsevier Science & Technology Books, Amsterdam, pp. 1253-1270.
- Kennerley, M., Mason, S., 2008. The use of information in decision making - literature review for the Audit Commission. Report available on www.auditcommission.co.uk
- Kolehmainen, K., 2010. Dynamic strategic performance measurement systems: balancing empowerment and alignment. Long Range Planning 43(4), 527-54.
- Krantz, D., Luce, R., Suppes, P., Tversky A., 1971 vol.1; 1989 vol.2; 1990 vol.3. Foundations of *Measurement*. Academic Press, New York.
- Kuhn, T.S., 1970. The Structure of Scientific Revolutions. University of Chicago Press, Chicago.
- Larson, E. W., Fukami, C. V., 1985. Employee absenteeism: the role of ease of movement. Academy of Management Journal 28, 464-471.
- Lenz, R.T., Lyles, M. A., 1985. Paralysis by analysis: is your planning system becoming too rational? Long Range Planning 18(4), 64-72.
- Lillis, A. M., 2002. Managing multiple dimensions of manufacturing performance - An exploratory study. Accounting, Organizations and Society 27(6), 497-529.
- Malina, M.A., Selto, F.H., 2001. Communicating and controlling strategy: an empirical study of the effectiveness of the Balanced Scorecard. Journal of Management Accounting Research 13(1), 47-90.

- March, J. G., Sutton, R. I., 1997. Organizational performance as a dependent variable. *Organization Science* 8(6), 698–706.
- Mari, L., 1997. The role of determination and assignment in measurement. *Measurement* 21(3), 79-90.
- Mari, L., 2000. Beyond the representational viewpoint: a new formalization of measurement *Measurement*, 27(2), 71-84.
- Mari, L., 2003. Epistemology of measurement. *Measurement* 34(1), 17-30.
- Mari, L., 2007. Measurability, in Boumans, M. (Ed.) *Measurement in economics*. Elsevier, London.
- Markham, S. E., 1985. An investigation of the relationship between unemployment and absenteeism: a multilevel approach. *Academy of Management Journal* 28, 228–234.
- Martinez, V., Pavlov, A., Bourne, M., 2010. Reviewing performance: an analysis of the structure and functions of performance management reviews. *Production Planning and Control* 21(1), 70-83.
- McKevitt, D., Lawton, A., 1996. The manager, the citizen, the politician and performance measures. *Public Money & Management* 16(3), 49-54.
- Meinong, A., 1904. Über gegenstandstheorie, in A. Meinong, R. Ameseder, and E. Mally (eds.), *Untersuchungen zur Gegenstandstheorie und Psychologie*, Leipzig: Barth, pp. 1–50. (Eng. trans. by I. Levi, D. B. Terrell, and R. M. Chisholm: ‘On the Theory of Objects’, in R. M. Chisholm (ed.), *Realism and the Background of Phenomenology*. The Free Press, New York, pp. 76–117).
- Micheli, P., Manzoni, J.-F., 2010. Strategic performance measurement: benefits, limitations and paradoxes. *Long Range Planning* 43(4), 465-76.

- Morgan, G., 1988. Accounting as reality construction: Towards a new epistemology for accounting practice. *Accounting, Organizations and Society* 13(5), 477-485.
- Nagel, E., 1961. *The Structure of Science: Problems in the Logic of Scientific Explanation*. Routledge, London.
- Neely, A., 1999. The performance measurement revolution: why now and where next. *International Journal of Operations and Production Management* 19(2), 205-228.
- Neely, A., Adams, C., Kennerley, M., 2002. *The Performance Prism: The Scorecard for Measuring and Managing Business Success*. Financial Times Prentice Hall, London.
- Neely, A., Bourne, M., 2000. Why measurement initiatives fail. *Measuring Business Excellence* 4(4), 3-7.
- Neely, A., Micheli, P., Martinez, V., 2006. Acting on information: lessons from theory and practice. Report prepared for the National Audit Office. Publication sponsored by the Advanced Institute of Management.
- Numagami, T., 1998. The infeasibility of invariant laws in management studies: a reflective dialogue in defense of case studies. *Organization Science* 9(1), 2-15.
- Osborne, D., Gaebler, T., 1992. *Reinventing Government: How the Entrepreneurial Spirit is Transforming the Public Sector*. Addison-Wesley: Reading, MA.
- Perera, S., Harrison, G., Poole, M., 1997. Customer-focused manufacturing strategy and the use of operations-based non-financial performance measures: A research note. *Accounting, Organizations and Society* 22(6), 557-572.
- Peters, T.J., Waterman, R.H. Jr, 1982. In *Search of Excellence: Lessons from America's Best-run Companies*. Harper & Row, New York.
- Pfeffer, J., Sutton, R., 2006. Hard facts, dangerous half-truths and total nonsense: profiting from

- evidence-based management. Harvard Business Publishing: Boston, MA.
- Pfeffer, J., 1998. Six dangerous myths about pay. *Harvard Business Review* 76(3), 108-119.
- Poister, T. H., 2010. The future of strategic planning in the public sector: linking strategic management and performance. *Public Administration Review* 70: s246–s254.
- Pollitt, C., 2006. Performance information for democracy: the missing link? *Evaluation* 12(1), 38-55.
- Power, M., 1997. *The Audit Society - Rituals of Verification*. Oxford University Press, Oxford.
- Power, M., 2004. Counting, control and calculation: reflections on measuring and management. *Human Relations* 57(6), 765-783.
- Richtmyer, F. K., Kennard, E. H., Cooper, J. N., 1969. *Introduction to Modern Physics*. McGraw-Hill, New York.
- Roberts, F., 1979. *Measurement Theory - With applications to Decision Making, Utility, and the Social Sciences*. Addison-Wesley, Reading, MA.
- Romme, A.G.L., 2003. Making a difference: organization as design. *Organization Science* 14(5), 558–573.
- Roos, G., Roos, J., 1997. Measuring your company's intellectual performance. *Long Range Planning* 30(3), 413-426.
- Rossi, G.B., 2007. Measurability. *Measurement* 40(6), 545-562.
- Royal Statistical Society, 2005. *Performance Indicators: Good, Bad and Ugly*. This document can be accessed from the Royal Statistical Society site at: <http://www.rss.org.uk>.
- Senge, P., 2006. *The Fifth Discipline: The Art and Practice of the Learning Organization*. Doubleday, New York, NY.
- Simon, H., 1996. *The Sciences of the Artificial*, third ed. MIT Press, Cambridge, MA.

- Simons, R., 1990. The role of management control systems in creating competitive advantage - new perspectives. *Accounting, Organizations and Society* 15(1-2), 127-143.
- Smith, P., 1995a. Performance indicators and outcome in the public sector. *Public Money & Management* 15(4), 13-16.
- Smith, P., 1995b. On the unintended consequences of publishing performance data in the public sector. *International Journal of Public Administration* 18(2-3), 277-310.
- Speckbacher, G., Bischof, J., Pfeiffer, T., 2003. A descriptive analysis on the implementation of Balanced Scorecards in German-speaking countries. *Management Accounting Research* 14(4), 361-388.
- Stevens, S.S., 1946. On the theory of scales of measurement. *Science* 103(2684), 677-680.
- Townley, B., Cooper, D.J., Oakes, L., 2003. Performance measures and the rationalization of organizations. *Organization Studies* 24(7), 1045-1071.
- Tsoukas, H., 1989. The validity of idiographic research explanations. *Academy of Management Review* 14(4), 551-561.
- von Hayek, F., 1989. The pretence of knowledge. *The American Economic Review* 79(6), 3-7.
- Williams, F. P. III, McShane, M. D., Sechrest, D., 1994. Barriers to effective performance review: the seduction of raw data. *Public Administration Review* 54(6), 537-542.
- Wouters, M., Wilderom, C., 2008. Developing performance-measurement systems as enabling formalization: a longitudinal field study of a logistics department. *Accounting, Organizations and Society* 33(4-5), 488-516.