



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Coordination and obsolescence

Citation for published version:

Wolff, JE 2023, 'Coordination and obsolescence: A response on behalf of measurement realism', *Synthese*, vol. 201, no. 3, 78. <https://doi.org/10.1007/s11229-023-04061-1>

Digital Object Identifier (DOI):

[10.1007/s11229-023-04061-1](https://doi.org/10.1007/s11229-023-04061-1)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

Synthese

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.





Coordination and obsolescence: a response on behalf of measurement realism

J. E. Wolff¹

Received: 13 April 2021 / Accepted: 24 January 2023

© The Author(s) 2023

Abstract

Measurement realism, the view that measurement targets quantitative attributes and that not all attributes are quantitative, has come under attack both from metrologists and philosophers. In this paper, I take a close look at two influential arguments against measurement realism: the argument from obsolescence and the argument from coordination. I concede that these arguments do challenge the epistemological position traditionally taken by measurement realists, but argue that the metaphysical core of measurement realism survives the challenge posed by these arguments. This metaphysical core is vital to maintaining a clear and ambitious standard for successful measurement.

Keywords Measurement · Realism · Coordination · Quantities · Conventionalism · Numerical representation · Coherentism · Operationalism

1 Introduction

My goal in this paper is to clarify two arguments against measurement realism—the argument from obsolescence and the argument from coordination—and to respond on behalf of measurement realism. These arguments play an important role in the recent resurgence of measurement antirealism (Chang, 2004; Tal, 2018; Van Fraassen, 2008).¹ As I show in this paper, however, the two arguments are less damning than they

¹ These are not the only arguments developed by measurement antirealists. Another set of arguments targets the idea of a true value for attributes (Teller, 2018). Since my paper is concerned primarily with the question of quantitative structure for attributes, not accuracy, I will focus only on the arguments from coordination and obsolescence here.

✉ J. E. Wolff
jo.wolff@ed.ac.uk

¹ School of Philosophy, Psychology, and Language Sciences, University of Edinburgh, 3 Charles Street, Edinburgh EH8 9AD, UK

appear to be. While they undermine traditional epistemological commitments commonly made by measurement realists, the more important metaphysical commitments of realism can survive the challenges.

Measurement realism and anti-realism come in a variety of flavours and disagree on a number of different points. I shall focus on the question of whether quantitateness is an objective feature of attributes, or whether quantitateness is created in the process of operationalising a concept. Measurement realists tend to hold a restrictive view of quantitateness, according to which only some attributes are quantitative and their quantitateness is a matter of having a certain kind of structure. Measurement anti-realists, by contrast, tend to hold a less restrictive view of quantitateness and often reject the idea that quantitateness is a mind-independent feature of attributes. Where realists tend to draw sharp distinctions between what is out there in the world (the attribute), what we do to find out about it (our measurement procedure), and how we represent our findings (our numerical representation of measurement outcomes), antirealists typically see these three elements as not clearly distinct. Measurement realists are committed both to the existence of quantitative attributes, and to our ability to find out, whether an attribute is quantitative. Two arguments in particular have been used to challenge these commitments of measurement realism: the argument from obsolescence and the argument from coordination. I argue that while these arguments pose a serious challenge to measurement realism, they ultimately undermine only the traditional epistemology of measurement realism, not the metaphysical thesis of realism.

The paper is organised as follows. Section 2 introduces the commitments of measurement realists and antirealists in more detail. Section 3 presents the argument from coordination. Section 4 responds to the argument from coordination on behalf of measurement realism, showing that the challenge is less severe than it seems to be. Section 5 presents the argument from obsolescence and investigates the prospects of a recently proposed compromise to resolve the argument from obsolescence. Section 6 proposes to revise the realist's epistemology while retaining the realist's metaphysical commitment to respond to the challenge from obsolescence. Section 7 concludes.

2 Measurement realism and antirealism

2.1 Measurement realism

Measurement realism can be understood as having metaphysical, semantic, and epistemic commitments.² Metaphysically, measurement realists hold that we measure attributes and that these attributes are independent of our methods of measuring them. In particular, these attributes have quantitative structure and definite magnitudes independent of our measurement of them (Mari & Giordani, 2014; Michell, 2004, 2005). Semantically, this suggests a commitment to the idea that theoretical quantity terms, such as 'temperature' or 'mass', purport to refer to attributes that are in themselves

² A similar differentiation of scientific realism into epistemic, semantic, and metaphysical components can be found in Psillos (1999, p. 19).

quantitative, independent of our measuring them or representing them numerically. Epistemically, measurement realists hold that we can find out whether an attribute is quantitative, as well as estimate the value of a quantity to increasing degrees of precision.

The epistemic commitments of measurement realism are usually understood along broadly empiricist lines (Campbell, 1920). Not only are we able to find out whether an attribute is quantitative, we are able to do so by empirical means: we should be able to test for the quantitateness of attributes. Such tests might proceed directly, as in the comparison of two lengths placed alongside one another, or they might proceed indirectly, exploiting the nomic relationships between several quantities, such as kinetic energy, momentum, and velocity (Michell, 1999).

Different measurement realists offer slightly different accounts of the exact ontological commitments that come with this picture of measurement. Some insist that realism commits us to universals (Mari & Giordani, 2014), or to numbers as well as universals (Michell, 1999). Wolff (2020) argues that these commitments go beyond what is required for measurement realism. The central commitment of measurement realism is that quantitative attributes form a distinctive class of attributes, and that these attributes have their quantitative structure independent of our ability to measure them or to represent them numerically (also see (Mari et al., 2017a, 2017b) for this way of presenting measurement realism). Measurement targets quantities, but not other attributes. This requires a sharp distinction between quantitative and non-quantitative attributes. A traditional way to characterise quantitative attributes is to say that these attributes have ratio structure, or permit ratio comparisons (Michell, 2004). For example, mass is a quantitative attribute, because it permits the formulation of mass ratios, e.g., a is three times as massive as b . By contrast, spiciness is not an attribute that permits ratios: a may be spicier than b , but a is not three times as spicy as b . This difference is usually captured by saying that spiciness has ordinal structure, but not ratio structure.

Restricting quantitateness to the presence of ratio structure, however, may seem overly restrictive. After all, some physical attributes, like temperature, were initially represented as having interval structure: differences between values on the Celsius scale are meaningful, but their ratios are not. The traditional realist might seem to face a dilemma, then: either admit that the distinction between quantitative and non-quantitative attributes is not sharp, because there are in between cases like temperature, or else deny that temperature and other attributes represented on interval scales are quantitative.³

Advances in formal measurement theory, however, have offered realists a third option: restrict quantitateness to attributes representable on interval and ratio scales, by contrast to attributes representable only on ordinal scales. The reason this has become an attractive option is the increasingly abstract nature of formal measurement theory. Formal results, like the Alper-Narens theorem (Alper, 1985; Narens, 1981, 2012/2002 ch 5) indicate that at least for continuous quantities, interval and ratio

³ The Kelvin scale is of course a ratio scale representation of temperature, as it has a meaningful zero point. Temperature today is hence unproblematic even on a ratio scale only view of quantities. The strength of earthquakes, by contrast, remains interval only, as the Richter scale is an interval scale. I'm sticking to temperature as the example, for its historical importance in this literature and its familiarity.

scale representations are more similar in their formal features to one another, than either is to ordinal scales. This suggests that there is a significant difference between ordinal scales on the one hand, and interval and ratio scales on the other, contra the suggestion that there is nothing distinctive about quantitative attributes since anything can be represented by numbers. What makes quantities distinctive is not representation by numbers, but having a particularly rich structure, which satisfies the conditions for representation at interval or ratio scale (Wolff, 2020). Not all attributes can be assumed to have such a structure. Formally at least, there seems to be a clear difference between quantitative and non-quantitative attributes. A measurement realist, then, is somebody who holds that quantitateness is a feature of attributes that can be represented on interval and ratio scales, but not a feature of attributes representable only on ordinal scales.

2.2 Measurement antirealism

Measurement antirealism today finds inspiration in the operationalist tradition. The key anti-realist element in these views is that what we measure and how we measure it do not come apart. An attribute and its structure are not entirely independent from the way in which we measure and represent the attribute and its structure. Sophisticated operationalists combine two traditional operationalist lessons. First, following Bridgman, they seek to limit the semantic reach of theoretical concepts—in Bridgman’s notorious formulation: “In general, we mean by any concept nothing more than a set of operations; *the concept is synonymous with the corresponding set of operations*” (Bridgman, 1958, p. 5). Since nothing can ensure that measurements taken with a mercury thermometer and measurements taken with an air thermometer take measurements of the same attribute, we should carefully limit the semantic reach of concepts such as *temperature* to particular measurement operations. A second pillar of operationalism, due to S.S. Stevens (1946), is the rejection of restrictivism: there is nothing special about attributes that permit a ratio representation. Instead, any systematic assignment of numbers to objects is to count as measurement.⁴ As a result, measurement is not restricted to a distinctive class of quantitative attributes. Ordering chilis according to their spiciness counts as a measurement just as much as using a spring balance to determine the weight of baking ingredients.

Combining these two elements of operationalism, one might suggest measuring student satisfaction by comparing the scores of different courses on a particular questionnaire. For an operationalist, the comparison, even if only ordinal, would count as a measurement. The worry that there might be more to student satisfaction than what is captured on a single questionnaire can be set aside first, by distinguishing the theoretical concept *student satisfaction* from its ordinary language counterpart, and secondly, by restricting the meaning of the theoretical concept *student satisfaction* to the particular measurement method. Since there might be more than one (proposed)

⁴ Of course, there is more to Bridgman’s and Stevens’ views than I have space to discuss here. For recent re-evaluations of operationalism, see (Chang, 2009; Feest, 2005; Hardcastle, 1995). I choose these two aspects of traditional operationalism as a starting point, because they form the origin of the contemporary antirealist position. The position and arguments I aim to elucidate in this paper are not Bridgman’s or Stevens’, but those of contemporary measurement anti-realists.

measurement method for student satisfaction in the pre-theoretical sense, this might lead to a plurality of theoretical concepts: *student satisfaction*₁, *student satisfaction*₂ and so forth, each indexed to a particular measurement method. For the operationalist, this is a welcome result, as it reminds us to be cautious in the conclusions we wish to draw from our measurements (Chang, 2004, pp. 141–158). We cannot simply assume that two questionnaires are both measures of *student satisfaction simpliciter*, nor that we can simply combine measurement methods across different subject areas or universities.

Sophisticated operationalists, like Chang, agree with Bridgman that theoretical concepts need to be given specific conditions for operationalisation, although for Chang the relationship between a concept and its operationalisation is typically mediated. Where Bridgman, at least at first pass, takes an operationalisation to be a specific physical measurement procedure (to be carried out in the laboratory), Chang takes the operationalisation of theoretical concepts to be their link to empirical concepts, a connection that is mediated by models (Chang, 2004, p. 207). Chang thereby takes on board two important commitments of operationalism: First, that theoretical concepts must be carefully distinguished from their ordinary or empirical counterparts, i.e., Kelvin's theoretical concept of *temperature*_k must be distinguished from any empirical concept of *temperature*_e involved in particular measurements, at least until reasons have been given for thinking that the two might be the same, or that the former can be operationalised using the latter. Second, operationalisations of theoretical concepts are part of the meaning of these theoretical concepts. Since Chang does not hold that the meaning of a theoretical concept is exhausted by its operationalisations, he can allow for the idea that the same concept may admit of different operationalisations. This allows Chang to sidestep to some extent the concern that there is nothing to unify the range of different theoretical concepts resulting from the different operationalisations. Theoretical concepts receive part of their meaning from their relationship to other theoretical concepts, in sharp contrast with empirical concepts. These relationships (may) remain unchanged under different operationalisations. The result would be a family of theoretical *temperature* concepts, for example, with a core meaning given by the relations between *temperature* and other theoretical concepts, such as *volume* and *pressure*.⁵ Chang's sophisticated operationalism offers a seemingly different semantics for theoretical concepts from that of the measurement realist. Where sophisticated operationalists emphasize the role of the operationalisation in providing (part of) the meaning of theoretical concepts, measurement realists presume, often without detailed argument, that theoretical concepts purport to refer to attributes, only some of which are quantitative.

Sophisticated operationalism further differs from measurement realism in that Chang and others adopt a coherentist epistemology of measurement. Whereas foundationalists take measurement to be justified ultimately through observation, the coherentist rejects such firm observational foundations for measurement as in principle

⁵ This conception of the meaning of theoretical concepts seems particularly appealing in the case of the physical sciences, where rich, law-like relations can often be found, like those between pressure, temperature, and volume. Borsboom has suggested that a similar move might be difficult to defend for concepts in psychology, where such relationships are usually insufficiently secure and rich to serve as partial definitions (Borsboom, 2005, pp. 156–157).

unavailable. As Chang has argued in his reconstruction of the history of the concept of *temperature*, the foundationalist epistemology adopted by many measurement realists cannot do justice to the iterative justification found in the practice of making *temperature* a measurable concept. Where the measurement realist seems to suggest that we first determine whether an attribute is quantitative and then develop a measurement procedure, measurement anti-realists worry that we are never in a position to make such a determination independently of particular operationalisations of a given theoretical concept. Instead of a single, conclusive determination of quantitateness, we find an iterative procedure of theoretical modelling of the target attribute, observational and experimental results on the basis of such modelling, followed by revisions to the initial model, followed by further refinements to the measuring procedure. When things go well, this iterative process leads to gradual improvement through revisions and refinements. The resulting justification for taking a particular concept to be quantitative is coherentist, not foundationalist, though. At no point is there a conclusive empirical test to establish the quantitateness of a concept independent of its operationalisation.

The metaphysical difference between measurement realists and anti-realists is more difficult to pinpoint. While realists have a clear metaphysical picture, distinguishing between the theoretical concept and the quantitative attribute to which the concept refers, anti-realists are less explicit about the metaphysical picture they endorse. Still, some of their remarks suggest that they reject the realist claim that there are distinctively quantitative attributes to which theoretical concepts purportedly refer. Since measurement anti-realists start from concepts that describe the target of measurement, the sense in which such concepts have referents remains underspecified. It seems clear, though, that we should understand Chang to be rejecting the metaphysical picture of the measurement realist, according to which there is an attribute out there, temperature, which already has a quantitative structure. Insofar as we have such quantitative attributes at all, their quantitative structure—or at least their values—are in some sense the result of a successful operationalisation of a concept. What are some of the arguments for this form of measurement antirealism?

3 The argument from coordination

The conclusion of what I will call the *argument from coordination* is that the representation of a physical attribute by means of a numerical structure should not be understood as matching the numerical structure to the structure of that attribute. The argument itself begins from the observation that the numerical representation of an attribute like temperature requires iterative stages of treating phenomena exhibiting the attribute as having a suitable structure. Along the way, numerous problems need to be resolved, as Chang demonstrates in his history of temperature (2004). We need to identify objects or processes of stable temperature, to produce fixed points as reference points for measurement, such as the melting point of ice or the boiling point of water. We need to identify physical processes that permit the observation of systematic temperature variation, like the expansion and contraction of mercury in a column. We need to find a way to represent this variation numerically and to identify, which numerical comparisons are meaningful, for example in deciding whether we should

treat temperature as having an absolute zero point. In each step, suggests the antirealist, we are building on previous choices, while revising them in light of theory, updated on the basis of our previous round of “measurement”.

Antirealists wish to draw a two-fold moral from this account of iterative development: first, we must reject a foundationalist epistemology that holds that we can somehow determine the status of an attribute as quantitative purely from empirical observations. Every step along the way of constructing a measurement procedure is theory-laden and indeed choice-laden: to set up a measurement structure requires choices not determined by the observations themselves (Tal, 2018; Van Fraassen, 2008). Second, we should resist the idea that in setting up a measurement procedure and in formulating our theoretical concepts, we are trying to match the structure of an attribute (or phenomenon) in the world, or as antirealists like to express this idea: matching the structure between our representation and the WORLD (Chang, 2004, pp. 206–207; Van Fraassen, 2008, p. 138). The all-caps ‘WORLD’ is meant to indicate a strictly mind-independent conception of the world. On the contrary, there is never a comparison to be had between ‘the structure of the attribute’ and the operationalisation. All we ever compare are the theoretical concept as we currently understand it and the outcomes of our measurement processes, as we currently model them.⁶ Where there is a (sufficient) match between these two, we can say that an operationalisation has been successful, but this match is created by our modelling as much as it is ‘found’ empirically. The success of an operationalisation therefore cannot be said to depend, antecedently, on there being an attribute with a quantitative structure. Such structure is created in the process of operationalising the abstract concept.

Both morals are aimed at measurement realism. Where the first moral proposes an alternative epistemology for measurement, the second moral rejects the metaphysical commitments of realism: that quantitative structure is a feature of attributes independent of our operationalisation of them.

To what extent are these morals supported by the accounts of the history of science, specifically the history of making temperature measurable? I take it antirealists do not put forward these morals as conclusions of a deductive argument. Rather, they are presented as alternatives that solve a problem that otherwise seems irresolvable. The problem, in short, is this: if we are to understand the success of our (numerical) representations of physical attributes in terms of a match between the structure of the numerical representation and the structure of the attribute, then this success is out of reach, in principle, since we cannot access the structure of the attribute independently of our (iteratively constructed) representations of it. To say that attributes themselves have a quantitative structure is to suggest that our aim should be to match said structure in the construction of our measurements. But since, according to antirealism, we can never compare the structure of the attribute ‘in itself’ to the structure of our representations of it, we cannot ascertain whether we’ve been successful in our attempts. Instead, we can only ever compare different representations, for example based on different measurement procedures, to one another, never to the (alleged) structure of

⁶ This, at any rate, is how I understand Chang, who insists that we cannot understand the success of an operationalisation procedure through a comparison between our abstract concept and the world (Chang, 2004, p. 216).

the attribute. Giving up the idea that it is attributes that have independent quantitative structure frees us from this futile attempt.

In the eyes of its proponents, the argument from coordination offers a principled difficulty for determining the quantitateness of an attribute. Since we cannot compare the structure of the attribute to our theoretical concept without first devising an operationalisation, that is, without first forming a concrete image of the theoretical concept in context with other theoretical concepts and then matching this image to a physical system, there is not much point in insisting that only quantitative attributes can be measured in the strict sense. Not only is there no determination of quantitateness prior to operationalisation, in some sense there is not even a fact of the matter about the quantitateness of an attribute prior to, or independent of, the operationalisation of its concept. If we were to accept that an attribute has quantitative structure independent of our operationalisations of its concept, then we should expect that in order to be successful, an operationalisation needs to match the structure of the attribute. But since we cannot, in principle, establish such a match, because each step along the way is mediated by a model of the attribute and of the measurement process, we must give up this standard of success for operationalisations. But once this standard has been given up, whether or not an attribute has quantitative structure becomes *irrelevant* to the success or legitimacy of an operationalisation. Indeed, it might seem to make sense to shift the seat of quantitateness from attributes to concepts. Concepts are made quantitative by operationalising them, and can be said to have quantitative structure only after the operationalisation has been carried out. This argument, then, targets the *realist* aspect of restrictive realism, not the restrictiveness.

4 Responding to the argument from coordination

The argument from coordination seems to suggest that we should adjust our metaphysics to fit our epistemology—a move generally resisted by realists of any type. Part of the point of being a realist is to insist that what we have evidence for and what is the case can come apart. There might be attributes whose quantitative structure we never discover, and there might be cases where we mistakenly attribute quantitative structure to an attribute, because we falsely believe that we've found a successful measurement procedure. To adjust our standards for success of operationalisations to match our apparent epistemic limitations is to give up the realist's key commitments. What is presented, by the antirealist, as an excellent solution to a serious problem is altogether unacceptable to the realist.

Can realists say more, beyond complaining that the argument is question begging? I think they can. Realists need to show that the challenge of coordinating measurement representation and world is not as daunting as the antirealist suggests, and they need to show that the realist proposal is not an epistemically risky, but optional add-on to the practice of measurement. Let's take these two aspects in turn.

The antirealist argument seems to depend rather heavily on painting the realist as upholding an unmeetable standard for the success of operationalisations, while restricting the means by which the success of operationalisations can be ascertained. The antirealist makes the realist project out to be hopeless, because they seek to

compare directly the structure of the attribute with the structure of the measurement outcomes resulting from the operationalisation. But this is not quite right.

In describing the realist's aim as establishing a match between our operationalisation and the all-caps 'WORLD', anti-realists are indicating a much deeper disagreement between *global* realists and antirealists, not merely between *measurement* realism and antirealism.⁷ A global realist holds that the world possesses some structure independent of our conceptualising it a certain way,⁸ whereas a global antirealist takes the world that is the target of our scientific investigations to be a world already conceptualised by us in some way or another. This disagreement between global realism and antirealism is of great philosophical interest and importance, but it's not a particularly precise weapon to wield against the measurement realist. After all, if *all* local forms of realism are mistaken because global anti-realism requires us to treat all structure as a result of our conceptualising the world in a certain way, then measurement realism simply looks like a regrettable casualty in a much larger war, not really worth singling out for separate refutation.

Instead of trying to resolve this global disagreement, let's focus instead on the specific problems that seem to arise for the matching of *measurement* representations to the world. Restrictive realists like Michell hold that quantitateness is not up to us: attributes either are quantitative, or they are not. An antirealist perspective on measurement looks attractive, because the process of developing measurement procedures and representations involves conventional choices. This raises the question of whether quantitative structure is really something we discover, or whether it isn't perhaps something we create through our choices in the process of measurement. Unlike the broad dispute over global realism, this question is specific to measurement, so I'll focus on this for the remainder of the paper.

Even realists like Michell will concede that our numerical representations of quantities contain excess structure, due to conventional choices. Most obviously this is the case for particular units. The choice of kilogram or pound as a unit for mass is conventional and not an attempt to match the structure of the attribute. There is no fact of the matter as to which of these scales uses the correct unit, even from a realist perspective.⁹

⁷ Some measurement antirealists would perhaps prefer to resist articulating a global metaphysical antirealism, on account of a general suspicion against metaphysics. A global antirealist undercurrent seems to run through both Chang's and van Fraassen's treatment of measurement, however. For a discussion of the latter, in connection with transcendental idealism as a global antirealism, see (Belot, 2010).

⁸ For a recent articulation of an ambitious global realism, see (Sider, 2011, 2020). According to this form of global metaphysical realism, there can in fact be a (ideally: unique) match between the structure of our (linguistic) representations and the mind-independent structure of the world. The extent to which science can help uncover the true structure of the world, let alone the extent to which science has succeed in this, is another matter. Interestingly, Sider singles out quantities as a problem case for his kind of realism (Sider, 2020 chapter 4).

⁹ This marks an interesting difference between the measurement realist and some forms of global realism. From Sider's perspective, the most attractive view given his realist commitments would be one according to which there is a preferred unit—this would be a kind of 'absolutism' about magnitudes. Sider himself does not take that to be the right interpretation of quantities—he opts for a mixed absolutism not dissimilar from that of the standard measurement realist, but concedes that this is not an optimal view given his other realist commitments; indeed his overall conclusion about the options for global realism in a fundamentalist guise are pessimistic when it comes to quantities (Sider, 2020, pp. 165–166).

But Chang, van Fraassen, and Tal want to go further. Our conventional choices not merely infect the particular numerical values we assign to particular objects. Even our interpretation of certain causal relationships as linear, or perhaps indeed the interpretation of two particular temperature indications as the same temperature, are subject to conventional choices (Tal, 2018). These conventional choices are hidden, because they present themselves as adopting the simplest form of the law (linear), or as the most natural choice of sameness (thermal equilibrium), but Tal insists that they should nonetheless be marked as choices. Empirical observations by themselves do not force them upon us.

Here realists face a choice: either they concede the antirealist point that there are additional conventional elements in measurement representations, or they insist that while units are conventional, these other aspects of the measurement representation are not. Unlike in the case of units, measurement realists might wish to hold that there is a fact of the matter as to whether two bodies are equal in temperature (let alone in mass or in length). How can a measurement realist defend this position, while allowing for conventionality in the assignment of units?

The task is perhaps less daunting than the antirealist makes it out to be. For starters, appeal to simplicity and naturalness make a lot of sense, even in the iterative process proposed by coherentists. We begin from what looks simple or natural, and deviate from this course only where subsequent results suggest that we need to make modifications to our procedures. The choices here are far less arbitrary than the choices of units, and we can use fairly general principles to guide those choices. Especially where we are formulating functional laws, what are simpler and what are more complicated formulations of laws will be clear from the type of mathematical equation used. Moreover, which choices can actually be implemented is empirically constrained: practically, we aren't confronted with a wide range of plausible options for what to take as sameness of magnitude for a given quantity, let alone how to formulate law-like relationships between different quantities. Indeed, the examples where a choice seems to be available, such as Brian Ellis' famous 'dinches' example (Ellis, 1966), where concatenation of lengths is implemented not by aligning rods end to end in a line, but end to end at right angles, strike most of us as clearly contrived. We do not, then, confront an obviously arbitrary choice, as we did in the case of units, which provides at least *prima facie* reason for thinking that a realist approach may be appropriate here, where it was inappropriate in the case of units. Like for realists elsewhere in the philosophy of science, this requires adopting an optimistic view of values like simplicity and naturalness—they need to be treated as being truth-conducive. A simple and natural theory, on this view, is not only nicer to work with for us, but more likely to be true. The measurement realists' response is here simply analogous to the general scientific realist response to problems of theory choice. Conversely, antirealists are unlikely going to be moved by this suggestion. Why should what looks simple and natural to us be truth-conducive?

Contemporary realists might hence opt for a different strategy: perhaps they can grant that measurement representations contain conventional elements beyond the choice of units, while holding on to the idea that quantitative structure is a feature of attributes independent of the measurement procedures? Traditional realists held that quantitateness was a matter of having a well-defined concatenation operation on an

attribute, corresponding quite literally to an addition over the real numbers (Campbell, 1920). Where such a concatenation operation was missing, or where multiple candidates for such operations were available, each suggesting a different numerical representation, the attribute could not really be understood to be quantitative. As the examples of temperature and length show, this standard is too restrictive even for physical quantities. Contemporary realists have rejected this very literal reading of the correspondence of our representational structure and the structure of the attribute. The structure an attribute needs to have to qualify as quantitative does not require that a concatenation operation is present, or that a unique such operation is present. Instead, the structure that makes an attribute quantitative can be described more abstractly, so as to allow for non-additive attributes, or attributes admitting of more than one way of being concatenated to count as quantitative (Michell, 1999; Wolff, 2019, 2020). Realists still hold that it is the attributes that have quantitative structure, but this structure can be found through a range of different measurement procedures and can be represented by a range of numerical representations containing arbitrary features. In modern measurement theoretic terms, Ellis' *dinches* are an instance of alternative numerical representations (Krantz et al., 1971, p. 99). While changes in units can be understood as different ways of mapping an empirical structure to the same numerical structure, by changing which empirical object is mapped to which number, alternative numerical representations are the mapping of an empirical structure to a different numerical structure. In the first case, the representations are homomorphically related, in the second case they are not so related. We have a choice, which numerical structure to use to represent a particular quantitative attribute, as well as a choice of unit, but whether an attribute is eligible to be thusly represented is still a matter of whether the attribute has a sufficiently rich structure. To characterise this more abstract structure, we do not specify particular relations of concatenation or multiplication, but rather insist that the structure must exhibit certain automorphisms.

Realists, on this strategy, do not demand a full match between our numerical representation and the attribute it represents. The representational structure is acknowledged to have arbitrary features that outstrip the structure present in the attribute and that are subject to choices. What the realist demands is merely that the attribute has *some* structure of its own. Not in all cases do attributes have quantitative structure—that is, structure representable on ratio or interval scales. Only where the attribute in fact has quantitative structure will the operationalisation yield a fully quantitative numerical representation. That such a representation can be achieved in different ways does not undermine the idea that the attribute already has some structure, even independent of our attempts to represent it numerically. The match that needs to be established is weaker. It is not that the structure of the representation must literally match the structure of the represented attribute; rather, the represented attribute must have enough structure in its own right to warrant the use of *some* quantitative representation or other. This makes room for some aspects of the representation being conventional, while retaining a realist view of attributes as quantitative.

The challenge for realism, then, is not as severe as antirealists make it out to be. Even so, if antirealism provides a better response to the challenge from coordination than realism, antirealism might still seem on balance more attractive. Perhaps realism is merely a risky, but optional add-on to the antirealist view of measurement? But this

is not so. What realism offers is a clear standard for success of an operationalisation: an operationalisation is successful if it correctly identifies an attribute as quantitative and finds a way of systematically and coherently representing it numerically. Coherentists, by contrast, wish to reject the idea that any aspect of operationalisation has to do with matching the structure of an attribute. But then what distinguishes successful from unsuccessful operationalisations? Successful operationalisation cannot merely mean having established, to one's own satisfaction, a systematic assignment of numbers to objects. Coordination, coherentists insist, is difficult. A successful operationalisation, Chang suggests, will be one where we achieve a convergence between different measurement procedures. But we shouldn't conclude from this that we've in fact identified a quantitative attribute, let alone that the convergent values are the true values of that attribute (Chang, 2004, p. 217).

Realists, of course, can readily agree that convergence provides evidence for a successful operationalisation, but unlike Chang, they take such convergence to be only *evidence* for success, not success itself. Convergence across different measurement methods gives realists reason to think that these measurement methods are successful in identifying a quantitative attribute, and indeed successful at re-identifying the same quantitative attribute across the different convergent methods. It is tempting to portray this realist attitude to convergence as a form of inference to the best explanation: realists wish to draw the further inference that the convergence of values shows that we've hit upon some quantitative structure, beyond the more epistemically cautious conclusion that the measurements indeed converge. On this reading, the convergence of values is explained by there being an attribute with the relevant quantitative structure. It is this explanatory power on which the realists' confidence in the existence of a quantitative attribute rests, and which perhaps even invites a stronger inference, to the existence of a true value of the attribute under measurement, which these converging values are in some sense 'approaching'. Such inference to the best explanation is open to objection: sometimes values of attributes appear to converge, yet later on the convergence is found to be spurious, that is, our evidence might have invited such inference to the best explanation where it would have been mistaken. More generally, antirealists tend to be sceptical of inference to the best explanation as a legitimate form of inference and will hence be unlikely to be persuaded by realist appeals to such modes of inference.

There is, however, a different way of understanding the realists' insistence that convergence is evidence for successful measurement, but doesn't make for successful measurement. Convergence does not licence a *new* inference to the existence of a quantitative attribute, but merely means that our evidence is consistent with an implicit presupposition of our measurement practice: that there is a quantitative attribute out there. So long as our measurements are converging, we can maintain this presupposition. This is because if there is a single attribute with a quantitative structure targeted by one or more measurement methods, the results are expected to converge (or perhaps more precisely: any apparent discrepancies must be accounted for). It is only when there is no or insufficient convergence that we have to question this presupposition. The inference is not an ampliative inference to the best explanation, but rather a hypothetico-deductive one: if our measurements are targeting a quantitative attribute, we should expect them to exhibit convergence. If they fail to exhibit convergence, we need to conclude that something has gone wrong. Of course, we don't immediately

know, which of the many assumptions that go into setting up a measurement might have been mistaken. Realists, on this understanding do not infer the existence of a quantitative attribute from the convergence of measurement results, but start from the assumption that there is such an attribute and only revise it in the light of contrary evidence (typically in the form of lack of convergence or difficulty in setting up a measurement procedure).

What is the reason for the realist presupposition that our measurement targets quantitative attributes? After all, since antirealists deny this presupposition, we seem to have reached an impasse, unless we can find further grounds for the realist position. Realists, it seems to me, presuppose that measurement targets quantitative attributes, because this offers a clear condition for success that outstrips our evidence. Chang's coherentist can neither explain why convergence sometimes obtains and sometimes does not, nor why it should be important. This is why we shouldn't simply change our standard for what makes an operationalisation successful in the face of difficulties. Realism provides an account of what *makes* for successful operationalisation, not merely what might indicate that an operationalisation was in fact successful. Giving up this standard seems to protect all attempts of operationalisation against criticism or scrutiny, since the criteria for success seem entirely internal to the particular operationalisation or narrowly defined theoretical concept. Indeed, it is not entirely clear what distinguishes successful from unsuccessful operationalisation attempts. By contrast, the realist requires that the attribute in question in fact has a quantitative structure not created by our operationalisation. Attributing quantitative structure to an property that lacks it, is a mistake, and conversely it is a mistake to deny that a property is quantitative when it has the requisite structure. Perhaps this is only a necessary condition for successful measurement, but it is one which explains why some efforts at operationalisation seem to work better than others: sometimes we really have latched on to a quantitative attribute. By contrast, the operationalist alternative seems like an illegitimate shifting of the goal posts: since the inference from convergence to the existence of an attribute with quantitative structure is fraught, let's dial down our success conditions for measurement.

In summary, realists can respond to the worries about conventionality in measurement representations, either by digging in their heels and insisting that simplicity and naturalness are truth-conducive virtues of representations, or by conceding that numerical representations of quantitative attributes contain conventional elements beyond the choice of unit, while insisting that quantitateness is still a matter of an attribute having a certain structure, quite independent of our measurement procedures. To hold on to this metaphysical commitment of a quantitative structure to be found, not created, is important, because it allows an ambitious standard for the success of operationalisation.

5 The argument from obsolescence

The argument from coordination, then, is not a fatal blow to measurement realism. But there is a second line of argument against measurement realism. It begins from the actual practice of science and targets the restrictiveness implied by measurement

realism. From the perspective of contemporary science and metrology, measurement realism looks a little old-fashioned. Even metrologists unsympathetic to Stevens' thoroughgoing permissivism find that much of what is called measurement today does not aim at attributes that meet the realist's strictures on what may count as a genuine quantity (Mari et al., 2017a, 2017b). Restricting measurement to quantities in this narrow sense would seem to be out of step with developments in science. Especially in the social sciences, 'measurement' of attributes with merely ordinal structure is common and fruitful. By insisting on a narrow conception of quantity, and by claiming that only concepts of quantitative attributes are genuinely measurable, measurement realists are insisting on an outdated model of measurement that is too focused on measurement of physical quantities. This physics centrism is unjustified and might even obstruct scientific progress. Let's call this the argument from obsolescence.

The argument from obsolescence points to a tension between narrow concepts of quantity and measurement on the one hand, and the actual practice of science on the other. To insist on a restrictive conception of measurement is to reject the claim of certain scientific practices to being measurement.¹⁰ To be a restrictive realist is to take a normative stance towards measurement, against what seems to be an overly permissive view taken by Stevens and his followers. The dialectic here is a familiar one in the philosophy of science, between those who wish to impose restrictions on scientific practice, and more naturalistically inclined philosophers, who reject the legitimacy of such philosophically motivated restrictions. Is the argument from obsolescence primarily a problem for restrictive measurement realism, or for measurement realism in general? At first glance, it's a problem for restrictivism, so a recent response to the argument from obsolescence has tried to offer a less restrictive form of measurement realism.

To address the mismatch between the requirements of (social) science, while acknowledging the special status of the particularly rich structures that Michell and others single out as quantitative, Luca Mari and different collaborators have recently proposed to split the difference: quantity remains a protected category, but measurement is allowed to target non-quantitative attributes as well (Mari et al., 2017a, 2017b). This would mean acknowledging that much of social science measurement might not be directed at quantitative attributes, that is, at attributes that have the rich structure exemplified by attributes like mass or temperature. Some of the attributes targeted by social sciences are perhaps merely ordinal, not truly quantitative, but according to Mari et al. that does not mean we should deny that they can be measured.

This is not meant to be a merely verbal fix to a philosophical problem. Mari et al. argue that there is no conceptual link between measurement and quantity: while the Euclidean concept of measure is tied to quantity, measurement, historically, is not. There is hence no conceptual requirement to restrict measurement to quantities only. Instead, measurement is a reliable, trustworthy and accurate tool for producing knowledge. The means by which measurement achieves this status are in principle open to investigations into non-quantitative attributes as well, or so Mari et al. claim. Indeed, they argue, *structurally* social science measurement does not differ much from measurement in the physical sciences (Mari et al., 2017a, 2017b). The structure they

¹⁰ This is indeed a central concern of Michell's critique of measurement in psychology (Michell, 2008).

identify as common in measurement across different sciences are the steps involved in the measurement process.

They identify a number of steps that are involved in the measurement process: “firstly, the specification of the object under measurement, the definition of the considered general property, and the definition of the measurand; second, the specification of the measuring system, including the choice of the measuring instruments and the design of the measurement procedure; third, the modelling activity underlying the measurement execution” (Mari et al., 2017a, 2017b, pp. 48–49). It is telling that the first two of these steps revolve around specifications and definitions: measurement produces reliable and trustworthy results in no small part because of the (sometimes intense) regimentation that goes into setting up the measurement process. Nothing in this structure suggests that such processes are limited to the physical sciences or to quantitative attributes. Mari et al. hence conclude that structurally, measurement of quantitative and non-quantitative attributes proceeds largely alike (Mari et al., 2017a, 2017b). Given such a structural characterisation of the measurement process, then, there is no reason to restrict measurement to quantities. If the means by which social science ‘measures’ intelligence, depression, social inclusion, or student satisfaction are sufficiently regimented, then perhaps there is no reason not to regard this practice as measurement, even if some of these attributes turn out not to be quantitative. Mari’s position allows measurement realists to hang on to their narrow conception of quantity, while opening up the term ‘measurement’ to cover the wide range of practices currently described as measurement in different sciences.

How acceptable is this compromise to a measurement realist? Measurement realists needn’t disagree with Mari et al. on the steps involved in measurement and can also concede that there is a way of describing ‘measurement’ in the social sciences in a way that matches this structure. But the structural characterisation of the measurement process might seem to leave out what measurement realists are most interested in: the numerical representation of measurement outcomes. The generous extension of the concept of measurement to cover both ordinal and quantitative attributes is unlikely to assuage Michell’s worries about the standing of certain practices in psychology (Michell, 1999). Mari and Michell here simply seem to disagree about the core of what measurement is. Where Mari focusses on measurement as a *process or method*, Michell’s interest is in measurement as a tool for producing *quantitative* representations. In Mari’s interpretation, the key feature of measurement is the care taken in specifying the target and the method by which the target is to be estimated. The important feature of measurement, on this view, is that it is careful and precise. On Michell’s view, the important feature of measurement is that it is quantitative. To resolve the dispute among realists over restrictivism, then, it seems we need to identify the features that are responsible for the epistemic virtues of measurement. Realists will generally agree that measurement is epistemically particularly valuable, but it’s not immediately clear why this should be so. If the epistemic benefits of measurement are mainly due to explicit and precise specifications and definitions, then there is indeed little reason to think that similar care and precision couldn’t be applied to non-quantitative attributes. If, on the other hand, what makes measurement epistemically valuable is closely tied to the quantitative nature of the attributes targeted, then we should think that even if the steps in the measurement process are similar in the case of targeting quantitative

and qualitative attributes, the epistemic benefits of measurement come primarily with measuring quantities.

The question of just what makes measurement epistemically valuable goes beyond what can be responsibly covered in this paper. One consideration in favour of the idea that the value of measurement is tied to having quantitative attributes as targets is that the measurement of quantitative attributes permits refinements and increasing precision, whereas it is not clear how we can increase the precision of measurement of non-quantitative attributes. While measurements of physical constants, like the speed of light, have increased so much over the past century that they are now used to define measurement units, it's difficult to see how we can increase precision for attributes like pain or student satisfaction, which are presently not even measured at interval scale level, unless they turn out to be genuinely quantitative. Of course, this consideration is not conclusive, but it does suggest a reason for thinking that there is something distinctive to quantitative measurement, which might be a reason to restrict the term measurement to quantities.

While this offers an sense of why restrictive realists might not wish to settle for Mari's compromise, it does not yet help to respond to the argument from obsolescence. How might restrictive realists assuage concerns that restrictivism is obsolete?

6 Quantitativeness as a presupposition

Realists might be able to draw on Chang's history of temperature to respond to this concern. For despite their disagreement over the problem of coordination, both realists and antirealists are motivated by a good deal of epistemic caution. Chang's operationalist is reluctant to assume that measures can be extended beyond their empirically tested range, or to assume that measurements of potentially related attributes are measurements of the same attribute. Similarly, operationalists won't simply equate a theoretical concept, like *temperature_k*, with an empirical concept like *temperature_e*, or indeed more precisely, *temperature_{e1}* if we want to speak of temperature measured in a particular fashion. But Chang is willing to concede that we might accrue reasons for linking theoretical concepts to empirical ones, and seems to think that we might be justified in extending the reach of our concepts beyond their initial range of application. Of course, such justification for Chang is coherentist, but as his history of temperature shows, it is not impossible.

Likewise, Michell is demanding epistemic caution. We must recognize that quantitativeness is an assumption that cannot simply be taken for granted. His complaint against psychometrics in particular is that the quantitativeness of attributes like intelligence is never called into question. It is presupposed by the very discipline, and, Michell suggests, for reasons having to do more with the perceived prestige of measurement, not on strong theoretical grounds (Michell, 2008).

This provides a suggestion for a response to the obsolescence objection, using the lessons from the history of temperature. Measurement realists should embrace naturalistic coherentism, according to which no single belief is confirmed or refuted by experience. Confirmation is achieved holistically; conversely, no belief is a priori exempt from revision. This requires accepting that we cannot perform an empirical

test that demonstrates conclusively that an attribute is quantitative. That a particular attribute is quantitative is an assumption which cannot be tested in isolation. Nor can we argue a priori for or against the quantitateness of a given attribute. The assumption that an attribute is quantitative is open to challenge and revision in light of empirical evidence. Measurement of any attribute may be attempted, but to do so is to take on the presupposition that the attribute possesses quantitative structure.

Perhaps there is further disagreement afoot here¹¹: what is the standing of the assumption that a particular targeted attribute is quantitative? One option is to treat it as one of many auxiliary hypotheses, any one of which might be subject to revision. For example, if two surveys on student satisfaction differing in their order of questions fail to provide sufficient agreement in the numerical scores for the courses rated, we might reject the auxiliary hypothesis that the order of questions does not matter for the respondents' evaluation, or we might reject the auxiliary hypothesis that student evaluation has sufficient structure to produce meaningful numerical assignments. Which of these auxiliary hypotheses we should reject, then, depends on the evidential support we have for the different auxiliaries. On this reading, we might take robust and reliable ordinal rankings produced by a particular measurement procedure to be preliminary evidence for the ultimately quantitative nature of an attribute like student satisfaction. Realists, by contrast, may find this evidence too weak to retain the hypothesis.

But there is a second way we might think about the status of the claim that a particular attribute is quantitative: it is a constitutive presupposition of the practice of *measuring* that attribute, because of the link between measurement and quantitateness asserted by measurement realism. Unlike the first option, this way of understanding the assumption makes it not one of several auxiliary hypotheses, but a commitment undertaken by virtue of engaging in measurement of, say, student satisfaction, in the first place. In this case, to reject the assumption that the attribute is quantitative would mean to give up the entire practice of attempting to measure the attribute, not merely to make changes to how the measurement is carried out or interpreted. Measurement realists may well prefer this second understanding, as it reaffirms the strong connection between measurement and quantities.

Even if we take this second interpretation, it seems to me the question of whether the measurement of a particular attribute is legitimate is neither a priori, nor to be settled empirically *prior* to attempts at measuring it. The coherentist lesson still applies. Early measurement of temperature, qua measurement, already presupposed that the targeted attribute had quantitative structure, but that this presupposition was warranted was justified only by the subsequent successful history of temperature measurement. It was neither a foregone conclusion, nor a matter of devising a decisive empirical test. If this is true for paradigmatic cases of measurement in physics, it seems unfair to demand that psychometrics or other forms of social science measurement do better. Instead, we should acknowledge that in attempting to measure an attribute, we presuppose that the attribute is quantitative. We may have to be content with orderings at first, but the aim is to arrive at interval and ratio representations. These presuppose that the targeted attribute is quantitative. This presupposition may be false. If we fail to find a

¹¹ Many thanks to an anonymous reviewer for inviting me to clarify and expand this point.

coherent measurement procedure that leads to interval or ratio representations, one of the conclusions we might draw is that the attribute targeted is not in fact quantitative.

The success of the iterative process Chang describes for the case of temperature provides good holistic reasons for taking temperature to be a quantitative attribute. Physical attributes like temperature provide an exemplar of quantification not because they are physical or because they conform to the traditional conception of fundamental measurement. They are exemplars because they show what a successful operationalisation looks like. Conversely, the empirical difficulties faced, as well as the theoretical challenges encountered in the measurement of wellbeing, intelligence, or student satisfaction provide reasons to be cautious about holding on to the assumption that these concepts must be measurable. Some might turn out to be measurable, others not. To insist that these attributes must be measurable, come what may, is to be dogmatic in a way incompatible with coherentist naturalism. It therefore remains possible to criticise certain measurement practices, albeit in a less a priori fashion than Michell proposes.

To suggest that intelligence, wellbeing, student satisfaction, and others more will go the way of temperature is to presuppose that these are indeed quantitative attributes. Pace Michell, to make such a presupposition is a legitimate move, but Michell is right that the presupposition needs to be marked as such. Measurement of attributes like intelligence, student satisfaction, or wellbeing has the same success conditions as the measurement of attributes like mass or temperature. Realism needn't stand in the way of risky measurement practices, then, but is right to insist that we should hold on to a narrow definition of measurement.

7 Conclusion

In this paper I've taken a close look at two important arguments against measurement realism: the argument from coordination and the argument from obsolescence. I've argued that the argument from coordination, while seemingly a principled objection to measurement realism, turns out to be less devastating than anticipated. The conventional aspects of measurement representations do not conclusively undermine the idea that our measurements target attributes with quantitative structure. Indeed, giving up this presupposition of measurement seems to leave us with no clear or compelling standard for successful measurement. The argument from obsolescence claims that traditional measurement realism, with its restriction of measurement to the measurement of quantitative attributes fails to do justice to measurement-like activities carried out in many social sciences. By insisting that measurement targets quantitative attributes only, measurement realists are out of touch with scientific practice. This argument, as well as a suggested compromise by Mari et al., might seem to be merely verbal, but actually points to the difficult question at the heart of the measurement realism debate: what is measurement? I've suggested that to answer this question we need to identify the presuppositions of measurement as well as the expected epistemic benefits of measurement. It is with respect to the latter that we find disagreement both between realists and antirealists, and within the realist camp. I've argued that measurement realists have good reason to hold on to their metaphysical commitment to quantitative attributes as the targets of measurement, but that they should modify their epistemic

commitments in light of these antirealist arguments. In particular, they should treat the assumption that an attribute is quantitative as either a hypothesis or as a constitutive presupposition, subject to revisions.

The crucial advantage of insisting that only some attributes have quantitative structure and that this is a feature of the attributes, not a matter of our representation or measurement procedure, is that it offers a standard for success beyond the evidence for or against success. We don't just want convergent measurements for the sake of convergence, but because a lack of convergence suggests that we are not tracking the same attribute. Similarly, if we fail to find a more than ordinal representation for a targeted attribute, this might be evidence for a lack in quantitative structure in the attribute. These problems can only become evidence of *failure*, if we take the target of measurement to be something beyond our measurement procedures.

Acknowledgements I would like to thank three anonymous reviewers for their helpful comments. I would also like to thank the editors of this special issue for their patience throughout this process.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alper, T. M. (1985). A note on real measurement structures of scale type $(m, m + 1)$. *Journal of Mathematical Psychology*, 29(1), 73–81. [https://doi.org/10.1016/0022-2496\(85\)90019-7](https://doi.org/10.1016/0022-2496(85)90019-7)
- Belot, G. (2010). Transcendental idealism among the Jersey metaphysicians. *Philosophical Studies*, 150(3), 429–438. <https://doi.org/10.1007/s11098-010-9546-8>
- Borsboom, D. (2005). *Measuring the mind: Conceptual issues in psychometrics*. Cambridge University Press.
- Bridgman, P. W. (1958). *The logic of modern physics*. Macmillan New York.
- Campbell, N. R. (1920). *Physics: The elements*. Cambridge University Press.
- Chang, H. (2004). *Inventing temperature: Measurement and scientific progress*. Oxford University Press.
- Chang, H. (2009). Operationalism. *Stanford Encyclopedia of Philosophy, Fall 2009*.
- Ellis, B. D. (1966). *Basic concepts of measurement*. Cambridge University Press.
- Feest, U. (2005). Operationism in psychology: What the debate is about, what the debate should be about. *Journal of the History of the Behavioral Sciences*, 41(2), 131–149.
- Hardcastle, G. L. (1995). S.S. Stevens and the origins of operationism. *Philosophy of Science*, 62(3), 404–424.
- Krantz, D. H., Suppes, P., Luce, R. D., & Tversky, A. (1971). *Foundations of measurement*. Academic Press.
- Mari, L., Carbone, P., Giordani, A., & Petri, D. (2017a). A structural interpretation of measurement and some related epistemological issues. *Studies in History and Philosophy of Science Part A*, 65–66, 46–56. <https://doi.org/10.1016/j.shpsa.2017.08.001>

- Mari, L., & Giordani, A. (2014). Modeling measurement: Error and uncertainty. In M. Boumans, G. Hon, & A. Petersen (Eds.), *Error and uncertainty in scientific practice* (pp. 79–96). Pickering & Chatto.
- Mari, L., Maul, A., Torres Irribarra, D., & Wilson, M. (2017b). Quantities, quantification, and the necessary and sufficient conditions for measurement. *Measurement*, *100*, 115–121. <https://doi.org/10.1016/j.measurement.2016.12.050>
- Michell, J. (1999). *Measurement in psychology: Critical history of a methodological concept*. Cambridge University Press.
- Michell, J. (2004). History and philosophy of measurement: A realist view. *Proceedings of the 10th IMEKO TC7 International Symposium on Advances of Measurement Science*, 3–9.
- Michell, J. (2005). The logic of measurement: A realist overview. *Measurement*, *38*(4), 285–294. <https://doi.org/10.1016/j.measurement.2005.09.004>
- Michell, J. (2008). Is psychometrics pathological science? *Measurement*, *6*(1–2), 7–24. <https://doi.org/10.1080/15366360802035489>
- Narens, L. (1981). A general theory of ratio scalability with remarks about the measurement-theoretic concept of meaningfulness. *Theory and Decision*, *13*(1), 1–70.
- Narens, L. (2012). *Theories of meaningfulness*. Psychology Press, Taylor & Francis Group. (Original work published 2002)
- Psillos, S. (1999). *Scientific realism: How science tracks truth*. Routledge.
- Sider, T. (2011). *Writing the book of the world*. Oxford University Press.
- Sider, T. (2020). *The tools of metaphysics and the metaphysics of science*. Oxford University Press.
- Tal, E. (2018). Naturalness and convention in the International System of Units☆. *Measurement*, *116*, 631–643. <https://doi.org/10.1016/j.measurement.2017.09.008>
- Teller, P. (2018). Measurement accuracy realism. In *The Experimental side of modeling* (pp. 273–298). University of Minnesota Press.
- Van Fraassen, B. C. (2008). *Scientific representation: Paradoxes of perspective*. Oxford University Press.
- Wolff, J. E. (2019). Representationalism in measurement theory. Structuralism or perspectivalism? In *Understanding Perspectivism*. Routledge.
- Wolff, J. E. (2020). *The metaphysics of quantities*. Oxford University Press.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.