

# Old and New Problems in Philosophy of Measurement

Eran Tal, Bielefeld University, July 2013

Forthcoming in *Philosophy Compass*

## Abstract

The philosophy of measurement studies the conceptual, ontological, epistemic and technological conditions that make measurement possible and reliable. A new wave of philosophical scholarship has emerged in the last decade that emphasizes the material and historical dimensions of measurement and the relationships between measurement and theoretical modeling. This essay surveys these developments and contrasts them with earlier work on the semantics of quantity terms and the representational character of measurement. The conclusions highlight four characteristics of the emerging research program in philosophy of measurement: it is epistemological, coherentist, practice-oriented and model-based.

## Introduction

Measurement – whether performed by operating thermometers, counting unemployment benefit claims or administering quality-of-life questionnaires – is an activity intended to produce knowledge about the state of an empirical system<sup>1</sup>. The philosophy of measurement sets out to characterize and classify measurement procedures and to clarify the conceptual, ontological, epistemic and technological conditions that make measurement possible and reliable. Rather than providing a comprehensive introduction to this field, the purpose of this essay is to survey recent developments and contrast them with earlier work<sup>2</sup>. In particular, I will focus on two areas where scholarship over the past decade has significantly departed from traditional philosophical approaches<sup>3</sup>. The first area is the coordination between theoretical quantity concepts like mass and length and the empirical procedures that measure them. Recent scholarship has highlighted the material and historical aspects of coordination, in contrast with the traditional focus on conventional definitions. The second area concerns

the role of representation in measurement. Here the emphasis has shifted from the adequacy of numerical representations to the theoretical and statistical modeling of measuring instruments and the means by which measuring instruments gather reliable information. This shift towards modeling and information has in turn generated new questions concerning the observational grounding of measurement and its relationship with computation, which will be discussed in section three. Finally, my conclusion will underscore four characteristics of contemporary philosophical approaches to measurement: they tend to be epistemological, coherentist, practice-oriented and model-based.

## **1. The semantics of quantity: from definition to realization**

### 1.1 The problem of coordination

Scientific theories and models are commonly expressed in terms of quantitative relations among parameters, bearing names such as ‘mass’, ‘acidity’ and ‘productivity’. Considered purely as elements of a mathematical formalism, such parameters are not yet associated with empirical content. It is only once linked, or ‘coordinated’, with one or more procedures for determining their values that such parameters acquire their empirical significance. A long-standing problem in the philosophy of measurement – as well as in philosophy of science more generally – concerns the proper way of coordinating theoretical quantity-terms with empirical measuring procedures (van Fraassen 2008, 115-139). The problem is that the empirical adequacy of the theory or model and the reliability of measuring procedures appear to presuppose each other in a circular way. To establish a theory of mass, for example, it is necessary to test its predictions, a task which requires a reliable method of measuring mass. Testing the reliability of measurements of mass, however, presupposes background theoretical

knowledge about mass and its relations with other physical properties such as force and motion.

The traditional philosophical approach to this problem holds that coordination is accomplished by specifying definitions, or definition-like statements, for some of the relevant quantity-terms. These definitions are thought to be analytic statements and to require no empirical testing, thereby avoiding the dangers of circularity or an infinite regress of coordinating statements.

The most straightforward representative of the traditional approach is the early work of Bridgman (1927), who offered to define quantity-concepts directly by the operations that measure them. Length, for example, would be defined as the result of the operation of concatenating rigid rods. Consequently, different operations measure different quantities: the quantities measured by using rulers and by timing electromagnetic pulses should, strictly speaking, be labelled 'length-1' and 'length-2'. Nevertheless, Bridgman conceded that as long as the results of different operations agree within experimental error it is pragmatically justified to label the corresponding quantities with the same name (ibid, 16).<sup>4</sup>

Operationalism, it was soon revealed, was riddled with problems. Among such problems were the automatic reliability operationalism conferred on measurement operations, the ambiguities surrounding the notion of operation, the overly restrictive operational criterion of meaningfulness, and the fact that many useful theoretical concepts lack clear operational definitions (Gillies 1972; Chang 2009)<sup>5</sup>. Accordingly, most writers on the semantics of quantity have avoided espousing an operational analysis<sup>6</sup>.

A more nuanced approach to the problem of coordination, still within the traditional strand, is known as conventionalism. It includes a variety of views originating from the late nineteenth century and up to the 1960s. These views admit a conventional, definition-like element to coordination, while resisting attempts to reduce the meaning of quantity terms to measurement operations. Conventionalist accounts differ in the particular aspects of

measurement they deem conventional and in the degree of arbitrariness they ascribe to such conventions.

It is undisputed that *some* aspects of measurement are conventional. Whether one measures temperature on the Celsius or Fahrenheit scales, or whether one uses the meter or inch as a unit of length, are choices that ultimately hang on consensus among humans rather than facts about nature. Conventionalism about measurement aims to additionally show that some *nontrivial* aspects of the application of quantity-concepts to measurement procedures, previously not thought to rest on human consensus, are in fact conventional, and that this conventionality explains the possibility of coordination.

Conventionalism about measurement was espoused by Ernst Mach, who coined the term ‘principle of coordination’ for the choice of a standard thermometric fluid (1986 [1896], 52). Mach noted that different types of fluid expand at different (and nonlinearly related) rates when heated, raising the question: which fluid expands most uniformly with temperature? According to Mach, there is no fact of the matter as to which fluid expands more uniformly, because the very notion of equality among temperature intervals has no determinate application prior to a conventional choice of standard thermometric fluid. The concepts of uniformity of time and space received similar treatments by Poincaré (1958 [1898]; 2007 [1905], Part 2). Poincaré argued that procedures used to determine equality among durations stem from scientists’ unconscious preference for descriptive simplicity, rather than from any fact about nature. Similarly, scientists’ choice to represent space with either Euclidean or non-Euclidean geometries is not determined by experience but by considerations of convenience.

Conventionalism with respect to measurement reached its most sophisticated expression in logical positivism. Logical positivists like Reichenbach and Carnap proposed ‘coordinative definitions’ or ‘correspondence rules’ as the semantic link between theoretical and observational terms. These a priori, definition-like statements were intended to regulate the use of theoretical terms by connecting them with empirical procedures (Reichenbach 1958

[1927], 14-19; Carnap 1995 [1966], Chap. 24). An example of a coordinative definition is the statement: ‘a measuring rod retains its length when transported’. According to Reichenbach, this statement cannot be empirically verified, because a universal and experimentally undetectable force could exist that equally distorts every object’s length when it is transported. In accordance with verificationism, statements that are unverifiable are neither true nor false. Instead, Reichenbach took this statement to express an arbitrary rule for regulating the use of the concept of equality of length, namely, for determining whether particular instances of length are equal (Reichenbach 1958 [1927], 16). At the same time, such statements were not taken to *define* concepts such as length or length-equality, thereby avoiding some of the problems associated with operationalism<sup>7</sup>.

## 1.2 Empirical constraints and epistemic iterations

During the second half of the twentieth century logical positivism was heavily criticized and eventually abandoned by mainstream philosophers of science. Although some of the criticisms had important implications for the philosophy of measurement (for example Kuhn 1977 [1961]), the decline of logical positivism was followed by a general decrease of philosophical interest in measurement<sup>8</sup>. Measurement continued to be discussed by philosophers in connection to its mathematical foundations (see Section 2), but was otherwise relegated to a secondary research topic and mentioned mainly by virtue of its relations to other areas of philosophical debate, such as observation, scientific realism, evidence, causality and experimentation, or in connection with measurement problems in the special sciences<sup>9</sup>.

The return of measurement to the forefront of philosophical research in the early 2000s was accompanied by renewed interest in the problem of coordination. A new strand of writing on the problem has emerged in the last decade, consisting most notably of the works of Chang (2001, 2004, 2007) and van Fraassen (2008, Ch. 5; 2009; 2012). These works take a historical

and coherentist approach to the problem. Rather than attempting to avoid the circularity of coordination completely, as their predecessors did, they set out to show that the circularity is not vicious. Chang argues that constructing a quantity-concept and standardizing its measurement are co-dependent and iterative tasks. Each ‘epistemic iteration’ in the history of standardization respects existing traditions while at the same time correcting them (Chang 2004, Chap. 5). The pre-scientific concept of temperature, for example, was associated with crude and ambiguous methods of ordering objects from hot to cold. Thermoscopes, and eventually thermometers, helped modify the original concept and made it more precise. With each such iteration the quantity concept was re-coordinated to a more stable set of standards, which in turn allowed theoretical predictions to be tested more precisely, facilitating the subsequent development of standards, and so on.

How this process avoids vicious circularity becomes clear when we look at it either ‘from above’, i.e. in retrospect given our current scientific knowledge, or ‘from within’, by looking at historical developments in their original context (van Fraassen 2008, 122). From either vantage point, coordination succeeds because it increases coherence among elements of theory and instrumentation. The questions ‘what counts as a measurement of quantity X?’ and ‘what is quantity X?’, though unanswerable independently of each other, are addressed together in a process of mutual refinement. It is only when one adopts a foundationalist view and attempts to find a starting point for coordination free of presupposition that this historical process erroneously appears to lack epistemic justification (ibid, 137).

The new literature on coordination shifts the emphasis of the discussion from the definitions of quantity-terms to the *realizations* of those definitions, and hence to *metrology* – the science of measurement and standardization. In metrological jargon, a ‘realization’ is a physical instrument or procedure that approximately satisfies a given definition (cf. JCGM 2012, 5.1). Examples of metrological realizations are the official prototypes of the kilogram and cesium fountain clocks used to standardize the second. The methods used to design,

maintain and compare these instruments have a direct bearing on the practical application of concepts of quantity, unit and scale, no less than the definitions of those concepts. Philosophers are now beginning to engage with the rich conceptual issues underlying metrological practice, and particularly with the inferences involved in evaluating and improving the accuracy of measurement standards (Boumans 2005a Chap. 5, 2005b, 2007; Frigerio, Giordani and Mari 2010; Tal 2011, 2012; Teller 2013, under review; Riordan under review). In so doing, philosophers are drawing on the work of historians and sociologists, who have been investigating the field of metrology for a longer period (Latour 1987 Chap. 6; Schaffer 1992; Porter 1995, 2007; Galison 2003), as well as on the history and philosophy of scientific experimentation (Hacking 1983, Franklin 1986, Cartwright 1999).

Contrary to conventionalists, the new studies take a practice-oriented approach to the problem of coordination and highlight the empirical constraints that inform historical choices of measurement standards. For example, the mid-nineteenth century choice to standardize temperature based on the expansion of air rested on detailed experimental work showing differences in the reproducibility of results among thermometers filled with different fluids (Chang 2001; 2004, Chap. 2). Similarly, contemporary choices among different ways of standardizing time are constrained by the results of robustness tests, which verify the mutual compatibility of uncertainties ascribed to different atomic clocks (Tal 2011). Conventionalists classified such choices as mere conveniences aimed at simplifying the mathematical form of physical laws, and neglected the intricate roles played by experimentation in real episodes of standardization.

Another disadvantage of the conventionalist position was the static nature of its proposed mechanisms of coordination. In the replacement of one principle of coordination with another, conventionalists saw a sudden, arbitrary shift in the accepted usage of a term rather than a mark of scientific progress. By contrast, historical cases of coordination tend to exhibit a gradual and cumulative character. For example, the current atomic scale used for

global timekeeping, known as Coordinated Universal Time, was painstakingly designed to be compatible with previous astronomical timescales while allowing a significant reduction of measurement uncertainty (Jones 2000, Chap. 3 and 5). This cumulative aspect of standardization is better accounted for by Chang's iterative approach than by conventionalism as espoused by Carnap and Reichenbach<sup>10</sup>.

The philosophical study of standardization is still in its nascent stages, and much work lies ahead. To provide a single example, the standardization of subjective measures of psychometric constructs currently poses new challenges. Attempts to validate questionnaires for measuring subjective well-being and quality of life raise something similar to the problem of coordination: are the questionnaires measuring what they should? Should the construct be defined in terms of the best-correlated questionnaires? It is doubtful whether these questions can be answered through a process of iterative stabilization similar to the one encountered in the standardization of physical quantities. As Alexandrova (2008) points out, ethical considerations bear on questions about construct validity no less than considerations of stability. Such ethical considerations are context sensitive, and can only be applied piecemeal. McClimans (2010) shows that uniformity is not always an appropriate goal for designing questionnaires. Indeed, the open-endedness of questions is both unavoidable and desirable for obtaining relevant information from subjects<sup>11</sup>. These insights highlight the need for a thorough analysis of the epistemic and ethical issues underlying psychometric construct validation.

## **2. Measurement as representation: from morphisms to information**

### 2.1 The adequacy of numerical representations

Measurement represents the empirical world, but the precise nature and means of representation involved in measurement are difficult to pin down. In particular, the use of *numbers* to represent empirical objects<sup>12</sup> has been a traditional topic of philosophical concern. Measurement outcomes are commonly expressed in numerical terms, e.g. ‘the mass of this object is 5 Kg’, ‘*x* and *y* are equal in length’, ‘*x* has twice the mass of *y*’ or ‘*x* is 30 degrees Celsius hotter than *y*’. Under what conditions is it adequate to use mathematical concepts such as number, equality, ratio and difference to represent empirical objects, and what do these representations tell us about those objects?<sup>13</sup> Modern engagement with these questions – known as ‘measurement theory’ – dates back to the late nineteenth century.<sup>14</sup> Broadly speaking, measurement theories attempt to specify the conditions under which empirical objects can be represented with numbers or other mathematical entities. This task is complicated by the fact that mathematical relations among numbers do not always correspond to empirical relations among measured objects. For example, 60 is twice 30, but one would be mistaken in thinking that an object measured at 60 degrees Celsius is twice as hot as an object at 30 degrees Celsius. The interval between the real numbers 3.1 and 3.2 is infinitely divisible, but the physical interval between 3.1 and 3.2 Coulomb is not, due to the existence of a fundamental electric charge. Equality among numbers is transitive (( $a=b$  &  $b=c$ ) implies  $a=c$ ) but empirical comparisons among magnitudes reveal only approximate equality, which is not a transitive relation.

Traditional measurement theories tackle these and similar difficulties by pursuing two complementary lines of inquiry. The first line of inquiry identifies algebraic structures in *qualitative* experience, that is, experience prior to numerical representation. Helmholtz (1930 [1887]), Hölder (1901) and Campbell (1920) proceed in this vein. They argue that certain empirical operations exhibit a qualitative structure that shares essential features with algebraic operations among numbers. For example, a qualitative ordering of rigid rods by their perceived lengths shares structural features with the mathematical relation ‘equal to or bigger

than'. Similarly, an empirical concatenation of rigid rods shares structural features with the mathematical operation of addition. Taken together, ordering and concatenation are sufficient for the construction of an *additive* representation of the relevant magnitudes, that is, a numerical representation in which addition is empirically meaningful (and hence also multiplication, division etc.) As the abovementioned authors stressed, additive representations exist for some magnitudes – like length, weight and duration – but not for others. The hardness of minerals, for example, admits of ordering (from softest to hardest) but not of concatenation, as there is no empirical procedure for combining the degrees of hardness of two minerals. As a result, numbers assigned to degrees of hardness have empirical significance only insofar as their order is concerned, whereas sums and ratios of these numbers have no such significance.<sup>15</sup>

The second and closely related line of inquiry is the classification of types of measurement scales. Stevens (1946) distinguishes between four types of scales: nominal, ordinal, interval and ratio. Nominal scales represent objects as belonging to classes that have no particular order, e.g. male and female. Ordinal scales represent order but no further algebraic structure, for example, the Mohs scale of mineral hardness. Celsius and Fahrenheit are examples of interval scales: they allow meaningful arithmetic operations on *intervals* of temperature, but not on temperature values themselves, because the zero points of these scales are arbitrary. The Kelvin scale, by contrast, is a ratio scale, as are the familiar scales representing mass in kilograms, length in meters and duration in seconds. Ratio scales represent magnitudes as having the same algebraic structure as the real numbers, e.g. as having meaningful sums and ratios. As Stevens notes, scale types are individuated by the families of transformations they can undergo. Empirical relations represented on ratio scales, for example, are invariant under multiplication by a positive number, e.g. multiplication by 2.54 converts from inches to centimeters. Interval scales allow both multiplication by a

positive number *and* a constant shift, e.g. the conversion from Celsius to Fahrenheit in accordance with the formula  $^{\circ}\text{C} \times 9/5 + 32 = ^{\circ}\text{F}$ .

These two lines of inquiry – the analysis of qualitative structures and the classification of types of scales – converge in the Representational Theory of Measurement (Suppes 1951; Krantz et al. 1971, 1989, 1990). RTM defines measurement as the construction of mappings from empirical relational structures into numerical relational structures (Krantz et al. 1971, 9). An empirical relational structure consists of a set of empirical objects (e.g. rigid rods) along with certain qualitative relations among them (e.g. ordering, concatenation), while a numerical relational structure consists of a set of numbers (e.g. real numbers) and specific mathematical relations among them (e.g. ‘equal to or bigger than’, addition). Simply put, a measurement scale is a mapping – a *homomorphism* – from an empirical to a numerical relational structure, and measurement is the construction of scales.<sup>16</sup> RTM goes into great detail in clarifying the assumptions underlying the construction of different types of measurement scales. Each type of scale – including the four mentioned above – is associated with a set of assumptions, or ‘axioms’, about the qualitative relations obtaining among empirical objects. From these assumptions the authors of RTM derive the representational adequacy of each scale type, as well as the family of permissible transformations making that type of scale unique.

## 2.2. An inferential conception of measurement

While the achievements of traditional measurement theories in axiomatizing measurement scales are undisputed, these achievements need to be placed in a larger context. Traditional thinkers such as Campbell, Stevens and Suppes took ‘measurement’ to be synonymous with either ‘number assignment’<sup>17</sup> or ‘scale construction’, and neglected the

‘applied’ aspects of measurement such as accuracy, precision, error, uncertainty and calibration (Kyburg 1992; Mets 2012)<sup>18</sup>. As philosophers have come to recognize in recent years, mathematical scales are only one of several means of representation involved in measurement, and often not the most epistemically problematic or interesting ones. Measurement involves a host of *theoretical* and *statistical* representations of measuring systems and the data they produce. As will be clarified below, the assumptions underlying such representations influence which measurement outcomes are obtained, how errors are detected and corrected and how accuracy is evaluated. Once the richness of representational means involved in measurement is acknowledged, the traditional conception of measurement as the construction of homomorphisms is revealed as overly restrictive (Mari 2000; van Fraassen 2008, 158-166). Additionally, the very distinction between ‘applied’ and ‘foundational’ concepts of measurement dissolves, as choosing a measurement scale and correcting for errors are seen as interdependent tasks (Tal 2012, 73-78).

Recent philosophical writing conceptually divides measurement procedures into two levels: (i) a concrete process involving interactions between an object of interest, an instrument, and the environment; and (ii) a theoretical and / or statistical representation of that process. The outputs of measurement are accordingly distinguished into two kinds. On the concrete level one speaks of instrument *indications*, namely the final states of an instrument after each measurement run is complete. Typical examples of indications are pointer positions, digits appearing on a display, and marks on a multiple-choice questionnaire. On the abstract level, measurement *outcomes* are knowledge claims about the state of the object of interest. Outcomes are often expressed in the form: “quantity  $Q$  associated with object  $O$  has value  $q$  with uncertainty  $U$ ”, although in general outcomes need not be expressed numerically.

This two-tier picture draws attention to the inferential nature of measurement. Measurement outcomes are obtained from indications by a chain of inferences, and the particular inferences drawn depend on the particular theoretical and statistical assumptions

with which the concrete measurement process is represented<sup>19</sup>. This way of viewing measurement raises a host of representational questions that have been either neglected or only partially addressed by traditional accounts, including:

1. How, and under what conditions, does a measurement outcome represent the state of an object of interest? Specifically, what roles do theoretical and statistical assumptions about a measurement process play in establishing the representational adequacy of outcomes?
2. Can measurement-related notions like accuracy, precision, error and uncertainty be clarified in terms of representational relations between the abstract and concrete levels of measurement?
3. What sorts of inference are involved in the calibration and standardization of measuring instruments, and how do theoretical and statistical assumptions about the measurement process feature in such inferences?
4. What is the relationship between measurement and other activities that involve theoretical and statistical representations, such as modeling, prediction, imaging and simulation?

Van Fraassen (2008, 141-185) focuses on the first question. As he argues, measurement is a means of gathering information about an object (ibid, 143). More specifically, “measurement is an operation that locates an item (already classified as in the domain of a given theory) in a logical space (provided by the theory to represent a range of possible states or characteristics of such items)” (ibid, 164). A measurement outcome is thus a *region* in parameter space where the relevant theory locates the actual state of the object based on the indications of an instrument (ibid 164, 172). Such region is considered an adequate representation of the object only when the theory provides a coherent story of the ways in

which possible indications of the apparatus reflect possible states of the object. Three crucial departures from traditional measurement theories are worth noting here: outcomes are not numbers but regions in parameter space; outcomes depend on theory not only for their interpretation but also for their capacity to represent measured objects in the first place; and the mapping of indications to outcomes is not a matter of shared algebraic structure but of information transmission.<sup>20</sup>

### 2.3 Model-based approaches and the epistemology of measurement

Classifying measurement as information-gathering is the starting point for a more comprehensive attempt to address all four questions above, a project sometimes referred to as the *epistemology of measurement*. The epistemology of measurement investigates the conditions under which measurement and standardization methods produce knowledge, the nature, scope, and limits of this knowledge, and the sources of its reliability (Mari 2003, 2005a; Leplège 2003; Tal 2012, 3-5). These aims have also informed the works of earlier thinkers including Mach, Poincaré, Campbell, Bridgman and Reichenbach. Nonetheless, current engagement with these issues is novel in at least two respects. First, contemporary scholars approach the topics of measurement accuracy and error with significantly more detail than their predecessors. Several studies have emerged in recent years that grapple in detail with the inferences involved in evaluating measurement uncertainties, calibrating instruments and establishing accuracy (Mari 2000, 2005b; Mari and Giordani 2013; Boumans 2005a, 2005b, 2007, 2012a; Tal 2011).

The second novelty in recent studies is their emphasis on the role of *models* – abstract, local, and simplifying representations – in underwriting claims to the epistemic reliability of measurement<sup>21</sup>. While the relevance of modelling to measurement has been acknowledged in earlier philosophical literature, its role was thought to be restricted to the statistical analysis of

measured data (Suppes 1962). By contrast, contemporary studies show that theoretical and statistical models of a measurement process are necessary preconditions for obtaining meaningful measurement outcomes in the first place.

For example, Boumans (2006, 2007, 2009) and Mari (2005b) discuss the possibility of reconstructing the state of a system under measurement from the readings of a measuring device. Such reconstruction requires obtaining a calibration function for the device, that is, a function that relates possible values of the quantity being measured, possible indications of the device, and possible values of external influencing variables. To obtain a calibration function, the device, measured object and environment must all be modelled in an abstract way. Different abstract representations of the same elements are possible that involve different assumptions, ranging from detailed analytical models to simple ‘black-box’ models. Which calibration function – and hence which measurement outcomes – will be associated with the device depends on the assumptions with which the measurement process is modelled.

Model-based analyses also shed light on measurement in economics. Like physical quantities, values of economic variables often cannot be observed directly and must be inferred from observations based on abstract and idealized models. The nineteenth century economist William Jevons, for example, measured changes in the value of gold by postulating certain causal relationships between the value of gold, the supply of gold and the general level of prices (Hoover and Dowell 2001, 155-159; Morgan 2001, 239). The considerable reliance on models in economic measurement has led some philosophers to view certain economic models as measuring instruments in their own right, analogously to rulers and balances. Boumans (2005c, 2006, 2007, 2009, 2012a) explains how macroeconomists are able to isolate a variable of interest from external influences by tuning parameters in a model of the macroeconomic system. This technique frees economists from the task of controlling the actual system, which would be practically impossible in this case. As Boumans argues, macroeconomic models function as measuring instruments insofar as they produce invariant

relations between inputs (indications) and outputs (outcomes), and insofar as this invariance can be tested by calibration against known and stable facts.

The studies mentioned in this section suggest that information and modelling are likely to be key concepts in future epistemological accounts of measurement. The precise relationship between these two notions is, however, still unclear. Do models of measurement processes carry information, and, if so, how? Indeed, the term ‘information’ itself is used imprecisely and variably across several of these studies. Further research is required to conceptually connect the epistemology of measurement with the wealth of writing on information in other fields, including information theory and the philosophy of information, as well as with the wider philosophical scholarship on modeling.

### **3. Computer simulations as measuring instruments? The problem of observational grounding**

Characterizing measurement as model-based information gathering has led to comparisons between measurement and computational methods. Similar to measuring instruments, some scientific computer simulations generate reliable information about empirical systems from observational data and theoretical or statistical models. Should such simulations be classified as measuring instruments? Several contemporary thinkers argue that under specific conditions the answer to this question may be ‘yes’. We already saw that Boumans classifies certain economic models as measuring instruments. The investigation and tuning of economic models is often performed by computer simulation, suggesting that such simulations themselves function as measuring instruments.

Morrison (2009) similarly argues that certain computer simulations in physics, such as particle methods used to simulate the evolution of many-body systems, function like measuring instruments due to their similarity with laboratory experiments. While

acknowledging the importance of materiality in laboratory experiments, she indicates that the crucial epistemic ‘work’ involved in measuring and experimentation is carried out on the level of abstract representation. Claims to measurement accuracy and experimental validity are ultimately justified by appealing to the theoretical models representing the experimental object and apparatus, rather than by direct appeal to material properties. As long as theoretical models and the algorithms that discretize them are well constructed, computer simulations have all the essential epistemic characteristics of laboratory experiments, and the outputs of such simulations should properly count as measurements<sup>22</sup>.

Regarding meteorological computer simulations, Parker (2013) has recently argued that under special circumstances some simulation results may be rightly considered measurements. She discusses data assimilation methods, including methods that estimate past weather conditions at places and times where traditional measurements are not available by finding a weather simulation that fits available measurements in the vicinity. Parker argues that the data assimilation procedure as a whole (encompassing both empirical and computational information gathering) can be placed on a spectrum with other measuring procedures. Nonetheless, what matters in the epistemic analysis of data is not the type of its source – be it measurement, observation or computation – but an understanding of “the limitations of a putative source of information and the uncertainties associated with its deliverances” (ibid, 12).

Regardless of whether one finds these arguments convincing, they raise new questions about the epistemic status of measurement. What exactly is at stake, epistemically speaking, in classifying an estimation procedure as a ‘measurement’ procedure? What role does observation play in measuring, and is this role essentially different from the role observation plays in other estimation methods, e.g. data-driven modelling and retrodiction?

Questions such as these suggest an inversion of the traditional problem of the theory-ladenness of measurement. Traditional philosophy of measurement, dating back to early

logical positivism and to Campbell's (1920) notion of fundamental measurement, took observation to be the ultimate source of justification for measurement claims. The observational basis was intended to guarantee the theory-independence and neutrality of measurement. Theory-ladenness, i.e. the pervasive use of theoretical assumptions in designing measurement apparatuses and interpreting their indications, was seen as the central traditional threat to the neutrality of measurements (Kuhn 1977 [1961], Shapere 1982, Franklin et al. 1989). In contrast to its historical counterparts, contemporary scholarship has come to view measurement as theory-dependent by default. Measurement is considered possible only against a theoretical background, with ever more accurate measurements requiring even richer theoretical backgrounds. This conception of measurement has given rise to something like an opposite to the problem of theory-ladenness, which may be called the problem of *observational grounding*. As measurement and computational methods have become more sophisticated, it is increasingly difficult to specify what sort of connection with observation is sufficient to grant a procedure the privileged epistemic status normally called 'measurement'. The problem is not merely terminological. The designation 'measurement procedure' supposedly implies suitability for producing scientific evidence, a distinction that is not shared by scientific procedures in general. This difference requires explaining (or explaining away). This problem is complicated further by the fact that the notion of observation is itself highly ambiguous and technology-laden.<sup>23</sup>

## **Conclusions**

A wave of scholarship has emerged in the past decade that views measurement from a novel perspective, bringing standards, artefacts, modeling practices and the history of science and technology to bear on philosophical problems concerning measurement. This recent work departs from the foundationalist and axiomatic approaches that characterized the philosophy

of measurement during much of the twentieth century. Inspired by developments in the philosophy of scientific modeling and experimentation, contemporary authors draw attention to scientific methodology and especially to metrology, the science of measurement and standardization. The increased focus on modeling and information has given rise to exciting new problems, such as the problem of observational grounding. At the same time, current discussions also return to traditional problems concerning e.g. the representational adequacy of measurement outcomes and the semantic relationship between quantity concepts and measurement operations. The ‘old’ and ‘new’ problems in philosophy of measurement therefore partially overlap, with more recent work forming a historical continuity with traditional strands.

The new scholarship has not yet coalesced into clear research programs or schools of thought. While it would be premature to predict the course of future work, a few preliminary trends have emerged from this survey. I conclude this article by highlighting four characteristics of recent philosophy of measurement as it has developed thus far:

1. **Epistemological** – recent work treats measurement as a knowledge-producing process and attempts to analyze the sources of its reliability, rather than the metaphysical or mathematical conditions of its possibility. Consequently, recent work is much more concerned with inferential validity, accuracy, error, and information quality than traditional philosophy of measurement.
2. **Coherentist** – recent work tackles questions about the epistemic reliability of measurement by appealing to coherence among elements of scientific methodology – such as instruments, models, statistical analysis tools and background theories – rather than to observational or a priori foundations of knowledge.

3. **Practice-oriented** – recent scholarship attempts to make sense of the concrete methods employed in making measurements and standardizing measuring instruments through analysis of both contemporary and historical examples.
4. **Model-based** – recent scholarship seeks to clarify the roles played by theoretical and statistical models in producing and validating – as opposed to merely interpreting – measurement outcomes. Models are seen as crucial for supporting inferences from instrument indications to measurement outcomes and for evaluating measurement error and uncertainty, among other roles.

---

**Acknowledgements:** The author would like to thank Sally Riordan and an anonymous reviewer for their helpful comments. Work on this article was supported by the Alexander von Humboldt Foundation.

**Author's Biography:** Eran Tal received an M.A. in History and Philosophy of Science from Tel Aviv University in 2006 and a Ph.D. in Philosophy from the University of Toronto in 2012. His doctoral dissertation, titled *The Epistemology of Measurement: A Model-Based Account*, analyzes the inferential structure of measurement and calibration procedures performed by contemporary standardization bureaus and argues that idealizations are necessary preconditions for the possibility of establishing the objectivity and accuracy of measurement outcomes. Tal is currently an Alexander von Humboldt Postdoctoral Research Fellow at Bielefeld University, Germany, where he works on the philosophical implications of 'virtual measurement', i.e. the use of computer simulations as enhancements or replacements for traditional measuring instruments in the natural sciences. His articles have appeared in *Philosophy of Science* and *Synthese*.

---

<sup>1</sup> As the discussion below will make clear, defining measurement is a difficult task. The present characterization is not meant as a definition.

<sup>2</sup> Particularly, this essay will not cover the rich literature on the metaphysics of measurement (Byerly and Lazara 1973; Swoyer 1987; Michell 1994) nor the topic of the relationship between measurement and causation (Cartwright 1989). Measurement problems that are special to particular scientific fields, like neuropsychology or quantum mechanics, are beyond the scope of this essay.

<sup>3</sup> When exactly the current wave of scholarship began is a matter of interpretation. Two early influential monographs are Hasok Chang's (2004) *Inventing Temperature* and Marcel Boumans' (2005a) *How Economists Model the World into Numbers*. However, several works belonging to the same strand were published earlier, including Chang (1995, 2001), Boumans (1999), Morgan and Klein (2001) and Mari (2000, 2003).

<sup>4</sup> Bridgman's empiricist caution was inspired by the success of Einstein's special relativity theory, which exposed the naive assumptions behind classical, absolute conceptions of space and time and replaced them with operational concepts. Bridgman's operational analysis was intended to "render unnecessary the services of the unborn Einsteins." (1927, 24)

<sup>5</sup> Bridgman later revised his account and no longer claimed operationalism was a comprehensive theory of meaning (1938; Chang 2009, section 2.1)

<sup>6</sup> For an exception, see Dingle (1950).

<sup>7</sup> Another influential conventionalist account of measurement is offered by Ellis (1966). Ellis takes quantity concepts to be cluster concepts and defines them on the basis of a *collection* of measurement operations, rather than a single operation.

<sup>8</sup> Many of the standard objections to logical empiricism are summarized in Uebel (2011, Section 3). A possible reason for the neglect of the problem of coordination by realist philosophers of science might be the anti-realist lessons conventionalists tended to draw from this problem. It must nonetheless be emphasized that the problem of coordination is a general semantic and epistemological problem that does not presuppose any particular metaphysics, and is therefore relevant for realists and anti-realists alike.

<sup>9</sup> For exceptions to the philosophical neglect of measurement during this period (apart from discussions of the Representational Theory) see references mentioned in footnotes 2 and 19, as well as Kyburg (1984).

<sup>10</sup> Earlier conventionalists like Mach and Poincaré were more sensitive to the historical dimensions of coordination than some of their logical positivist successors.

<sup>11</sup> See also McClimans & Browne (2012) and Angner (2013). For additional challenges to measurement in the social sciences see Chang and Cartwright (2008, 372-5).

<sup>12</sup> In what follows I will use the word 'object' to refer to a system under measurement. This designation is meant to cover processes and events as well.

<sup>13</sup> Helmholtz formulated the question thus: "what is the objective meaning of expressing, through denominate numbers, the relations of real objects as magnitudes, and under what conditions can we do this?" (1930 [1887], 4)

<sup>14</sup> See Helmholtz (1930 [1887]), Hölder (1901) and Russell (1903, ch. 12). Discussions about the nature of quantity and magnitude date back to ancient Greece. I limit my survey of modern measurement theories to the period commonly associated with such theories in the literature. For historical surveys of measurement theory see Savage and Ehrlich (1992), Michell and Ernst (1996), Diez (1997a, 1997b), Luce and Suppes (2002) and Michell (2003).

<sup>15</sup> These comments concern the possibility of a fundamental measurement of hardness. Other, 'indirect' methods of measuring hardness, like those proposed by Brinell, Vickers and Knoop, produce results that are represented on ratio scales. See Tabor (1970) for discussion.

<sup>16</sup> As the same number may represent several empirical objects, e.g. different rods of the same length, RTM focuses on many-to-one rather than one-to-one mappings, and therefore talks of homomorphisms rather than more specifically about isomorphisms (cf. Krantz et al 1971, 8 ff. 1).

<sup>17</sup> Campbell defined measurement as "the process of assigning numbers to represent qualities" (1920, 267). Stevens, paraphrasing on Campbell, defined measurement as "the assignment of numerals to objects or events according to rules" (1946, 677).

<sup>18</sup> Describing the aims of measurement theory, Roberts notes: "We are not interested in a measuring apparatus and in the interaction between the apparatus and the objects being measured. Rather, we attempt to describe how to put measurement on a firm, well-defined foundation." (1979, 3; qtd. in Boumans 2012b, 401)

<sup>19</sup> Precursors to the inferential view of measurement are Bogen and Woodward's discussion of the measurement of the melting point of lead (1988, 307-310), Rothbart and Slayden's analysis of absorption spectrometry (1994) and Franklin's analysis of calibration (1997).

---

<sup>20</sup> As van Fraassen notes, the mapping of indications to outcomes does not have to be deterministic. In the case of Quantum Mechanical measurements, mappings are stochastic functions that correlate probabilities of object and indicator states conditional upon their measurement (2008, 151-2).

<sup>21</sup> Unlike the authors of RTM, who take the term ‘model’ to mean a set-theoretical structure that interprets a formal language, the authors just cited use the term ‘model’ to denote an abstract representation of a system constructed from simplifying assumptions. This notion of ‘model’ is developed by Morrison (1999), Morrison and Morgan (1999) and Cartwright (1999), among others.

<sup>22</sup> This summary oversimplifies Morrison’s discussion, which covers the epistemic functions of materiality, the measurement-calculation distinction and the different levels of modelling involved in computer simulation in detail. For additional discussion on the distinction between experiment and simulation see Morgan (2005), Parker (2009) and Winsberg (2010, chapter 4).

<sup>23</sup> The problem of observational grounding should not be confused with van Fraassen’s (2012) criterion of empirical grounding for scientific theories. Van Fraassen’s criterion implicitly supposes a clear distinction between empirical and theoretical activities, whereas the problem of observational grounding arises precisely because this distinction is increasingly difficult to maintain.

## Works cited

- Alexandrova, Anna. ‘First Person Reports and the Measurement of Happiness.’ *Philosophical Psychology* 21.5 (2008): 571-583.
- Angner, Erik. ‘Is it Possible to Measure Happiness? The argument from measurability.’ *European Journal for Philosophy of Science* 3 (2013): 221-240.
- Audoin, Claude and Bernard Guinot. *The Measurement of Time: Time, Frequency and the Atomic Clock*. Trans. Stephen Lyle. Cambridge: Cambridge University Press, 2001.
- Bogen, James, and James Woodward. ‘Saving the Phenomena.’ *The Philosophical Review* 97.3 (1988): 303-352.
- Boumans, Marcel. ‘Representation and Stability in Testing and Measuring Rational Expectations.’ *Journal of Economic Methodology* 6 (1999): 381-401.
- . *How Economists Model the World into Numbers*. Routledge, 2005a.
- . ‘Truth versus Precision.’ *Logic, Methodology and Philosophy of Science: Proceedings of the Twelfth International Congress*. Eds. P. Hájek, L. Valdés-Villanueva and D. Westerstaahl. King’s College Publications. 2005b. 257-269.
- . ‘Measurement outside the laboratory.’ *Philosophy of Science* 72 (2005c): 850-863.
- . ‘The difference between answering a ‘why’ question and answering a ‘how much’ question.’ *Simulation: Pragmatic Construction of Reality*. Eds. Johannes Lenhard, Günter Küppers, and Terry Shinn. Dordrecht: Springer. 2006. 107-124.
- . ‘Invariance and Calibration.’ *Measurement in Economics: A Handbook*. Ed. Marcel Boumans. London: Elsevier. 2007. 231-248.
- . ‘Grey-Box Understanding in Economics.’ *Scientific Understanding: Philosophical Perspectives*. Eds. Henk W. de Regt, Sabina Leonelli and Kai Eigner. Pittsburgh: University of Pittsburgh Press. 2009. 210-229.
- . ‘Modeling Strategies for Measuring Phenomena In- and Outside the Laboratory.’ *The European Philosophy of Science Association Proceedings* 1 (2012a): 1-11.
- . ‘Measurement in Economics.’ *Philosophy of Economics*. Vol. 13 of Handbook of the Philosophy of Science. Ed. Uskali Mäki. Oxford: Elsevier, 2012b. 395-423.
- Bridgman, Percy W. *The Logic of Modern Physics*. New York: Macmillan, 1927.
- . ‘Operational Analysis.’ *Philosophy of Science* 5 (1938): 114-131.
- Byerly, Henry C. and Lazara, Vincent A. ‘Realist Foundations of Measurement.’ *Philosophy of Science* 40.1 (1973): 10-28.
- Campbell, Norman R. *Physics: the Elements*. London: Cambridge University Press, 1920.
- Carnap, Rudolf. *An Introduction to the Philosophy of Science*. 1966. Ed. by Martin Gardner. NY: Dover. 1995.
- Cartwright, Nancy. *Nature’s Capacities and Their Measurement*. New York: Oxford University Press, 1989.
- . *The Dappled World: A Study of the Boundaries of Science*. Cambridge: Cambridge University Press, 1999.
- Chang, Hasok. ‘Circularity and Reliability in Measurement.’ *Perspectives on Science* 3.2 (1995): 153-172.

- . ‘Spirit, air, and quicksilver: The search for the "real" scale of temperature.’ *Historical Studies in the Physical and Biological Sciences*, 31.2 (2001): 249-284.
- . *Inventing Temperature: Measurement and Scientific Progress*. Oxford University Press, 2004.
- . ‘Scientific Progress: Beyond Foundationalism and Coherentism.’ *Royal Institute of Philosophy Supplement* 61 (2007): 1-20.
- . ‘Operationalism.’ *The Stanford Encyclopedia of Philosophy*. Ed. E.N. Zalta. 2009. <<http://plato.stanford.edu/archives/fall2009/entries/operationalism/>>
- Chang, Hasok, and Nancy Cartwright. ‘Measurement.’ *The Routledge Companion to Philosophy of Science*. Ed. Psillos, Stathis. and Curd, Martin. New York: Routledge, 2008. 367-375.
- Diez, José A. ‘A Hundred Years of Numbers. An Historical Introduction to Measurement Theory 1887-1990 – Part 1.’ *Studies in History and Philosophy of Science* 28.1 (1997a): 167-185.
- . ‘A Hundred Years of Numbers. An Historical Introduction to Measurement Theory 1887-1990 – Part 2.’ *Studies in History and Philosophy of Science* 28.2 (1997b): 237-265.
- Dingle, Herbert. ‘A Theory of Measurement.’ *The British Journal for the Philosophy of Science*, 1.1 (1950): 5-26.
- Ellis, Brian. *Basic Concepts of Measurement*. Cambridge University Press, 1966.
- Franklin, Allan. *The Neglect of Experiment*. Cambridge University Press, 1986.
- . ‘Calibration.’ *Perspectives on Science* 5.1 (1997): 31-80.
- Franklin, Allan, et al. ‘Can a Theory-Laden Observation Test the Theory?’ *The British Journal for the Philosophy of Science* 40.2 (1989): 229-231.
- Frigerio, Aldo, Alessandro Giordani and Luca Mari. ‘Outline of a general model of measurement.’ *Synthese* 175 (2010): 123-149.
- Galison, Peter. *Einstein’s Clocks, Poincaré’s Maps: Empires of Time*. New York and London: W.W. Norton, 2003.
- Gillies, D.A. ‘Operationalism.’ *Synthese* 25.1 (1972): 1-24.
- Hacking, Ian. *Representing and Intervening*. Cambridge University Press, 1983.
- Helmholtz, H. von. *Counting and measuring*. 1887. Trans. Charlotte Lowe Bryan. New Jersey: D. Van Nostrand, 1930.
- Hölder, Otto. ‘Die Axiome der Quantität und die Lehre vom Mass.’ *Berichte über die Verhandlungen der Königlich Sächsischen Gesellschaft der Wissenschaften zu Leipzig, Mathematische-Physische Klasse* 53 (1901): 1–64.
- Hoover, Kevin and Michael Dowell. ‘Measuring Causes: Episodes in the Quantitative Assessment of the Value of Money.’ *The Age of Economic Measurement*. Annual supplement to vol. 33 of *History of Political Economy*. Eds. Judy Klein and Mary Morgan. 2001. 137-161.
- JCGM (Joint Committee for Guides in Metrology). *International Vocabulary of Metrology - Basic and general concepts and associated terms (VIM)*. 3rd ed. Sèvres: JCGM, 2008. <<http://www.bipm.org/en/publications/guides/vim.html>>
- Jones, Tony. *Splitting the Second: The Story of Atomic Time*. Bristol and Philadelphia: Institute of Physics Publishing, 2000.
- Krantz, David, H. R. Duncan Luce, Patrick Suppes and Amos Tversky. *Foundations of Measurement*. Vol 1. New York and London: Academic Press, 1971. Suppes et al. Vol. 2. 1989. Luce et al. Vol. 3. 1990.
- Kuhn, Thomas S. ‘The Function of Measurement in Modern Physical Sciences.’ 1961. Reprinted in *The Essential Tension: Selected Studies in Scientific Tradition and Change*. Chicago: University of Chicago Press. 1977. 178-224.
- Kyburg, Henry H. Jr. *Theory and Measurement*. Cambridge: Cambridge University Press, 1984.
- . ‘Measuring Errors of Measurement.’ *Philosophical and Foundational Issues in Measurement Theory*. Eds. Savage and Ehrlich. New Jersey: Lawrence Erlbaum, 1992. 75-91.
- Latour, Bruno. *Science in Action*. Cambridge: Harvard University Press, 1987.
- Luce, R. Duncan and Patrick Suppes, ‘Representational Measurement Theory.’ *Stevens' Handbook of Experimental Psychology*. Eds. J. Wixted and H. Pashler. 3rd Ed. Vol. 4: Methodology in Experimental Psychology. New York: Wiley, 2002. 1-41.
- Lepître, Alain. ‘Epistemology of Measurement in the Social Sciences: Historical and Contemporary Perspectives.’ *Social Science Information* 42 (2003): 451-462.
- Mach, Ernst. *Principles of the Theory of Heat*. 1896. Trans. Thomas J. McCormack. Dordrecht: D. Reidel, 1986.
- Mari, Luca. ‘Beyond the representational viewpoint: a new formalization of measurement.’ *Measurement* 27 (2000): 71-84.
- . ‘Epistemology of Measurement.’ *Measurement* 34 (2003): 17-30.
- . ‘The problem of foundations of measurement.’ *Measurement* 38 (2005a): 259-266.
- . ‘Models of the Measurement Process.’ *Handbook of Measuring Systems Design*. Vol 2. Ed. P. Sydenman and R. Thorn. Wiley, 2005b. Chap. 104.

- Mari, Luca and Alessandro Giordani. 'Modeling measurement: error and uncertainty.' *Error and Uncertainty in Scientific Practice*. Eds. Marcel Boumans, Giora Hon and Arthur Petersen. Pickering & Chatto, forthcoming in 2013.
- McClimans, Leah. 'A theoretical framework for patient-reported outcome measures.' *Theoretical Medicine and Bioethics* 31 (2010): 225-240.
- McClimans, Leah and Browne, John P. 'Quality of life is a process not an outcome.' *Theoretical Medicine and Bioethics* 33 (2012): 279-292.
- Mets, Ave. 'Measurement theory, nomological machine and measurement uncertainties (in classical physics).' *Studia Philosophica Estonica* 5.2 (2012): 167-186.
- Michell, Joel. 'Numbers as Quantitative Relations and the Traditional Theory of Measurement.' *British Journal for the Philosophy of Science* 45 (1994): 389-406.
- . 'Epistemology of Measurement: the Relevance of its History for Quantification in the Social Sciences.' *Social Science Information* 42.4 (2003): 515-534.
- Michell, Joel and Catherine Ernst. 'The Axioms of Quantity and the Theory of Measurement.' *Journal of Mathematical Psychology* 40 (1996): 235-252.
- Morgan, Mary. 'Making measuring instruments.' *The Age of Economic Measurement*. Annual supplement to vol. 33 of *History of Political Economy*. Eds. Judy Klein and Mary Morgan. 2001. 235-251.
- . 'Experiments versus Models: New Phenomena, Inference, and Surprise.' *Journal of Economic Methodology* 12.2 (2005): 317-329.
- Morgan, Mary and Judy L. Klein (eds.) *The Age of Economic Measurement*. Annual supplement to vol. 33 of *History of Political Economy*, 2001.
- Morrison, Margaret. 'Models as Autonomous Agents.' *Models as Mediators: Perspectives on Natural and Social Science*. Eds. Mary Morgan and Margaret Morrison. Cambridge: Cambridge University Press, 1999. 38-65.
- . 'Models, measurement and computer simulation: the changing face of experimentation.' *Philosophical Studies* 143 (2009): 33-57.
- Morrison, Margaret, and Mary Morgan. 'Models as Mediating Instruments.' *Models as Mediators: Perspectives on Natural and Social Science*. Eds. Mary Morgan and Margaret Morrison. Cambridge: Cambridge University Press, 1999. 10-37.
- Parker, Wendy. 'Does Matter really Matter: Computer Simulations, Experiments, and Materiality' *Synthese* 169.3 (2009): 483-496.
- . 'Models, Measurement and the Construction of Global Climate Datasets.' Draft of paper presented at the Dimensions of Measurement conference, Bielefeld 2013.
- Poincaré, Henri. 'The Measure of Time. 1898. *The Value of Science*. New York: Dover, 1958. 26-36.
- . *Science and Hypothesis*. 1905. Trans. W.J. Greenstreet. New York: Cosimo, 2007.
- Porter, Theodore M. *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. New Jersey: Princeton University Press, 1995.
- . 'Precision'. *Measurement in Economics: A Handbook*. Ed. Marcel Boumans. London: Elsevier. 2007. 343-356.
- Riordan, Sally. 'The Objectivity of Scientific Measures'. Manuscript under review.
- Reichenbach, Hans. *The Philosophy of Space and Time*. 1927. Courier Dover Publications, 1958.
- Roberts, Fred S. *Measurement Theory with Applications*. Providence: American Philosophical Society, 1979.
- Rothbart, Daniel and Suzanne W. Slayden. 'The Epistemology of a Spectrometer.' *Philosophy of Science* 61 (1994): 25-38.
- Russell, Bertrand. *The Principles of Mathematics*. New York: W.W. Norton, 1903.
- Savage, C. Wade and Philip Ehrlich. 'A brief introduction to measurement theory and to the essays.' *Philosophical and Foundational Issues in Measurement Theory*. Eds. Savage and Ehrlich. New Jersey: Lawrence Erlbaum, 1992. 1-14.
- Schaffer, Simon. 'Late Victorian metrology and its instrumentation: a manufactory of Ohms.' *Invisible Connections: Instruments, Institutions, and Science*. Ed. Robert Bud and Susan E. Cozzens. Cardiff: SPIE Optical Engineering. 1992. 23-56.
- Shapere, Dudley. 'The Concept of Observation in Science and Philosophy.' *Philosophy of Science* 49.4 (1982): 485-525.
- Stevens, S. S. 'On the theory of scales of measurement.' *Science* 103 (1946): 677-680.
- Suppes, Patrick. 'A set of independent axioms for extensive quantities.' *Portugaliae Mathematica* 10 (1951): 163-172.
- . 'Models of Data.' *Logic, methodology and philosophy of science: proceedings of the 1960 International Congress*. Ed. Ernest Nagel. Stanford: Stanford University Press. 1962. 252-261.
- Swoyer, Chris. 'The Metaphysics of Measurement.' *Measurement, Realism and Objectivity*. Ed. John Forge. Reidel, 1987. 235-290.
- Tabor, D. 'The hardness of solids.' *Review of Physics in Technology* 1.3 (1970): 145-179.
- Tal, Eran. 'How Accurate Is the Standard Second?' *Philosophy of Science* 78.5 (2011): 1082-96.

- . ‘The Epistemology of Measurement: A Model-Based Account.’ PhD Dissertation. University of Toronto, 2012. <<http://hdl.handle.net/1807/34936>>
- Teller, Paul. ‘The concept of measurement-precision.’ *Synthese*, 190 (2013): 189-202.
- . ‘Measurement accuracy realism.’ Manuscript under review.
- Uebel, Thomas. ‘Vienna Circle.’ *The Stanford Encyclopedia of Philosophy*. Ed. E.N. Zalta. 2011. <<http://plato.stanford.edu/archives/sum2012/entries/vienna-circle/>>.
- van Fraassen, Bas C. *Scientific Representation: Paradoxes of Perspective*. Oxford University Press, 2008.
- . ‘The perils of Perrin, in the hands of philosophers.’ *Philosophical Studies*, 143 (2009): 5-24.
- . ‘Modeling and Measurement: The Criterion of Empirical Grounding.’ *Philosophy of Science*. 79.5 (2012): 773-784.
- Winsberg, Eric. *Science in the Age of Computer Simulation*. Chicago and London: University of Chicago Press, 2010.