

Evaluations as value-measurement links: Exploring metrics and meanings in science

Social Science Information

2019, Vol. 58(2) 282–300

© The Author(s) 2019



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/0539018419850771

journals.sagepub.com/home/ssi



Felicitas Hesselmann 

Humboldt-Universität zu Berlin and German Centre for Higher Education Research and Science Studies, Berlin, Germany

Cornelia Schendzielorz

German Centre for Higher Education Research and Science Studies and Humboldt-Universität zu Berlin and Centre Marc Bloch

Abstract

This contribution seeks to provide a more detailed insight into the entanglement of value and measurement. Drawing on insights from semiotics and a Bourdieusian perspective on language as an economy of linguistic exchange, we develop the theoretical concept of value-measurement links and distinguish three processes – operationalisation, nomination, and indetermination – as forms in which these links can be constructed. We illustrate these three processes using (e)valuation practices in science, particularly the journal impact factor, as an empirical object of investigation. As this example illustrates, measured values can function as building blocks for further measurements, and thus establish chains of evaluations, where it becomes more and more obscure which values the measurements actually express. We conclude that in the case of measured values such as impact factors, these chains are driven by the interplay between the interpretative openness of language and the seeming tendency of numbers to fixate meaning thus continually re-creating, transforming and modifying values.

Keywords

Bourdieu, evaluation, journal impact factor, measurement, quantification, semiotics, value

Corresponding author:

Felicitas Hesselmann, Humboldt-Universität zu Berlin – Institut für Sozialwissenschaften, Universitätsstrasse 3b, Berlin, 10117, Germany.

Email: hesselmann@dzhw.eu

Résumé

Cet article entend apporter une connaissance plus détaillée de l'imbrication existant entre valeurs et mesures. À partir de connaissances provenant de la sémiotique et partant d'une perspective bourdieusienne qui conçoit le langage comme un système économique d'échanges linguistiques, nous développons le concept théorique de liaisons valeur-mesure et distinguons trois processus – l'opérationnalisme, la dénomination et l'indétermination – comme des formes à partir desquelles ces liaisons peuvent être construites. Nous illustrons ces trois processus en convoquant les pratiques d'évaluation utilisées en science, et plus particulièrement le facteur d'impact, pris comme objet d'étude empirique. Comme cet exemple le montre, les valeurs mesurées peuvent fonctionner comme des éléments de base menant vers plus de mesures, et construisent de cette façon des chaînes d'évaluations, où il devient de moins en moins clair quelles valeurs ces mesures expriment exactement. Nous en concluons que dans le cas de valeurs mesurées tel que le sont les facteurs d'impact, ces chaînes sont déterminées par l'interaction entre l'ouverture interprétative du langage et l'apparente tendance qu'ont les chiffres à fixer un sens, recréant, transformant, modifiant continuellement les valeurs en jeu.

Mots-clés

Bourdieu, évaluation, facteur d'impact de revue, mesure, quantification, sémiotique, valeur

Introduction

Measurements and evaluations are pervasive in our day-to-day lives (Lamont, 2012). When buying a book from Amazon, for example, we can read what feels like a flood of user-written reviews of books in order to guide our decision-making. For our convenience, these reviews themselves are ranked by how many other users evaluated them as 'helpful', and at the bottom of each review we, too, are asked: 'Was this review helpful for you?' In this article we take this pervasive tendency of measurements to spawn further measurements that aim to build on, elaborate, transform, or qualify previous measurements as the initial point of our argument. Here, we want to provide a more detailed insight into the entanglement of measurements and values, using valuation practices in science as an object of investigation.

Measurement is always an act of value discovery or creation and re-creation, insofar as it rests on the assertion that something is potentially of value or valuable: 'Measuring presupposes that we approach something as worthy or valuable, even if we eventually find it wanting – for instance, under-specified, or wrongly specified' (Brighenti, 2018: 6). Without such a claim to represent and express an abstract value and hence be charged with meaning, measurements would not be recognisable; they would be mere calculations, graphics or texts. This creation of value, or valuation (see Krüger & Reinhart, 2018) just as mechanisms of value discovery (Vormbusch, 2018a), however, does not yet say anything about the actual value it refers to, it asserts *that* something has value, but not, *which* value it has. Thus, it establishes a potential link between values and measures without substantiating which measures link to which values and what the

specific value-measurement link does actually express. In line with Brighenti's diagnosis of a fundamental problem concerning 'measure' (Brighenti, 2012: 16) we note phenomenologically that the multiplication of measurements and their uses in subsequent evaluations¹ breeds considerable value diffusion: by blurring their designation these processes ceaselessly re-create, transform and modify values. Hence values remain in constant flux and are steadily overthrown in measuring practices. In this contribution, our aim is to develop a conceptual proposition to make sense of this highly dynamic and complex interrelation of measurements, values, and evaluations. To address the question of how value and measurement are interrelated we first elaborate our main theoretical reference points from which we derive our analytical tools. These then enable us to define our key terms, clarifying what we understand by measurements, values and evaluations. Subsequently we will discuss the question of how measurement and value relate to each other from two perspectives: the first perspective focuses on how values and measurements are linked and develops an analytical heuristic by distinguishing three forms of linkage and discussing exemplary variants of linkage and their practical consequences; the second perspective focuses on chains of evaluations in the course of subsequent value-measurement links. We finally summarise the outcomes of our investigation of value-measurement links by outlining possible sociopolitical implications of our approach, considering value-measurement links as tools for governance.

Linking values and measurements: Analytical tools

To ask how value and measurement become connected means to employ a post-positivist perspective: accordingly, we do not understand measurements to merely express a specific underlying, ontologically stable attribute (value) of the object under measurement, as many positivist approaches do (for example Borsboom et al., 2004). In line with post-structuralist ideas about validity (e.g. Lather, 1993) we assume in contrast that there is no necessary or internal connection between a value and a corresponding measurement. Therefore, we investigate how the world and its valuability are captured by linking actual values and actual measurements, both by way of numeric representations, as well as through wording. Hence we focus on how the words chosen encompass the world they intend to name.

In conceptualising this language-based relation between values, measurements and the corresponding value-measurement links we build on the semiotic theory of Charles Sanders Peirce, and in particular his triadic model of the sign. In the most general sense, a sign is a 'thing which stands for another thing' (Peirce, 1986: 76). Peirce understands a sign to be comprised of three distinct elements, the sign-vehicle, the object, and the interpretant. The three elements are related so that the sign-vehicle, such as the word 'chair', represents the object, which can be a physical object like a chair, but also a person, an event, or an idea, through the interpretant, which can be seen as the specific understanding by which the vehicle represents the object: 'Thus, in looking at a map, the map itself is the vehicle, the country represented is the natural object, and the idea excited in the mind is the interpretant' (Peirce quoted in Deledalle, 2001: 38). As Peirce emphasises, all three elements are only defined by their relation, a map is only a vehicle when it is understood and used as a representation of a piece of land (real or imaginary), not if it is used

as decoration, or as wrapping paper. Likewise, the piece of land is only an object when represented through a sign-vehicle, not when it is simply used as soil to grow crops. With the element of the interpretant, Peirce furthermore highlights the fact that something only becomes a sign through its specific use: '[A sign] is not actually a sign unless it is used as such; that is unless it is interpreted to thought and addresses itself to some mind' (Peirce, 1986: 76). Thus, signs are only produced through interpretation, rather than existing on their own.

Building on Peirce's model, social semiotics (e.g. Fiske, 1989; Hodge & Kress, 1988; Vannini, 2004; 2007) develop a more sociologically inclined reading of Peircean semiotics. While Peirce himself locates the interpretant within the individual mind, social semiotics believe interpretation to take place 'within the process of context-bound and conflict-laden interpersonal interaction' (Vannini, 2007: 115). Interpretations of a word or sign are embedded in a social context, that is, in precisely structured social relationships in which words are charged with field-specific connotations: 'The meaning of a sign is [...] in the social context of its use' (Chandler, 2002: 13). From this assumption, social semiotics conclude two things. First, that the sign is malleable in principle: because links between signs and objects are negotiated in social situations, they are always to some degree open for re-negotiation and hence re-interpretation. What emerges as the meaning of a sign in one situation can vary from previous and subsequent situations. Multiple interpretations can also coexist at once, with no interpretation being essentially more correct or more valid than others (see also Hall, 2006: 167ff.). Second, that interpretation is a thoroughly political process and is both shaped by and in turn shapes existing power relations (Kress, 1993; Vannini, 2004). As such, meaning-making can be used strategically to gain power and resources, and to further specific interests. These interests will then be inscribed into the particular sign:

That is, in relation to a particular object or event, 'interest' leads the producer of the sign to focus on a particular characteristic of an object or event [...] to make that the criterial characteristic of the object or event, that is, make it the basis of the production of a signified. (Kress, 1993: 174)

Hence, as signs and their meanings are negotiated in social interactions, they are highly sensitive to power structures, and the ability to produce 'acceptable' interpretations is not distributed evenly among all members of a given community, but rather depends on symbolic resources and interpretative authority (Van Leeuwen, 2005: 48ff.).

Drawing an analogy from the triadic model of the sign to the relation of value and measurement, we argue that a specific measurement can be seen as a sign-vehicle, which represents the value (i.e. Peirce's object), through a value-measurement link, which parallels Peirce's interpretant. Against the background of social semiotics we regard linking actual values and actual measurements as a language-based practice. However, in our understanding this does not necessarily result in an opposition and detachment of reality from its sign-mediated representation. Rather, it is important to stress that the variety of meanings and various ways of speaking do not cohabit in equal coexistence. Consequently, we need to take the importance of terminology and nomenclature seriously. Following Bourdieu (Bourdieu, 2005: 41ff.), we assume that these manifold attributions of meaning

always go hand in hand with a positioning on the so-called linguistic market, which is formed and played on with rhetorical means and strategies. In this concept different forms of talk are an expression of the struggles for symbolic power congealed in publicity, recognition, reputation, prestige, stardom, dominance etc., fought out through the medium and the very object of language. The social position of the various ways of speaking is determined in relation to the standardised official language, which claims legitimacy as the 'formal language' (Bourdieu, 2005: 48–50; Schendzielorz, 2011: 29–33). The linking of signs is hence not entirely contingent or free-floating, but rather influenced by existing systems of signs that in turn are shaped by social relations of power. In this sense, we conceive linking values and measurements as a language-based practice whose execution can be grasped and deciphered more congruently with Bourdieu's notion of a 'linguistic market' as a field structured by power relations than for example with Wittgenstein's concept of 'language game'. Whereas Wittgenstein's 'language games' are an ordered, rule-based practice, in which rules are followed 'blindly' (Wittgenstein, 1971: 219, 351),² in Bourdieu's linguistic market strategies are at work. The concept of strategy implies a higher degree of consciousness and choice than that of a rule. Accordingly, in a Bourdieusian sense, a language-based practice functions as a means in symbolic struggles, among other things, to maintain and improve the respective social position. Likewise it can serve as a reference point for contestation and resistance. Thus, the question of linking of measures, signification and values can also be understood as a field of struggle for legitimate wording and language and thus for power of and over interpretation and evaluation. Furthermore, both measures and particularly their corresponding values can remain quite vague and ambiguous even in the cases when a link is established. As such, the terms chosen in the course of denomination only lay a rough trail for connotation, by limiting the scope of possible meaning and interpretation, while still leaving room for association and different interpretations. From a Bourdieusian point of view, 'the word has no social existence as a neutralized product of the practical references in which it actually operates' (Bourdieu, 2005: 43).³ Hence values are produced in the process and practice of naming in which the interplay of denotation and connotation is traced. Despite this designation of values, their meaning remains unstable, as it is interminably negotiated and disputed in the juxtaposition of denotation and connotation that runs through every language-based practice and may indicate struggles in power relations. At the same time, the process configuring the value-measurement link can be carried out by selection procedures based on current opportunities, so that it does not necessarily need to refer to a qualitative definition of value. A current example for the assignment of meaning combined with measurement from the field of science is the so-called 'impact' of research. It can be suggested to stand for almost everything and therefore to be not very meaningful. Only preceding adjectives such as 'societal' name a certain field and rudimentary limit the range of the stated impact.

Defining our key terms

Drawing on these analytical tools we can now further conceptualise the relationship between values and measurements. To do so, we will first define the terms and put them in a conceptual relation to each other. Therefore, we locate them in relation to core

practices of (e)valuation in which measures and values are linked, such as categorisation, commensuration and quantification processes.

Measurements are precisely related to an artefact. By claiming to capture this artefact in one or more specific characteristics, measurements can operate with signs, and/or words as well as numbers. In case they consist of countable signs they enable mathematical operations with and quantification of the measured values.⁴ In case of non-countable signs, such as words, mathematical operations cannot be straightforwardly done. Quantitative measurements transfer the description of an artefact in a numeric representation through designating countable signs presuming specific taxonomies (see Vormbusch, 2012: 223). However, counting is not always a practice of measuring: counting sheep, for example, is not a measurement of sleeplessness. Likewise, not all measuring necessarily results in quantification, as examples such as the litmus test or the voltage tester illustrate. By means of units of measurement – centimetres, cubic metres, litres, and other measuring scales – measurements are always integrated into systems of classification and hence systems of order (Bowker & Star, 1999) that shape their significance and contribute to their continuous development. Nevertheless, as Power puts it, measurements ‘can be, and often are, imagined before the availability of reliable instrumentation’ (Power, 2004: 768). Thus, measurements are always empirically connected with interests ranging from the interest of a functional fit of the measurement to the measured artefact to normative interests of ordering and classification. These can then be asserted purposefully and powerfully in the context of conflicts of interest and symbolic struggles for one signification or another.

Values can be considered as abstract differentiations which, unlike measurements, do not necessarily require a concrete reference to an artefact. This understanding derives upon Dewey’s (1939) critique of value as a substance and considers a praxeological conception of the value-genesis, in which ‘values are created by projections of the creative imagination’ (Quéré, 2015: 173) as Quéré puts it, drawing on Dewey as well. These include, for example, categorisations that are not tied to a closely specified empirical (object-) reference, such as quality, freedom, originality, truthfulness or responsibility.⁵ Nevertheless, values are always implicitly charged with meaning through their naming in the medium of language, which need not necessarily to be explicit, beyond the specific notion of value. To a certain extent, it can resonate passively in the nomination and, if necessary, be activated for specific purposes and then be used and guided by specific interests. Following Quéré’s reading of Dewey and his ‘adverbial approach’, ‘values are alternately ends and means’ (Quéré, 2015: 174), given that ‘the distinction between ends and means is temporal and relational’ (Dewey, 1939: 43). The fact that values can be both ends and means converges with our idea that the measurements do not need to be in an instrumental relation to the value. They do not have to be a means to the end of value determination or valorisation, but can remain indeterminate like an empty signifier.

In *evaluations*, measures and values are linked insofar as concrete empirical measurements and implicit abstract horizons of interpretation are translated into differentiated meanings. Here, meaning is explicitly assigned and weighed up so that meaning is specifically set (Krüger & Reinhart, 2017: 277–278; 2018: 12ff.). Empirically, evaluations often contain and make use of commensurative practices and comparative assessments. This range of evaluation is also compatible with Quéré’s praxeological concept

in which evaluation implies a more cognitive aspect, as opposed to ‘de facto valuing [..] for they consist of judgments formed through an inquiry, a reflection, or a measurement’ (Quéré, 2015: 166). In addition, as ascriptions of meaning, evaluations always produce power effects.

Consequently, we do not understand the linking of measurements and values in evaluations as a simple means-ends relation in which measurements are merely the execution and representation of values. This would mean thinking of measurements as mere descriptions and overlooking measurements as powerful medium and instrument, given the fact that the design of measurement procedures themselves are always linked to interests, as Angermüller convincingly demonstrates with the example of science as a ‘numerocratic power-knowledge complex’ (Angermüller, 2011: 176). As value-measurement links, evaluations and their power effects are also not fixed and static, but variable: firstly because the potential references to abstract terminology and its value-laden terminology are diverse and the nomenclature remains variable; secondly, because the objects that are recorded and described in measurement in some characteristics, can affect the meaning of evaluations. Value-measurement links as well as meaning are ‘a two-way relation’ (Abbott, 2001: 19).⁶ Therefore, we understand value-measurement linkages as a dynamic relationship in which (symbolic) struggles are conducted in the medium of language.

Three forms of value-measurement links: A heuristic

These struggles relate to the matter of how values and measurements are linked. In this perspective, three forms of value-measurement links can be distinguished, which manifest themselves in the medium of language: *operationalisation*, *nomination*, and *indetermination*. To be able to distinguish these three forms, they must be thought of as processes, rather than states: their difference thus primarily lies in the way the connection between values, measurements, and evaluations is established. These three forms will be discussed below, using the ‘impact’ of scientific research, and particularly the journal impact factor, as an exemplary case.

Operationalisation

In the first form, a given value is translated into specific empirical measures. That means that starting from a value as an object, a specific link is established to a sign-vehicle that represents the value, thereby constituting meaning. The creation of the journal impact factor by Eugene Garfield, one of the founding fathers of bibliometrics, presents a classical example of operationalisation. Initially, Garfield wanted to create a measure that would help him decide which journals to include in his newly founded *Science Citation Index* (Garfield, 2006). In addition to this general aim, Garfield also developed a range of further conceptual considerations, such as not unduly discriminating against smaller journals with lower overall publication counts. The arithmetic measurement was then developed in order to fit this goal.

However technical this operationalisation might appear, it is also clearly a value-laden process: the idea of the journal impact factor was tied to the specific practical interest of establishing a criterion for inclusion and exclusion, and thus to confer the value of

good (enough) for inclusion (or too bad for inclusion) to the objects it would then be applied to. Today, the question of whether a journal is included in the *Science Citation Index* (or any other index of that kind) is hotly debated (Dannenberg, 2017; Flaherty, 2017, Batagelj, Ferligoj & Squazzoni, 2017, Wissenschaftsrat, 2017), and it is easy to see that this question of inclusion and exclusion is a highly political one that is intimately connected to power struggles in the field of scientific publishing. Furthermore, the additional considerations about the representation of smaller journals reveal that, from its beginning, Garfield's index was supposed to reflect a very specific segment of the published literature, one that was deemed valuable by presupposition, even before measurement. Intentionally giving weight to smaller journals is a normative consideration, one that does not result from any technical or arithmetic necessity. These can also be seen as more conventionally normative questions of fairness and representation, which accompanied the creation of the measurement from the start.

Nomination

In the second form, an existing measure is given a label that indicates a specific value. Nomination happens when a given measurement, such as the journal impact factor, becomes linked to a value for the first time, or to a different value than before. Nomination comprises denotation and connotation. Denotation refers to what is grammatically and semiotically defined and postulated as a 'constant and common' basis. Connotation refers to what is socially conditioned, variable, i.e. the variety of meaning that arises only in the manifold social relations in which the words are charged with connotations specific to their field (Bourdieu, 2005: 43). Connotation is accordingly permeated and conditioned by the experiences and contextualisation of the actors using these words (Schendzielorz, 2011: 32–35). Building on our theoretical framework, the distinction between denotation and connotation is one of usage and of the power to define authoritative interpretations, not one of 'reality' in language' (Hall, 2006: 168). There are myriad examples of situations in which the journal impact factor becomes repurposed and redefined. Highly contested (Callaway, 2016; Seglen, 1997), but still widely employed (Wilsdon et al., 2015), is the practice of linking the journal impact factor to the value of quality.⁷ This particular value-measurement link is typically fashioned in discourses that take place outside the bibliometrics community, in areas such as science policy, or the wider scientific community. Below are two examples:

The Impact Factor and SCImago Journal Ranking (SJR) are two common measurements or rankings of journal quality.⁸ (Curtin University, 2019)

The most commonly used measure of journal quality is the Journal Impact Factor (JIF). The Impact Factor attempts to measure the quality of a journal in terms of its influence on the academic community. (University of New England, 2019)

Nomination gives meaning to a measurement, and in doing so sets a route for possible uses. The journal impact factor these two examples talk about does not help the decision-making about which journals to include in the future *Science Citation Index*

anymore (in fact, it is now the other way around: journals not indexed in the Web of Science do not have a journal impact factor, because they are not included in the measurement). Rather, it can be used in a range of other ways, including in the allocation of funding or tenure (see also Wilsdon et al., 2015). An impact factor that is read as a measure of quality awards those who score high with particular interpretative authority. Especially through its influence on resource allocation, it additionally affects power relations within science, by helping those with high scores gain formal positions of authority (such as professorships). Such authority in turn can be used to stabilise particular interpretations of the impact factor.

Indetermination

The third way leaves the measurement-value linkage unspecified. Much like the other two, it must be thought of as the active process of fabricating an indeterminate link, rather than as a state of indetermination. This can be done in two ways: either by an underdetermined link to a measurement that cannot assure to capture a thing as something and to function as sign-vehicle for a value, or by producing and using measurements leaving the corresponding value altogether unspecified. Following the argument detailed above, these measures still presume to express *some* value, but without providing indication which value(s). Hospital rankings for example (Dorn, this issue), often do not provide an unequivocal ranking in the strict sense, but rather a collection of good practices. These measurements thus particularly invite and require interpretation in order to make sense of them. At the same time, they are also particularly versatile, lending themselves to a wide variety of uses.⁹ The San Francisco Declaration on Research Assessment (DORA, 2013), for instance, discusses appropriate uses and interpretations of publication metrics, especially the journal impact factor. It can be seen as a document dedicated to problematising established value-measurement links and potentially factoring new ones. Under ‘General Recommendation’ it reads:

Do not use journal-based metrics, such as Journal Impact Factors, as a surrogate measure of the quality of individual research articles, to assess an individual scientist’s contributions, or in hiring, promotion, or funding decisions. (DORA, 2013)

This general recommendation represents an example of indetermination. It opposes a number of existing value-measurement links (most prominently the link between the journal impact factor and the quality of individual articles) and thus seeks to release both the journal impact factor from this specific interpretation and subsequent uses, and the notion of quality of a research article from its bind to a specific existing measure. As an argumentation strategy, it makes sense to start a list of recommendations with dissolving an existing value-measurement link by criticising the operationalisation or nomination in order to subsequently establish a new possible link. Interestingly, the declaration is mostly dedicated to indetermination, by cautioning against specific uses and interpretations of metrics, or urging publishers to proceed by ‘presenting the metric in the context of a variety of journal-based metrics (for example, 5-year impact factor, EigenFactor,

SCImago, h-index, editorial and publication times, etc.) that provide a richer view of journal performance' (Recommendation 6, DORA, 2013).

Such recommendations problematise existing value-measure links without providing any clear alternative links: providing a range of different measurements together generates uncertainty about what value each measurement stands for, as well as what value the overall assemblage of measurements should represent. With the mention of a range of possible metrics, the DORA also exhibits indetermination: while the interpretation of the (2-year) journal impact factor is extensively discussed in the document, the other metrics are only mentioned once, with no additional information as to how they should be interpreted, what they should stand for, or how they should be used. In general, metrics such as the Eigenfactor or the 5-year impact factor are not used very widely and a conventional interpretation for them has yet to emerge. On the one hand, the DORA also fails to offer a particular interpretation, and rather leaves the meaning of those metrics indeterminate. Measurements and values thus appear as free-floating, inviting and demanding new interpretations and conceptualisations. On the other hand, these metrics also exemplify issues of interpretative authority: because the (2-year) journal impact factor has gained such prominence and is used so widely, it proves very difficult to establish other metrics, even though these metrics might be more precise, reliable, or inclusive. Here, the production of new and alternative measurements is seriously impeded through existing patterns of measurement use.

To sum up, besides the different ways these forms of value-measurement links are established they can produce varying outcomes, which may pre-set how they are chained and promote in which way they play a part in evaluation practices.

Unfolding chains of evaluation

As Peirce argues, a triadic sign does not exist on its own, but only in connection with other signs, which it interprets. Signs are thus linked up into chains in a process Peirce terms 'infinite semiosis':

The meaning of a representation can be nothing but a representation. In fact, it is nothing but the representation itself conceived as stripped of irrelevant clothing. But this clothing never can be completely stripped off; it is only changed for something more diaphanous. So there is an infinite regression here. Finally, the interpretant is nothing but another representation to which the torch of truth is handed along; and as representation, it has its interpretant again. Lo, another infinite series. (Peirce, 1974: 171)

Firstly, the multiplication of evaluations may create a situation that is similar to a *simulation* in the sense of Baudrillard (1994). For Baudrillard, a simulation consists of signs that point back to other signs, that again point back only to more signs, so that the signs no longer reference any real or actual object. The simulation is thus 'substituting the signs of the real for the real' (Baudrillard, 1994: 2). According to Baudrillard, reality is not so much clouded or distorted, as made to disappear entirely, to collapse into its sign (see also Eco, 1985: 39).

We argue that this characteristic of signs to multiply and to link to further signs can also be observed with regard to measurements, which also frequently create new measurements

and evaluations. Seen from this angle, in ongoing chains of value-measurement linkages, the linkages also begin to only reference preceding links and lose their reference to any pre-existing 'reality'. That is, their measurements become more and more detached from any particular artefact (e.g. a journal, a scientific paper, a person, an apple) they were supposed to convey information about. Ultimately, they can only be understood as relating to other value-measurement links.

Value-measurement links are continuously modulated in actual practices, when filled in and stated more precisely in concrete cases. Specific uses of the former measurements can serve to clarify and substantiate previously ambiguous values. However, they can also contribute to obscure the values linked to measurements by adding new layers of interpretation and new possible and actual links between measures and values. For example measured values, such as 'impact factors' are frequently used and re-used in other, ensuing practices of measurement and evaluation. They can function as building blocks for further measurements, by using quantitative outcomes of preceding measurement processes as input variables in further calculations. They can also themselves be subject to further evaluations and measurements that aim at determining, for example, reliability, precision, or the predictive value of the original measures. Furthermore, measurements can themselves be turned into values in subsequent evaluations, as they are in processes of benchmarking. These ensuing uses in turn also affect the former connection between measures and values by creating new links and interpretations for the former measures. This is evident for example in the multiplicity of evaluation and measurement procedures associated with the peer review system in science: journal peer review can be seen as one of the cornerstones of evaluation in science (Weller, 2001; Biagioli, 2002; Reinhart, 2012). However, the value measured in this practice remains somewhat ambiguous and is variously interpreted as scientific quality, suitability for a specific journal, or as a measure of the ability of the respective article to be consensual. This value-measurement link is usually discussed along the terms fairness (bias), reliability and validity (Cicchetti, 1991; Daniel, 1993; Reinhart & Sirtes, 2006). Further evaluative practices however frequently make use of peer-reviewed articles as a measure of scientific quality, for example when the number of peer-reviewed papers serves as a selection criterion in hiring decisions or in the allocation of funds (Biester & Flink, 2015; De Rijcke et al., 2016). At the same time, peer-reviewed articles serve as building blocks for more elaborated measures such as citation indices, which themselves again give rise to ensuing measurements. Such is the case when journal impact factors are used to create journal rankings (for example the SCImago Journal Rank). Journal impact factors are also frequently the object of ensuing measurements and calculations concerning questions such as the precision and reliability of such metrics (for example Chen, Jen & Wu, 2014; Donner, 2018; Moed et al., 1996; Van Leeuwen & Moed, 2005). In this case, the value-measurement links progressively lose sight of their original artefact and substitute it with other value-measurement links. In this example, the elimination is complete when the term 'journal' is dropped entirely, and the value-measurement link is only described as 'impact factor'.

Hence the chains of measurement can cloud the processes of operationalization and nomination. For example, when the relative amount of rejections is measured, it then serves to enable an inter-journal comparison that merges into rankings. Subsequently, the rankings are used to estimate the quality of a publication. Hence the resulting topoi is a

culmination of preceding measurement and evaluation practices that are again associated with values, such as importance, reputation and impact. This shows how measurements can transform and overstrike the former processes of nomination and value attribution and dissolve the evaluation.

This tendency of measurements to multiply, to create new measurements and evaluations, and thus to move from one environment into another environment is both pervasive and highly consequential. In connections and chains of measure-value linkages, it frequently happens that existing and previous interpretations influence the situational production of linkages:

In other words, as a sign producer moves into greater facility with existing semiotic systems, the production of signs takes place in a situation of ever increasing tension between the meanings of existing signs, [...] and the producer's need or wish to produce new signs. (Kress, 1993: 173).

Consequently, terms themselves are always already charged with value, carry certain connotations with them and evoke associations that can vary depending on the current fashion and how they are situated in current discourses and symbolic struggles. Terms can hence be formed, become broader or narrower, controversial or one-sided (e.g. pejorative). The inception of the journal impact factor provides an instructive case here: by choosing the term '(journal) impact factor', the newfound metric carried with it connotations of specific values and meanings from its very beginning. Impact is a concept commonly used in mechanics (e.g. Stronge, 2004), to describe the force or shock resulting from the collision of two or more entities. Through its connection to physics as the motherland of measurement, this term carries with it connotations of accuracy and objectivity, and claims a specific authority.¹⁰ At the same time, the metaphors of force, shock, and collision invoke a somewhat violent imagery, for example when compared to the much more benignly termed 'Article Influence Score', and might even be read as military metaphors (Sontag, 2013). In trying to maximise the force or shock with which they hit their targeted audience, journals and authors almost seem to want to cause serious damage to their readership.¹¹ It appears curious as to how such a violent metaphor is turned into something good and desirable in the field of scientific publishing. Presumably, a different nomenclature, such as 'citation average factor' would have brought with it a different tone for the metric, and might have also differently influenced its further uses.

Power effects of chains of evaluations

As argued above, the power relations at play in practices of wording and defining limit the possibilities of interpretation and the ways of speaking. Power thus constitutes the trace of a reality that is difficult to eliminate from the precession of signs and that may keep chains of evaluations from becoming, in the language of Baudrillard, simulations. At first glance, it appears that the practices of measurement – as first steps towards a calculative representation – have a momentum of their own, when they are decoupled from the artefact or field they seek to describe by measuring and counting (Vormbusch, 2012: 223ff). This tendency is reinforced by the advancing technological development

that includes an overabundance of data and that facilitates and exponentially accelerates calculating and computing practices. Multiple measurements are used in politically and economically contested fields where they can be used to allocate resources competitively. For example, they answer to the need for key figures and indicators to categorise and rank organisations such as universities, research achievements, etc. on the basis of comparative measurements. They thus lend themselves to the requirements of science policies rather easily. Thus, these dynamics of measurements can feed into a spiral of evaluations. This way, as long as this connection exists, is maintained and rhetorically nurtured, alternative measurements are attractive and requested, since they provide different value calculations and thus have the potential to modify the description and diagnosis of the competitive situation. Even if in scientific peer review, executed by leading national or supranational research centres, funding agencies, and other departments concerned with science policy, such as the European Research Council (ERC), or the European Science Foundation (ESF), these correlated measurements are mostly used in combination with qualitative evaluation practices, they provide value-measurement links that are effective political tools for governance in and of science. However, taking indetermination and the interpretative scope of nomination and operationalisation seriously, there is also a chance to produce linkages with breathing space and the possibility of alternative and perhaps resistant re-readings, re-uses and re-configurations of measurements and the values they are supposed to link to.

Shifts in interpretations will appear particularly pronounced when value-measurement links move into new environments, and chains of evaluations begin to span multiple fields, such as academia, science policy, and economy. Just as value-measurement links are always shaped by existing power relations and feed back into them, the unfolding chains of evaluations are also influenced by aspects of power. Re-negotiations of evaluations may be motivated by shifts in political interests or by struggles for interpretative authority. Different fields exhibit different power structures, different interests, and different pre-existing interpretations of specific terms. They will thus use and interpret measurements differently from previous fields, sometimes leading to considerable re-negotiation of existing value-measurement links.

Conclusion

In this paper, we have argued that evaluations link measures and values in at least three different forms, described above as operationalisation, nomination and indetermination. Operationalisation provides a specific measure of a value and thereby produces a meaning as particular as the measure. Nomination attributes a meaning to a value-measurement link, whose specificity depends on the terminology, and the interplay of its denotation and the respective spectrum of connotation at a given historical point in time. Indetermination in contrast provides under-determination, it opens up a horizon of diverse possible interpretations: either by leaving the link unspecified, or by blurring the value and / or the measurement in the course of their linkage. Hence the three forms slightly vary in respect to their particular evaluative potential. Thereby the three forms of value-measurement links facilitate different ways of functioning in political use. Whereas operationalisation may often aim at producing a precise meaning, indetermination when

used as a tool for governance leans towards intended blur. This political potential of indetermination is similarly highlighted for the topos of quality by Dahler-Larsen as well, when he emphasizes in which way the vagueness of a concept can be used as a gateway for measurements that promise to fill a leeway by calibrating the undetermined space. He claims that: ‘Conceptual unclarity does not make measurement difficult. Instead, measurement controls quality in a way that the concept itself is too confused to do. In fact, the social function of measurement expands exactly because measurement plays a key role in regulating the meanings and social implications of an otherwise elusive phenomenon’ (Dahler-Larsen, 2019: 13). This diagnosis provides an example of how indeterminate value-measurement links as language-based practices can operate as tools for governance. Nomination however can serve both aims –precision and vagueness, depending on the choice of words and the rhetorical skills of the actors in the symbolic power game on the linguistic market. In addition, these forms can be strung together and chained up in diverse ways and thereby create, dissolve and re-create further evaluation practices and procedures that can take effect in political use.

We consider this paper as a conceptual contribution, which tries to develop theoretically fruitful analytical instruments. Building on this conceptualisation, three specific investigation perspectives can be imagined in which our instruments could be used and tested as heuristics for exploration. First, a historical-genealogical analysis perspective could explore the genesis of measures and their attained meaningfulness or meaninglessness. Second, a sociologically pragmatic analysis perspective can take measurements as the starting point and from this opens up specific contexts perceived as ‘environments’, ‘constellations’ ‘situations’ ‘configurations’, ‘territories’, ‘assemblages’ or else in relation to values and valuation, calculation, quantification, narration, representation, simulation and many other conceivable aspects in order to identify mutual condition factors. Third, a socio-theoretical analysis perspective could aim at studying the effects, performances, and socio-political and socio-economic consequences of measurements. Such a perspective will try to investigate how and which dynamics the measures and metrics performatively create in their respective fields of use and in which way measurements contribute to transform and configure power relations.

Furthermore, value-measurement links comprise traits of counting and calculation, that can produce numerical self-description and self-observations, which can provide a basis for controlling interventions in an environment perceived as contingent (Cevolini, 2014; Vormbusch, 2018b). They thus include the ability to produce numerocratic and highly competitive regimes of governance. Furthermore, numbering and quantification can be seen as devices for inscription and stabilisation that aim at fixating facts and consequently power effects. However, value-measurement links also embody characteristics of language, in particular the malleability of meaning and the inherent potential for interpretation and re-interpretation. As Quéré puts it: ‘Behind words and concepts there are not essences, but, to use William James’ vocabulary, the concrete pulsations of experience’ (Quéré, 2015: 165). This openness of language persistently undercuts the tendency of numbers to inscribe and fixate. It is precisely this interplay between words and numbers that functions as the driver of the genesis and movement of chains of evaluation.

As we have shown, value-measurement links are inherently open for re-negotiation and can be seen as technologies that enable the production of new and ever changing

representations and interpretations of the world (Vormbusch, 2018a: 103). It is thus this openness of language and meaning-making that establishes a route for contestation and resistance against the looming power effects implied in quantification and measurement. Creative uses of language in the context of value-measurement links always include the possibility for ‘the coexistence of multiple matrices of evaluation’ (Lamont, 2012: 202) and thus resist simple hierarchisation and a seamless, totalising ‘governance by numbers’ (Heintz, 2008). Given the relevance of counting, calculation, and the importance of language in the struggle over divergent interpretations and meanings, the political dimension of evaluations also needs to be considered as weighty decision mechanisms and tools for governance.

Consequently, ‘language and action need to be released from strict formalization’ (August, 2018: 149) for only then can we enjoy the chances and the liberating potential of contingency, which are inherent to enduring power struggles, and keep power relations open for revision. Therefore, with a view to the numerous analyses of value, measures, rankings, metrics and evaluations, it would be desirable to regard language as a gateway that is always open, and as a central instrument for configuring the contested relations of power.

Acknowledgements

The authors would like to thank Stephan Gauch for invaluable feedback on an earlier version of this manuscript and Jens Ambrasat for important remarks and discussions. We would also like to thank the participants of the Workshop ‘Evaluation Practices in Science and Higher Education’ as well as Andrea Brighenti, the participants of the Session ‘Theorising Measures, Rankings and Metrics’ and the anonymous reviewers for helpful suggestions.

Funding

This research was partly supported by the Bundesministerium für Bildung und Forschung, grant no. 01PQ16003.


Notes

1. An exemplary case for this pervasive tendency of measurement is the increasing proliferation and use of science metrics, such as citation impact metrics, whose problematic effects bibliometric research is also aware of (Jappe, Pithan & Heinze, 2018:,2; 21-23) and which reflexive bibliometric is concerned with (Dr. Stephan Gauch, personal communication).
2. Wittgenstein’s ‘language games’ are endlessly diverse and variable; their common denominator lies in their character of a game as an ordered practice in which rules are followed.
3. Translated from German by the authors.
4. This understanding also coincides with Abbott’s comprehension that measurement creates ‘a formal relation between differences in some aspect of reality and either an ordered set of numbers or a set of categories’ (Abbott, 1997: 358).
5. Unlike Boltanski and Esquerre in their recent book *Bereicherung* (2018), we are not concerned with the value of merchandise, which obviously refers to an empirical object. Besides they focus on the relation of price and value, in which the value is an instrument to justify the price and therefore is downstreamed to the price. Moreover, according to their conception, the relation of value and price for them is mediated by the metaprice, which provides a numerical estimate of the value (Boltanski & Esquerre, 2018: 145–150, 161–162ff., 184–189).

Therefore, in contrast to what the nomenclature may suggest, the four forms of valuing objects ('Wertermittlungsformen'; 'forme de mise en valeur des objets'), distinguished by Boltanski & Esquerre, indicate how the so-called metaprice is justified and thereby resort again to the orders of justification developed in 'On Justification' (Boltanski & Thévenot, 2006), but do not specify the concept of value in a way we could draw on for our conceptualisation of value-measurement links.

6. See also Abbott's following explanation: 'Meaning is a two-way relation. Concept denotes content but at the same time contents define concepts' (Abbott, 2001: 19).
7. The critical investigation on quality by Peter Dahler-Larsen points in a similar direction. He states: 'quality is not a thing with an internal, essential substance. Instead it is a concept the meaning of which is derived from the social use made of it in specific contexts' (Dahler-Larsen, 2019: 154). In order to take this praxeological dimension of meaning into account, he advocates, also inspired by Dewey (1931), to focus on 'verbs and social processes' instead of nouns, and pleads for the term 'qualitization' that 'deals not only with measurement of quality but also with various assumptions about what makes measurement trustworthy, and even more importantly, what consequences should the resulting notion of quality have for various categories of people' (Dahler-Larsen, 2019: 154).
8. It is interesting to note how the journal impact factor, which in itself is not necessarily competitive, is mentioned alongside a ranking and thus serves to reinforce the notion of publishing as a (highly) competitive field.
9. Thus, indetermined measures might also be said to bear a resemblance to floating signifiers (Laclau & Mouffe, 1985).
10. It is interesting to note that other metrics, such as the Eigenfactor, also uphold this terminological connection to the natural sciences, especially physics.
11. We, of course, wholeheartedly wish our readers well.

ORCID iD

Felicitas Hesselmann  <https://orcid.org/0000-0002-5902-0511>

References

- Abbott A (1997) Seven types of ambiguity. *Theory and Society* 26(2/3): 357–391.
- Abbott A (2001) *Chaos of Disciplines*. Chicago: University of Chicago Press.
- Angermüller J (2011) Wissenschaft zählen. Regieren im digitalen Panopticon. In: Hempel L, Krasmann S, Bröckling U (eds) *Sichtbarkeitsregime: Überwachung, Sicherheit und Privatheit im 21. Jahrhundert*. Wiesbaden: VS Verlag, 174–190.
- August V (2018) *Transparenz: Schlüsselbegriff einer politischen Anthropologie der Gegenwart*. Berlin: Panama Verlag.
- Batagelj V, Ferligoj A, Squazzoni F (2017) The emergence of a field: A network analysis of research on peer review. *Scientometrics* 113(1): 503–532.
- Baudrillard J (1994) *Simulacra and Simulation*. Ann Arbor: University of Michigan Press.
- Biagioli M (2002) From book censorship to academic peer review. *Emergences: Journal for The Study of Media & Composite Cultures* 12(1): 11–45.
- Biester C, Flink T (2015) The elusive effectiveness of performance measurement in science: Insights from a German university. In: Welpel IM, Wollersheim J, Ringelhan S, et al. (eds) *Incentives and Performance*. Cham: Springer International Publishing, 397–412.
- Boltanski L, Thévenot L (2006) *On Justification: Economies of Worth*. Princeton: Princeton University Press.
- Boltanski L, Esquerre A (2018) *Bereicherung. Eine Kritik der Ware*. Berlin: Suhrkamp.

- Borsboom D, Mellenbergh GJ, Van Heerden J (2004) The concept of validity. *Psychological Review* 111(4): 1061–1071.
- Bourdieu P (2005) *Was heißt sprechen? zur Ökonomie des sprachlichen Tausches*. 2., erw. und überarb. Aufl., unveränd. Nachdr. der 2. Aufl. Wien: New Academic Press.
- Bowker GC, Star SL (1999) *Sorting Things Out: Classification and its Consequences*. Cambridge, London: The MIT Press.
- Brighenti AM (2012) Der Aufstieg indexierter Sichtbarkeiten. *Nebulosa – Zeitschrift für Sichtbarkeit und Sozialität* 1: 16–32.
- Brighenti AM (2018) The social life of measures: Conceptualizing measure–value environments. *Theory, Culture & Society* 35(1): 23–44.
- Callaway E (2016) Beat it, impact factor! Publishing elite turns against controversial metric. *Nature News* 535(7611): 210–211.
- Cevoloni A (2014) *Die Ordnung des Kontingenten Beiträge zur zahlenmäßigen Selbstbeschreibung der modernen Gesellschaft*. Dordrecht: Springer.
- Chandler D (2002) *Semiotics: The Basics*. London, New York: Routledge.
- Chen K-M, Jen T-H, Wu M (2014) Estimating the accuracies of journal impact factor through bootstrap. *Journal of Informetrics* 8(1): 181–196.
- Cicchetti DV (1991) The reliability of peer review for manuscript and grant submissions: A cross-disciplinary investigation. *Behavioral and Brain Sciences* 14(1): 119–135.
- Curtin University (2019) Measure research impact and quality. Available at: <http://libguides.library.curtin.edu.au/c.php?g=202403&p=1332898>
- Dahler-Larsen P (2019) *Quality. From Plato to Performance*. Cham: Palgrave Macmillan.
- Daniel HD (1993) *Guardians of Science*. New York: VCH.
- Dannenberg PA (2017) Auf der Suche nach der verlorenen Qualität. *duz - unabhängige deutsche Universitätszeitung - Magazin für Forscher und Wissenschaftsmanager*. Available at: <http://www.duz.de/duz-magazin/2017/06/auf-der-suche-nach-der-verlorenen-qualitaet/434>
- De Rijcke S, Wouters PF, Rushforth AD, et al. (2016) Evaluation practices and effects of indicator use. A literature review. *Research Evaluation* 25(2): 161–169.
- Deledalle G (2001) *Charles S. Peirce's Philosophy of Signs: Essays in Comparative Semiotics*. Bloomington: Indiana University Press.
- Dewey J (1931) *Philosophy and Civilization*. New York: G.P. Putnam's Sons.
- Dewey J (1939) *Theory of Valuation*. Chicago: University of Chicago Press.
- Donner P (2018) Effect of publication month on citation impact. *Journal of Informetrics* 12(1): 330–343.
- DORA (2013) San Francisco Declaration on Research Assessment (DORA). Available at: <https://sfedora.org/>
- Eco U (1985) *Über Gott und die Welt: Essays und Glossen*. München: Car Hanser Verlag.
- Fiske J (1989) *Understanding Popular Culture*. London: Routledge.
- Flaherty C (2017) Peer review's give-and-take. *Inside Higher Education*. Available at: <https://www.insidehighered.com/news/2017/10/24/maybe-there-isnt-peer-review-crisis-least-terms-quantity>
- Garfield E (2006) The history and meaning of the journal Impact Factor. *Journal of the American Medical Association* 295(1): 90–93.
- Hall S (2006) Encoding/decoding. In: Durham MG, Kellner DM (eds) *Media and Cultural Studies: Keywords*. Malden: Blackwell Publishers, 163–173.
- Heintz B (2008) Governance by numbers. Zum Zusammenhang von Quantifizierung und Globalisierung am Beispiel der Hochschulpolitik. In: Schuppert GF, Voßkuhl A (eds) *Governance von und durch Wissen. Schriften zur Governance-Forschung*. Vol 12. Baden-Baden: Nomos, 110–128.

- Hodge B, Kress GR (1988) *Social Semiotics*. Ithaca: Cornell University Press.
- Jappe A, Pithan D, Heinze T (2018) Does bibliometric research confer legitimacy to research assessment practice? A sociological study of reputational control, 1972–2016. *PLOS ONE* 13(6): e0199031.
- Kress G (1993) Against arbitrariness: The social production of the sign as a foundational issue in critical discourse analysis. *Discourse & Society* 4(2): 169–191.
- Krüger AK, Reinhart M (2017) Theories of valuation – Building blocks for conceptualizing valuation between practice and structure. *Historical Social Research* 42(1): 263–285.
- Krueger AK, Reinhart M (2018) Emotional value attribution and comparative value assessment – Analytical elements for a sociology of valuation and evaluation. Paper, Humboldt Universität zu Berlin, Germany, July. Available at: <https://osf.io/preprints/socarxiv/huwk3/>
- Laclau E, Mouffe C (1985) *Hegemony and Socialist Strategy: Towards a Radical Democratic Politics*. New York: Verso.
- Lamont M (2012) Toward a comparative sociology of valuation and evaluation. *Annual Review of Sociology* 38(1): 201–221.
- Lather P (1993) Fertile obsession: Validity after poststructuralism. *The Sociological Quarterly* 34(4): 673–693.
- Moed HF, Van Leeuwen TN, Reedijk J (1996) A critical analysis of the journal impact factors of *Angewandte Chemie* and the journal of the American Chemical Society inaccuracies in published impact factors based on overall citations only. *Scientometrics* 37(1): 105–116.
- Peirce CS (1974) *Collected Papers of Charles Sanders Peirce, Volume I and II: Principles of Philosophy and Elements of Logic*. Cambridge: Harvard University Press.
- Peirce CS (1986) *Writings of Charles S. Peirce: A Chronological Edition, Volume 3: 1872–1878*. Bloomington: Indiana University Press.
- Power M (2004) Counting, control and calculation: Reflections on measuring and management. *Human Relations* 57(6): 765–783.
- Quérel L (2015) Value as a social fact: An adverbial approach. *Human Studies* 38(1): 157–177.
- Reinhart M (2012) *Soziologie und Epistemologie des Peer Review*. Baden-Baden: Nomos.
- Reinhart M, Sirtes D (2006) Wieviel Intransparenz ist für Entscheidungen über exzellente Wissenschaft notwendig? *IfQ Working Paper* (1): 27–36.
- Schendzielorz C (2011) Anerkennung im Sprechen. Eine theoretische und empirische. Analyse der sozialen Dimension des Sprechens. Frankfurt a. M. Available at: <https://freidok.uni-freiburg.de/data/149356>
- Seglen PO (1997) Why the impact factor of journals should not be used for evaluating research. *British Medical Journal* 314(7079): 498–502.
- Sontag S (2013) *Illness as Metaphor and AIDS and its Metaphors*. London: Penguin.
- Stronge WJ (2004) *Impact Mechanics*. Cambridge: Cambridge University Press.
- University of New England (2019) Journal Quality. University of New England. Available at: <https://www.une.edu.au/library/support/eskills-plus/mastering-the-academic-literature/journal-quality>
- Van Leeuwen TN (2005) *Introducing Social Semiotics*. London; New York: Routledge.
- Van Leeuwen TN, Moed HF (2005) Characteristics of journal impact factors: The effects of uncitedness and citation distribution on the understanding of journal impact factors. *Scientometrics* 63(2): 357–371.
- Vannini P (2004) Toward an interpretive analytics of the sign: Interactionism, power, and semiosis (and George W. Bush). *Studies in Symbolic Interaction* 27. Bingley, UK: Emerald Group Publishing Limited, 149–174.
- Vannini P (2007) Social semiotics and fieldwork: Method and analytics. *Qualitative Inquiry* 13(1): 113–140.

- Vormbusch U (2012) *Die Herrschaft der Zahlen: zur Kalkulation des Sozialen in der kapitalistischen Moderne*. Frankfurt: Campus Verlag.
- Vormbusch U (2018a) Performative Entdeckungsverfahren und die Krise von Wert. In: Beyer J, Senge K (eds) *Finanzmarktsoziologie*. Springer VS, Wiesbaden, 93–106.
- Vormbusch U (2018b) Quantifizierung, Autonomie und die Zukunft des Sozialen. *Soziologische Revue* 41(2): 278–286.
- Weller AC (2001) *Editorial Peer Review: Its Strengths and Weaknesses*. ASIST monograph series. Medford: Information Today.
- Wilsdon J, Allen L, Belfiore E, et al. (2015) *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*. Technical Report. DOI: 10.13140/rg.2.1.4929.1363
- Wissenschaftsrat (2017) *Begutachtungen im Wissenschaftssystem. Positionspapier*. Berlin. Available at: <https://www.wissenschaftsrat.de/download/archiv/6680-17.pdf>
- Wittgenstein L (1971) *Philosophische Untersuchungen I*. Frankfurt: Suhrkamp.

Author biographies

Felicitas Hesselmann studied sociology at the TU Dresden and the University of Mannheim. She works at the German Centre for Higher Education Research and Science Studies and is a research assistant at the Humboldt University of Berlin.

Cornelia Schendzielorz studied sociology at the University of Freiburg and the University of Bordeaux. Her doctoral thesis was on her work on subjectivation in vocational soft-skill trainings conducted at the Centre Marc Bloch in Berlin. She is a research associate at the German Center for Higher Education Research and Science Studies and a research assistant at Humboldt University of Berlin.