Stevens' forgotten crossroads: the divergent measurement traditions in the physical and psychological sciences from the mid-twentieth century

Joshua A. McGrane*

Pearson Psychometric Laboratory, Faculty of Education, The University of Western Australia, Perth, WA, Australia

The late nineteenth and early twentieth centuries saw the consolidation in physics of the three main traditions that predominate in discussions of measurement theory. These are: (i) the systematic tradition pioneered by Maxwell (1873); (ii) the representational tradition pioneered by Campbell (1920); and (iii) the operational tradition pioneered by Bridgman (1927). These divergent approaches created uncertainty about the nature of measurement in the physical sciences and provided Stevens (1946) with an opportunity and rationale to, in effect, reinvent the definition of scientific measurement. Stevens appropriated the representational and operational traditions as the sole basis for his definition of measurement, excluding any place for the systematic approach. In committing to Stevens' path, the psychological sciences were blinded to the advances made in metrology, the establishment of the International System (SI) and the standard units contained within this system. These advances were only possible due to the deep conceptual and instrumental connections between the system of physical units and the body of physical theory and laws developed over the preceding centuries. It is argued that if the psychological sciences are to ever achieve equivalent methodological advances, they must bridge this "metrological gap" created by Stevens' measurement crossroads and understand the ways in which the systematic approach advanced measurement. This means that psychological measurement needs to be de-abstracted, rid of operational rules for numerical assignment and set upon a foundation of quantitative theory, definition and law. In the absence of such theoretical foundations, claims of measurement in the psychological sciences remain a methodological chimera.

Keywords: measurement, metrology, S. S. Stevens, operationalism, representationalism, measurement systems, psychometrics

Introduction

The late nineteenth century and the early twentieth century, leading up to Stanley Smith Stevens' measurement crossroads, saw the consolidation in physics of the three main theoretical approaches to measurement that still predominate to varying extents in both the physical and psychological sciences. These are the systematic, representational and operational approaches to measurement. This discussion will address the body of writing of three of the major contributors of each of

these approaches; Maxwell, Campbell, and Bridgman. It will outline how these approaches differ with respect to their definition of measurement, their demarcation of measurable and non-measurable phenomena and their relation to the wider body of scientific theory, definition and law. It will become clear that each of these theorists, in fact, endorsed a systematic approach to measurement in some form. However, Stevens selectively appropriated aspects of the representationalist and operationalist approaches to justify a liberalized definition of scientific measurement where a systematic approach was excluded. The publication of this liberalized definition represents a crossroads where the development of psychological measurement fundamentally diverged from, and was blinded to, developments in metrology. The following historical overview is intended to illuminate this crossroads and the "metrological gap" it created.

The Systematic, Representationalist and Operationalist Approaches to Measurement

One of the modern forefathers of the systematic approach was James Clerk Maxwell, one of the most influential physicists of all time. Maxwell, in his seminal *Treatise on Electricity and Magnetism*, defined measurement as consisting of two essential factors, "One of these is the name of a certain known quantity of the same kind as the quantity to be expressed, which is taken as a standard of reference [i.e., the unit]. The other component is the number of times [i.e., the ratio] the standard is to be taken in order to make up the required quantity" (Maxwell, 1873, p. 1). As a consequence, physical quantities are the subject matter of measurement and a particular kind of quantity can only be measured if there is a scientifically established unit to measure it by. As Maxwell states, "there must be as many different units as there are different kinds of quantities to be measured" (Maxwell, 1873, p. 1). His view would form the foundation of the standard scientific definition of measurement and the establishment of an international system of units (SI; de Boer, 1995).

Maxwell was a pioneer of a system of units built on the foundation of what he referred to as fundamental units (now known as base units); so called as they could not be expressed as relations between or powers of other fundamental units. Non-fundamental units, or what he referred to as derived units (also the contemporary terminology), were established by quantity equations that specified relations between and powers of the fundamental units on the basis of physical laws. Thus, the link between measurement and the wider body of physical theory was fundamental and explicit. This resulted in the system being invariant under different national sets of measurement standards and meant that as science progressed and new quantities were discovered, derived units for measuring them could be easily reconciled with and integrated into the system, or the system could be extended to include new kinds of fundamental units.

Furthermore, Maxwell pioneered the movement away from the use of standard objects as units, such as France's meter bar, to units defined in terms of the behaviors of microscopic molecules that are invariant across objects, locations and time. "If…we wish

to obtain standards of length, time, and mass which shall be absolutely permanent, we must seek them not in the dimensions, or the motion, or the mass of our planet, but in the wavelength, the period of vibration, and the absolute mass of these imperishable and unalterable and perfectly similar molecules" (Maxwell, 1870, p. 421). This approach cleared the path for more accurate and precise measurement by tying the systems of units to advances in physical theory and scientific practices.

The most renowned and influential advocate of the representational approach around this era was the English physicist and philosopher of science, Norman Campbell, who introduced his highly influential ideas on measurement in *Physics: The Elements* (Campbell, 1920). Although Campbell demonstrated knowledge of Maxwell's work (Campbell, 1921, p. 155), he was heavily influenced by recent mathematical advances, particularly those of Bertrand Russell, which lead him to define measurement more liberally as "the assigning of numbers to represent properties" in order to "enable the powerful weapon of mathematical analysis to be applied to the subject matter of science" (Campbell, 1920, pp. 267–268). In this context, Campbell intended "number" in the mathematical sense, i.e., a member of the "real" number line. Consequently, measurement was understood as a process of discovering the relation between empirical properties and mathematical objects so that numbers could be meaningfully assigned to represent the former, i.e., the empirical property could be treated analogously as a mathematical object.

Campbell argued that the possibility of this representation was dependent upon the experimental satisfaction of a number of "rules" of measurement (Campbell, 1921). In essence, these rules were intended to determine "the possibility or impossibility of finding in connection with the properties a physical significance for the [mathematical] process of addition" (Campbell, 1920, pp. 277–278). Fundamental measurement is possible for empirical properties where the "rules of addition" may be "directly," experimentally examined, e.g., we can "directly" examine the measurability of length by the relations between rigid rods and their concatenation. Derived measurement, on the other hand, relies upon the existence of what he termed "numerical laws" between fundamental measures (Campbell, 1921). Although Campbell explicitly acknowledged the relationship between scientific laws and derived measurement, he also argued that these laws could only be understood in a mathematical sense; e.g., a mass cannot actually be divided into a number of volumes, "though that there may be such a [numerical] relation has been suggested to us by the study of the physical property [density]" (Campbell, 1921, p. 138).

Campbell goes on to make a distinction between "arbitrary" and "true" measurement, and provides the assessment of Hardness by Moh's scale, and the assessment of temperature by the Celsius thermometer, as examples of the former (Campbell, 1920). The categorization of the Celsius measure of temperature as "arbitrary" is particularly intriguing. Campbell argued that the division of the Celsius scale into equal-interval units between the fixed points (i.e., the freezing and boiling points of water) was entirely arbitrary as it was "unconnected with any *laws of temperature* [emphasis added]" (Campbell, 1920, p. 358). In more general terms, he argued this distinction centered on whether you have a true system of measurement, that is, one where the

"numerical laws" define "true derived magnitudes" (Campbell, 1920, p. 360), which he describes as, "one of the most fruitful sources of scientific progress" (Campbell, 1920, p. 361). In other words, "true" measurement is only possible with a "true" system of measurement, where measurement units are founded upon and their relations reflect quantity-specific laws.

The Nobel Prize winning physicist Percy Bridgman introduced operationalism as a broad philosophy of science in his classic book, *The Logic of Modern Physics* (Bridgman, 1927), but was explicit that his earlier publication, *Dimensional Analysis* (Bridgman, 1922), was a precursory application of his operational theory (Bridgman, 1959). In this earlier publication, Bridgman stated that, "for each different kind of quantity we have a different rule of operation by which we measure it, that is, associate the quantity with a number" (Bridgman, 1922, p. 17). This definition of measurement is then elaborated with respect to "primary" and "secondary" quantities, which are distinguished in the way that the numbers are operationally associated with them. For primary quantities, "certain rules of operation must be set up, establishing the physical procedure by which it is possible to measure a length in terms of a particular length…or in general…by which it is possible to measure any primary quantity directly in terms of units of its own kind" (Bridgman, 1922, p. 18).

Measurement of secondary quantities is performed by, "making measurements of certain quantities of the first [primary] kind associated with the quantity under consideration, and then combining the measurements of the associated primary quantities according to certain rules which give a number that is defined as the measure of the secondary quantity in question" (Bridgman, 1922, p. 19). Moreover, these "certain rules" to define secondary quantities must satisfy "…the same requirement that we [specified] for primary quantities, namely, that the ratio of the numbers measuring any two concrete examples of a secondary quantity shall be independent of the size of the fundamental units used in making the required primary measurements" (Bridgman, 1922, p. 19). In elaborating his definition of measurement, Bridgman is describing a systematic approach to measurement in the vein of Maxwell, albeit couched in operational terms.

This operationalist reinterpretation of Physics culminated in his infamous operational principle that any "concept is synonymous with a corresponding set of operations" (p. 5), which Bridgman would later rue ever stating (Bridgman, 1959). This principle led physicist Herbert Dingle to propose an even more reductively operational definition of measurement as, "any precisely specified operation that yields a number" (Dingle, 1950, p. 11). Whilst such a definition could be argued to be fairly benign as a description of existing physical measurement, where "precisely specified operations" already involved the sorts of physical laws and systems of units described as necessary for measurement to varying extents in the approaches above, the same could not be said for new disciplines attempting to establish a measurement base. The primacy given to sets of operations meant that, ultimately, the only restraint on claims of measurement for these new disciplines was the imaginations of their scientists to invent number yielding sets of operations.

The above overview provides a context to Stevens' claim that "measurement exists in a variety of forms" (Stevens, 1946, p. 677).

Around this era, some of the most prominent physicists were inconsistent and at times unclear on what they took measurement to be. Maxwell's systematic approach reserved measurement for physical quantities and ratios with their specific units, which were integrated into a coherent system reflecting physical laws and their quantity equations. Campbell's representationalist approach provided a somewhat more liberalized view of measurement where numbers are assigned to empirical properties, thus removing the requirement for physical quantities and their ratios. Although, "true measurement" involved a system of measurement not dissimilar to Maxwell's, where numerical laws between fundamentally measured properties correctly specify derived magnitudes/properties. Finally, Bridgman's operationalist approach defined measurement as an operational association between number and quantity. Whilst his explications of the operational rules of association describe a systematic approach, the confounding of concepts with operations led others to reduce measurement to number yielding operations. Thus, it was not only professionally convenient for Stevens to promote a liberal understanding of measurement, but arguably scientifically defensible given the leading thought of the day. However, his adoption of the more liberal path excluded the common thread throughout Maxwell, Campbell, and Bridgman's writings, i.e., some level of systematic understanding of measurement, and thus created a divergence between the physical and psychological sciences that continues today.

Stevens' Crossroads and the Path Beyond

The origins of psychological measurement were intertwined with a Physics tradition, which is unsurprising given that many of the earliest quantitative psychologists were themselves researchers in physical as well as psychological sciences, including Ernst Weber, Wilhelm Wundt, and Gustav Fechner (Gescheider, 2013). Early "psychophysical" research was explicitly concerned with establishing empirical relations and laws between physical and perceptual magnitudes and measuring the latter in units established by just-noticeable-difference (JND) experiments. Whilst such research was underpinned by questionable dualist assumptions and the dubious separation between JND "units" and their physical counterparts (see McGregor, 1935; Michell, 1997), it appears that the tradition shared commonalities with the systematic approach. The psychophysical approach to psychological measurement spread to a range of different perceptual modalities, and firstly entailed establishing a kind of perceptual quantity by way of empirical relations and laws, and secondly establishing a specific unit to measure that quantity. However, no attempt was made to systematize these laws, their corresponding equations or the units of perceptual magnitude.

In terms of Stevens' contemporaries, Louis Thurstone, a major originator of modern psychometrics, described measurement as allocating an attribute of an object "to a point [i.e., a number] on an abstract [linear] continuum… which requires some point at which counting begins, called the origin, and some unit of measurement in terms of which the counting is done" (Thurstone, 1931, p. 259). Douglas McGregor, a colleague of Stevens at Harvard University, defined measurement as, "the process of assigning numbers to

represent quantities" (McGregor, 1935, p. 249; which he adopted from Cohen and Nagel, 1934). In both of these definitions, one can recognize the muddying of the measurement concept that happened in the twentieth century, and in particular, the influences of the representationalist approach. However, parallels with the physical understanding of scientific measurement remain with the requirement for quantity and/or units of measurement. Stevens would liberalize measurement from either of these requirements.

Stevens was explicit about his influence by Campbell when redefining measurement "as the assignment of numerals to objects or events according to rules" (Stevens, 1946, p. 677). He had an intimate knowledge of Campbell's work, as the latter played a dominant role in the Ferguson Committee that was set up by the British Association for the Advancement of Science to evaluate claims of measurement by psychophysicists, including Stevens' Sone scale of auditory sensation (Ferguson et al., 1940). However, unlike Campbell, Stevens did not restrict the "rules" to those that found an empirical link between physical and mathematical properties by satisfying the criterion of additivity. Rather, he introduced the now ubiquitous notion of "scales of measurement." Stevens explained that "scales are possible in the first place only because there is a certain isomorphism between what we can do with the aspects of objects and the properties of the numeral series" (Stevens, 1946, p. 677). However, his inclusion of the nominal and ordinal scales, or what he later ambivalently referred to as "the weaker forms of measurement" (Stevens, 1958, p. 384), liberated Campbell's approach to the point where any consistently followed rule that resulted in a numerical assignment could be called measurement (Stevens, 1946).

But Campbell was not the only influence upon Stevens, as his ultimate liberation of measurement was inspired by his interpretation of Bridgman's operationalism (Stevens, 1939). Specifically, Stevens stated that "the type of scale [of measurement] achieved depends upon the character of the empirical operations performed…once selected, the operations determine that there will eventuate one or another of the scales" (Stevens, 1946, p. 677). So, the demarcation between "weak" and "strong" forms of measurement is simply the researcher's choice of operations. It is little wonder that this definition has led to the proliferation of tens of thousands of so called "measures" in the psychological sciences, many of which are interpreted as analogous to the physical measures of the SI, despite the latter numbering less than 100, each of which is underpinned by explicit physical theory and law. In fact, Stevens seemed to take pride in this proliferation in a later paper when he stated, "the variety of rules invented thus far for the assignment of number has already grown enormous, and novel means of measuring continue to emerge" (Stevens, 1968, p. 850).

Despite the dominance of Stevens' definition in the psychological sciences (Michell, 1999), a number of mathematical psychologists have acknowledged that his approach diverged too far from the understanding of measurement in the physical sciences (Krantz, 1972; Luce and Narens, 1987). Their proposed solution reconciled Stevens' scales of measurement with a more strictly Campbell influenced, representational approach. So, like Campbell and Stevens, measurement was taken to be the assignment of numbers, but, this assignment depended upon the mapping of an empirical, qualitative relational system on to the mathematical,

quantitative number system. Thus, unlike Stevens, the possible empirical mapping or "representation" determined the type of measurement, but unlike Campbell, this mapping did not exclusively depend upon the discovery of an empirical analog of addition. The proponents of such an approach to measurement developed it in to an extensive abstract theory of measurement, which they referred to as the *Foundations of Measurement* (Krantz et al., 1971).

This abstract theory has been held in high esteem by contemporary philosophers of measurement, yet it has had little practical influence outside of mathematical psychology and economic utility theory (Cliff, 1992; Kyngdon, 2013). Stevens, ironically, provided insight into its lack of influence in science when he stated, "measurement models sometimes drift off into the vacuum of abstraction and become decoupled from their concrete reference…A full theory of measurement cannot detach itself from the empirical substrate that gives it meaning" (Stevens, 1968, p. 854). The irony of Stevens' comment is that his own re-definition broadened measurement beyond any commitment to an empirical substrate, i.e., an actual quantitative property that we attempt to estimate, and relies on nothing more than a consistent choice of empirical operations that may be constructed at the will of the researcher (Luce, 1997). Nonetheless, his comment does strike at a pertinent point regarding the abstract theory. In creating what they believed to be a logical basis for measurement, the abstract theorists' mathematical account of measurement was grounded in mathematical formalisms and axioms instead of empirical theory and law (Berka, 1982).

By perpetuating the representationalist path forged by Campbell, Stevens and their intellectual forebears, the abstract theorists missed the advances that had been made by the developers of the SI in the mathematical representation of measurement. This representation was developed to be empirically grounded in physical quantities, laws and the equations that express them. The development of the SI has "indisputably [become] the basis of all aspects of modern metrology" (Quinn and Kovalevsky, 2005, p. 2313), and in turn, all aspects of measurement in the physical sciences. Over the course of the twentieth and twenty-first centuries, this development has required the clarification of conceptual confusions, the development of empirically-grounded symbolic representations and the deepening of the interconnection between measurement and physical theory.

The Developments of the SI and Metrology

A thorough historical overview of the development is beyond the scope of this paper (see Silsbee, 1962; Quinn, 2011, for such a review), and thus I will concentrate on a number of key developments that occurred around the time of Stevens' crossroads and beyond. Briefly, as previously mentioned, Maxwell, in collaboration with William Thomson under the auspices of the British Association for the Advancement of Science, pioneered the development of a coherent system of units, the centimeter-gram-second (CGS) system. Under this system, a number of derived units for the measurement of various mechanical and electromagnetic quantities are defined in terms of the base units for the physical quantities, length, mass and time. The CGS system was

superseded by the meter-kilogram-second (MKS) system in the 1940s, which evolved into the first SI in 1960.

Between 1948 and 1950 an additional base unit was added to the MKS system, the Ampere, for the physical quantity, electrical current (giving the MKSA system), and the quantity equations for electromagnetic phenomena were changed to their "rationalized" form (Silsbee, 1962). Whilst the technical details of this change are unimportant for the current paper, debates concerning it highlighted a number of conceptual confusions inherent in the metrology community, providing further evidence of the evolving nature of the concept, which would result in a more complete explication of the systematic approach as the basis of the SI specifically, and scientific measurement more generally (de Boer, 1995). These debates centered around two (fictitious) camps, Realist and Systematist. The former were typically applied scientists and engineers that routinely dealt with concrete (i.e., physically instantiated) quantities, concrete measurement units and numbers from concrete measurement procedures. The latter were typically more theoretical physicists who dealt with the abstract quantities and units that were used in mathematical statements of quantity equations (de Boer, 1995).

Similar to the abovementioned view of Campbell, the Realist camp were inclined to interpret quantity equations as simply numerical-value equations, thus ignoring the status of the quantities in the equations. This view was influenced by the concrete consideration that a quantity of one kind cannot literally be interpreted as being divided or multiplied by a quantity of another kind and nor can this product/quotient be realistically interpreted as equivalent to a quantity of a third kind. Furthermore, there is even doubt under this Realist view as to whether it is meaningful to express a magnitude of a quantity as equivalent to the product of a unit quantity of the same kind and a numerical value, as this is not reflective of a concrete process of measurement.

In reply, the Systematist camp argued that the Realist camp fundamentally misunderstood the symbolism of systems of measurement, which led to the widespread push for the acceptance of quantity calculus (QC) as a formal language to express the quantity equations that underpin the system of units (de Boer, 1995). QC was first introduced by Maxwell and further developed by Wallot as a form of mathematical representation for physical phenomena that gives primacy to physical quantities and their relations, rather than numbers, as per ordinary algebra (Humphry, 2011). Under this symbolism, abstract quantities are always paired with their abstract units (abstract quantities of the same kind), thus making explicit that any mathematical operation is between numerical values obtained from the ratio of a quantity and a unit. The main rationale for the use of QC was that the quantity equations (i.e., the mathematical expressions of physical laws) remained invariant across choices of units. Moreover, QC was an algebraically efficient, but still empirically grounded expression of the more empirically complete statement of the proportionality of ratios that quantity equations represent (Humphry, 2013b).

A concurrent development in the history of the SI was the evolution of unit definitions, as per Maxwell's aforementioned vision, from concrete, material prototypes to theoretical expressions of ontologically invariant physical phenomena. For example, the unit of measurement for time is theoretically defined in terms of the invariant frequency of radiation emission by a caesium atom during a particular physical transition (Tal, 2014). This definition is theoretical in the sense that the stipulated conditions for the absolute invariance of this frequency are not empirically realizable. Moreover, in the next iteration of the SI, the kilogram prototype is to be discarded and replaced with a unit definition based upon an application of the quantum-Hall and Josephson effects, which ties it to the Planck constant, an ontologically invariant physical relation (Quinn, 2011). Therefore, by the next SI, the complete integration of the representational systems of units, quantity equations and physical theory and law will be complete.

The abstract nature of the quantity equations and idealized definitions of units that underpin the systematic approach to measurement raises the question of how these may be empirically realized as concrete standards and measurement systems. This realization is the primary task of applied metrology, or the Realist camp to use de Boer's (1995) nomenclature, and its epistemology has come under increasing philosophical scrutiny in recent years (Mari, 2003; Tal, 2013). Whilst a thorough overview of this epistemology goes beyond the scope of this paper, the realizations of primary, secondary, etc., metrological standards are just as contingent upon physical theory and laws as their abstract representations within a systematic approach to measurement. Specifically the metrologically realized standards serve as concrete instantiations of physical theories and laws by approximately instrumentalizing a standardized magnitude of a quantitative property or process and providing the means to estimate the ratio of an unknown magnitude of the same quantity to the standardized magnitude (Quinn, 1997; Humphry, 2013a).

For example, the theoretical definition of the unit of time is approximately physically realized in a primary standard known as a caesium fountain clock, which is able to reproduce a specific duration of time, the second, with an extremely high level of precision by instantiating a cyclical physical process that draws upon theoretical knowledge of atomic structure, radioactivity, thermodynamics and gravitational force, amongst other numerous quantitative and non-quantitative physical phenomena (Tal, 2014). Because the caesium fountain clocks are only stable for a relatively short period of time, they are used to calibrate the standardized magnitude inherent in more stable secondary standards, atomic clocks, which rely upon more experimentally controllable electromagnetic effects. These atomic clocks provide the "ticks," i.e., one cycle of the physical process that instantiates the standard duration, of Coordinated universal time that may be counted to estimate the ratio of an unknown duration to the standardized duration, i.e., measure the unknown magnitude of time. Whilst the counting of these "ticks" may be thought of as simply a rule to assign numbers to temporal duration on a ratio scale, that would deeply trivialize the fundamentally theoretical nature of the concrete and abstract measurement systems that underpin atomic clocks and time measurement in general.

The Systematic Approach and Psychological Measurement

As the systematic approach has come to predominate in physical scientific measurement, the psychological sciences have

entrenched their separate path, complete with an array of atheoretical methodology, including (but certainly not limited to) operational definitions, statistical distributions and mathematical probability theory. In the physical sciences, the quantity status of some physical properties [e.g., temperature (Sherry, 2011)] was theoretically and empirically fought for over decades, if not centuries. In contrast, the psychological sciences have adopted practices where psychological quantities may be invented at the will of the researcher and attention is then focused upon ever more creative and technical means to impose "real number" mathematics upon psychological attributes with little to no theoretical justification for doing so. There can be little debate that the psychological sciences severely lack substantive quantitative theory, laws and equations, and that there are no scientifically established measurement units, let alone an integrated, coherent system. Moreover, there has been very little mainstream academic dialog and debate about this absence, although a groundswell has been building the past two decades (for examples see Michell, 1999; Trendler, 2009; Humphry, 2011; Sherry, 2011; Kyngdon, 2013). This is attributable in no short measure to Stevens' redefinition of measurement and the divergence with the physical tradition that it did not create, but deeply entrenched into the discipline. As Newman put it, since its inception, Stevens' definition has "stood like the Decalogue" (Newman, 1974, p. 137).

Despite the devout (and often implicit) following of so many psychological researchers, the divergence of Stevens' definition has not gone entirely unnoticed within the psychological sciences (Michell, 1997). Specifically, the abstract measurement theorists discussed above have attempted to redress its effects. As Duncan Luce put it, "No measurement theorist I know accepts Stevens' broad definition of measurement…the only sensible meaning for 'rule' is *empirically testable laws about the attribute* [emphasis added]" (Luce, 1997, p. 395). Whilst the above discussion is critical of such theorists for their complete abstraction of their measurement theory from any empirical grounding, the abstract structures identified by these theorists may be helpful to psychological researchers if re-grounded in empirical theory and experimentation (Krantz, 1972; Kyngdon, 2013).

Until such theorizing and experimentation is done, and the psychological sciences' measurement agenda adopts a systematic approach, any claims of measurement are premature and potentially misleading if taken to be comparable to measurement practices in metrology. This is not a trivial point, as, for example, Item Response Theorists regularly argue that their methods provide "interval-level" measures of psychological attributes akin to the measurement of temperature using a Celsius thermometer and present their findings using spatial representations (e.g., Bond and Fox, 2007), despite the glaring absence of any quantitative theory or ontologically defined unit[1] to justify such practices (Sherry, 2011). This claim of "interval-level" measurement is then used as justification for making further quantitative interpretations of test results, such as academic growth over time, which are substantively unclear and potentially meaningless in the absence of a clearly defined unit of measurement (Zwick, 1992). There

may, however, also be examples where the application of a more systematic approach bears little practical implication for current practices in the psychological sciences (Briggs and Weeks, 2009; Briggs, 2013). Clearly, further research is required to understand and elaborate the implications of a systematic approach.

One such implication may be that no psychological quantities and, therefore, no basis for a system of units are uncovered at all, which would perhaps be unsurprising given the complexity of psychological phenomena (Trendler, 2009). Given this, under a systematic understanding, the psychological sciences could not claim to measure anything (Michell, 2012). But, the prospect of such a finding should not be viewed pessimistically as diminishing the scientific status of the psychological sciences, as the physical sciences are filled with non-quantitative properties and methods (Sechrest and Sidani, 1995).

As an illustrative point, only a single unit of measurement has been established outside of Physics, the Mole, which is used in Chemistry to measure amount of chemical substance, and even this has been a contentious addition given its interrelationship with the continuous physical quantity, mass (Johansson, 2010; Cooper and Humphry, 2012). Moreover, scientists of various disciplines, including physics, chemistry, health science, clinical laboratory sciences, biology, engineering, biochemistry, food science, and molecular biology, which routinely deal with properties of systems that are only amenable to nominal examination have developed a vocabulary of nominal properties and examinations (VIN; Nordin et al., 2010). This is intended to provide a common and standardized language for scientific methods concerned with non-quantitative properties and relations of natural systems [and as a sister document to the International Vocabulary of Measurement (VIM)]. All physical sciences routinely apply a range of rigorous observational methodologies and their scientific validity is determined by the nature of their subject matter, not by normatively elevating one methodology above all others (Hibberd, 2014). This point was simply put by Johnson (1936, p. 351), "Those data should be measured which can be measured; those which cannot be measured should be treated otherwise."

Some psychological researchers have begun to heed this message by adapting complex network models from biological and ecological sciences to investigate psychological phenomena without any necessary assumption of underlying psychological quantities or measurement (Cramer et al., 2010; Borsboom and Cramer, 2013). Complex networks were created to scientifically model natural systems that are inherently dynamic, non-linear and show structural complexity (Strogatz, 2001); the kinds of properties that seem likely of psychological systems. These models may include both quantitative and non-quantitative elements, and as a consequence, are amenable to measurements and other forms of rigorous scientific observation. So whilst measurement substantiated by quantitative theory and laws, and systems of quantity equations and units may arguably remain a pipedream in the psychological sciences, a more general systematic approach stresses the fundamental nature of theory and observation to their scientific progress. This does not necessitate redefining measurement as rule-based numerical assignment, but rather a commitment to rigorously investigate the ontological statuses of psychological systems by methods that are theoretically and empirically

---

[1]Humphry (2013a) describes the unit inherent in IRT models, e.g., the "logit," as a "quasi-unit" because it is mathematically and not substantively defined.

substantiated (Barrett, 2003; Sherry, 2011; Maul, 2013; Hibberd, 2014).

Conclusion

Leading up to Stevens' crossroads, three main approaches to measurement were apparent which, at face value, diverged in their definitions of measurement, their demarcation of measurable phenomena and their emphasis of mathematical versus substantive theory in claims of measurement. It has been argued in a brief overview of the writings of three of the main initiators of each of these approaches that such differences may have been overstated, as each endorsed a systematic approach in some form. Such a conclusion was contingent upon a broader reading both within and across their body of work, rather than "reading" each position in terms of a "single slogan taken out of the context of the very paragraph in which it occurred" (Koch, 1992, p. 261).

Nonetheless, each approach had key differences and these differences provided a historical and conceptual precedence for Stevens to present an entirely liberalized definition of measurement. This definition created a significant divergence between the methodological developments of the physical and psychological sciences. The psychological sciences, for the most part, embarked on a "measure" proliferation exercise at will and seemingly *ad infinitum*, rather than scientifically examining the nature of psychological phenomena and determining appropriate methodology on that basis. Meanwhile, the physical sciences clarified key conceptual confusions concerning measurement, created a representational language that gives primacy to physical quantities and quantitative relations rather than numbers, and further entrenched the theory-measure nexus through the implementation of theoretical definitions of units. It has been argued that if the psychological sciences' measurement practices are to gain similar scientific credibility, this metrological gap must be spanned by the abandonment of Stevens' liberal definition and the adoption of a systematic approach to measurement. Quantitative theory and laws, as well as the system of physical equations and units that they determine, are the actual foundations of scientific measurement. Whilst these foundations remain largely unconsidered and unquestioned in the psychological sciences, claims of measurement remain a methodological chimera.

References

Barrett, P. (2003). Beyond psychometrics: measurement, non-quantitative structure, and applied numerics. *J. Manage. Psychol.* 18, 421–439. doi: 10.1108/02683940310484026

Berka, K. (1982). *Measurement: Its Concepts, Theories and Problems*. Dordrecht: Kluwer Academic Publishers.

Bond, T. G., and Fox, C. M. (2007). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. Mahwah, NJ: Lawrence Erlbaum.

Borsboom, D., and Cramer, A. O. (2013). Network analysis: an integrative approach to the structure of psychopathology. *Annu. Rev. Clin. Psychol.* 9, 91–121. doi: 10.1146/annurev-clinpsy-050212-185608

Bridgman, P. W. (1922). *Dimensional Analysis*. New Haven, CT: Yale University Press.

Bridgman, P. W. (1927). *The Logic of Modern Physics*. New York, NY: MacMillan.

Bridgman, P. W. (1959). P. W. Bridgman's "The Logic of Modern Physics" after Thirty Years. *Daedalus* 88, 518–526.

Briggs, D. C. (2013). Measuring growth with vertical scales. *J. Educ. Meas.* 50, 204–226. doi: 10.1111/jedm.12011

Briggs, D. C., and Weeks, J. P. (2009). The impact of vertical scaling decisions on growth interpretations. *Educ. Meas. Issues Pract.* 28, 3–14. doi: 10.1111/j.1745-3992.2009.00158.x

Campbell, N. R. (1920). *Physics: The Elements*. Cambridge: Cambridge University Press.

Campbell, N. R. (1921). *What is Science?* London: Methuen & Co. Ltd.

Cliff, N. (1992). Abstract measurement theory and the revolution that never happened. *Psychol. Sci.* 3, 186–190. doi: 10.1111/j.1467-9280.1992.tb00024.x

Cohen, M. R., and Nagel, E. (1934). *An Introduction to Logic and Scientific Method*. London: Routledge & Sons.

Cooper, G., and Humphry, S. M. (2012). The ontological distinction between units and entities. *Synthese* 187, 393–401. doi: 10.1007/s11229-010-9832-1

Cramer, A. O., Waldorp, L. J., van der Maas, H. L., and Borsboom, D. (2010). Comorbidity: a network perspective. *Behav. Brain Sci.* 33, 137–150. doi: 10.1017/S0140525X09991567

de Boer, J. (1995). On the history of quantity calculus and the international system. *Metrologia* 31, 405–429. doi: 10.1088/0026-1394/31/6/001

Dingle, H. (1950). A theory of measurement. *Br. J. Philos. Sci.* 1, 5–26. doi: 10.1093/bjps/I.1.5

Ferguson, A., Myers, C. S., Bartlett, R. J., Banister, H., Bartlett, F. C., Brown, W., et al. (1940). Quantitative estimates of sensory events: final report of the committee appointed to consider and report upon the possibility of quantitative estimates of sensory events. *Adv. Sci.* 1, 331–349.

Gescheider, G. A. (2013). *Psychophysics: The Fundamentals*, 3rd Edn. Hoboken, NJ: Taylor and Francis.

Hibberd, F. J. (2014). The metaphysical basis of a process psychology. *J. Theor. Philos. Psychol.* 34, 161. doi: 10.1037/a0036242

Humphry, S. M. (2011). The role of the unit in physics and psychometrics. *Measurement (Mahwah N. J.)* 9, 1–24. doi: 10.1080/15366367.2011.558442

Humphry, S. M. (2013a). A middle path between abandoning measurement and measurement theory. *Theory Psychol.* 23, 770–785. doi: 10.1177/0959354313499638

Humphry, S. M. (2013b). Understanding measurement in light of its origins. *Front. Psychol.* 4:113. doi: 10.3389/fpsyg.2013.00113

Johansson, I. (2010). Metrological thinking needs the notions of parametric quantities, units and dimensions. *Metrologia* 47, 219. doi: 10.1088/0026-1394/47/3/012

Johnson, H. M. (1936). Pseudo-mathematics in the mental and social sciences. *Am. J. Psychol.* 48, 342–351. doi: 10.2307/1415754

Koch, S. (1992). Psychology's Bridgman vs Bridgman's Bridgman an essay in reconstruction. *Theory Psychol.* 2, 261–290. doi: 10.1177/0959354392023002

Krantz, D. H. (1972). Measurement structures and psychological laws. *Science* 175, 1427–1435. doi: 10.1126/science.175.4029.1427

Krantz, D. H., Luce, R. D., Suppes, P., and Tversky, A. (1971). *Foundations of Measurement*, Vol. 1, *Additive and Polynomial Representations*. New York, NY: Academic Press.

Kyngdon, A. (2013). Descriptive theories of behaviour may allow for the scientific measurement of psychological attributes. *Theory Psychol.* 23, 227–250. doi: 10.1177/0959354312468221

Luce, R. D. (1997). Quantification and symmetry: commentary on Michell, quantitative science and the definition of measurement in psychology. *Br. J. Psychol.* 88, 395–398. doi: 10.1111/j.2044-8295.1997.tb02645.x

Luce, R. D., and Narens, L. (1987). Measurement scales on the continuum. *Science* 236, 1527–1532. doi: 10.1126/science.236.4808.1527

Mari, L. (2003). Epistemology of measurement. *Measurement* 34, 17–30. doi: 10.1016/S0263-2241(03)00016-2

Maul, A. (2013). On the ontology of psychological attributes. *Theory Psychol.* 23, 752–769. doi: 10.1177/0959354313506273

Maxwell, J. C. (1870). Sectional Proceedings: Section A.—Mathematical and Physical Science. *Nature* 2, 419–422. doi: 10.1038/002419a0

Maxwell, J. C. (1873). *A Treatise on Electricity and Magnetism*. Oxford: Clarendon Press.

McGregor, D. (1935). Scientific measurement and psychology. *Psychol. Rev.* 42, 246. doi: 10.1037/h0053856

Michell, J. (1997). Quantitative science and the definition of measurement in psychology. *Br. J. Psychol.* 88, 355–383. doi: 10.1111/j.2044-8295.1997.tb02641.x

Michell, J. (1999). *Measurement in Psychology: A Critical History of a Methodological Concept.* New York, NY: Cambridge University Press.

Michell, J. (2012). Alfred Binet and the concept of heterogeneous orders. *Front. Psychol.* 3:261. doi: 10.3389/fpsyg.2012.00261

Newman, E. B. (1974). "On The Origin of 'Scales of Measurement',' in *Sensation and Measurement: Papers in Honor of S. S. Stevens*, eds H. R. Moskowitz, B. Scharf, and J. C. Stevens (Dordrecht: D. Reidel Publishing), 137–145.

Nordin, G., Dybkaer, R., Forsum, U., Fuentes-Arderiu, X., Schadow, G., and Pontet, F. (2010). An outline for a vocabulary of nominal properties and examinations–basic and general concepts and associated terms. *Clin. Chem. Lab. Med.* 48, 1553–1566. doi: 10.1515/CCLM.2010.325

Quinn, T. (1997). Primary methods of measurement and primary standards. *Metrologia* 34, 61. doi: 10.1088/0026-1394/34/1/9

Quinn, T. (2011). *From Artefacts to Atoms: The BIPM and the Search for Ultimate Measurement Standards.* Oxford: Oxford University Press.

Quinn, T., and Kovalevsky, J. (2005). The development of modern metrology and its role today. *Philos. Trans. A Math. Phys. Eng. Sci.* 363, 2307–2327. doi: 10.1098/rsta.2005.1642

Sechrest, L., and Sidani, S. (1995). Quantitative and qualitative methods: is there an alternative? *Eval. Program Plann.* 18, 77–87. doi: 10.1016/0149-7189(94)00051-X

Sherry, D. (2011). Thermoscopes, thermometers, and the foundations of measurement. *Stud. Hist. Philos. Sci. A* 42, 509–524. doi: 10.1016/j.shpsa.2011.07.001

Silsbee, F. B. (1962). System of electrical units. *J. Res. Natl. Bur. Stand. C Eng. Instrum.* 66C, 137–178. doi: 10.6028/jres.066C.014

Stevens, S. S. (1939). Psychology and the science of science. *Psychol. Bull.* 36, 221. doi: 10.1037/h0056886

Stevens, S. S. (1946). On the theory of scales of measurement. *Science* 103, 677–680. doi: 10.1126/science.103.2684.677

Stevens, S. S. (1958). Measurement and man. *Science* 127, 383–389. doi: 10.1126/science.127.3295.383

Stevens, S. S. (1968). Measurement, statistics, and the schemapiric view—like the faces of Janus, science looks two ways—toward schematics and empirics. *Science* 161, 849–856.

Strogatz, S. H. (2001). Exploring complex networks. *Nature* 410, 268–276. doi: 10.1038/35065725

Tal, E. (2013). Old and new problems in philosophy of measurement. *Philos. Compass* 8, 1159–1173. doi: 10.1111/phc3.12089

Tal, E. (2014). Making time: a study in the epistemology of measurement. *Br. J. Philos. Sci.* doi: 10.1093/bjps/axu037 [Epub ahead of print].

Thurstone, L. (1931). The measurement of social attitudes. *Psychology* 26, 249–269. doi: 10.1037/h0070363

Trendler, G. (2009). Measurement theory, psychology and the revolution that cannot happen. *Theory Psychol.* 19, 579–599. doi: 10.1177/0959354309341926

Zwick, R. (1992). Statistical and psychometric issues in the measurement of educational achievement trends: examples from the National Assessment of Educational Progress. *J. Educ. Stat.* 17, 205–218. doi: 10.2307/1165170