Qualitative research meets the ghost of Pythagoras

Joel Michell

Theory Psychology 2011 21: 241
DOI: 10.1177/0959354310391351

Article

Theory & Psychology
21(2) 241–259

DOI: 10.1177/0959354310391351
tap.sagepub.com

Qualitative research meets the ghost of Pythagoras

Joel Michell
University of Sydney

Abstract

The issue of qualitative versus quantitative methods is rooted first and foremost in the character of the phenomena investigated and not in an investigator's methodological preferences. If the phenomenon under investigation is non-quantitative, then it cannot be studied successfully by attempting to use quantitative methods because trying to impose quantitative concepts upon qualitative phenomena misrepresents them. If the target articles provide any guide, these truths are ignored as much by psychologists wanting to mix quantitative with qualitative methods as by mainstream quantitative researchers. These articles display both the power of the modernist fantasy that measurement is always a discretionary choice of any investigator and the power of the persistent delusion that psychological attributes must be measurable. In psychology, as ever, the ghost of Pythagoras rules.

Keywords

measurement, order, qualitative, quantitative, scientific methods

The situations studied in any science possess both *form* and *content*. *Content* defines the subject matter: that is, whether we are studying, say, psychology or something else, like, say, geology. *Form*, on the other hand, relates to the logical structure of situations, and formal issues are invariant across subject matters. When scientists err logically, they fall into confusion and their conclusions are unsafe. The distinction between quantitative and qualitative attributes is formal. This means that controversies over quantitative versus qualitative methods need to address certain logical questions, clarification of which is a necessary condition for success. Therefore, my discussion attends only to formal issues. Beyond those, I do not go. Providing researchers avoid errors of logic, I offer nothing but encouragement: "Let a thousand flowers bloom."

The mainstream of psychology is committed to projecting its own fabricated image of successful quantitative science. While this implies that the mainstream has, at least, a

Corresponding author:
Associate Professor Joel Michell, School of Psychology, University of Sydney, Sydney, NSW 2006, Australia.
Email: joelm@psych.usyd.edu.au

superficial idea of what quantitative science is, it does not follow that the mainstream knows what quantitative science really is. Neither does it follow that the mainstream knows what measurement is. Indeed, the mainstream is deeply confused about measurement (Michell, 1997)[1] and, so, it is not surprising that psychologists who aspire to mix quantitative with qualitative methods are likewise confused.

While the psychological mainstream remains loyal to its post-Second World War methodological consensus (Michell, 2002), which made quantitative methods mandatory in the leading psychology journals, over recent decades a preference for qualitative methods has resurfaced at the fringes of the discipline (Marchel & Owens, 2007). It should not be forgotten that from its beginnings in the 19th century, qualitative methods occupied a place in psychology, being recommended by some of the discipline's founders (e.g., Freud, 1900/1953–1975; James, 1902; Wundt—see Haeberlin, 1916) and qualitative methods were still employed well into the 20th century (e.g., Adorno, Frenkel-Brunswik, Levinson, & Sanford, 1950). However, the post-Second World War period differed from earlier decades through psychology's participation in the politically constructed economy of the Big Science Era (Solovey, 2004) and its conscious adoption of a façade of rigor, adopted also in other human sciences (Schorske, 1997). This "rigorism" meant aping genuinely quantitative sciences (like physics and biology) and imposing arbitrary, but rigid, methodological imperatives. It also meant espousing what their opinion leaders thought of as avant-garde scientific philosophy (viz., logical empiricism).[2] During the 1930s, S.S. Stevens (e.g., 1936, 1939) had exposed psychologists to his interpretation of operationalism and logical empiricism, and after the Second World War psychological researchers paid lip service to these philosophies and, along with the required courses on quantitative methods, they entered the mainstream curriculum. Perhaps in reaction to this, some of those recently advocating qualitative methods adopted alternative philosophical identities. This deepened tensions between mainstream and qualitative researchers because for a time advertising alternative philosophical credentials was *de rigueur* amongst the latter (e.g., Denzin & Lincoln, 1994).

Of course, one can only productively discuss research methods philosophically, but if one is discussing scientific research, realism is the preferred philosophy because only realism makes sense of science as a cognitive enterprise. Realism, first, is the idea that the various kinds of systems investigated in research exist logically independently[3] of the investigators studying them and, second, it is the idea that it is possible for investigators to come to know something about these systems. Realism does not entail methodological imperatives and the fact that the mainstream has translated its methodological preferences into non-negotiable imperatives puts it at odds with realism. The proposition that qualitative methods might be useful in psychology is unobjectionable. It is advocacy of exclusively quantitative or exclusively qualitative methods that is anti-science. Certainly, some of the practices that Westerman, Yanchar, and Osatuke and Stiles describe are unobjectionable and their common commitment to methodological pluralism is laudable. On the other hand, some of their interpretations of their methods do not fit the phenomena investigated and their understanding of the concepts of quantity and measurement echo ingrained confusions.

According to the realist view (not to be confused with positivism;[4] see Hibberd, 2005; Michell, 2003), science is the attempt to discover the characteristic ways of working of natural systems. For example, physicists investigate physical systems, such as thermal or electrical phenomena; biologists investigate biological systems, such as reproduction or digestion; and psychologists investigate psychological systems, such as motivation or cognition, or, more narrowly, social interactions involving conversation (Westerman), work-related activity systems (Yanchar), and processes of psychological integration (Osatuke & Stiles). Natural systems overlap and interpenetrate and there are no free-standing systems. Since we cannot investigate everything at once, investigating any system involves ignoring others and, where possible, attempting to minimize the influence of others upon the system studied.[5] Investigating any natural system requires methods, and these might be loosely classified as (a) methods of observation, (b) methods of inference, and (c) methods of criticism.

Methods of observation

Observation is a form of interaction between the scientist and the system investigated, and from the realist perspective there is no question of the scientist being removed from the system investigated, of being a mere "spectator." Westerman's (2011, p. 156) complaint that the (unidentified) "philosophical tradition" presumes this kind of "uninvolved subject" does not apply to realism. Furthermore, we can only find out about natural systems by investigating specific instances of the system occurring in particular spatially and temporally located things (e.g., persons). Typically, observation results in claims about the properties these things have (e.g., person *a* is a teacher) or relations in which they stand to other things (e.g., *a* teaches arithmetic to *b*).

Because observation in science is of particular things, number is always already present in the situations observed. A particular thing is always one thing (of some kind), and if we observe more than one thing, then we must observe two or three or some other number.[6] A number is a relation between a magnitude (say, an aggregate) and a unit (say, a kind of thing)[7] and, so, whenever any thing of some kind is observed, also present in that situation is the number one because there is an aggregate (composed just of the thing of the relevant kind) and a unit (the kind of thing it is) and, so, there is the relation between them of that aggregate containing just one thing of that kind. Note, however, that the fact that number is always present does not mean that it must be consciously attended to. It only means that if numerical facts are relevant to the investigator's concerns, as they were, for example, to Engeström's (as cited in Yanchar, 2011, p. 192), they can always be counted or estimated and no investigator need ever apologize for doing so, as Westerman (2011, p. 164) feels he must when describing the 1993 study by Horowitz et al. on dyselaboration in discourse analysis.[8] Believing that numerical facts should never be attended to is as pathological as believing that only numerical facts should be attended to.

A further way in which number and quantity are always present follows from the fact that things are always observed in some region of space and time. Thus, we always have the option of observing the spatial and temporal relations that these things stand in. For example, were it relevant to the point of our investigations, distances from other things or time elapsed since other events could be observed. Observing such relations would bring us into contact with continuous quantities and, through measurement, ratios of continuous quantities (i.e., real numbers) in the situations investigated. That is, distance, for example, is a continuous quantity (i.e., it has the kind of structure described by

Hölder, 1901; see also Michell & Ernst, 1996) and the relative magnitude of one distance to another (i.e., the ratio between them) is a positive real number. These quantities and numbers are just as much part of the situations observed as any other features, and, again, no apology is needed for noting them if they are deemed relevant to the investigator's concerns, or for ignoring them otherwise. Indeed, to the extent that the situations observed involve physical objects or processes, any physical, quantitative attribute may be measured.

Even though we might ignore them in some investigations, quantity and number are ubiquitous features of every real situation. This is the germ of truth that gives the Pythagorean vision of reality *prima facie* plausibility. The Pythagorean vision derives from the 6th century BC Greek philosopher, Pythagoras, who, it is said taught that all things are made of numbers.[9] In modern dress, this is the view that all attributes are quantitative. It is no arcane view confined only to antiquity. It is one of the most enduring ideas in human history. It captivated the minds of philosophers (such as Plato, 1971), theologians (such as Augustine, as cited in McEvoy, 1987), and, of course, scientists (such as Galilei, 1623/1960) and it dominates the modern world, in part because of the false equation of numerical data with objectivity (Porter, 1995). It has had a constant presence throughout the trajectory of Western culture from ancient times to the present (Joost-Gaugier, 2006; Kahn, 2001; Riedweg, 2002). Most psychologists are exposed to it in their studies, although rarely by name, but often through slogans like Thorndike's credo, "whatever exists at all exists in some amount" (1918, p. 16), and elaborations of the same theme (e.g., McCall, 1922). Although it is a major factor driving the quantitative imperative in psychology (Michell, 2003), Pythagoreanism is a mistaken view. While number and quantity are present in every situation, it is not true that every attribute encountered in such situations is quantitative. It is around this logical issue that tensions over quantitative and qualitative methods really turn.

As noted earlier, every observation made in science concerns either the properties that things have or the relations in which they stand. Certain classes of properties and of relations constitute attributes, and attributes always possess structure. Some properties and relations constitute categorical attributes. For example, for any person, *X*, either *X* is pregnant or not; or for any specific pair of people, *Y* and *Z*, either *Y* is the father of *Z* or not. In the case of pregnancy, the attribute consists of two mutually exclusive properties: *being pregnant*; and *not being pregnant*. In the case of fatherhood, the attribute consists of two mutually exclusive relations: *being the father of Z*; and *not being the father of Z*. There is no issue of degrees of pregnancy or degrees of fatherhood. The structure of a categorical attribute is determined by an equivalence relation (i.e., a reflexive, symmetric, and transitive relation), and such an attribute is a simple classificatory system (Suppes & Zinnes, 1963). Numerically coded, it is more familiar under the now conventional designation of "nominal scale" (Stevens, 1951). Despite Stevens, however, numerical coding is not measurement and a nominal scale is not a measurement scale.

However, some attributes admit of degrees. For example, some minerals are hard, some are soft, and hardness comes in degrees. Or, to take another example, the elderly differ in the degree to which they are functionally independent in performing physical tasks of ordinary life: one person might be highly independent (e.g., able to live alone without help from others), while another might be moderately independent (e.g., require help with some day-to-day tasks), and another might only possess a very low level of independence (e.g., depend on help in most things). That is, functional independence comes in degrees. When possession of a property or participation in a relation is a matter of degree, the different degrees of the attribute are generally ordered from less to greater.[10] For example, the hardness of diamond is greater than that of lead or one person's level of functional independence may exceed another's. Ordinal attributes, coded numerically, are familiar to us as "ordinal scales," so called (Stevens, 1951). Again, Stevens erred in thinking of ordinal scaling as measurement (Michell, 1997). "Ordinal scaling" is numerical coding, not measurement.

Furthermore, some attributes admitting degrees are not only ordered, they are quantitative. For example, temperature is a quantitative property and the force one thing exerts upon another is a quantitative relation. The different magnitudes[11] of a quantity stand in additive relations to one another, and it is this additive structure that entails the ratios between them that makes measurement of magnitudes possible.

Quantitative structure is not an "essence" (Yanchar, 2011, p. 185) or "abstract entity" (Westerman, 2011, p. 158). It is a mundane, discoverable fact about attributes. In science, it is only safe to conclude that an attribute is quantitative on the basis of observational evidence for quantitative structure[12] and never on the basis of wishful thinking. There is nothing unworldly about mathematical structure. Both categorical and ordinal attributes possess structure and, given that mathematics is the science of structure (e.g., Resnik, 1997), categorical and ordinal structures are mathematical.[13] Calling such structures "abstract" is misleading if it is meant to imply that they belong to a realm of essences or abstract entities beyond space and time. Strictly speaking, it is not mathematical structures that are *abstract*; rather, it is we who *abstract* them from the spatiotemporal locations within which they reside by focusing our attention just upon the relevant properties and relationships constituting them and ignoring other features of the situations in which they occur. That is, *being abstract* is not a quality of structures; rather, *abstracting* is a mental act engaged in when considering them. By this act, we note features common across a variety of different situations, just as, for example, we note that different things are the same color. So, quantitative structure is present in each and every situation involving quantitative attributes (such as length, mass, temperature, etc.) despite other differences in those situations.

If an attribute is not quantitative, it cannot be measured. Therefore, if anyone aspires to measure some particular attribute, the first step, logically speaking, is to investigate whether it possesses quantitative structure. It is in relation to this step that mainstream psychology has been egregiously negligent and the stance it has adopted is scientifically anomalous. *There is no evidence that the attributes that psychometricians aspire to measure (such as abilities, attitudes and personality traits) are quantitative.* Mainstream psychologists presumed that they are and that psychometric tests are suitable to measure them. All the evidence is that these attributes are merely ordinal, and while psychologists are free to speculate about whether they *might be* quantitative, the fact is that, after a century, such speculations remain empty. If these speculations are mistaken for fact, it is only because people want to believe them (for non-scientific reasons), not because they can be backed up in the normal scientific way.

Instead of candidly admitting this, psychologists deflected attention from the issue by using Pythagorean slogans, such as Thorndike's, quoted above. If all attributes are quantitative, then psychological attributes, even apparently qualitative ones, must also be quantitative and, so, it appears, there is no need to search for evidence of quantitative structure. In this spirit, measuring psychological attributes would seem to be merely a matter of constructing tests composed of items satisfying conventional psychometric criteria. However, in science, the truth of substantive hypotheses cannot be deduced from metaphysical slogans, no matter how beguiling.

A second way of deflecting attention from this issue was through muddying the meaning of measurement. In science, the concept of measurement has a definite meaning: measurement is estimation of the ratio between a magnitude of a quantitative attribute and a unit of the same attribute. However, from the 1950s, as part of their post-Second World War methodological consensus, mainstream psychologists endorsed Stevens' vacuous definition of measurement: namely, that measurement is the assignment of numerals to objects or events according to rules (Michell, 1997, 2007a). Stevens (1946, 1951) derived this definition from his own blend of operationalism and logical empiricism (Michell, 1999). It was thought to provide a philosophical justification for regarding psychometric practices as measurement. It also reinforced the practice of referring to test scores as measurements. Generally, test scores are frequencies (viz., the number of items correctly answered). Transforming them into factor scores (via factor analysis) or ability scores (via application of item response models) does not transform them into measures of anything when there is no evidence that the hypothesised factors or abilities are quantitative.

A quite different line of argument had initially been used to support this practice. During the 19th century, attempting to justify psychophysical measurement, psychologists resorted to "the constantly recurring argument that *S* must be measurable because it is a magnitude, because we can speak, intelligently and intelligibly, of 'more' and 'less' of it" (Titchener, 1905, p. lxiii [where *S* means *sensation*]). That is, they held that if an attribute is ordered (i.e., displays degrees), then it must be quantitative and, thus, measurable. This inference from order to quantity played such a decisive role in psychometrics that I call it "the psychometricians' fallacy" (Michell, 2006, 2009). That the authors of the target articles are also its victims illustrates how deeply embedded it is in the culture of modern psychology.

Despite attempts to prove that order entails quantity (Michell, 2007b), it is easily shown that while quantity entails order, there is no implication in the opposite direction. Johannes von Kries (1882) presented one good reason (see Michell, 2009, for another). He asked whether "it makes sense (and what sense it does make) to say that a change in sensation from R_1 to R_2 may be equal to a change in sensation from R_K to R_L" and answered that it "makes absolutely no sense" because "the one increment is something quite different from the other" (Niall, 1995, p. 291).[14] What he meant was that for any given attribute, differences between sensation intensities are qualitatively different to one another: that is, qualitatively heterogeneous. Whether he was correct as far as sensation differences go is here irrelevant. In framing his objection, he saw that if differences between degrees of any ordered attribute are heterogeneous, then that attribute is not quantitative.[15]

Von Kries was not saying that the sensations themselves are qualitatively heterogeneous, only that the differences between them are. What he recognized was that the differences between such degrees need not be mutually homogeneous. If these differences are not mutually homogeneous (i.e., if they differ in kind), then the differences will not be intrinsically greater than, equal to, or less than one another and, so, the differences cannot stand in quantitative relations to one another. Since the existence of quantitative relations between differences between degrees is a necessary condition for an attribute to be quantitative, it follows that in such circumstances the attribute involved cannot be quantitative. But is it possible to have an ordered attribute in which differences between degrees are qualitatively heterogeneous?

Consider the Assimilation of Problematic Experiences Scale described by Osatuke and Stiles (2011). It consists of an ordered series of levels of "mutual assimilation" of conflicting "voices" within a given person as displayed in discourse in psychotherapy sessions. As described in their Table 1 (p. 208), each level of assimilation is composed of a complex array of cognitive and affective features, and it is hypothesized that the levels are ascending stages of assimilation, each stage being a degree of assimilation greater than those before it. If so, then the stages are homogeneous, in the sense that they are all degrees of assimilation.

However, if attention is focused upon the *differences* between consecutive degrees or stages, it is clear that they are not all differences of one kind. To take just one feature of the array of characteristics comprising the different degrees and confining attention to stages 0, 1, and 2 only, the authors say that at stage 0, the client is "unaware of the problem" (i.e., unaware that there are unassimilated "voices" in the way experiences are expressed), while at stage 1 the client "prefers not to think about" the unassimilated expressions, and at stage 2, the client is "aware of the problem but cannot formulate it … reflect on it." At least one difference between stages 0 and 1 is that between *being unaware of something* and *being aware of it (but avoiding it)*; and one difference between stages 1 and 2 is that between *being aware of something and avoiding it* and *being aware of something but being unable to formulate it (reflect upon it)*. In the first case, the difference is at least that of minimal *awareness*, while in the second case it is at least that of minimal *acceptance* of what one is aware of. Minimal *awareness* and minimal *acceptance* are qualitatively different mental acts. As stages of assimilation ascend, we are not dealing with some homogeneous stuff in quantitative amounts that are successively added to each stage. Hence, there is no obvious sense in which one of these differences is intrinsically greater than, equal to, or less than the other. As a person progresses from stage 0 to stage 1 to stage 2 (and so on), the level of assimilation increases by an increment, but the increments are not quantitatively related because they are qualitatively different. The authors have confused *qualitative* increase or difference with *quantitative* increase or difference. In a similar way, I would suggest that the ratings of coordination processes that Westerman (2011) refers to are not "measures of coordination" (p. 166), but at best simply relate to an attribute that is experienced only as ordinal and not known to be quantitative.

To confuse differences between degrees of a quality with differences (or, as they might be called, "distances") between magnitudes of a quantity is to commit an error identified long ago by David Hume (1888), who noted the habit by which "any great *difference* in the degrees of any quality is call'd a *distance* by a common metaphor. . . . The ideas of distance and difference are, therefore, connected together. Connected ideas are readily taken for each other" (p. 393). In this case, when *taken* one for the other, they are, really, *mistaken*.

This observation is not a criticism of the Assimilation of Problematic Experiences Scale. It is a criticism of one interpretation of it and it illustrates the point that with some ordinal attributes, differences between degrees are qualitatively heterogeneous. This means that some ordinal attributes are necessarily non-quantitative. The authors are not alone. If any psychometric scale is considered, typically, the differences between the degrees of the relevant attribute, as these are manifest in the differences between the test items concerned, reveal exactly this kind of heterogeneity. *Prima facie*, it appears, all psychological attributes are ordinal, non-quantitative attributes.

Therefore, Osatuke and Stiles (2011) are mistaken in thinking that their scale "quantifies an underlying theoretical dimension of client experience" (p. 209). The character of their error needs to be identified clearly, for it is one typical of mainstream psychology. In itself, there is no error involved in assigning numerals to the stages of assimilation that the scale identifies, thereby coding increasing degrees of assimilation by numerals denoting numbers of increasing magnitude. Yet while Osatuke and Stiles correctly note that when numerically coded their scale is "no less quantitative than other ordinal measurement scales" (p. 209), they fail to see that this is only because their scale, like all other ordinal scales, is not quantitative at all. Their error is to think that an ordinal attribute is quantified by making numerical assignments. If an attribute is not known to be quantitative, no amount of numerical coding will make it so.

Failing to see this, they proceed to report the results of an analysis of covariance applied to the numerical codes. Analysis of covariance is, as its name implies, concerned with relationships between variances, and variances are numerical measures of *differences* between measurements. That is, Osatuke and Stiles, in making statistical inferences treat the differences between the degrees of their ordinal attribute *as if* they were quantitative differences. Since these differences are not known to be quantitative, they do not present us with a valid argument for the conclusions they infer from their observations.

Methods of inference

This brings us to a discussion of the second category of methods of investigation: namely, methods of inference. Inferences from observations take the observations as premises and, possibly in conjunction with other premises (deemed true), proceed to deduce conclusions. The worth of the conclusions depends not only upon the accuracy of the observations themselves, but also upon the validity of the methods of inference used. Unfortunately, the teaching of methods within mainstream psychology has become so dominated by instrumental values that the underlying logic of inference is now almost universally overlooked. Yanchar (2011, p. 181) alludes to the problems identified in relation to conventional methods of hypothesis testing. The problem is especially serious with statistical methods, which have been sold as packets of inference tickets to be applied to numerical data, regardless of their character, as if "the numbers don't remember where they came from" (Lord, 1953, p. 751). There are many logical issues here, but the one that most concerns my critique of Osatuke and Stiles goes by the name of the "problem of permissible statistics," famously discussed by Stevens (1946, 1951).[16]

Despite the appeal of Lord's slogan,[17] numbers do not need to "remember" where they came from because numbers themselves never leave the situations in which they occur.

The slogan confuses numbers with numerals. Although numerical data are typically presented as tables of free-standing numerals, each datum is not a numeral or even a number, it is a proposition (Rozeboom, 1961), such as, *the weight of person P in kilograms is 90* or *the score of person P on test Q is 25*. Tabulated numerals are shorthand for a set of statements recording each number's provenance. Furthermore, deductions from data are inferences from the propositions involved, not inferences from mere numbers or numerals.

Stevens recognized that conclusions regarding means and variances do not validly follow from ordinal-scale data, but his message was mixed. With one hand he forbade calculation of means and variances with ordinal data, but, with a nod and wink, with the other, he waived caution away because, he claimed, using those statistics can lead to "fruitful results" (Stevens, 1946, p. 679). However, the logic of the situation is clear: conclusions about means and variances would only ever validly follow from ordinal-scale data if it were the case that these ordinal-scale values actually denoted quantitative measures (i.e., if they were not ordinal-scale values). Numerically coding observations, as Osatuke and Stiles have done, may make them look quantitative, but to see them as quantitative is to fall victim to an illusion. It would be sad indeed if qualitative researchers, in their attempts to also use quantitative methods, followed the same, illusory path as the befuddled mainstream.[18]

While the analytical procedures used by Osatuke and Stiles conform to established mainstream practice, their conclusions do not follow validly from their observations. It is ironic that their analysis of covariance was used to derive merely ordinal conclusions (e.g., that clients with relatively well-assimilated problems did better in one form of therapy than another, while clients with poorly assimilated problems did equally well in both). However, to test whether even ordinal conclusions follow validly from data, one would need to use a logically valid chain of inference (i.e., one using methods of analysis suitable to ordinal data (e.g., Cliff, 1996).

In the current research culture, where errors of the sort identified above are typical and excused by appealing to their "fruitfulness," in the spirit of Stevens, my critique may be ignored with social impunity. Nonetheless, fruitfulness is a relative concept, and while it is the case that the invalid inferential practices of the mainstream are often fruitful relative to the aim of promoting the illusory façade of a quantitative science and, possibly, advancing researchers' careers, such practices can never advance genuinely scientific aims. Evidence is fundamental to scientific conclusions, and where there are insufficient grounds in data for the conclusions derived, objectively, the conclusions are unsupported.

Of course, any researcher is free to speculate about what might be concluded were the attributes involved quantitative (rather than ordinal) and were the scale values used actually measurements of those attributes (rather than ordinal indices) and the sorts of analyses referred to above then done. However, in that case, the conclusions would follow from a combination of observation and speculation. Presenting such conclusions as valid inferences from the observations alone would reflect, at best, profound ignorance. Those who support scientific research financially and those whose lives are affected by the conclusions scientists arrive at are entitled to expect scientists to alert them when conclusions stand, even partly, on speculation, and they have a right to be assured that all other conclusions are supported exclusively by evidence.

While, in terms of social consequences, the mistakes made by Osatuke and Stiles are minor, they are, nonetheless indicative of a discipline-wide attitude. It will be objected, perhaps, that the domain of real psychological research is inevitably messy and no country for purists. However, the fact that psychological research is messy simply means that errors will sometimes occur. Using that as a reason to deliberately make further, "fruitful" errors is like a robber justifying theft on the grounds that crime happens.

The same kind of error occurs in Yanchar's paper (2011, p. 190) when he describes how Engström first numerically coded an ordering of cleaners' performances and then used means[19] to draw essentially ordinal conclusions from the data. Because the sum of the differences between each number and the arithmetic mean is zero in any distribution, the value of the mean is sensitive to the magnitude of these differences. When an ordering upon anything is numerically coded, the magnitude of the differences between each number and the mean is an arbitrary feature, signifying nothing; so, the value of the mean is likewise to some extent an artifact; and the validity of even ordinal inferences from means may be unsafe. For this reason, ordinal conclusions do not generally follow validly from means in such circumstance, but they do follow validly from medians.

The distinction between merely ordinal and quantitative attributes marks an important boundary between qualitative and quantitative science. One of the features that gives quantitative sciences, such as physics, sweeping explanatory power is the fact that quantitative attributes are yoked together in a network of interrelationships uniform across the range of the attribute, such as the relationship between density, mass, and volume (i.e., density is the ratio of mass to volume). A necessary condition for the existence of such relationships is the homogeneity of the differences between magnitudes within each attribute. However, merely ordinal attributes lack this feature whenever the differences between degrees are mutually heterogeneous, as seems to be the case with the Assimilation of Problematic Experiences Scale (APES) described by Osatuke and Stiles (2011). If so, then their attempt (p. 212) to use statistical methods to detect an S-shaped relationship between the attribute assessed by this scale and others cannot succeed—which is not to say that merely ordinal relationships may not be detected.

The heterogeneity of differences between degrees of a merely ordinal attribute significantly increases the complexity of the situation. Because these differences are qualitatively different to one another, it is always possible that they enter into quite different causal relationships with other attributes at their different levels. For example, factors that might produce an increase at one level of the APES might not be important at other levels and vice versa. Such a possibility works against the possibility of finding uniform relationships across the entire range of ordinal attributes. Dealing with this possibility, of course, is one of the challenges of qualitative science, and it is a possibility which it might have been expected qualitative researchers would have been especially attuned to.

Methods of criticism

The third class of methods mentioned above is that of criticism. In the present context, the criticism is of the logic of methodological practices and their philosophical underpinnings. The above discussion already provides illustrations, but I want to say something about the topic of philosophical underpinnings because methods of criticism always presume a logical basis.

At the beginning I said that when discussing scientific methods, the preferred philosophical basis is realism. I will attempt to explain why and what I mean by realism. Opinion leaders amongst qualitative researchers often confuse realism with positivism. For example, Guba and Lincoln (1994) equate the "positivist paradigm" with "realism (commonly called 'naïve realism')" (p. 109). Naïve realism is our natural philosophical stance, the one that we adopt, largely unreflectingly, as children, and it is the one most people retain throughout their lives, whatever they might profess to the contrary. It is the view that "our ordinary perception of physical objects is direct, unmediated by awareness of subjective entities, and that, in normal perceptual conditions, these objects have the properties they appear to have" (Dretske, 1995, p. 602). It is worth stressing that "although this theory bears the name 'naïve,' and is often said to be the view of the person on the street, it need not deny or conflict with scientific accounts of perception" (p. 602), and it is worth stressing that it does not rule out the possibility of perceptual (or, more generally, cognitive) error.

One thing it does entail, however, is the possibility of true belief (i.e., the possibility of knowing), where by "true" is meant things being as we believe them to be (Mackie, 1973). Such a view means that scientific knowledge is possible and it enables us to construe scientific methods as procedures for observation, analysis, and criticism undertaken in the belief that they increase the likelihood of attaining knowledge of how the systems under investigation work. This is why realism is the preferred philosophical orientation for discussions of scientific method.

Of course, I am conscious of the fact that some philosophers (and qualitative researchers) have reservations about realism, and while I respect all serious attempts to come to grips with epistemological and metaphysical issues (because I believe that none of these matters will ever be neatly sewn up to everyone's satisfaction), I cannot begin to consider all such reservations. My general stance is that no matter how beguiling alternative philosophical visions appear, in the end they rest upon inevitably fallible philosophical arguments, which means that they can never firmly establish an alternative to realism, delivering at best the following disjunction: either realism is false or the philosophical alternative is false. A disjunction merely leaves us with a choice and the attractions of realism; particularly its premise that knowledge is possible cannot be coherently rejected.

None of this is to underestimate the difficulties intrinsic to attaining scientific knowledge in practice and especially the difficulties involved in psychological research. Leaving aside the significant problems engendered by what Bacon (1620/1960) called "the dullness, incompetency and deceptions of the senses" (p. 50) and the fact even the best of our methods may only deliver what Locke (1690/1959) referred to as the "twilight of probability" (p. 360), psychological research presents its own special difficulties, not the least of which is the fact that the investigator always brings social, psychological, and philosophical baggage to any investigations, which may influence the observations reported, the speculations proposed, and the conclusions arrived at; and that in psychology, investigations are always into mental phenomena, a category of experience still awaiting a satisfactory characterization. The philosophical preambles given in the target articles point to some of the implications of these special difficulties, the most important being the fact that "psychological phenomena are irreducibly meaningful" (Westerman, 2011, p. 171). Without disagreeing with that, I agree with Petocz (1999, 2001) that the phenomenon of meaning poses no insurmountable problems for a realist understanding of psychology.

However, one apparent difficulty that Yanchar (2011) claims to identify is that resulting from "the translation of meaningful human phenomena and experience into one or another set of theoretical abstractions that often make little meaningful connection with practical involvement in ordinary human life" (p. 181), especially as this is thought to apply to the kind of concepts identified within quantitative science, such as " 'averages,' 'variances,' and so on" (p. 181). While the distance between the quantitative concepts of mainstream psychology and direct experience is a significant problem, the cause resides not in abstraction *per se*, but in the failure of quantitative concepts to fit psychological attributes as experienced by us. As I said earlier, quantitative structure is not an abstraction in any pejorative sense. Where quantity occurs, quantitative structure is present and, if our methods are suitably tuned, detectable by observation. Furthermore, if quantity is present, then the reductions that occur in, say, statistical data analysis involving averages, variances, and so on, identify existing quantitative relationships—relationships that have a real presence in the relevant situations. This is why engineers are able to successfully build bridges and plan successful space missions. They use quantitative abstractions and reductions successfully because in the situations investigated they deal with real quantitative structures.

Some people have a horror of abstraction, holding that it falsifies experience. This position was developed most rigorously by F.H. Bradley (1893), who thought that putting experience into words falsifies it because it always involves the use of concepts and concepts are "an organ of misunderstanding" (James, 1910, p. 29). Such a position is literally unspeakable because uttering it requires concepts and, so, on its own premise, it must always be stated falsely. The error here is that of thinking of concepts as organs of any kind, whether of misunderstanding (e.g., Bradley) or of understanding (e.g., Kant). Concepts are not tools, somehow already present in our minds and able to be imposed upon a non-conceptual reality, for good or ill. Rather, concepts are features of reality. For example, when I see that this rose is red, whatever it is *to be a rose* and *to be red* (i.e., the concepts involved) are real qualities, present in indefinitely many particular plants and indefinitely many physical objects. Otherwise I could never judge veridically that this rose is red—something I can surely do.

Any act of cognition involves judging something veridically and brings the cognizing person into contact with instantiated concepts (viz., those that are features of the situation cognized). Furthermore, all concepts are abstract, in the sense that we can attend just to them, ignoring other features of any situation in which they occur. Abstraction does not distance us from our experience of situations; rather it brings us closer to certain features because in abstracting we focus attention just on them. Abstraction is like increasing the magnification of a microscope and bringing certain features into clearer view. This connects with what was said earlier about quantitative structure. When such structure is present in situations that we experience, we know it more thoroughly through attending just to its structural features and ignoring features extraneous to quantitative structure.

It is not abstraction that distances us from our experience of situations. Instead, what distances us is falsely attributing a quality or structure to a situation when it is not experienced therein. When the concepts we judge as present in a situation are there, we are in veridical contact with it, but when these concepts are not present, our judgment misses the mark and our contact with reality is loosened. The feeling of alienation that Yanchar reports derives, not from abstraction, but from the fact that quantitative structure is not experienced in the attributes concerned. It is thus ironic that Yanchar (2011, p. 190) praises Engeström for "quantifying" the performances of the cleaners he investigated, because the performance categories reported ("home cleaning," "rationalized cleaning," and "consciously mastered orientation"), while conceivably experienced as merely ordered, are certainly not experienced as quantitative in structure. Treating these categories as quantitative loosens the investigator's contact with the situations studied.

It is doubly ironic that each target article in its own way is implicated in this kind of mistake. In each, an attribute experienced only as ordinal is treated as if quantitative. Treating them thus, when they are only experienced as qualitative, involves overlooking the very feature to which qualitative researchers are trained to be sensitive. In their haste to marry quantitative methods with qualitative, qualitative researchers commit some of the same errors for which the quantitative mainstream is notorious.

The underlying cause here is the failure of mainstream psychology to sustain a culture of uncompromising critical inquiry. In a talk in 1974, the physicist Richard Feynman (1985) drew attention to the fact that what psychology lacked was "a kind of scientific integrity, a principle of scientific thought that corresponds to a kind of utter honesty—a kind of leaning over backwards" (p. 341) to accommodate criticisms. Feynman was not referring to personal qualities of individual psychologists. Psychologists, in general, have as much personal integrity as other scientists, being drawn from the same population. Feynman was referring to a cultural phenomenon: the failure of mainstream psychology as a social movement to sustain institutional mechanisms of rigorous criticism. Since the post-Second World War methodological consensus, when psychologists dreamt of "riding natural scientists' coattails onto the endless frontier" (Solovey, 2004, p. 393) of quantitative science, the mainstream, as a social movement, turned a deaf ear to any criticism that might have been seen as deflecting psychology from its imitation of quantitative science, and over the past half-century centers of rigorous methodological criticism, where they existed, withered. Such a criticism-free culture is an inevitable consequence of a discipline in which attempts at application and window-dressing outstrip the relevant knowledge base. More than in other sciences, psychology is always conducted under the dark shadow of wishful thinking, and, as is evident from the papers discussed here, the same shadows threaten the vision of those outside the dominant paradigm. Wherever two or three psychologists are gathered together, there is the ghost of Pythagoras amidst the shadows of their dreams.

Conclusion

The distinction between quantitative and qualitative methods resides, ultimately, in the structure of the attributes investigated and, firstly, in the kind of structure experienced in these attributes. Since, at present, *all* of the attributes psychometricians aspire to measure are experienced only as non-quantitative, ordinal attributes (i.e., as qualitative), were investigators to reflect their experience faithfully, they would acknowledge that these attributes at present require only qualitative methods. This is not to say that quantitative methods might not be legitimately used investigating these attributes. They might, but only in the counting and analysis of associated frequencies; in the measurement and analysis of physical attributes, such as distance and time, as related to them; and possibly, in the future, should evidence eventually support the hypothesis that they are quantitative.

In adopting a different approach, the authors of the target articles have succumbed to the positivist illusion that the application of quantitative methods in science depends not upon the existence of quantitative structure in the attributes investigated, but upon the quantitative aspirations of the investigator. It may seem ironic that researchers espousing post-positivist philosophies should agree with positivists, but it is not surprising. Much of the novelty of post-positivist philosophy derives from extending positivistic conventionalism beyond the quantitative domain to the remainder of science (Hibberd, 2005).

As long as it is held that there is a hiatus between *what is said* and *what is*, it will seem that all discourse, including scientific discourse, is merely a product of conventions. On the other hand, if it is accepted that it is possible for discourse to engage real situations; that it is possible for what is said to be literally true; that it is possible for things *to be* as *stated*; then it will be seen that quantitative discourse is only suited to quantitative aspects of reality. Where quantity and number already exist in the situations investigated, quantitative methods are appropriate; where the attributes investigated are non-quantitative, quantitative methods are inappropriate for understanding those attributes. Qualitative methods are currently essential for investigating attributes experienced only as ordinal, and pretending to measure them represents a victory for Pythagorean whimsy over scrupulous scientific endeavor.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Notes

1. I have argued that the so-called "quantitative methods" developed in psychology are little more than pretence (e.g., Michell, 1990, 1997, 1999, 2000, 2008). This led me to consider the place of qualitative methods (e.g., Michell, 2003, 2004). I welcome this opportunity to reconsider the logic of the quantitative/qualitative distinction and to illustrate my discussion by commenting on the target articles.
2. The term "logical empiricism" was adopted by logical positivists associated with the Vienna Circle upon arrival in the United States as a preferred label for their philosophical doctrines (Reisch, 2005).
3. Logical independence does not rule out causal dependence or interaction. What it does rule out, however, is that the existence of the system studied is somehow constituted by being studied or known. As the point is made by the realist philosopher John Anderson (1962): when, for example, "I know this paper, 'I know' in no way constitutes this paper, nor does 'know this paper' in any way constitute me, nor does 'know' in any way constitute either me or this paper" (p. 27).
4. There is *positivism*, the *image of positivism held by post-positivists* (see, e.g., Richardson, 2007) and there is *realism*. Confusing these does not advance discussion of methodological issues. Positivism, as developed by the logical positivists is anti-realist, not least in its

commitment to *conventionalism* in the philosophy of mathematics. Thoroughgoing realism treats the categories of number and quantity as features of the natural world (Anderson, 1962; Armstrong, 1997).

5. Furthermore, scientific investigators, themselves, contain and are contained within an indefinite number of other systems. For example, investigators are psychological systems, such as motivational and cognitive systems, and they belong to various social systems involving economic and cultural relations. All of these, to a greater or lesser extent, may influence any scientific investigations made. Science is not above the flux.
6. The statement of Osatuke and Stiles (2011) that "numbers are words" (p. 200) is not just false; it amounts to a debilitating confusion. There are, of course, words *for* numbers, just as there are words for other things, such as colors, but numbers are no more words than colors are. The numbers are what the words "one," "two," "three," and so on, stand for, and, it may be noted, what these words stand for are not the signs, "1," "2," "3," and so on, for these signs (which also denote numbers) are not numbers; they are *numerals*. This confusion of *number words*, *numerals*, and *numbers* is ubiquitous. It derives from the logical empiricist view that numbers are elements within a formal, symbolic system, rather than features present in every situation.
7. This conception of number covers Euclid's definition of (natural) number as "a multitude composed of units" (Elements, Bk. VII, Definition 2, as cited in Heath, 1908, p. 277) and Frege's understanding of (real) number as "ratio of magnitudes" (1903, p. 155). Forrest and Armstrong (1987) show that this realist conception covers natural, integral, rational, and real numbers (see also Michell, 1993, 1994). In the 20th century it was overturned in the popular mind by logical positivism. Nonetheless, the realist view is still presumed in traditionally quantitative sciences, such as physics.
8. Westerman (2011) makes a common mistake when claiming that "dyselaboration processes were *measured* [emphasis added] using quantitative techniques" (p. 163), when all he means is *that frequencies were counted*. The term "measurement" has many colloquial meanings but, in science, it means "assessment of continuous quantities." The researchers involved did not *measure* anything. They counted frequencies of events. Counting is a quantitative method, but not measurement (although it may be involved in processes of measurement), and its use in psychology does not imply that psychological attributes are quantitative.
9. Our knowledge of Pythagoras's teachings depends upon second-hand reports. For example, "Again, the Pythagoreans, because they saw many attributes of numbers belonging to sensible bodies, supposed real things to be numbers— not separable numbers, however, but numbers of which real things consist" (Aristotle, *Metaphysics*, Bk. XIV, Ch. 3, 1090[a], 20–23, as cited in McKeon, 1941, p. 918).
10. The structure involved may not necessarily be a simple order, but only a partial order (Michell, 1990).
11. When the degrees of an ordered attribute possess additive structure, they are commonly referred to as *magnitudes* of that attribute.
12. Some kinds of evidence for quantitative structure in attributes are classified under the headings of extensive and conjoint measurement (Michell, 1990, 2005), and doubtless there are other kinds as well.
13. The class of mathematical structures is wider than the class of numerical structures (a numerical structure being a mathematical structure involving numbers).
14. There is more to von Kries's critique than this. Much of his argument, however, misses the mark because it contains misunderstandings about measurement of intensive magnitudes in physics. These were corrected by von Helmholtz (1887), by Campbell (1920), and, most importantly, by development of the theory of conjoint measurement (see Krantz, Luce, Suppes, & Tversky, 1971).

15. Likewise, John Maynard Keynes (1921), influenced by von Kries, argued from the same premise that probabilities are, in general, non-quantitative. Apparently quite independently, the English philosopher R.G. Collingwood (1933) made the same general point with respect to his "scales of forms."
16. See Michell (1986) for an analysis. Stevens was not the first to recognize it, even in psychology. For example, see Boring (1920) and Johnson (1936).
17. To be fair to Lord, he recognized the dangers of treating ordinal-scale data as if quantitative (Lord, 1954), and in his earlier paper a fictional character uttered this slogan.
18. This is not to suggest that ratings, rankings, or numerically coded ordinal observations are not legitimate data (as Yanchar, 2011, p. 185 seems to imply that I might be suggesting). It is to point out that they are not *measures* of anything. Inferences premised upon the false proposition that these kinds of numerical data are measures are, therefore, not safe.
19. Yanchar also reports that Engström calculated correlations using these data, but it is not clear whether Pearson's product-moment correlation coefficient (which presumes quantitative data) or Kendall's tau coefficient (an ordinal statistic) was used.

References

Adorno, T.W., Frenkel-Brunswik, E., Levinson, D.J., & Sanford, R.N. (1950). *The authoritarian personality.* New York, NY: Harper & Row.

Anderson, J. (1962). *Studies in empirical philosophy*. Sydney, Australia: Angus & Robertson.

Armstrong, D.M. (1997). *A world of states of affairs*. Cambridge, UK: Cambridge University Press.

Bacon, F. (1960). *New organon* (Vol. 1; F. Anderson, Ed.; J. Spedding, Trans.). New York, NY: Bobbs-Merrill. (Original work published 1620)

Boring, E.G. (1920). The logic of the normal law of error in mental measurement. *American Journal of Psychology*, *31*, 1–33.

Bradley, F.H. (1893). *Appearance and reality*. Oxford, UK: Clarendon Press.

Campbell, N.R. (1920). *Physics, the elements.* Cambridge, UK: Cambridge University Press.

Cliff, N. (1996). *Ordinal methods for behavioural data analysis*. Mahwah, NJ: Erlbaum.

Collingwood, R.G. (1933). *An essay on philosophical method*. Oxford, UK: Clarendon Press.

Denzin, N.K., & Lincoln, Y.S. (1994). *Handbook of qualitative research*. London, UK: Sage.

Dretske, F. (1995). Naive realism. In T. Honderich (Ed.), *The Oxford companion to philosophy* (p. 602). Oxford, UK: Oxford University Press.

Feynman, R.P. (1985). Cargo cult science. In R.P. Feynman & R. Leighton (Eds.), *Surely you're joking Mr. Feynman: Adventures of a curious character* (pp. 339–346). New York, NY: Norton.

Forrest, P., & Armstrong, D.M. (1987). The nature of number. *Philosophical Papers*, *16*, 165–186.

Frege, G. (1903). *Grundgesetze der Arithmetik* [Basic laws of arithmetic] (Vol. 2). Hildesheim, Germany: Georg Olms.

Freud, S. (1953–1975). The interpretation of dreams. In J. Strachey (Ed. & Trans.), *The standard edition of the complete psychological works of Sigmund Freud* (Vols. IV and V). London, UK: Hogarth. (Original work published 1900)

Galilei, G. (1960). *Galileo Galilei et al., the controversy of the comets of 1618* (S. Drake, Trans.). Philadelphia: University of Pennsylvania Press. (Original work published 1623)

Guba, E.G., & Lincoln, Y.S. (1994). Competing paradigms in qualitative research. In N.K. Denzin & Y.S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 105–117). London, UK: Sage.

Haeberlin, H.K. (1916). The theoretical foundations of Wundt's folk psychology. *Psychological Review*, *23*, 279–302.

Heath, T.L. (1908). *The thirteen books of Euclid's Elements* (Vol. 2). Cambridge, UK: Cambridge University Press.

Hibberd, F. (2005). *Unfolding social constructionism*. New York, NY: Springer.
Hölder, O. (1901). Die Axiome der Quantität und die Lehre vom Mass [The axioms of quantity and the theory of measurement]. *Berichte über die Verhandlungen der Königlich Sächsischen Gesellschaft der Wissenschaften zu Leipzig, Mathematisch-Physische Klasse*, *53*, 1–46.
Hume, D. (1888). *A treatise of human nature*. Oxford, UK: Clarendon Press.
James, W. (1902). *The varieties of religious experience: A study in human nature*. New York, NY: Longmans, Green & Co.
James, W. (1910). Bradley or Bergson? *Journal of Philosophy, Psychology and Scientific Methods*, *7*, 29–33.
Johnson, H.M. (1936). Pseudo-mathematics in the social sciences. *American Journal of Psychology*, *48*, 342–351.
Joost-Gaugier, C.L. (2006). *Measuring heaven: Pythagoras and his influence on thought and art in antiquity and the Middle Ages*. New York, NY: Cornell University Press.
Kahn, C.H. (2001). *Pythagoras and the Pythagoreans*. Indianapolis, IN: Hackett.
Keynes, J.M. (1921). *A treatise on probability*. London, UK: Macmillan.
Krantz, D.H., Luce, R.D., Suppes, P., & Tversky, A. (1971). *Foundations of measurement* (Vol. 1). New York, NY: Academic Press.
Locke, J. (1959). *An essay concerning human understanding* (Vol. 2). Oxford, UK: Oxford University Press. (Original work published 1690)
Lord, F.M. (1953). On the statistical treatment of football numbers. *American Psychologist*, *8*, 750–751.
Lord, F.M. (1954). Further comments on "football numbers." *American Psychologist*, *9*, 264–265.
Mackie, J.L. (1973). *Truth, probability and paradox: Studies in philosophical logic*. Oxford, UK: Clarendon Press.
Marchel, C., & Owens, S. (2007). Qualitative research in psychology: Could William James get a job? *History of Psychology*, *10*, 301–324.
McCall, W.A. (1922). *How to measure in education*. New York, NY: Macmillan.
McEvoy, J. (1987). The divine as the measure of being in Platonic and scholastic thought. In J.F. Wippel (Ed.), *Studies in medieval philosophy* (pp. 85–116). Washington, DC: Catholic University of America Press.
McKeon, R. (1941). *The basic works of Aristotle*. New York, NY: Random House.
Michell, J. (1986). Measurement scales and statistics: A clash of paradigms. *Psychological Bulletin*, *100*, 398–407.
Michell, J. (1990). *An introduction to the logic of psychological measurement*. Hillsdale, NJ: Erlbaum.
Michell, J. (1993). Numbers, ratios, and structural relations. *Australasian Journal of Philosophy*, *71*, 325–332.
Michell, J. (1994). Numbers as quantitative relations and the traditional theory of measurement. *British Journal for Philosophy of Science*, *45,* 389–406.
Michell, J. (1997). Quantitative science and the definition of *measurement* in psychology. *British Journal of Psychology*, *88*, 355–383.
Michell, J. (1999). *Measurement in psychology: A critical history of a methodological concept.* Cambridge, UK: Cambridge University Press.
Michell, J. (2000). Normal science, pathological science and psychometrics. *Theory & Psychology*, *10*, 639–667.
Michell, J. (2002). Stevens's theory of scales of measurement and its place in modern psychology. *Australian Journal of Psychology*, *54,* 99–104.
Michell, J. (2003). The quantitative imperative: Positivism, naïve realism and the place of qualitative methods in psychology. *Theory & Psychology*, *13,* 5–31.

Michell, J. (2004). The place of qualitative methods in psychology. *Qualitative Research in Psychology*, *1*, 307–319.

Michell, J. (2005). The logic of measurement: A realist overview. *Measurement*, *38*, 285–294.

Michell, J. (2006). Psychophysics, intensive magnitudes, and the psychometricians' fallacy. *Studies in the History and Philosophy of Biological and Biomedical Sciences*, *17*, 414–432.

Michell, J. (2007a). Measurement. In S. Turner & M. Risjord (Eds.), *Handbook of the philosophy of science: Philosophy of anthropology and sociology* (pp. 71–119). Amsterdam, The Netherlands: Elsevier.

Michell, J. (2007b, June). *Bergson's and Bradley's versions of the psychometricians' fallacy argument*. Paper presented at the First Joint Meeting of ESHHS and CHEIRON, University College Dublin, Ireland.

Michell, J. (2008). Is psychometrics pathological science*? Measurement: Interdisciplinary Research and Perspectives*, *6,* 7–24.

Michell, J. (2009). The psychometricians' fallacy: Too clever by half? *British Journal of Mathematical and Statistical Psychology*, *62*, 41–55.

Michell, J., & Ernst, C. (1996). The axioms of quantity and the theory of measurement, Part I. An English translation of Hölder (1901), Part I. *Journal of Mathematical Psychology*, *40*, 235–252.

Niall, K.K. (1995). Conventions of measurement in psychophysics: Von Kries on the so-called psychophysical law. *Spatial Vision*, *9*, 275–305.

Osatuke, K., & Stiles, W.B. (2011). Numbers in assimilation research. *Theory & Psychology*, *21*, 200–219.

Petocz, A. (1999). *Freud, psychoanalysis, and symbolism*. Cambridge, UK: Cambridge University Press.

Petocz, A. (2001). Psychology in the twenty-first century: Closing the gap between science and the symbol. In J.R. Morss, N. Stephenson, & H. van Rappard (Eds.), *Theoretical issues in psychology: Proceedings of the International Society for Theoretical Psychology 1999 Conference* (pp. 367–377). Boston, MA: Kluwer.

Plato. (1971). *Timaeus and Critias* (B. Radice, Ed.; D. Lee, Trans.). Harmondsworth, UK: Penguin.

Porter, T.M. (1995). *Trust in numbers: The pursuit of objectivity in science and public life.* Princeton, NJ: Princeton University Press.

Reisch, G.A. (2005). *How the Cold War transformed philosophy of science to the icy slopes of logic.* Cambridge, UK: Cambridge University Press.

Resnik, M.D. (1997). *Mathematics as a science of patterns*. Oxford, UK: Oxford University Press.

Richardson, A. (2007). "The sort of everyday image of logical empiricism": Thomas Kuhn and the decline of logical empiricist philosophy of science. In A. Richardson & T. Uebel (Eds.), *The Cambridge companion to logical empiricism* (pp. 346–369). Cambridge, UK: Cambridge University Press.

Riedweg, C. (2002). *Pythagoras: His life, teaching and influence*. New York, NY: Cornell University Press.

Rozeboom, W.W. (1961). Ontological induction and the logical typology of scientific variables. *Philosophy of Science*, *28*, 337–377.

Schorske, C.E. (1997). The new rigorism in the human sciences, 1940–1960. In T. Bender & C.E. Schorske (Eds.), *American academic culture in transformation: Fifty years, four disciplines* (pp. 309–329). Princeton, NJ: Princeton University Press.

Solovey, M. (2004). Riding natural scientists' coattails onto the endless frontier: The SSRC and the quest for scientific legitimacy. *Journal of the History of the Behavioral Sciences*, *40*, 393–422.

Stevens, S.S. (1936). Psychology: The propaedeutic science. *Philosophy of Science*, *3*, 90–103.

Stevens, S.S. (1939). Psychology and the science of science. *Psychological Bulletin*, *36*, 221–263.

Stevens, S.S. (1946). On the theory of scales of measurement. *Science*, *103*, 677–680.

Stevens, S.S. (1951). Mathematics, measurement and psychophysics. In S.S. Stevens (Ed.), *Handbook of experimental psychology* (pp. 1–49). New York, NY: Wiley.

Suppes, P., & Zinnes, J.L. (1963). Basic measurement theory. In R.D. Luce, R.R. Bush & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 1, pp. 1–76). New York, NY: Wiley.

Thorndike, E.L. (1918). The nature, purposes, and general methods of measurements of educational products. In G.M. Whipple (Ed.), *Seventeenth yearbook of the National Society for the Study of Education* (Vol. 2, pp. 16–24). Bloomington, IL: Public School Publishing.

Titchener, E.B. (1905). *Experimental psychology: A manual of laboratory practice: Vol. II. Quantitative experiments: Part II. Instructor's manual*. London, UK: Macmillan.

von Helmholtz, H. (1887). Zählen und Messen erkenntnistheortisch betrachtet [Numbering and measuring from an epistemological viewpoint]. *Philosophische Aufsätze Eduard Zeller zu seinem fünfzigjährigen Doktorjubiläum gewidmet* [Philosophical essays dedicated to Eduard Zeller on the occasion of the fiftieth anniversary of his doctorate] (pp. 17–52). Leipzig, Germany: Fues' Verlag.

von Kries, J. (1882). Über die Messung intensiver Grössen und über das sogenannte psychophysische Gesetz [On the measurement of intensive magnitudes and the so-called psychophysical law]. *Vierteljahrsschrift für wissenschaftliche Philosophie, 6*, 257–294.

Westerman, M.A. (2011). Conversation analysis and interpretive quantitative research on psychotherapy process and problematic interpersonal behavior. *Theory & Psychology*, *21,* 155–178.

Yanchar, S.C. (2011). Using numerical data in explicitly interpretive, contextual inquiry: A "practical discourse" framework and examples from Engeström's research on activity systems. *Theory & Psychology*, *21*, 179–199.

Joel Michell is an honorary member of the School of Psychology at the University of Sydney, Australia, where he studied and taught psychometrics and the history and philosophy of psychology for more than forty years. He has published in psychology and philosophy journals, written books on the history and philosophy of measurement in psychology, and composed numerous articles for encyclopaedias, handbooks and other collections. He is currently investigating the history and philosophy of the distinctions between the categories of kind, degree, and measure. Address: School of Psychology, University of Sydney, Sydney, NSW 2006, Australia. [email: joelm@psych.usyd.edu.au]