CHAPTER 3

THE QUANTITY/QUALITY INTERCHANGE

A Blind Spot On The Highway Of Science

Joel Michell

Science has but one subject and only one fundamental method. Its subject matter is *the array of interconnected and interacting systems we find in the universe around us*, including, of course, the various social, psychological, and biological systems of primary concern to psychologists. Science's fundamental method is that of *critical inquiry*, whereby hypotheses about the structure and ways or working of systems are subjected to scrutiny relative to the observations we make and logic, as we understand it. While science has but one fundamental method, realization of that method takes any one of indefinitely many different forms in the potentially infinite array of contexts of investigation. In particular, in subjecting hypotheses to critical scrutiny, various methods of observation are employed, such methods being tailored to both the *substantive* and *formal* character of the attributes under investigation.

Any system is always composed of things of various kinds standing in various relations to one another and, so, involves attributes. Attributes always come in ranges. That is, to take a property as an example, there is not just one possible *length*; there is the infinite array of different possible lengths. Or, to consider a relation, there is not just one possible *nationality*; there is the finite array of different possible nationalities. Such arrays of properties

Methodological Thinking in Psychology: 60 Years Gone Astray?, pages 45–68

or relations are what I mean by *attributes*. Attributes differ from one another *substantially*, in the sense that length is a different kind of attribute to nationality, involving extension in space rather than connection to a nation state. Furthermore, attributes may differ from one another *formally*, in the sense that length is a *quantitative* attribute and nationality, a merely *classificatory* (or categorical) attribute. Formally, amongst the indefinitely large array of formal possibilities, attributes may be *quantitative*, *ordinal*, or *classificatory*, the latter two kinds being instances of what is meant by non-quantitative or *qualitative* attributes. It is important to note that observational methods differ depending upon the formal character of the attribute investigated, in particular, upon whether the attribute is quantitative or qualitative.

Examples of categories of quantitative attributes are *frequencies* and *continuous quantities*. For example, the number of people living in a city is a frequency. Frequencies are generally observed by methods of counting and the frequency so observed is always a natural number. Continuous quantities cannot be counted and they are best observed by methods of measurement. Although counting may be involved in the process of measurement, measurement itself is the estimation of the ratio between a particular magnitude of the attribute and a unit magnitude, such a ratio always being a positive real number.

It was through the discovery and measurement of continuous quantitative attributes that science, and in particular, physical science, enabled the spectacular progress in our understanding of the world around us that transformed human life and society over the past five hundred years. Not only has this progress been an inspiration for many later scientists but unfortunately, physics is now sometimes taken as the paradigm of what any science ought to be like by investigators in disciplines like psychology, which emerged after many of the triumphant successes of physics had occurred. Thus, in the tradition of British and American psychometrics, modern psychology is thought to have commenced with G. T. Fechner's (1860) development of quantitative psychophysical methods and the progress of the discipline is gauged by the steadily increasing sophistication of quantitative theories and methods (e.g., Rao & Sinharay, 2007).

At first, attempts at psychological measurement coexisted with non-quantitative methods of observation. Some of the founding fathers of modern psychology (e.g., Freud, 1900; James, 1902; Wundt [see Haeberlin, 1916]) employed qualitative methods in their investigations and some researchers questioned the suitability of the subject matter of psychology to methods of measurement (e.g., von Kries, 1882). Until the middle of the twentieth century, especially in Europe, but also in the United States, strong traditions of qualitative research developed (e.g., the psychoanalytic tradition, the phenomenological tradition, as well as Köhler, Piaget, Vygotsky, Adorno, and others) and these made enduring contributions to the development of psychology as a science.

However, after the Second World War, the balance changed. In the US, new economic conditions regulating research emerged, conditions documented by Schorske (1997) and Solovey (2004). These changes caused the human sciences increasingly to ape the quantitative rigor of the physical sciences. Given the collapse of psychological research in Europe during the war and the influence of the US in the post-war period, these changes affected psychology, first, in countries, like Britain, Canada, and Australia; and, later, in Europe and Asia as well. Now the mainstream wanted to present psychology as methodologically "rigorous" (i.e., quantitative), a strategy, which it was thought would maximize funding opportunities under the new, post-war dispensation and as a result the use of qualitative methods declined. It is notable that these changes occurred in the context of a much wider identification of quantification with objectivity that affected other movements and institutions in Western society (Porter, 1995), so these developments in psychology were interpreted as riding the crest of a wave to the scientific destiny of the discipline. While qualitative methods have recently reappeared at the fringes of the discipline (Marchel & Owens, 2007), the mainstream remains committed exclusively to quantitative methods. Is there now room for qualitative methods in psychology and are any of the criticisms of quantification offered during earlier times valid?

There are many issues here, but the most striking is the fact that when the sorts of attributes psychometricians aspire to measure (e.g., mental abilities, personality traits, social attitudes, and sensory intensities) is considered; these attributes are always only experienced as qualitative attributes, never as quantitative. For example, one person gets a test item correct, while another is incorrect and, so, the first is observed to be *more intellectually able* than the second with respect to that item; one person answers "yes" when asked do they like to go to parties, while another says "no" and, so, the first is observed to present him or herself as being *more sociable* than the second; similarly, one attitude is observed to be *more supportive of euthanasia* than another; or one cup of coffee is reported to taste *sweeter* than another. These are observations of ordinal (i.e., qualitative) relations between degrees of the attributes involved, not of quantitative relations and there is nothing in these sorts of observations *per se* that entails that the relevant attributes are quantitative.

While aggregations over many such qualitative observations are quantitative, what is directly observed in counting such aggregations are frequencies (i.e., the number of times that a response of a given kind is made), not continuous quantities, and such frequencies are not, by themselves, *measures* of the attributes investigated (i.e., of intellectual abilities, personality traits, social attitudes, or sensory intensities). Such frequencies, of course, may be *interpreted* as measures of such attributes, but doing so always requires invoking the hypothesis that the relevant attributes are continuous quantities and are linked by some specific mathematical function to such frequencies, as is postulated in factor analytic or item response theories. That is, when psychometricians interpret such frequencies as measures of such attributes, they do so only via the use of hypotheses presuming that such attributes are quantitative. That abilities, etc., are quantitative is never observed and as I have pointed out (Michell, 1990, 1997, 1999, 2000, 2006, 2007a, 2008, & 2009), in the century and a half since Fechner (1860), no scientific evidence has been collected capable of sustaining the hypotheses that such attributes are quantitative. To understand the character of the presumption that psychological attributes are quantitative and where it has come from it is necessary to consider in some detail the difference between continuous, quantitative (i.e., measurable) and qualitative (i.e., non-measurable) attributes.

(i) Measurable Attributes

Otto Hölder (1901)[1] specified the scientific sense of *quantitative* and later expositions have used his axioms (e.g., Behrend, 1953; Huntington, 1902; Kyburg, 1997; Mundy, 1987; Nagel, 1931; Niederée, 1992; Suppes, 1951; Swoyer, 1987; & Whitney, 1968). In the following, by "a quantity" is meant a measurable, determinable property or relation, such as length, mass or duration (Armstrong, 1997). Specific, determinate properties or relations falling under a given quantity are called "magnitudes of that quantity." For example, length, in general, is a quantity, while a length of two meters is a magnitude of the attribute, length. Letting Q denote a quantity; letting magnitudes of Q be designated by $a, b, c, \ldots$ and for any three magnitudes, a, b, and c, of Q, letting $a + b = c$ mean that c is entirely composed of discrete parts a and b, Hölder's axioms for an *unbounded, continuous quantity* are as follows:

1. Given any two magnitudes, a and b, of Q, one and only one of the following is true:
   i. a is identical to b (i.e., $a = b$ and $b = a$);
   ii. a is greater than b and b is less than a (i.e., $a > b$ & $b < a$); or
   iii. b is greater than a and a is less than b (i.e., $b > a$ & $a < b$).
2. For every magnitude, a, of Q, there exists a b in Q such that $b < a$.
3. For every pair of magnitudes, a and b, in Q, there exists a magnitude, c, in Q such that $a + b = c$.
4. For every pair of magnitudes, a and b, in Q, $a + b > a$ and $a + b > b$.

5. For every pair of magnitudes, *a* and *b*, in *Q*, if $a < b$, then there exists magnitudes, *c* and *d*, in *Q* such that $a + c = b$ and $d + a = b$.
6. For every triple of magnitudes, *a*, *b*, and *c*, in *Q*, $(a + b) + c = a + (b + c)$.
7. For every pair of classes, ϕ and ψ, of magnitudes of *Q*, such that
   i. each magnitude of *Q* belongs to one and only one of ϕ and ψ;
   ii. neither ϕ nor ψ is empty; and
   iii. [illegible]s less than each magnitude in ψ, there ex- [illegible]*Q* such that for every other magnitude, *x*', [illegible] ψ and if $x' > x$, then $x' \in \psi$ (depending on the particular case, *x* may belong to either class).

replace 'ψ' with 'ϕ'

I stress that these axioms do not describe the *behavior of objects*: of, say, rods when concatenated or marbles of different weights combined in balance pans. They specify the *structure of attributes*. Whether this specification obtains for any given attribute may sometimes be tested via observations of the behavior of objects, but it is only in the case of *extensive* quantities, like length or weight, that such tests directly reflect the hypothesized structure. In the case of *intensive* quantities, like temperature or density, evidence is always indirect. However, in both cases, relevant attributes are hypothesized to possess the above structure and for the attributes measured in physics, there is overwhelming evidence, both direct and indirect, that they do.

The significance of these axioms is that they show how the possibility of measurement unfolds from the structure of the attribute involved. If an attribute has this kind of structure then ratios of its magnitudes equal positive real numbers and, so, if any specific magnitude is adopted as the unit of measurement, the *measure* of each magnitude is its ratio to the unit. The *measurement* of any magnitude, *a*, in units of another, *b*, is then an estimate of the ratio of *a* to *b*. Quantitative structure is one thing; measurement another, but the latter always depends on the former. Measurement is not something that can be imposed upon any attribute. *Only quantitative attributes can be measured.*

This understanding of measurement has its roots in Book V of Euclid's *Elements* (Heath, 1908), assembled around the 4th century, BC. Euclid defined the concept of *ratio* and showed how ratios of magnitudes (including incommensurable magnitudes) relate to ratios of numbers (i.e., in modern terms, rational numbers), a treatment, which in some respects anticipated the modern concept of real number (Stein, 1990). It was left to Hölder (1901) to specify the kind of structure that an attribute must have for ratios of its magnitudes to equal real numbers.

Because Euclid's *Elements* was a constant component of education from ancient times until the twentieth century (with the possible exception of the Early Middle Ages), his concept of ratio became a key component of the paradigm of measurement informing the Western Intellectual Tradition and, crucially, informing the Scientific Revolution of the seventeenth century, which gave centre stage to quantitative methods. This paradigm still informs modern quantitative science. To claim that an attribute is measurable is to imply that it is quantitative.

During the twentieth century, a different paradigm, the *representational*, came to dominate philosophy of science. Its trajectory stretches from Russell (1903) to Narens (2002) and contributions to measurement theory carried out under its auspices (e.g., Krantz, Luce, Suppes, & Tversky, 1971) are best understood within the Euclidean paradigm as indicating the sorts of empirical situations that may be used to gather evidence relevant to the issue of whether attributes are quantitative (Michell, 2007a, b). Krantz, et al. (1971) clearly display how physical, extensive quantities, such as length, weight, and time satisfy those of Hölder's axioms that are directly testable; and how physical, intensive quantities, such as density and temperature satisfy their axioms of conjoint measurement, which provide indirect evidence of quantitative structure. Not all attributes of interest to physicists have been shown to be quantitative. An example is *hardness*. Whether the structure of this attribute satisfies Hölder's axioms remains, at present, unknown (Luce, 2005) and, while different degrees of hardness may be ordered, hardness is not presently measurable.

(ii) Qualitative Attributes

What about qualitative attributes? In so far as the sorts of attributes that psychologists aspire to measure are considered, because these are experienced as ordinal, I will confine discussion to ordinal attributes. From a mathematical viewpoint, Bertrand Russell (1901) pioneered the idea that order is of interest in its own right[2] and this led to a specification of ordinal structure. Mathematicians now distinguish many varieties of order, ranging from the *strict simple* to the *partial* (Michell, 1990). Every quantity entails a strict simple ordering of its magnitudes as, for example, the class of all different lengths is ordered, each length being greater or less than each other. A strict simple order is one in which the relevant order relation is *transitive*, *asymmetric* and *connected*. Letting xRy mean that x stands in relation R to y, these three properties are as follows.

1. A relation, R, is *transitive* upon a class if and only if for every x, y, and z in that class, if xRy and yRz, then xRz.
2. A relation, R, is *asymmetric* upon a class if and only if for every x and y in that class, if xRy, then not yRx.

3. A relation, R, is *connected* upon a class if and only if for every x and y in that class such that $x \neq y$, either xRy or yRx.

These axioms describe what is common to all of those sequences we encounter in life when, figuratively speaking, things can be lined up in such a way that with respect to any two in the sequence, one always comes before the other (connexity), if one comes before another, the second does not precede the first (asymmetry) and whenever one things comes before another and the later comes before something else, then the first always comes before the third (transitivity). Of course, putting the matter in common parlance, this way, shows how dependent we are upon both spatial and temporal metaphors when describing order. We speak of things being *lined up* and of *one thing coming before another*. The fact that space and time are not only ordinal in structure, but quantitative as well means that we easily slide from these metaphors into thinking that the sequences they are used to describe are also quantitative. Likewise, sequences that are correlated with spatial or temporal orders, such as the sequence of intellectual, developmental stages, may be assumed to be quantitative because of this association. Thus, psychologists easily came to think that intellectual development parallels chronological age in being quantitative and, thus, saw IQ as a measurement of intelligence.

However, comparing Hölder's seven axioms with these three conditions for a strict, simple order, it is clear that the concept of quantity involves more than order. Quantity is order plus something else, viz., additive structure. Mere order lacks additive structure. Nonetheless, since order alone does not rule out the possibility of quantitative structure, those who want to measure qualitative attributes often feel free to speculate that the attributes of interest to them are really quantitative. Psychologists have gone further and *concluded* that the attributes they aspire to measure are quantitative, even though they have no scientific evidence sustaining this conclusion. In concluding this they have used a number of forms of argument. It is instructive to review these, for while they fail to establish the desired conclusion, reviewing them will sharpen our focus on the relevant issues.

THE INFERENCE FROM ORDER TO QUANTITY

(i) Pythagoreanism

There is a subterranean, Pythagorean current beneath the flow of Western culture (Kahn, 2001; Riedweg, 2002), one presuming that the underlying structure to reality is quantitative. Filtered through Platonism in philosophy (see, for example, Joost-Gaugier, 2006) and via Augustinian theology (McEvoy, 1987), it is manifest in our preference for quantitative conceptions over non-quantitative ones in circumstances where direct experience shows otherwise. The world of ordinary experience is Aristotelian in the sense that we seem to find both quantitative and qualitative attributes therein. It is as if the Pythagorean inhabits another world, one in which reality is composed entirely of quantities and in which qualitatively ordered attributes are always seen as latently quantitative. If the Pythagorean doctrine that "all things are numbers" (Burnet, 1914, p. 52) is understood to mean that all attributes are quantitative, then it follows deductively that all qualitatively-ordered attributes are quantitative, despite appearances, and, therefore, in principle, measurable. For some, this Pythagorean world is taken to be the world of science. This current is discernable in the historical course of psychology (Murphy, 1967) and has often found verbal expression, as, for example, in Thorndike's much used *credo*, that "whatever exists at all exists in some amount" (1918, p. 16).[3] Despite its hold over our minds, Pythagoreanism is contrary to the scientific spirit. Whether any attribute is quantitative is an empirical issue and, so, from a scientific point of view, must be judged on the basis of evidence and not philosophical dogma.

(ii) History of Quantitative Science

Amongst attributes now successfully measured in physics, some, in the first instance were experienced only as qualitative orders. This has affected the thinking of scientists and non-scientists alike, making it seem credible that quantitative structure hides behind apparently ordinal attributes. For the layperson, the most potent examples are temperature and pressure because household instruments are used to measure them. These instruments captured the public's imagination in the eighteenth century[4] and since then have provided ready metaphors for discussing qualitative attributes (see Castle, 1995). We speak of "barometers" and "thermometers" for many qualities, which although unmeasured are often discussed in quantitative terms. However, these metaphors are theoretically loaded, exhibiting qualitative attributes through metrical lenses. One might not believe that some putative "barometer" for one or another quality measures it, but that metaphor primes us to believe in the attribute's measurability.

Within science, the slippery slope from order to quantity became a methodological principle, the *quantitative imperative* (Michell, 1990). As Bunge noted, "The thesis that every ordinal "scale" can be replaced by a quantity or magnitude … has been taken for granted in physics ever since Galilei: it is an unspoken methodological principle" (1973, p. 111). By the nineteenth century, "The world [was] seen as constituted by numerical magnitudes" (Hacking, 1983, p. 242; see also Wise, 1995) and inevitably aspiring sciences absorbed this imperative. Considering the history of psychophysics, Titchener remarked on the "constantly recurring argument that *S* [sensation intensity] must be measurable because it is a magnitude, because we can speak, intelligently and intelligibly, of 'more' and 'less' of it" (1905, p. lxiii). The fact that other human sciences followed psychophysics (Michell, 2007a) encouraged wider public acceptance of the idea that measurement of qualitative, ordinal attributes is an inevitable feature of "progress" in disciplines like economics and management, and domains such as public administration (Porter, 1995) and justice (Spigelman, 2002).

While the history of science may seem to suggest that all ordinal attributes are really quantitative, this is an illusion. The history of science actually teaches the fallibility of our judgment, particularly in cases where we try to second-guess nature. The success of quantitative physics implies no more about the appropriateness of quantitative concepts for psychological attributes than the success of any specific observational method in one sphere implies about its appropriateness for any other. We need to know whether the two contexts are similar in relevant ways. Physics and psychology deal with different attributes and, so, before we can know whether quantitative concepts are appropriate for psychological attributes, we need to know whether such attributes are quantitative.

(iii) Explanatory Simplicity

Russell noted that: "Number is of all conceptions, the easiest to operate with, and science seeks everywhere for an opportunity to apply it" (1896/1983, p. 301), meaning that number enables applications of quantitative mathematics with its limitless supply of numerical methods, inference rules, and theorems. However, quantification entails deeper benefits, enabling parsimonious explanations of change within attributes and relations between them.

Quantitative attributes possess two features enabling this parsimony: additive structure and homogeneity. For example, the same length difference may supplement any given length, as, say, lengths of 10 meters and 100 meters may each be increased by the same length, say one metre. In this respect, the attribute of hardness seems to be different. It has been assessed via Mohs' ordinal scale since 1824 (Jerrard & McNeill, 1992) and it may yet prove to be quantitative, but, as currently understood, hardness *differences* seem non-homogeneous. That is, such differences might not be differences in *amount*; instead, they might be differences in *kind*. That is, whatever constitutes differences in hardness might be a different kind of thing at different degrees of hardness. Then explanations of change of hardness would be more complex than explanations of change of length. The former might even require different laws for different degrees.

When relations between attributes are considered, quantitative concepts again enable simpler explanations. For example, because they are quantitative, the relationship between density, mass, and volume, viz., that *mass* = *density* × *volume*, has a simple mathematical form. However, if ordinal assessment were the only possibility, the relationship would appear more complex:

a. With volume fixed, mass increases with density and vice versa;
b. With density fixed, mass increases with volume and vice versa; and
c. With mass fixed, volume increases with decreases in density and vice versa.

Furthermore, complex as these three laws are, they do not track the way specific changes in any one attribute are linked to specific changes in others. Again, that might require separate, specific laws for individual degrees. With ordinal attributes, a web of complexity stands where quantification promises simplicity.

Intensive Attributes, Additive Structure and Homogeneity

Simplicity, however, comes at a cost. It requires that differences between degrees of the relevant attribute must be both mutually homogeneous and additively structured. As long as quantitative science dealt only with extensive attributes, like length and weight, these requirements were unproblematic because then, as Bradley neatly put it, "the components fall asunder and are visible side by side" (1895, p. 11). However, when first applied to intensive attributes it was not clear how additive structure should be understood. Cross (1998) notes that the medieval philosopher, John Duns Scotus resolved this problem. Scotus showed how qualitative increase or diminution could be thought of as involving the addition or subtraction of homogeneous parts *of the attribute*, even if not parts of the object possessing the attribute. With extensive attributes, the quantitative structure of the attribute is displayed in the behavior of objects possessing magnitudes; as, for example, the behavior of rigid, straight rods of different lengths displays the additive structure of length. Scotus's analysis required distinguishing *attributes* from *objects* and grasping the logical independence between the behavior of objects and the structure of attributes. Just as there is no logical necessity that the concatenation of rods must reflect additive relations between lengths (because how a thing behaves in any situation may, in principle, be a function of any of its attributes), so there is no logical necessity that the existence of additive structure within any quantitative attribute must depend upon the possibility of our being able to appropriately concatenate relevant objects. That is, it is possible to conceive of quantitative increase in attributes even when there is no known additive operation involving relevant objects. This was Scotus's insight and he expressed it in relation to heat as follows:

> Sometimes a tepid [degree] added to a tepid [degree] in diverse subjects does not increase [heat]; [but] this is accidental, on account of the extension and dispersal of the parts. If [the new tepid degree] were in the same [extended] part of a subject with the pre-existing tepidness, then [the subject] would certainly be increased and hotter. (As quoted in Cross (1998, p. 190).

At the time, the dominant view was that hot and cold are qualities because they are experienced as merely ordinal and there are no concatenation operations upon hot or cold objects that directly reflect additive structure. Despite this, Scotus saw that we are able to conceive of increase in heat, by analogy with increase in length, as involving addition of homogeneous parts of the attribute, except that such addition is not via concatenation of separate objects, but occurs *within* one object as, for example, when a container of water is heated by a flame, thereby increasing the quantity of heat.

This insight implies that the distinction between extensive and intensive quantities is a relative one and not intrinsic to them *as quantities*. It has to do with processes involved in discovering quantitative structure, rather than with the character of that attribute's structure itself, although, at first sight, it seems the other way around. This is because extensive magnitudes are experienced directly as both ordered and additive, while intensive magnitudes are experienced directly only as ordered. The quantitative structure of intensive magnitudes is only ever known indirectly.[5]

Scotus successfully augmented the Euclidean paradigm of measurement, including intensive quantities and this breakthrough endured. Consider Kant's understanding of measurement and quantity. As with Scotus, Kant's treatment of these issues was infused with the Euclidean paradigm (Sutherland, 2004), but he had the advantage of four more centuries of scientific progress, including the scientific revolution of the seventeenth century and the subsequent quantification of intensive attributes, like density and force. His understanding of the intensive-extensive distinction, however, is remarkably similar to that of Scotus. Kant thought of intensive quantities as, like extensive quantities, composed of additively interrelated homogeneous magnitudes, each magnitude being a part of those greater than it and having as parts all lesser magnitudes.

Those who thought of qualitative, ordinal attributes as quantitative, now had in Scotus's augmented Euclidean paradigm, a conceptual framework within which such a way of thinking was intelligible. This package was so successful and its simplicity so seductive that by the end of the nineteenth century it seemed to many psychologists to be the only scientific option. For example, the founders of the British and American psychometric tradition, right from the outset foreclosed discussion on the matter: Francis Galton insisted that "until the phenomena of any branch of knowledge have been submitted to measurement and number, it cannot assume the status and dignity of science" (1879, p. 147) and James McKeen Cattell, that "Psychology cannot attain the certainty and exactness of the physical sciences, unless it rests on a foundation of experiment and measurement" (1890, p. 373). The augmented Euclidean paradigm seemed the ideal vehicle for this vision of the future of psychology.

It should be noted in passing, however, in keeping with the theme of this book, that other traditions of psychological thought, particularly in continental Europe were not seduced by the attractions of this package and preferred qualitative approaches. This showed a greater sensitivity, not only to the directly-experienced character of the phenomena under investigation but also to the logic of science, for while the explanatory advantages that quantitative concepts deliver are real and significant, accrual of these advantages in psychology depends directly upon its attributes being quantitative. There is no advantage consistent with the aims of science to be gained by applying quantitative concepts to psychological attributes where these attributes are not quantitative. Our overarching aim in science is to find out how things are in nature, the structure of the attributes investigated and how those attributes interrelate, both causally and otherwise. Quantitative physics derives advantages from quantitative concepts only because it deals with quantitative attributes and it can never advance science to conceptualize attributes as quantitative when they are not. If psychological attributes are actually non-quantitative, then science is only advanced by conceptualizing them as they are, not as we might want them to be. The preference of these continental schools for qualitative methods was entirely justified by what was then known about psychological attributes.

(iv) Homogeneity and the Deduction of Quantity from Order

The reasons considered in the sections above persuaded many, but a few scholars wanted more. It seemed that quantitative structure might be deduced from the character of merely qualitative, ordinal attributes alone. The degrees of any such attribute, all being degrees of the same kind are thereby mutually homogeneous. Whenever one degree is greater than another, must it not contain the lesser degree as a part and, so, be composed additively of homogeneous parts? Thus, must it not possess additive structure and, so, be quantitative? This reasoning is a bridge too far. It not only fails logically to deliver its conclusion; it reveals the reason why that conclusion cannot be delivered by indicating conditions under which qualitative, ordinal attributes are not quantitative.

The philosopher, F. H. Bradley (1895), reasoned along the above lines.[6] He was interested in the intensity of mental states (or, as he called them, *psychical states*) and the issue of whether mere order entails quantitative structure. From the observation that psychical states are ordered according to their degree of intensity, he concluded, "all psychical states … if not measurable in fact are at least measurable in principle" (p. 11). His argument was general, hinging on his premise that "Degrees not resting on units in the end are meaningless" (p. 11). Here is his summary:

> In every perceived series of degrees the second member can be seen to involve the first plus something else of the same sort. All psychical states, apparently without exception, stand to one another as greater or less or again as equal amounts of my personal being. All psychical states therefore (it would seem to follow) must contain a certain amount of common units. It is only in spatial and temporal wholes that the components fall asunder and are visible side by side. And, apart from this, measurement in the proper sense seems to be impracticable. But, apart from this, measurement does not appear in principle and in abstract to be impossible. (p. 11).

Abstracting from Bradley's psychological context, his argument is that any two different degrees of the same kind, call them x and y (where, say, x is greater than y) are such that x is entirely composed of y plus something else of the same kind. (The *something else*, of course, is the difference between x and y). Thus, x and y are just different *amounts* of the same kind of thing, that is, are composed of different numbers of common units. Hence, they are quantitative. Therefore, order entails quantity.

This argument relies upon the suppressed premise that degrees x and y (where $x > y$) are qualitatively homogeneous with the difference between x and y. However, not everyone would have agreed with this. For example, Russell thought that:

> We have, at the outset, a fundamental division of quantities into two kinds, extensive and intensive, according as a change of quantity is, or is not, a quantity of the same kind as the quantity changed. A change of length is itself a length, but a change of temperature or illumination is not itself hot or bright. (Russell, 1990/1897, p. 75).

replace '1990/1897' with '1897/1990'

One might disagree about whether this is the basis of the extensive/intensive distinction, but it appears, for at least some attributes, there is a real issue as to whether differences between degrees are qualitatively homogeneous with the degrees themselves. Russell (1987/1990) later added the example of pleasure (i.e., is the difference between two degrees of pleasure itself a degree of pleasure?) and it, together with other examples suggests

replace '1987/1990' with '1897/1990'

that degrees and differences between degrees might not always be mutually homogeneous. For example, two legal judgments may differ in the degree of justice involved, but is the difference between any two degrees of justice the same kind of thing as a degree of justice? Or two prisons may differ in the degree of freedom that inmates have, but is this difference the same kind of thing as a degree of freedom?

However, when Bradley's argument is reframed in terms of *differences between degrees*, the significance of this issue disappears. The reframed argument is as follows. Let x, y, and z be any three degrees of an ordered attribute such that $x > y > z$. The difference between x and z is entirely composed of two discrete parts, viz., the difference between x and y, and the difference between y and z. Hence, it is an additive sum of parts and a similar conclusion applies with any three degrees. Therefore, the range of degrees is quantitative. But this argument in turn presumes that the class of differences between all pairs of degrees is homogeneous.

In considering this argument, it is tempting to think of degrees as like points on a straight line and differences as like intervals between points. Via this analogy, the argument seems valid. Geometric analogies have been applied to ordered qualities since Nicole Oresme in the fourteenth century (Clagett, 1968) and are implicit in using thermometers or barometers as metaphors for qualitative attributes and, of course, in using rating scales for assessment of psychological attributes. Nonetheless, such analogies are systematically misleading.

Bradley's argument, as reframed presumes two things. First, that for any given ordinal attribute, the differences between degrees are mutually homogeneous and, second, that the class of these differences is appropriately structured, viz., mutually interrelated by analogy with intervals between points on a straight line. Neither is necessarily true. I will not dwell upon the second because I have treated it elsewhere (Michell, 2009). However, in support of qualitative methods in psychology, the first presumption reveals a neglected window of opportunity.

The philosopher, R. G. Collingwood, while an admirer of Bradley, contradicted his argument when, in discussing certain kinds of ordinal attributes, denied that "these things must be in principle measurable, on the ground that whatever exceeds another thing must exceed it by a definite and therefore measurable amount" (1933, pp. 71–72). To illustrate his point, Collingwood considered the psychological attribute, *sensation of heat* and claimed that:

> As I move my hand nearer to the fire, I feel it grow hotter; but every increase in the heat I feel is also a change in the kind of feeling I experience; from a faint warmth through a decided warmth it passes to a definite heat, first pleasant, then dully painful, then sharply painful; the heat at one degree soothes me, at another excites me, at another torments me. I can detect as many dif-

> ferences in kind as I can detect differences in degree; and these are not two sets of differences but one single set. I can call them differences in degree if I like, but I am using the word in a special sense, a sense in which differences of degree not merely entail, but actually are, differences of kind. (1933, pp. 72–73)

Observing that differences between degrees of sensations of heat "are differences of a peculiar type, which are differences at once in degree and kind" (p. 73), Collingwood concluded that "This is why they cannot be measured, for measurement applies only to pure differences of degree" (p. 73). By "pure differences of degree" he meant those that are not also differences in kind. In other words, he insisted that a necessary condition for an ordinal attribute to be quantitative is that *the differences between its degrees be mutually homogeneous*. He concluded that the attribute, *sensations of heat*, is not measurable even though it is ordinal because the class of differences between its degrees is not homogeneous, thereby invalidating Bradley's argument from order to quantity.

The crucial point is that if the differences between the degrees of an ordinal attribute are not homogeneous (i.e., not differences of the same kind), then that attribute is not quantitative, and, thus, not measurable. Such an attribute cannot satisfy Hölder's axioms of quantity. If Hölder's axioms apply to any attribute, then they also apply to the differences between its magnitudes, and if differences between degrees are heterogeneous, they do not satisfy Hölder's axioms because, for example, they fail axiom 1, being not intrinsically greater than, less than or equal to one another. Hence, the relevant attribute must be non-quantitative.

Collingwood's argument is directly relevant to psychology, but psychologists have ignored it. While this is not remarkable, given that his argument appeared in a philosophical work and that he is a minor figure in twentieth century philosophy; it is remarkable when we consider that a similar argument had been presented fifty years earlier by von Kries (1882)[7] in his widely-cited critique of psychophysical measurement. Von Kries[8] (1882) reasoned that within any increasing series of sensations, sensation differences can never be equal because with respect to any two increases from different points in the series, "one increment is something quite different from the other; … they admit of no comparison" (Niall, 1995, p. 291). This is, von Kries claimed that psychophysical measurement is not possible because the increments are mutually heterogeneous. That this objection was pretty much passed over and had little impact cannot be taken as any reflection upon its intrinsic merits. Later, von Kries "developed a logical interpretation of probability based on the very arguments he used to reject the measurement of sensation" (Heidelberger, 2004, p. 229) and the same argument applied in that new context (von Kries, 1886) influenced the dissertation on probability being written a generation later by the young John Maynard Keynes (see Fioretti, 2001).

Keynes (1921) thought of probabilities as degrees of partial entailment holding between sets of premises and conclusions. In cases where the premises do not entail the conclusion deductively, they may still *support* the conclusion, this relation of support being thought of by Keynes as objective (i.e., in the argument) and not merely subjective (i.e., in the mind of the arguer), as critics maintained. Following von Kries, Keynes recognized that a "degree of probability is not composed of some homogeneous material, and is not apparently divisible into parts of like character with one another" (1921, p. 30) and, so, concluded that across the class of arguments providing good but inconclusive reasons for a conclusion, differences between degrees of partial entailment may be qualitatively heterogeneous. So this kind of objection to the measurement of ordinal attributes was in the air prior to the Second World War and influenced one of the century's greatest minds.

While the idea that homogeneity is necessary for quantity is as old as Euclid,[9] it is quite another matter to recognize the possibility of non-homogeneity in the differences between degrees of an ordinal attribute. Nonetheless, this possibility is only a small step away from the related idea of heterogeneous differences between determinate qualities falling under a general determinable,[10] as, for example, the determinate qualities of *being a dog*, *being a cat*, *being a human*, etc. fall under the general determinable, *being an animal*. The difference between *being a dog* and *being a cat*, for instance, is qualitatively different to that between, say, *being a dog* and *being a human*. Furthermore, sometimes the determinates falling under a general determinable are hierarchically ordered, as, for example, the qualities of *being an animal*, *being an ungulate*, *being a perissodactyl*, *being a horse* are ordered because all horses are perissodactyls, all perissodactyls are ungulates, all ungulates are animals. However, such an ordering is not also quantitative because the differences between the consecutive qualities are not mutually homogeneous.

These were matters of general knowledge. For example, according to Sutherland (2004), Kant contrasted the part-whole relation obtaining between magnitudes of the same quantity and the part-whole relation obtaining between hierarchically ordered concepts, that is, all pairs of concepts *X* and *Y*, where *all X are Y*. The important contrast is between the fact that with quantities, relations of equality hold between the differences between some pairs of magnitudes and the fact that equality relations never hold between the differences between concepts so ordered. This is because in the latter case the differences are heterogeneous. That is, for example, the difference between *being ungulate* and *being a perissodactyl* is qualitatively different to the difference between, say, *being animal* and *being ungulate*. Such a hierarchically ordered set of concepts constitutes a counterexample to

Bradley's argument and, so, displays the fact that mere order does not entail quantity.

Given the influence of Kant's philosophy upon late nineteenth and early twentieth century German philosophy and, therefore, upon Continental psychology during that period, Kant's observations on this point may also help explain the widespread preference within Continental psychology for qualitative methods in contexts similar to those where British and American psychologists favoured quantitative methods. Their Kantian heritage inoculated them against the temptation to see every ordinal attribute as quantitative. However, Kant's influence was weaker at that time in Britain and the US.

The failure of these psychologists to take account of these matters also reflects the path taken by the representational paradigm of measurement because it dominated American psychology after the 1930s (Michell, 1999). Whereas the Euclidean paradigm locates the measurability or otherwise of an attribute in the character of the attribute's structure, in American psychology the emphasis for most of the twentieth century, under the influence of the representational paradigm, was upon the numerical representation of empirical relations between objects (Michell, 2007b). The character of the attribute was no longer taken to be an issue. By the time attention focused again on attributes and quantitative structure (beginning with Mundy (1987) and Swoyer (1987)), the issue of the heterogeneity of differences was long forgotten and, so, not attended to. The issue is fundamental, however, not only because it is logically possible that differences between degrees of an ordinal attribute could be mutually heterogeneous, but also because it is possible that at least some of the attributes that psychologists and others aspire to measure possess this characteristic, thereby making those attributes unmeasurable in principle.

An Example of an Ordinal Attribute with Heterogeneous Differences Between Degrees

As an example, consider the Functional Independence "Measure" mentioned by Embretson (2006). It is used to assess how independently the elderly function physically in their daily lives and consists of the following series of items, listed in order, from most to least difficult:

1. Climbing stairs;
2. Transferring to bathtub;
3. Bathing;
4. Walking;
5. Dressing upper body;
6. Independent toileting;
7. Transferring to bed;

8. Dressing lower body;
9. Mobility without a wheelchair;
10. Bladder control;
11. Performing personal grooming; and
12. Bowel control.

It appears that elderly people generally lose the ability to climb stairs unaided before losing the ability to transfer unaided to a bathtub and that ability, in turn, is generally lost before the ability to bathe independently, and so on. These sorts of facts order the items. A person's pattern of physical capabilities, as assessed via this scale indicates a specific degree of functional independence.[11] The idea is that a person able to complete all of the activities that another person can plus at least one further activity on the list is the more independent of the two. Thus, degrees of functional independence are homogeneous (because each degree includes all lower degrees).

However, *differences between degrees* seem to be mutually heterogeneous. For example, the difference between being able, on the one hand, and being unable, on the other, to climb stairs, and the difference between being able and being unable to transfer to a tub are differences of qualitatively diverse kinds, and each of these in turn is of another kind to that between being able and being unable to bathe independently, and so on for the other differences between degrees of this attribute. Because of this, it is not obvious that *the differences between degrees* of functional independence stand in intrinsic relations of *greater than*, *less than* or *equality* to one another.

If we try to think of such differences as being thoroughly homogeneous, we face the question, what could a decrease in functional independence be other than an inability to do some kind of specific daily activity that one was previously able to do independently? As far as I can see, there is no homogeneous stuff, *independence*, adhering in various amounts to each person; there is only the set of distinct capacities to do the range of different daily activities constituting total functional independence from others, which, in being lost one by one, mark decreasing degrees of that attribute. Before it would be safe to conclude that *differences between degrees* of functional independence are ordered, evidence of non-arbitrary order relations between such differences would be required. Otherwise, the safest conclusion is that functional independence is a merely ordinal attribute.[12]

To generalize from this example, given an attribute that is experienced only as ordinal, it is never safe to conclude that it is quantitative because it is possible that differences between its degrees may be heterogeneous. For those interested to investigate the possibility that such an attribute is quantitative, *the first task is produce evidence of non-arbitrary order relations upon differences between its degrees.*

CONCLUSIONS

If there are, as Collingwood argues, ordered attributes in which differences between degrees are qualitatively heterogeneous then such attributes are, as Collingwood noted, intrinsically non-quantitative and, therefore, unmeasurable. I will refer to any such attribute as one possessing *intrinsically non-metric degrees* (or *IND*). Recognizing the possibility of *IND* alters the complexion of not only Bradley's argument from order to quantity, to which it stands as an invalidating counterexample, but also the arguments looked at earlier in this paper. This is because it raises the possibility that psychological attributes have this form. That is, it raises the possibility of *mental IND* (or *MIND*). Given the possibility of *MIND*, qualitative methods are at least as scientifically reasonable as quantitative methods. Let me expand upon these conclusions.

The fact that *MIND* appears to exist is of enormous significance for psychology. First of all, it has implications for the reasonableness of all the arguments that psychologists have given for preferring quantitative to qualitative methods. It means that in the contest between the Pythagorean and Aristotelian visions of reality, it is the Aristotelian that is to be preferred. All of the attributes that psychometricians aspire to measure are directly experienced only as ordinal and it is a genuine empirical question whether in each case the differences between degrees are heterogeneous. Considering the character of the items used in psychological tests, the psychological processes involved in producing test responses would appear to favour the hypothesis of heterogeneity. For example, to take the case of tests of mathematical ability, it seems that a person of high mathematical ability, say, does not differ from a person of merely moderate ability by possessing the same kind of knowledge, skills and strategies that distinguish the moderately able from persons of low ability. Instead such a person has a high degree of ability precisely because of the qualitatively different, superior mathematical knowledge, skills and strategies possessed. If such attributes are quantitative, then not only is their quantitative structure yet to be discovered, but also their character as different amounts of some homogeneous quantity is yet to be specified.

Furthermore, the possibility of *MIND* means that it is not reasonable to infer from the history of science alone that psychological attributes must be quantitative. It may be that at least some psychological attributes possess a positive feature (i.e., mutually heterogeneous differences between degrees) that distinguishes them from the kind of physical attributes that have proved to be quantitative. Therefore, before concluding that psychological attributes are quantitative, it is necessary to investigate the possibility that these attributes possess *MIND*.

For the same reason, it is not reasonable to infer that psychological attributes must be quantitative just from the fact that quantification offers simpler forms of explanation. Because of the possibility of *MIND*, that possibility must be first investigated even though, if true, it may require much more complex forms of investigation. As Galileo stressed,

> We must not ask Nature to accommodate herself to what might seem to us the best disposition and order, but must adapt our intellect to what she has made, certain that such is best and not something else. (as quoted in Crombie, 1994, p. 45).

Secondly, the possibility of *MIND*, combined with two facts, viz., that all of the attributes that psychometricians aspire to measure are experienced only as ordinal, and that 150 years of psychometric research has failed to produce hard evidence that such attributes are quantitative, jointly imply that the use of qualitative methods in psychological research is not only justified, it implies that qualitative methods are, on all purely scientific grounds, not to be preferred any less than quantitative methods.

In its turn, this implication makes clear the fact that the mainstream's obsession with the exclusive use of quantitative methods is not good scientific practice. In its turn, this implies that this obsession serves extra-scientific interests. Indeed, it adds more weight to my contention that psychometrics is a pathological science (Michell, 2000, 2004, 2008). In the excitement of founding a new discipline, Fechner's methodological turn was seen as a superhighway through the wilderness, a mistake magnified 60 years ago when quantification was enforced as the only straight and narrow way to the kingdom of science. The possibility of *MIND* invites us to look again, more soberly, at the place of qualitative methods in psychology and at the wisdom of the methodological preferences of those schools of psychological thought that flourished in Continental Europe prior to World War Two.

NOTES

1. See Michell & Ernst (1996, 1997) for an English translation.
2. Indeed, Russell (1897/1990, 1903) adopted the idiosyncratic view that order alone, rather than order plus additivity defines quantity (Michell, 1997).
3. McCall (1922) gives a disquisition expounding Thorndike's Pythagoreanism.
4. For histories of these devices and their popularisation see Middleton (1964, 1966) and Frängsmyr, Heilbron, and Rider (1990).
5. The relativity of the extensive/intensive distinction is clearly demonstrated by showing that had our sensori-motor capacities been different, the additive structure of a quantity like length, which we know as extensive, might only have been known indirectly (i.e., as intensive). See Michell (1993).

6. Bergson had proposed a similar argument in 1889 (see Bergson, 1913) and while Bradley makes no reference to him, there are striking points of contact between their discussions.
7. Niall (1995) provides an English translation of von Kries's paper.
8. Von Kries was a sensory physiologist who formulated the famous "quantity objection" to Fechner's psychophysics. He is also of some importance due to his writings on the concept of probability.
9. Mueller (1981) discusses its central role in his commentary.
10. On the distinction between determinates and determinables see Johnson (1921).
11. Strictly speaking, functional independence is not a psychological attribute. It is a social attribute, which the scale indexes via a range of physical abilities, absence of any one of which contributes to a person's dependence upon helpers. Thus, the attribute assessed is actually a rather complex socio-physical one.
12. Embretson notes that the Rasch model fits Functional Independence scale data "reasonably well" (2006, p. 52). However, this cannot be taken as evidence that the relevant attribute is quantitative. The statistical tests used are not sensitive to the issue of whether the relevant attribute is quantitative and data involving non-quantitative ordinal attributes may fit quantitative psychometric models.

REFERENCES

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

Behrend, F. A. (1953). A system of independent axioms for magnitudes. *Journal and Proceedings of the Royal Society of NSW, 87*, 27–30.

Bergson, H. (1913). *Time and free will.* (F. L. Pogson, trans.). London: George Allen & Co.

Bradley, F. H. (1895). What do we mean by the intensity of psychical states? *Mind, 13*, 1–27.

Bunge, M. (1973). On confusing 'measure' with 'measurement' in the methodology of behavioral science. In M. Bunge (Ed.), *The methodological unity of science* (pp. 105–122). Dordrecht-Holland: Reidel.

Castle, T. (1995). *The female thermometer: eighteenth-century culture and the invention of the uncanny.* Oxford: Oxford University Press.

Cattell, J. McK. (1890). Mental tests and measurements. *Mind, 15*, 373–380.

Clagett, M. (1968). *Nicole Oresme and the medieval geometry of qualities and motions.* Madison, WI: University of Wisconsin Press.

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

Crombie, A. C. (1994). *Styles of scientific thinking in the European tradition.* London: Duckworth.

Cross, R. (1998). *The physics of Duns Scotus: The scientific context of a theological vision.* Oxford: Clarendon Press.

Embretson, S. E. (2006). The continued search for non-arbitrary metrics in psychology. *American Psychologist, 61*, 50–55.

Fechner, G. T. (1860). *Elemente der Psychophysik*. Leipzig: Breitkopf and Hartel.
Fioretti, G. (2001). Von Kries and other 'German logicians': non-numerical probabilities before Keynes. *Economics and Philosophy, 17*, 245–273.
Frängsmyr, T., Heilbron, J. L., & Rider, R. E. (1990). *The quantifying spirit of the 18th century*. Berkeley: University of California Press.
Freud, S. (1900). *The interpretation of dreams. Standard edition, vols. IV & V*. London: Hogarth.
Galton, F. (1879). Psychometric experiments. *Brain, 2*, 147–162.
Hacking, I. (1983). *Representing and intervening: introductory topics in the philosophy of natural science*. Cambridge: Cambridge University Press.
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: Cambridge University Press.
Heidelberger, M. (2004). *Nature from within: Gustav Theodor Fechner and his psychophysical worldview*. Pittsburgh: University of Pittsburgh Press.
Hölder, O. (1901). Die Axiome der Quantität und die Lehre vom Mass. *Berichte über die Verhandlungen der Königlich Sächsischen Gesellschaft der Wissenschaften zu Leipzig, Mathematisch-Physische Klasse, 53*, 1–46.
Huntington, E. V. (1902). A complete set of postulates for the theory of absolute continuous magnitude. *Transactions of the American Mathematical Society, 3*, 264–284.
James, W. (1902). *The varieties of religious experience: A study in human nature*. New York: Longmans, Green & Co.
Jerrard, H. G., & McNeill, D. B. (1992). *Dictionary of scientific units including dimensionless numbers and scales*. London: Chapman & Hall.
Johnson, W. E. (1921). *Logic, part 1*. Cambridge: Cambridge University Press.
Joost-Gaugier, C. L. (2006). *Measuring heaven: Pythagoras and his influence on thought and art in antiquity and the Middle Ages*. New York: Cornell University Press.
Kahn, C. H. (2001). *Pythagoras and the Pythagoreans*. Indianapolis: Hackett Publishing Co.
Keynes, J. M. (1921). *A treatise on probability*. London: Macmillan.
Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (1971). *Foundations of measurement, vol. 1*. New York: Academic Press.
Kyburg, H. E. (1997). Quantities, magnitudes, and numbers. *Philosophy of Science, 64*, 377–410.
Luce, R. D. (2005). Measurement analogies: comparisons of behavioural and physical measures. *Psychometrika, 70*, 227–251.
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: 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.
Michell, J. (1990). *An introduction to the logic of psychological measurement*. Hillsdale NJ: Erlbaum.

Michell, (1993). Numbers, ratios, and structural relations. *Australasian Journal of Philosophy, 71*, 325–332.
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: Cambridge University Press.
Michell, J. (2000). Normal science, pathological science and psychometrics. *Theory & Psychology, 10*, 639–667.
Michell, J. (2004). Item response models, pathological science and the shape of error: reply to Borsboom and Mellenbergh. *Theory & Psychology, 14*, 121–129.
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: Elsevier.
Michell, J. (2007b). Representational theory of measurement. In M. Boumans (Ed.), *Measurement in economics: A handbook* (pp.19–39). Amsterdam: Elsevier.
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.
Michell, J., & Ernst, C. 1997). The axioms of quantity and the theory of measurement, Part II, An English translation of Hölder (1901), Part II. *Journal of Mathematical Psychology, 41*, 345–356.
Middleton, W. E. K. (1964). *The history of the barometer.* Baltimore: The Johns Hopkins Press.
Middleton, W. E. K. (1966). *A history of the thermometer and its use in meteorology.* Baltimore: The Johns Hopkins Press.
Mueller, I. (1981). *Philosophy of mathematics and deductive structure in Euclid's Elements.* New York: Dover.
Mundy, B. (1987). The metaphysics of quantity. *Philosophical Studies, 51*, 29–54.
Murphy, G. (1967). Pythagorean number theory and its implications for psychology. *American Psychologist, 22*, 423–431.
Nagel, E. (1931). Measurement. *Erkenntnis, 2*, 313–333.
Narens, L. (2002). *Theories of meaningfulness.* Mahwah, NJ: Erlbaum.
Niall, K. K. (1995). Conventions of measurement in psychophysics: von Kries on the so-called psychophysical law. *Spatial Vision, 9*, 275–305.
Niederée, R. (1992). What do numbers measure? A new approach to fundamental measurement. *Mathematical Social Science, 24*, 237–276.
Porter, T. M. (1995). *Trust in numbers: the pursuit of objectivity in science and public life.* Princeton: Princeton University Press.
Rao, C. R. & Sinharay, S (2007). *Psychometrics.* Amsterdam: Elsevier.

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

Russell, B. (1896/1983). The a priori in geometry. In K. Blackwell, et al. (Eds.), *The collected papers of Bertrand Russell, volume 1: Cambridge essays, 1896–99* (pp. 291–304). London: Allen & Unwin.

Russell, B. (1897/1990). On the relations of number and quantity. *Mind, 6*, 326–341. Reprinted in N. Griffin & A. C. Lewis (Eds.), *The collected papers of Bertrand Russell, volume 2: philosophical papers 1896–99* (pp. 70–82). London: Routledge.

Russell, B. (1901). On the notion of order. *Mind, 10*, 30–51.

Russell, B. (1903). *Principles of mathematics.* Cambridge: Cambridge University Press.

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: 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.

Spigelman (2002). Quality in an age of measurement. *Quadrant, 66*, 9–15.

Stein, H. (1990). Eudoxos and Dedekind: on the ancient Greek theory of ratios and its relation to modern mathematics. *Synthese, 84*, 163–211.

Suppes, P. (1951). A set of independent axioms for extensive quantities. *Portugaliae Mathematica, 10*, 163–172.

Sutherland, D. (2004). Kant's philosophy of mathematics and the Greek mathematical tradition. *Philosophical Review, 113*, 157–201.

Swoyer, C. (1987). The metaphysics of measurement. In J. Forge (ed.), *Measurement, realism and objectivity* (pp. 235–290). Dordrecht: Reidel.

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: Macmillan.

von Kries, J. (1882). Über die Messung intensiver Grössen und über das sogenannte psychophysische Gesetz. *Vierteljahrsschrift für wissenschaftliche Philosophie, 6*, 257–294.

von Kries, J. (1886). *Die Principien der Wahrscheinlichkeits-Rechnung: Eine logische Untersuchung.* Tübingen: Mohr.

Whitney, H. (1968). The mathematics of physical quantities. Part I: Mathematical model for measurement. *American Mathematical Monthly, 75*, 115–138.

Wise, M. N. (1995). *The values of precision.* Princeton: Princeton University Press.