Human Studies **27:** 429–454, 2004.

Meaning and Method in the Social Sciences[1]

WILLIAM P. FISHER, JR.
MetaMetrics, Inc., 1000 Park 40 Plaza, Suite 120, Durham, NC 27713, U.S.A.
(E-mail: wfisher@lexile.com)

Abstract. Academia's mathematical metaphysics are briefly explored en route to an elaboration of the qualitatively rigorous requirements underpinning the calibration and unambiguous interpretation of quantitative instrumentation in any science. Of particular interest are Gadamer's emphases on number as the paradigm of the noetic, on the role of play in interpretation, and on Hegel's sense of method as the activity of the thing itself that thought experiences. These point toward and overlap with (1) Latour's study of the metrological social networks through which technological phenomena are brought into language as modes of being that can be understood, and (2) the way that Rasch's models for measurement comprise a potential beginning for metaphysically astute, qualitatively and quantitatively integrated, mathematical methods in the social sciences. The paper closes with observations on the general problem that is philosophy, the need to remain open to multiplicities of meaning even as clear understandings are sought and obtained.

Key words: Gadamer, Hans-Georg; metaphysics; metrology; Rasch models; quantitative methodologies

Can We Tell Sense from Nonsense?

Following Wittgenstein (1980: 56), to be able to pay attention to our nonsense, we have to be able to distinguish nonsense from sense. But even a cursory examination of the extensive literatures on epistemology, philosophy, discourse analysis, etc., reveals little or no agreement on how to proceed. As is often the case, there may be more opinions on this matter than there are thinkers and theorists.

What then can we do? Where do we start? Is there any way even to begin to unravel the gridlock of competing criteria by which we might know when nonsense threatens violence, and when sense and sufficient understanding are in hand?

The problem is complicated by the fact that, at the same time as the choice in favor of meaning is made, the potential for nonsense also emerges, since "there can be no philosophy without presuppositions" (Ricoeur, 1974: 96). No matter how deeply they are mined, "no discourse can claim to be free of presuppositions for the simple reason that the conceptual operation by which a region of thought is thematized brings operative concepts into play, which cannot themselves be thematized at the same time" (Ricoeur, 1977: 257). In other words, "the attempt to escape metaphysics is no sooner put in the form of a proposition than it is seen to involve highly significant metaphysical postulates" (Burtt, 1954: 228). Or again, in other words,

> there is no sense in doing without the concepts of metaphysics in order to shake metaphysics. We have no language – no syntax and no lexicon – which is foreign to this history; we can pronounce not a single destructive proposition which has not already had to slip into the form, the logic, and the implicit postulations of precisely what it seeks to contest (Derrida, 1978: 280–281).

Accordingly, "the only way to avoid becoming a metaphysician is to say nothing" (Burtt, 1954: 227). If philosophy's primary presupposition is the choice in favor of meaning (Ricoeur, 1974: 75–76), then philosophy's metaphysics is a metaphysics of meaningfulness, and "'intellectuals' most basic prejudice' [is] that meaning lies in 'writable meaning' or representation" (Wilshire, 1982: 89, in Crease, 1993: 127). Academic discourse inherently assumes, then, that ". . . the sense aimed at through these figures [of metaphor] is an essence rigorously independent of that which transports it, which is an already philosophical *thesis*, one might even say philosophy's *unique thesis*" (Derrida, 1982: 229). Ricoeur (1977: 293) concurs: "no philosophical discourse would be possible, not even a discourse of deconstruction, if we ceased to assume what Derrida justly holds to be 'the sole thesis of philosophy.'"

The dialectic between belonging together in an experience of a whole, on the one hand, and feeling separate and apart in an experience of distanciation, on the other, embodies the creation of meaning. Insofar as any linguistic expression effectively communicates, it does so only to the extent that the metaphorical figures embodied in words, numbers, images, syntax, and semantics both belong to, and separate from, their respective meanings. So it is that Ricoeur (1981: 105) holds that "the first declaration of hermeneutics is to say that the problematic of objectivity presupposes a prior relation of inclusion which encompasses the allegedly autonomous subject and the allegedly adverse object." The possibility of rational objectivity is not in question; the problem is in locating it within the historically situated context in which the experiencing subject finds her- or himself.

This sense of the problem is shared by a wide variety of philosophers who differ markedly in their appraisals of its solvability, as well as in the solutions they offer (Fisher, 2003a, 2003b). Caputo (1997: 80; also see Kearney, 1984: 123–124), for instance, points out that "Derrida is not trying to bury the idea of 'objectivity' . . . [since] it is not that texts and languages have no 'referents' or 'objectivity' but that the referent and objectivity are not what they pass themselves off to be, a pure transcendental signified." So "it is important to see that the kind of negative conclusion that Derrida would constantly enact does not produce anarchy" (Risser, 1989: 184). Or, as Derrida (2003: 63) put it recently, "When I take liberties, it's always by measuring the distance from the standards I know or that I've been rigorously trained in."

Bernstein (1983) similarly seeks to situate standards and objectivity within "a suitably rich notion of practical rationality" that can adjudicate only within a given paradigm (Page, 1987: 83–84). And, as will be shown, Gadamer's (1989) emphases on number as the paradigm of the noetic, on the fundamental role of play in interpretation, and on the revival of Hegel's sense of method, lead to a firm connection with science's and engineering's most rigorous interpretive practices, the calibration, writing, and reading of instruments via the discipline of metrology. In establishing this connection, I explore the limits of Gadamer's and Bernstein's "modern presuppositions about the possibility of theory; in particular that it must be constructive in a manner ultimately subordinate to will" (Page, 1987: 100). That is, Gadamer's rejection of the Cartesian sense of method is founded in a fuller sense of method that is (1) better integrated with theory within the paradigm of mathematical reason, but which is also (2) less subordinate to will than the play of language games, and the invention-discovery of the rules to new ones that can come about only by being open to lived engagement with the process.

Broadly Mathematical Criteria for Meaningfulness

Since discourse is constituted in works or dialogues of finite length (Ricoeur, 1974), reduction is necessarily the first phase in the three phases of the phenomenological method (Heidegger, 1982), the other two being construction (or application) and deconstruction, followed by a spiraling return to a new reduction. "The first and most elementary work of interpretation" is then "to produce a relatively univocal discourse with polysemic words, and to identify this intention of univocity in the reception of messages" (Ricoeur, 1981: 44).

Similarly, "the first concern of all dialogical and dialectical inquiry is a *care for the unity and sameness* of the thing that is under discussion" (Gadamer, 1991: 64). The form of accountability required by dialectical inquiry is one in which "testing sets up the proposition to be tested not as something for one person to defend, as belonging to him or her, but as something 'in the middle,'" with everyone involved "testing the *logos* [the abstract ideal] to see whether it is refutable" and everyone involved "agreeing with regard to its eventual refutation or confirmation" (Gadamer, 1991: 65). Only when everyone agrees as to the unity and sameness of the thing at issue can it legitimately be taken up and applied. Even deconstruction's focus on the lack of unity and sameness in a text itself requires metaphysical assumptions of sameness and unity.

At this point the stage is set for a methodological insight of fundamental value. Over the course of the history of science it has often been said that a field of study is only as scientific as it is mathematical; Michell (1990: 3–8) documents an extensive collection of such assertions. Heidegger takes up Kant's particular expression of this theme and traces its origins (Heidegger, 1967: 68, 75) to the legend over the entrance to Plato's Academy: "Let no one untrained in geometry enter here." Heidegger explains at length that the point of Plato's prerequisite had nothing to do with a capacity for quantitative calculation, as is usually assumed.

Rather, in complete accord with the thesis of philosophy as the rigorous independence of figure from meaning, Plato stressed geometrical training because of its clear demarcation of the difference between the abstract whole to which something belongs and the concrete particulars representing it. Gadamer (1980: 100), paraphrasing Plato's *Republic* (510d, 527a-b) and Seventh Letter (342b), explains:

> Plainly no previous knowledge of the doctrine of ideas or of the dialectic of concepts is required to see that a circle is something different from the circular things which we call round, curved, oval, orbicular, etc., and which we can see with our eyes. It is clear to us that the figure which we draw to illustrate a mathematical relationship visually is not the mathematical relationship itself, and clearer still that the circular objects in nature are not to be confused with the circle of mathematics. . . .
>
> Geometry requires figures which we draw, but its object is the circle itself. . . . Even he who has not yet seen all the metaphysical implications of the concept of pure thinking but only grasps something of mathematics . . . knows that in a manner of speaking one looks right through the drawn circle and keeps the pure thought of the circle in mind.

Geometric figures then prepare students for philosophical study since they stand as models of "*all* those things which one can know through thought alone" (Gadamer, 1980: 101). Thus, Gadamer (1989: 412) says,

> we see that it is *not word but number* that is the real paradigm of the noetic: number, whose name is obviously pure convention and whose 'exactitude' consists in the fact that every number is defined by its place in the series, so that it is a pure structure of intelligibility, an *ens rationis*, not in the weak sense of a being-validity but in the strong sense of perfect rationality.

Mathematical rigor was "an indispensable preliminary to the study of philosophy" (Scott, 1960: 20; Gadamer, 1980: 101) not only for Plato, but for Husserl as well:

> The mathematical object seems to be the privileged example and most permanent thread guiding Husserl's reflection. . . . [on phenomenology]

> because the mathematical object is *ideal*. Its being is thoroughly transparent and exhausted by its phenomenality. (Derrida, 1989: 27)

It then follows that its "universality and objectivity make the ideal object into the 'absolute model for any object whatsoever'" (Bernet, 1989: 141; Derrida, 1989: 66).

In other words, numbers are justifiably "the most striking of always-already-knowns" because they are so thoroughly transparent: "numerical signs [are coordinated] with particular numbers, and they are the most ideal signs because their position in the order completely exhausts them" (Gadamer, 1989: 413). Although clarity in everyday speech necessarily relies on some basic degree of univocal sign-thing coordination, it was numeric figures' rigorous independence from the meaning they carry that became the metaphysical ideal, first in commerce, then for the basic astronomy needed for calendars, then for geometric figures, and then for the figures of any field that took itself to be scientific. In this broad sense, then, the mathematical is "the fundamental presupposition of all 'academic' work" (Heidegger, 1967: 76) and "of the knowledge of things" (Heidegger, 1967: 75).

What then is our criterion for telling sense from nonsense? As Gadamer (1980: 34; 1986: 102) puts it, "all logical confusion is a consequence of failing to distinguish and separate the *eidos* [the abstract ideal *logos* of a kind of thing] from what merely participates in it." In modern fashion, positivists, in effect, test the hypothesis of the *eidos* against its presumed empirical consequences, prioritizing the subjectivity of the investigator in an evaluation of the capacity of the *eidos* to account for the data comprised of particular observations.

In other words, modern, positivist methods presume that concepts can be defined in concrete terms alone, as constituted only in and of a defined set of actually existing examples. But, as Gadamer (1980: 33–34; 1986: 101–102) says,

> Such a procedure would be totally absurd in respect to a postulated *eidos*: that which constitutes being a horse could never be proved or disproved by a particular horse. Instead, the test which is to be applied in respect to the *eidos* is a test of the immanent, internal coherence of all that is intrinsic to it. One should go no further until one is clear about what the assumption of the *eidos* means and what it does not mean. It should be noted that consequently the hypothesis is not to be tested against presumed empirical consequences, but conversely the empirical consequences are to be tested against the hypothesis, i.e., that from the start everything empirical or accidental which the *eidos* does not mean and imply is to be excluded from consideration. This means above all that the particular which participates in an *eidos* is of importance in an argument *only* in regard to that in which it may be said to participate, i.e., only in regard to its eidetic content.

Play is of special importance in this context (Gadamer, 1989: 101–134). The positivist is dead serious and not at all playful in mechanically dissecting and assembling particulars, which, not surprisingly, cannot be brought to life. Positivism avoids references to metaphysical wholes but nonetheless inevitably makes such references. The problem then arises that one's metaphysics is implicit, not critically evaluated, and subtly transmitted to others via insinuations instead of via direct argument (Burtt, 1954: 229), all of which increase the risk of violent alienation or forced socialization.

The dialectician, in contrast, holds that nothing is serious and all is in play. In taking nothing seriously, the dialectician allows nothingness a role in the test of the *eidos*, since there is the chance that nothing will happen and the questioner's investment in a particular outcome of the test will be refuted. But when that refutation is not obtained, the Socratic midwife succeeds in facilitating the birth of an *eidos* as a form of life. As a living entity, the *eidos* is capable of creating audiences, and of referring to particular individuals participating in its mode of being, that were not imagined or intended by its author. In doing so, the *eidos* makes particulars mathematically coherent, in the sense of belonging together as parts of a whole while at the same time being separate from it.

Accordingly, "modern natural science, modern mathematics, and modern metaphysics sprang from the same root of the mathematical in the wider sense" (Heidegger, 1967: 97). That is, philosophy's metaphysical choice in favor of discourse over violence is a choice in favor of mathematical clarity, broadly conceived in the qualitative sense of facilitating a transparent view of the object of reference. The extent to which that clarity is obtained depends on the extent to which the particulars taken to represent a whole actually do. It may come as a surprise to many that mathematical clarity in the sciences is typically obtained via playful tests of the hypothesis of the *eidos*.

Mathematically Meaningful Constructions

When we read a clock, a ruler, a weight scale, a thermometer, or a voltmeter, we take advantage of a long history of two kinds of experiments (Mandel, 1978; Wernimont, 1978; Latour, 1987; O'Connell, 1993; Wise, 1995). The first kind, referred to by metrologists (measurement scientists) as intralaboratory ruggedness tests, establish a consistent, stable relationship between the structure of number and additive amounts of the thing measured within individual laboratories. In these experiments, the hypothesis of the *eidos* is tested in order to test the strength of a form of life as the extent to which a phenomenon consistently and repeatedly asserts its independent real existence, compelling agreement across investigators as to a capacity for consistent and unambiguous signification of its nonarbitrary properties. Mathematical theories of meaningfulness (Mundy, 1986; Narens, 2002) typically focus then on scale invariance as the primary criterion of a stable, constant expression of quantity.

The second kind of experiment, known as an interlaboratory round robin trial, is virtually unrecognized within any of the literature on measurement theory, which focuses nearly exclusively on mathematical criteria and local determinations of invariance. But interlaboratory trials are the means by which metrologists negotiate the arbitrary conventions of unit size and range, and data quality, through which they will compel agreement across expressions of the phenomenon in different laboratories. Given a viable *eidos*, the point now is to arrive at the signs by which we will know amounts of it when we see them anywhere and at any time, as has been achieved for variables such as time, temperature, length, etc.

Via this two-stage process, natural scientists arrive at universally available and accepted mathematical expressions capable of functioning as the language they think, act, and build in. Researchers in the social sciences, in contrast, generally accept numbers representing nonlinear, ordinal relations of unknown magnitude and consistency, and embodying interactions peculiar to particular samples. Numbers of unevaluated quality are commonly accepted even when statistical comparisons assume the existence of linear, interval relations of known magnitude and consistency that are free of sample-dependent interactions (Embretson, 1996; Stucki et al., 1996; Michell, 1990, 2000). Social scientists thus focus on common statistical procedures at the expense of unified metrics. Natural scientists, in contrast, focus more often on universal uniform measurement and the metaphysically astute process of experimental tests of the hypothesis that a variable is both qualitatively and quantitatively mathematical.

When failures of invariance emerge, as they often do, metrologists do not easily give up on calibrating a mathematically transparent measurement system. Instead, they investigate anomalies for what they might reveal about the way in which a view on the abstract ideal is clouded or distorted by the particulars of the instrument used, the person using it, the sample measured, the local environment, or another variable. This procedure is so pervasive that the history of science has been read as the history of approximations to abstract ideals (Pledge, 1939: 144). Errors are corrected, instruments are improved, theory changes, and new phenomena are discovered via these experiments that focus on achieving the closest possible generalizable referential connections between symbols and things.

The attitude informing these experiments accepts that numbers do not exist, pre-made, in the universe itself, and that the process of quantification in the natural sciences is neither as obvious nor as easy as many in the social sciences assume it to be. It is important to realize that, contrary to the popular perceptions of many in the humanities and social sciences, the history of natural science has many examples of productive qualitative investigations resulting in broadly mathematical insights (Kuhn, 1977; Heilbron, 1993; Roche, 1998), such as was required, for instance, in determining what such things as temperature or electrical resistance are, accomplishments nearly completely defined by the measurability of the constructs.

At the opposite, quantitative, extreme, also unrecognized for their full mathematical value, are generalized construct-specific metrics, such as the volt, the meter, degrees Celsius, etc. Just as much as is the case for the qualitative concept of mathematical clarity, these overtly mathematical universal uniform metrics are a consequence of science's mathematical metaphysics, as Heidegger (1967: 93) points out, and, as such, may be expected to emerge in the social sciences. The question as to how reference standard metrics might emerge is now taken up.

Situating Method and Theory

Gadamer (1989: 460–465) revives Hegel's sense of method as the activity of the thing itself that thought experiences, and compares the extent to which clear theoretical conceptualizations of things are developed in Plato and Aristotle (Gadamer, 1991, 1986, 1980). Insofar as science's success follows from phenomenologically and hermeneutically valid insights, metrology's two-stage process of metric calibration ought to be related to hermeneutic senses of method and concept formation.

What is meant by method as an activity of the thing itself is that method, in the root ancient Greek sense of *meta-odos*, is a following after (*meta*) the thing along its path (*odos*). When those willing to give themselves over to an experience of a conversational object's language game are carried along by the flow of the exchange, one question follows another, prompted by the always changing dynamic of the dialogue, seemingly of its own accord.

Something analogous to Socratic dialogue occurs in measurement efforts that are sensitive to the need to care for the unity and sameness of the thing being measured. These efforts are notable for the way in which the researchers give their hypothesis of an *eidos* over to the dialectical interplay of question and answer, and then check the internal coherence of the resulting observations. The checks on the observations constitute a deliberate and explicit following through on the prejudice or the guess that the questions asked and answers received do in fact participate in a single conversation, and are absorbed into a single, relatively univocal interpretation of the same thing. The extent to which a dialogue actually ensues with something real existing as its own independent form of life is the extent to which an appropriate medium for its manifestation has been provided to it via the research. "For implicit in the essential nature of all genuine method as a path toward the disclosure of objects is the tendency to order itself always toward that which it itself discloses" (Heidegger, 1982: 328). Hence, here we see the aptness of Gadamer's phrase construing method as an activity of the thing itself, and the real meaning of the motto of Husserlian phenomenology, "back to the things themselves."

To date, hermeneutical analyses of method and measurement have stopped short of explicitly pursuing the implications of academia's mathematical metaphysics for the social sciences in a way that recognizes and builds on commonalities with the natural sciences. Philosophers who mindfully attend to the hermeneutics of scientific instrumentation (Heelan, 1983; Ihde, 1991, 1998) do not follow these implications back into the social sciences. Conversely, in the social sciences, it is much more common for the implications of mathematical metaphysics not only to be left unstated and implicit within explicit hermeneutical concerns, but, in addition, a hermeneutical attitude is assumed to be needed only in the social, and not in the natural, sciences (Martin and Sugarman, 2001).

But the broader way in which method dominates the playful absorption of attention into the flowing activity of the things themselves was already noted by Nietzsche (1967, nos. 466, 469) and commented on by Heidegger (1971: 74). In fact, by definition, nowhere is poetic openness to the authentic activity of things themselves more apparent than in the process of scientific discovery, when a new phenomenon is revealed and enters language as a form of being that can be understood (Gerhart and Russell, 1984; Hallyn, 2000). Latour's (1987: 103–144) description of the situation brings out not only the real truth of Gadamer's sense of genuine method, but also recalls the original Greek sense of *theoria* as participatory involvement.

Latour construes the first phase of metrological experimentation as a process through which the discoverers of a phenomenon become enrolled as an audience for that phenomenon to the extent that the investigators' interests are captured. The phenomenon, and the technology through which it is manifest, become real, in Gadamer's sense of historically effective consciousness, to the extent that they enroll an audience of interested persons willing to commit resources to their production and application.

Latour (1987, 1999) and others (Ihde, 1991, 1998; O'Connell, 1993; Wise, 1995; Stengers, 2000) present relevant examples of various aspects of this process. In general, when a new effect shows itself to be stable across multiple independent experiments conducted in different laboratories using different instruments on different samples, it has enrolled its first audience by functioning as an *agent compelling agreement* as to its independent real existence. That audience, in turn, then collaborates together to turn the effect into a *product of agreement*, so as to facilitate the enrollment of other persons. The expansion of the audience requires that the initial audience be able to appeal to others' interests as persuasively as possible, and so is aided by the way in which full mathematization supports transparent reference to the new effect, its means of production, and its application to specific tasks.

Insofar as a phenomenon speaks to the interests of the investigators and enrolls them as an audience, the investigators behave with a kind of mindful passivity, giving themselves and their experimental equipment over to the phenomenon. Even when they collaborate among themselves to effect the "full and intimate . . . consummation" (Kuhn, 1961: 191) of the "true union of mathematics and measurement" (Roche, 1998: 145), the investigators still act on behalf of the thing itself and its entry into language as a form of life that can be clearly understood.

Just as genuine method is a matter of following along with the activity of the thing itself, genuine theory is rooted in the Greek *theoria*, as "true participation, not something active but something passive (*pathos*), namely being totally involved in and carried away by what one sees" (Gadamer, 1989: 125). As more individuals become involved in the production of any technologically produced phenomenon, such as fire, mechanical transportation, the measurement of time, temperature, or kilowatts, or the assessment of human performance, society invests more resources in its dissemination, so as to make it available wherever it is needed, and in readily interpretable form.

Latour (1987: 249) accordingly holds that every successful application of a science entails the progressive extension of a network of interconnected members of a phenomenon's audience whose interests are embodied by the thing itself. That the audience is carried away by its experience is expected insofar as Gadamer is correct in his sense of method and theory. Instruments, as readable technologies or inscription devices (Heelan, 1983; Ihde, 1991, 1998; Latour, 1987), require just the same kind of participatory involvement as that required for the understanding of texts. It then follows as much for an instrument-based text as for any other kind of text that "what is meaningful in it captivates us" (Gadamer, 1989: 490).

It must be understood that the passive involvement in the thing's self-representative activity "is in no way a matter of weakly allowing things to slide and drift along" (Heidegger, 1966: 61). Passivity in this sense is rather an openness to allowing one's horizon of experience to fuse with that of another form of life. There is a fundamental way in which the fusion of horizons (Gadamer, 1989) is constitutive of the meaning of a significant experience. The direction taken up by the arrow of meaning (Ricoeur, 1981: 193) is at its root the locutionary force with which the interplay of thing and thought goes of its own accord. Socratic dialogue and dialectical inquiry test the strength of the *eidos* by facilitating absorption into the flow of conversation. Such tests aim to achieve transparent mathematical representations, as is demonstrated, for instance, in the geometrical outcome of the dialogue with the slave boy in the *Meno*.

Initial scientific discoveries of new phenomena follow from an open attitude to unexpected anomalies and an associated ability to imagine what it is that unanticipated observations might point toward, and to metaphorically construe them as something (Kuhn, 1961, 1979). But the process of discovery only begins with the ability to recognize something new as having a potential value. Its significance as a discovery depends on following through from the first intuitions of what accidental or unexpected evidence might mean to the elaboration of that meaning in a distributed sign system.

Captivation in a charmed world is then an apt way of seeing metrology, as the means through which amounts of things come to be signified via universal uniform metrics. Caught up as they are in the play of ideas (Holton, 1988), refutation of the hypothesis of the *eidos* is taken by scientists as a sign of failure only after hundreds or perhaps thousands of trials. As Kuhn (1961: 171; 1977: 193) says, in this context, facts cease to be taken as given. Instead, scientists bent on demonstrating the existence of a particular thing struggle with facts, trying to make them conform to theoretical ideas and presupposed cultural values concerning what should be if anything is to be either qualitatively or quantitatively signifiable, in the sense of a mathematical convergence and separation of signifier and signified.

Researchers in the mainstream paradigm in the social sciences, in contrast, if they test the hypothesis of the *eidos* at all, typically take the first failure as conclusive. Mathematical models imposing the qualitatively mathematical requirements necessary for testing the hypothesis of the *eidos* are often viewed as inflexible (Hambleton and Swaminathan, 1985; van der Linden and Hambleton, 1997). These models are then often abandoned, along with the vigorous questioning of the hypothesized *eidos*, in favor of a statistical description of the data. In doing so, researchers cease to care for the unity and sameness of the thing investigated, and give up on a univocal interpretation that can be shared as a common language.

But when data are persistently questioned, and the hypothesis of the *eidos* is not refuted, multiple independent investigations of the same *eidos*, even when using different brands or configurations of instruments with different samples, will arrive at comparable results. Many may be surprised that this convergence occurs not only in the natural sciences, but also in the social sciences (Fisher, 1997a).

As will be shown in an example, these common results are expressed in the way that similar questions involving similar aspects of the matter of interest fall in similar places on the different tools calibrated, indicating the invariant stability, or the unity of an aspect, of something real. The first metrological phase of intralaboratory studies thus embodies Gadamer's sense of method as an activity of the thing itself. That is, research effectively provides a medium for a phenomenon's self-representation only when mathematical clarity is obtained across intralaboratory experiments. By making tests of the strength of an *eidos* central, metrologists accept Gadamer's requirement of not proceeding with interlaboratory research until they are clear about what the *eidos* means.

Fortunately, metrologists in the natural sciences learned in the first half of the 19th century (Roche, 1998) how to be clear about an *eidos*, and, also fortunately, they knew how to proceed from there. Measurement practitioners in the social sciences, however, are largely uninformed about metaphysically astute mathematical clarity. Those who have intuited some of what is required for approximating transparent reference have in fact produced results that constitute the realization of metrology's first, intralaboratory phase. These results are largely built on the work of Rasch (1960, 1961, 1977). Researchers tend to stop at this point, however, and remain negligent with respect to implementing metrology's second phase. There seems to be no comprehension of the importance or of the possibility of proceeding forward, generalizing locally clear expressions of the *eidos* into discipline-wide systems of shared signification.

The situation is remarkably similar to aspects of Gadamer's (1991: 4–8) contrast of Plato and Aristotle. Plato's dialogues repeatedly arrive at demonstrations of the way in which dialectical inquiry can test the hypothesis of the *eidos* and give birth to living ideas. None of the dialogues, however, ever take possession of that possibility as a general concept itself, but instead point "precisely, away from all supposed possession and toward the possibility of a possession which is always in store for it" (Gadamer, 1991: 6–7). Only in Aristotle is the invariant generality of conceptual work itself grasped and decisively possessed, making Aristotle "the originator of a philosophical science of ethics" and "the first theoretician" (Gadamer, 1991: 5).

Without being explicitly aware of the metaphysical mathematics generally structuring scientific research values, Rasch measurement practitioners nonetheless intuitively appreciate the mathematical beauty and inferential stability of their models, methods, and results. Like Plato's dialogues, their research reports are accounts of dialectical inquiries testing the hypothesis of the *eidos* and giving birth to living ideas. Also like Plato's dialogues, it is rare for any of these reports or subsequent applications of the research results to take possession of or deploy the full value obtained. Instead, like Plato's dialogues, they only point away from the process by which actual possession would be taken, and toward the possession that is always yet in store.

Despite repeated invitations (Fisher, 1997b, 2000), little enthusiasm has been shown for taking up the second phase of metrological research, through which a full grasp and application of the generality obtained would be achieved. There is plainly a need for a new theoretician, a metrological Aristotle, who would be the originator of a new philosophical science of ethics and measurement generalizable across all broadly mathematical disciplinary formations. And there is also a need for practical demonstrations of the viability and advantages of undertaking the extensive efforts that will be required.

Implications for Research in the Social Sciences

Discourse is inevitably metaphysical in never getting to the bottom of its presuppositions and in always risking nonsense. Discourse, actions, and their products are especially metaphysical to the extent that particulars are not examined as to their representativeness of the whole to which they are supposed to belong, and are simply assumed to represent it.

Research methods embodying integrated qualitative–quantitative mathematics have two crucial advantages over currently popular methods in the social sciences. The first is that strict criteria exist for knowing when mathematically transparent representation, and so, quantification, has been provisionally achieved. The second is the foundation this achievement lays for creating and maintaining universal uniform metrics as a common language for the exchange of value. Unfortunately, "vulnerability to falsification is commonly deemed by psychologists to be a fault rather than a virtue" (Michell, 1990: 130), even though the hermeneutic value of questioning depends on remaining open to unexpected answers: "the fruitfulness of scientific questioning is defined in an adequate manner if it is really open to answers in the sense that experience can refuse the anticipated confirmation" (Gadamer, 1981: 164).

In testing empirical consequences against a model specifying the ideal form that a construct would take in a perfect world, the measurement paradigm prescribes the requirements that must be met for data to come to life in a mode of being that can be nurtured and brought to maturity in the right environment. When test or survey questions are tested against a supposed *eidos* and some are found to not participate in this form of life, it often happens that substantive and theoretically compelling reasons for the refutation can be found in close examination of the question, the characteristics of those answering it in different ways, the administrative context, or other features of the situation.

Openness to refutation in the playful testing of the *eidos* occurs when psychological measurement proceeds in a metaphysically cognizant way. In this procedure, the particulars of data obtained as responses to questions supposedly pertaining to some one thing are evaluated for the extent to which they approximate an abstract, ideal, mathematical model of what a pure expression of that thing would be. To restate Gadamer's point on the order of things in the testing of the hypothesis of the *eidos*, models are not fit to data, but vice versa, data are fit to models. This difference has been a point of considerable controversy among measurement theoreticians and practitioners (Wright, 1984; Andrich, 2002; Fisher, 1994).

In contrast with mainstream contemporary practice in psychological measurement, the procedure does not begin with the attitude that a wide variety of models differing in the rigor of their qualitative and quantitative requirements are to be fit to data, with the model that best describes the data and all of its empirical and accidental vagaries chosen as the one with which to proceed. In this context, models are devised with the goal of describing the particular interactions and sample-dependencies characteristic of the data in hand. Such models are the tools used to achieve the positivist goal of defining the *eidos* as that which refers to a particular assemblage of unique, ephemeral, and irreproducible results. Because no effort is put into testing the observations in terms of the *eidos* and determining the extent to which a generalizable metrological concept emerges, the words and numbers describing the results remain tied to those results, and confusion is the inevitable consequence.

In order to mount a test of the immanent, internal coherence of all that is intrinsic to an *eidos*, we must first specify a model that is not mathematical in a merely quantitative sense but which is fully mathematical in requiring that the *eidos* be distinguished and separable from the particular questions and answers participating in it. This is the central point at which the vital, crucial importance of Rasch's (1960: 121–125; 1961: 325; 1977) analysis of "mutual conformity and separability," and his formulation of a separability theorem, becomes apparent.

Table I illustrates Rasch's notion of mutual conformity and separability. The data are a subsample of those shown in Table 2.4.1 in Wright and Stone's (1979: 33) analysis of the Knox Cube Test (KCT), a widely used assessment of attention span and short-term memory. Examinees observe and repeat sequences of taps on wooden blocks. As the number of taps, the number of different blocks tapped, and the reversals in the order of the taps increase, the difficulty of reproducing the pattern increases. Successful reproductions of the pattern presented are scored 1, and failures, 0.

Mutual conformity is evident in the conjoint order of persons and items in Table I. Convergence on a common object is illustrated in the table by the overlapping triangles of 1s and 0s, which shows an internally consistent pattern of success and failure. Easy items are the ones most likely to provoke a successful reproduction, no matter how many successes a person has. Difficult items are those least likely to be successfully reproduced, no matter how many successes are produced. Conversely, less able people are less likely to reproduce the series of taps, no matter how hard or easy the item is, and more able people are more likely to succeed, no matter which item is involved. The constancy of the joint relationship is an instance of the invariance of an *eidos*, since the relative values of the response probabilities do not change with the particular samples of persons or items involved. Following Gadamer (1991: 65), Table I illustrates a way in which everyone asking or answering KCT questions participates in testing the *eidos* of this form of cognitive functioning to see if it is refutable, and agrees with the failure to refute it.

Invariance follows from the mutual separability of the estimated person and item parameters. The basic questions asked in this "crucial experimental

Table I. Knox Cube Test data displaying conjoint order (selected from Wright and Stone, 1979: 33)

	Items (# taps/reversals; the number of blocks tapped is 4 for all items but 5, which had 3)														
Persons	5 (3/1)	7 (4/2)	9 (4/1)	8 (4/1)	10 (4/2)	11 (5/1)	13 (5/2)	12 (5/2)	14 (6/2)	15 (6/2)	16 (6/2)	17 (6/2)	Score	P[a]	Logit [b]
25	1	0	0	0	0	0	0	0	0	0	0	0	1	.08	−2.44
4	0	1	1	0	0	0	0	0	0	0	0	0	2	.17	−1.59
27	1	1	0	0	0	0	0	0	0	0	0	0	2	.17	−1.59
11	1	1	1	1	0	0	0	0	0	0	0	0	4	.34	−0.66
3	1	1	1	1	0	0	0	1	0	0	0	0	5	.42	−0.32
29	1	0	1	0	1	1	0	0	1	0	0	0	5	.42	−0.32
20	1	1	1	1	1	0	1	0	0	0	0	0	6	.50	0.00
34	1	1	1	1	1	0	0	1	1	0	0	0	7	.58	0.32
15	1	1	1	1	1	1	1	1	0	0	0	0	8	.67	0.71
7	1	1	1	1	1	1	1	1	0	1	0	0	9	.75	1.10
24	1	1	1	1	1	1	1	0	0	0	1	1	9	.75	1.10
Item															
Score	10	9	9	7	6	4	4	4	2	1	1	1			
P[a]	.91	.82	.82	.64	.55	.36	.36	.36	.18	.09	.09	.09			
Logit[b]	−2.31	−1.52	−1.52	−0.58	−0.20	0.58	0.58	0.58	1.52	2.31	2.31	2.31			

[a] P equals the observed score divided by the maximum possible score (12 for persons, 11 for items).
[b] The logit is the log-odds unit: the natural logarithm of the response odds, which are $P/(1 - P)$ for the persons, and $(1 - P)/P$ for the items.

test" (Thurstone, 1959: 228) are whether (1) the person measures would be reproduced by any sample of KCT items, and (2) the item calibrations would be reproduced by any sample of persons. For instance, how similar are the item difficulty calibrations shown in Table I with those given in Wright and Stone's Table 2.4.1? Given that the subsample in Table I comprises about one-third of the Wright and Stone data, a more stringent test of the invariance of the KCT items would be to ask about the invariance of the subsample estimates relative to estimates derived from an entirely different sample of data. Such data were gathered from 101 persons when new items had been added to the KCT in an effort to improve it (Wright and Stone, 1979: 89–93).

Table I subsample's item calibrations correlate .98 with the total sample calibrations, and .96 with the different sample calibrations. Figure 1 is a plot of the latter relationship, and shows the extent to which the items' order and positions on the variable have remained invariant over samples, and so also, the extent to which the *eidos* of a living conceptual entity has been found to separate from the contingencies of its origins and take on a life of its own. Paraphrasing Gadamer (1980: 100), we can say that psychology requires that we ask questions and record answers, but its object is the thing itself defined mathematically as that which remains invariant across the particulars serving as the

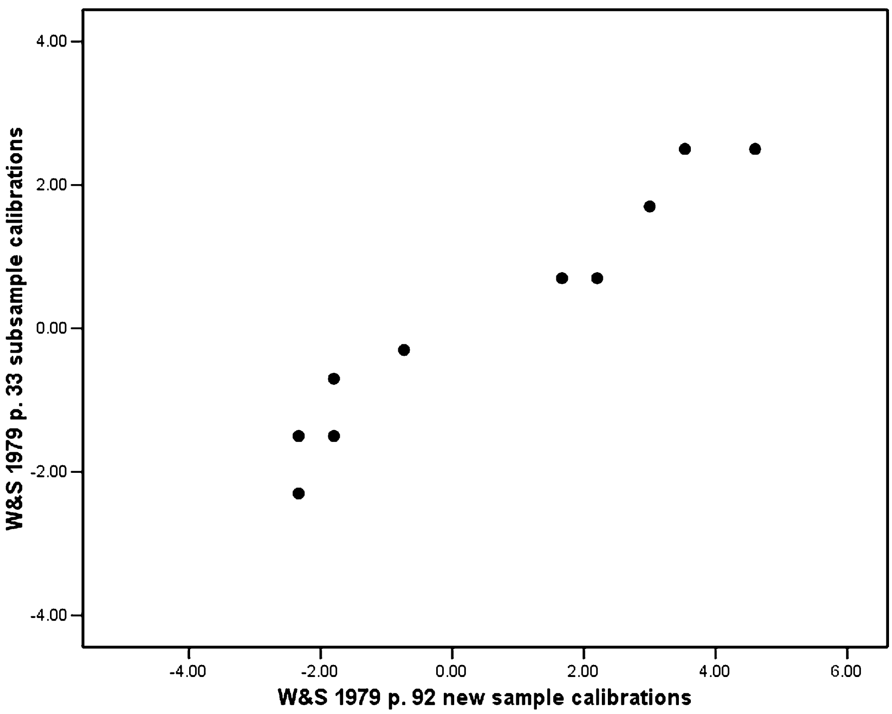

Figure 1. Knox Cube Test separate sample calibrations.

medium for the construct's self-representative play. In a manner of speaking, we need to devise techniques, analogous to geometrical figures, that enable us to look right through the particular questions asked and answered so as to keep the pure thought of the thing itself in mind. Repeated failures to falsify the hypothesis of the *eidos* for a particular variable provide evidence supporting the viability of a reference standard universal uniform metric for that variable.

We can look through the questions and answers in Table I, at the cognitive variable measured, to the extent that the scores function as sufficient statistics (Andersen, 1977). Scores are sufficient statistics to the extent that they extract all of the information available in the data. An object's measure of 24 cm on a meterstick, for instance, is a sufficient statistic because we can reproduce, from the score alone, the Yes/No pattern of responses to the question, "Is the object longer than this?" asked for every hash mark on not only the ruler actually used to make this measure, but any other one as well.

The scores shown in Table I are plainly not as mathematically transparent as centimeters, kilograms, or volts. Centimeters and inches, for instance, would fall in a very narrow straight line in a plot like that shown in Figure 1. Within the ranges along the measurement continuum where the probability of a correct response is 50/50, some play in the 1s and 0s is expected, and provides more information than is available in data lacking such play (Engelhard, 1994). More unexpected anomalous responses, where response probabilities are more or less than 50/50, open up opportunities for further qualitative investigation, as addressed below.

Were we to fill the cells in Table I with the observations expected on the basis of the overall pattern, we would have created a concrete image of the KCT attention span/short-term memory *eidos* as it bears on these items and people. Each score would then be associated with one unique set of responses and so would be a fully sufficient statistic. The comparison of these expected values with the observed values forms the basis for data quality evaluation and the estimation of model fit statistics (Smith, 2000). But because the scores are ordinal, qualitative, locally dependent pre-mathematical measures, and are not interval, quantitative, invariant mathematical measures, they support inferences and comparisons only with respect to rank order, not amount. In addition, the meaning of the scores varies depending on which questions are asked and of whom they are asked, when the point is to obtain transparent measures of amount that remain constant across samples of persons and items in a distributed system of signification.

The logits are measures capable of providing that transparency. A person with a KCT measure of 0.32, for instance, has an attention span enabling a high likelihood of succeeding on 3- and 4-tap items, a 50/50 chance on 5-tap items, and a low likelihood of success on 6-tap items, no matter which particular instances of those items are administered. Conversely, the order of the items is highly like to remain the same (within the range of error) no matter which sample of relevant persons is tested.

The relationships depicted in Table I and Figure 1 are not rare, isolated instances of very unusual phenomena. On the contrary, research shows that many instruments, such as the KCT, that were not designed with the intention of meeting the requirements of fundamental measurement theory for mutual conformity and separability, nonetheless do produce such data (Bond and Fox, 2001; Fisher et al., 1995, 1997; Fisher and Wright, 1994). In doing so, however, they also highlight the paucity of mathematically transparent substantive theory. And even when existing data do not provide the needed conjoint order and parameter separation, falsification of the hypothesis that the variable is broadly mathematical and specifically quantitative often provokes creative theorizing and modifications to the data, the ways in which the questions are asked, and/or the context in which they are asked, so that invariance is achieved.

For instance, consider the possibility of a hypothetical item in Table I that has a score of 2, but both of the 1s are at the top, as part of the scores achieved by persons 25 and 4. These two observations would be highly improbable, given the overall pattern of observations, and would stand out as the only 1s in a sea of 0s. Readily available software (Andrich et al., 2003; Linacre, 2003; Wu et al., 1998) provides model fit statistics that would flag these observations for further examination, since the measures and calibration based on the unexpected observations would mean something substantively different from what the rest of the measures and calibrations mean. The sensitivity of the fit statistics to situations in which “the thing signified is no longer easily separable from the signifier” (Derrida in Wood and Bernasconi, 1988: 88–89) illuminates dependencies on particulars that leave the *eidos* stillborn and opens the door to the deconstructive phase of research.

When failures of invariance such as this occur, care for the unity and sameness of the data text and its interpretation might lead the investigator to re-examine the data gathering process. This re-examination is an exploration of the possible questions to which the anomalous observations are answers, a destruction of the original research question motivated by the researchers’ persistent questioning. The “historical recursion to the tradition” (Heidegger, 1982: 23) of questions relevant in the given context is the means through which new meanings are appropriated, and an explanation, or at least a hypothetical conjecture, for the signifier-dependency is formulated. For instance, should examination of the original scoring sheets from the KCT administration reveal a data entry error, such that the 1s and 0s were reversed, the hypothetical item would then change positions in the order and its pattern would conform to the overall pattern. Other improbable patterns may occur, with other possible solutions (Bond and Fox, 2001: 173–186; Wright and Stone, 1979: 165–190).

When model fit statistics reveal failures of invariance, the hermeneutic persistence in pursuing the question compels the researcher to investigate and overcome barriers to understanding and inferential stability. Far from constituting an unethical tampering with data, correcting or omitting meaningless observations is commonplace in the natural sciences, as has been widely recognized at least since Kuhn's introduction of the concept of the paradigm and its influence on what counts as a legitimate observation. In contrast, the mainstream paradigm in the social sciences is positivistic in the sense of considering its mission one of describing the facts of test and survey data. The metaphysics of observation in the positivist framework deny the influence of the paradigm as a factor influencing which questions are asked, and so answers to test and survey questions are accepted at face value with limited recourse to resources for considering whether those answers and questions make sense as particulars participating in a common whole. Rasch's models have then been controversial because, instead of evaluating models representing an *eidos* in terms of their capacity to describe particular observations, they evaluate particular observations in terms of their capacity to participate in the *eidos* represented by a model (Wright, 1984; Fisher, 1994; Andrich, 2002).

The possibility that fundamental measurement theory and practice could contribute to a paradigm shift in the theory and methods of the social sciences is supported not only by its connection with the broadly mathematical metaphysics of meaning outlined here, but by at least six other factors.

- First, in their multiplicative form, Rasch's models have the same structure as the laws of natural science (Rasch, 1960: 109–120).
- Second, a wide variety of ways of formulating requirements for objectivity in measurement have been shown to reduce to Rasch models (Fischer, 1995; Wright, 1997).
- Third, the validity of the axioms of additive conjoint measurement, and of Rasch's separability theorem, have been demonstrated by the necessary mathematical proofs, as for instance by Suppes et al. (1989).
- Fourth, when (1) a ruler with unevenly spaced ordinal units (Fisher, 1988), (2) intuitively judged paired comparisons of weight (Choi, 1997), or (3) ratings of distance away (Moulton, 1993) are fit to a Rasch model, the resulting measures plot linearly with the respective centimeters, grams, and meters.
- Fifth, multiple independent studies of the same variable show that constructs have markedly robust broadly mathematical properties across investigations using different instruments with different respondents (Fisher, 1997a), different instruments with the same respondents (among many others, Fisher et al., 1995, 1997), or the same instrument with different respondents (among others, Fisher, 1997b, 1999).

- Sixth, hundreds of published studies employing Rasch's models populate the scientific literature in dozens of fields and demonstrate the versatility and applicability of the models relative to a wide variety of tests, surveys, and assessments.

Rasch-calibrated tests and surveys address a wide variety of constructs, such as the understanding of irony in poetry (Smith, 1990), learning-related attitudes and behaviors (Waugh, 2003), professional certification (Kelley and Schumacher, 1984), and health status and functionality assessments (Fisher, 1997a, 1997b, 1999; Fisher et al., 1995, 1997). The technical literature on measurement models (Wright and Mok, 2000), estimation methods (Linacre, 1999), instrument equating (Wolfe, 2000), and the statistical study of failures of invariance (Smith, 2000, 2002) is extensive. Resources on a broad array of relevant textbooks, journals, conferences, consultants, software, professional associations, and training seminars can be found at http://www.rasch.org.

Given the results obtained in the application of his models, there is little reason to doubt the veracity of Rasch's (1960: 115) contention that "the reading accuracy of a child . . . can be measured with the same kind of objectivity as we may tell its weight." Though those invested in positivist methods of fitting models to data may differ, the results already obtained via application of Rasch's models implies that the time for arguing theory is long past. But of urgent interest to those invested in the hermeneutic methods of fitting data to models are further analyses of the meaning of the new opportunities opened up by Rasch's implementations of fundamental measurement theory.

E Pluribus Unum

It seems reasonable to hope that we may effect a general improvement in the quality of research and applications in the social sciences to the extent that we more fully recognize and implement the meaning of the mathematical as "the fundamental presupposition of all 'academic' work," work that will remain unfinished "as long as we take ourselves seriously" (Heidegger, 1967: 76). Persisting in questioning the broad mathematical validity of qualitative and quantitative research alike is to persist in clarifying meaning, and in revealing biases and inconsistencies.

As with any bite from the apple of knowledge, such persistence creates as many new possibilities for fair and equitable opportunities for all as it does new opportunities for nonsensical oppression and violence. That is, even should we succeed in overcoming positivism's insufficient reductions and metaphysical blindness, there remains the constant danger

> of the systematic problem of philosophy itself: that the part of lived reality that can enter into the concept is always a flattened version – like every

> projection of a living bodily existence onto a surface. The gain in unambiguous comprehensibility and repeatable certainty is matched by a loss in stimulating multiplicity of meaning. (Gadamer, 1991: 7)

In other words, “all interpretation makes its object univocal and, by providing access to it, necessarily also obstructs access to it” (Gadamer, 1991: 8). This brings us to the central difficulty in paying attention to our nonsense, that, in choosing discourse over violence, the very act of trying to express the whole of an experience in words risks violence.

Husserl (1970) located the historical moment at which science lost its meaning for life in Galileo’s pragmatic leap to the utility of mathematical expressions whose justification and ground were left unarticulated. At the same instant that mathematization was achieved, the conditions of its possibility were erased by its very success. With a limited grasp of why or how science and technology work as they do, there have been few options for steering, guiding, selectively limiting and enhancing, or controlling their success in breeding ever more powerful forms of life and death.

The process is now culminating in genetic technologies that literalize the metaphor of forms of life, making it all the more imperative that we ask whether we might again experience the Pyrrhic victory of a new mathematization. Will we be able to pay sufficient attention to our broad mathematical goals while also diligently attending to the pre-mathematical particulars (Ballard, 1978: 186–190) from which mathematical ideals are abstracted?

If we are to live up to our historical standard of achieving unity from diversity, we must. It is encouraging that most software programs for implementing Rasch’s measurement models (Wright and Linacre, 2003; Andrich, et al., 2003; Wu, et al., 1998) typically provide more qualitative information on the pre-mathematical particulars than quantitative information on the unidimensional construct. That cannot lessen the fact that, following Gadamer (1980: 200), there would seem to be much to be gained from close consideration of the contrast between Plato’s insight into the nature of number, and Aristotle’s insight into the nature of what lives.

Notes

1. Thanks to Jack Stenner for his support of this work, and to many colleagues over the years for the engaging give and take leading to these thoughts. May we always keep the question open. Thanks also to two anonymous reviewers for helping to make this a better paper. As always, special thanks to Benjamin D. Wright.
2. Previous versions of this paper were presented at the meetings of the American Educational Research Association in Chicago in 2003, and of the American Psychological Association in Boston in 1999.

References

Andersen, E.B. (1977). Sufficient Statistics and Latent Trait Models. *Psychometrika* 42: 69–81.

Andrich, D.A. (1988). *Sage University Paper Series on Quantitative Applications in the Social Sciences. Vol. series no. 07-068: Rasch Models for Measurement*. Beverly Hills CA: Sage Publications.

Andrich, D.A. (2002). Understanding Resistance to the Data-Model Relationship in Rasch's Paradigm: A Reflection for the Next Generation. *Journal of Applied Measurement* 3: 325–359.

Andrich, D.A., Lyne, A., Sheridan, B. and Luo, G. (2003). *RUMM: Rasch Unidimensional Models for Measurement.* Perth, Australia: RUMM Laboratory Pty Ltd.

Ballard, E.G. (1978). *Man and Technology: Toward the Measurement of a Culture.* Pittsburgh: Duquesne University Press.

Bernet, R. (1989). On Derrida's 'Introduction' to Husserl's Origin of Geometry. In H.J. Silverman (Ed.), *Derrida and Deconstruction*. New York: Routledge.

Bernstein, R.J. (1983). *Beyond Objectivism and Relativism: Science, Hermeneutics, and Praxis.* Philadelphia: University of Pennsylvania Press.

Bond, T. and Fox, C. (2001). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. Mahwah NJ: LEA, Inc.

Burtt, E.A. (1932/1954). *The Metaphysical Foundations of Modern Physical Science*, rev. ed. Garden City, NY: Doubleday Anchor.

Caputo, J.D. (1997). A Commentary. In J.D. Caputo (Ed.), *Deconstruction in a Nutshell: A Conversation with Jacques Derrida*. New York: Fordham University Press.

Choi, S.E. (1997). Rasch Invents "Ounces". *Rasch Measurement Transactions* 11: 557 [http://www.rasch.org/rmt/rmt112.htm#Ounces].

Derrida, J. (1978). Structure, Sign and Play in the Discourse of the Human Sciences. In *Writing and Difference*. Chicago: University of Chicago Press.

Derrida, J. (1982). *Margins of Philosophy*. Chicago: University of Chicago Press.

Derrida, J. (1989). *Edmund Husserl's Origin of Geometry: An Introduction*. Lincoln: University of Nebraska Press.

Derrida, J. (2003). Interview on Writing. In G.A. Olson and L. Worsham (Eds.), *Critical Intellectuals on Writing*. Albany: State University of New York Press.

Embretson, S.E. (1996, September). Item Response Theory Models and Spurious Interaction Effects in Factorial ANOVA Designs. *Applied Psychological Measurement* 20: 201–212.

Engelhard, G., Jr. (1994). Resolving the Attenuation Paradox. *Rasch Measurement Transactions* 8: 379.

Fischer, G.H. (1995). Derivations of the Rasch model. In G. Fischer and I. Molenaar (Eds.), *Rasch Models: Foundations, Recent Developments, and Applications*. New York: Springer-Verlag.

Fisher, W.P., Jr. (1988). Truth, Method, and Measurement: The Hermeneutic of Instrumentation and the Rasch model [diss]. *Dissertation Abstracts International*, 49, 0778A, Department of Education, Division of the Social Sciences: University of Chicago (376 pages, 23 figures, 31 tables).

Fisher, W.P., Jr. (1994). The Rasch Debate: Validity and Revolution in Educational Measurement. In M. Wilson (Ed.), *Objective Measurement: Theory into Practice*, Vol. II. Norwood, NJ: Ablex Publishing Corporation.

Fisher, W.P., Jr. (1997a). Physical Disability Construct Convergence Across Instruments: Towards a Universal Metric. *Journal of Outcome Measurement* 1: 87–113.

Fisher, W.P., Jr. (1997b, June). What Scale-Free Measurement Means to Health Outcomes Research. *Physical Medicine & Rehabilitation State of the Art Review* 11: 357–373.

Fisher, W.P., Jr. (1999). Foundations for Health Status Metrology: The Stability of MOS SF-36 PF-10 Calibrations Across Samples. *Journal of the Louisiana State Medical Society* 151: 566–578.

Fisher, W.P., Jr. (2000). Objectivity in psychosocial measurement: What, Why, How. *Journal of Outcome Measurement* 4: 527–563.

Fisher, W.P., Jr. (2003a). The Mathematical Metaphysics of Measurement and Metrology: Toward Meaningful Quantification in the Human Sciences. In A. Morales (Ed.), *Renascent Pragmatism: Studies in Law and Social Science*. Burlington, VT: Ashgate Publishing Co.

Fisher, W.P., Jr. (2003b, December). Mathematics, Measurement, Metaphor, Metaphysics: Part I. Implications for Method in Postmodern Science. *Theory & Psychology* 13: 753–790.

Fisher, W.P., Jr., Eubanks, R.L. and Marier, R.L. (1997). Equating the MOS SF36 and the LSU HSI Physical Functioning Scales. *Journal of Outcome Measurement* 1: 329–362.

Fisher, W.P., Jr., Harvey, R.F., Taylor, P., Kilgore, K.M. and Kelly, C.K. (1995, February). Rehabits: A common Language of Functional Assessment. *Archives of Physical Medicine and Rehabilitation* 76: 113–122.

Fisher, W.P., Jr. and Wright, B.D. (1994). Introduction to Probabilistic Conjoint Measurement theory and Applications. *International Journal of Educational Research* 21: 559–568.

Gadamer, H.-G. (1980). *Dialogue and Dialectic: Eight Hermeneutical Studies on Plato*. Trans. P.C. Smith. New Haven, CT: Yale University Press.

Gadamer, H.-G. (1981). *Reason in the Age of Science.* Trans. F.G. Lawrence. Cambridge, MA: MIT Press.

Gadamer, H.-G. (1986). *The Idea of the Good in Platonic–Aristotelian Philosophy*. Trans. P.C. Smith. New Haven: CT: Yale University Press.

Gadamer, H.-G. (1989). *Truth and Method* (rev. ed.). Trans. J. Weinsheimer and D.G. Marshall. New York: Crossroad.

Gadamer, H.-G. (1991). *Plato's Dialectical Ethics: Phenomenological Interpretations Relating to the Philebus.* Trans. R.M. Wallace. New Haven: CT: Yale University Press.

Gerhart, M. and Russell, A. (1984). *Metaphoric Process: The Creation of Scientific and Religious Understanding*, Foreword by Paul Ricoeur. Fort Worth: Texas Christian University Press.

Hallyn, F. (Ed.). (2000). *Metaphor and Analogy in the Sciences*. Hingham, MA: Kluwer.

Hambleton, R.K. and Swaminathan, H. (1985). *Item Response Theory: Principles and Applications*. Hingham MA: Kluwer.

Heelan, P. (1983, June). Natural Science as a Hermeneutic of Instrumentation. *Philosophy of Science* 50: 181–204.

Heidegger, M. (1966). *Discourse on Thinking: A Translation of Gelassenheit*. Trans. J.M. Anderson and E.H. Freund. New York: Harper & Row.

Heidegger, M. (1967). *What is a Thing?* Trans. W.B. Barton, Jr. and V. Deutsch. South Bend, IN: Regnery/Gateway.

Heidegger, M. (1971). *On the Way to Language*. Trans. P.D. Hertz. New York: Harper & Row.

Heidegger, M. (1982). *The Basic Problems of Phenomenology*. Ed. J.M. Edie, Trans. A. Hofstadter. Bloomington: Indiana University Press.

Heilbron, J.L. (1993). Weighing Imponderables and Other Quantitative Science Around 1800. *Historical Studies in the Physical and Biological Sciences* (Vol. 24 (Supplement), Part I, pp. 1–337). Berkeley: University of California Press.

Holm, K. and Kavanagh, J. (1985). An Approach to Modifying Self-Report Instruments. *Research in Nursing and Health* 8: 13–18.

Holton, G. (1988). *Thematic Origins of Scientific Thought*, rev. ed. Cambridge, MA: Harvard University Press.

Husserl, E. (1970). *The Crisis of European Sciences and Transcendental Phenomenology*. Evanston, IL: Northwestern University Press.

Ihde, D. (1991). *Instrumental Realism: The Interface Between Philosophy of Science and Philosophy of Technology*. Bloomington: Indiana University Press.

Ihde, D. (1998). *Expanding Hermeneutics: Visualism in Science*. Evanston, IL: Northwestern University Press.

Kearney, R. (1984). *Dialogues with Contemporary Continental thinkers: The Phenomenological Heritage*. Manchester: Manchester University Press.

Kelley, P.R. and Schumacher, C.F. (1984, December). The Rasch Model: Its Use by the National Board of Medical Examiners. In *Evaluation and the Health Professions*.

Kuhn, T.S. (1961). The Function of Measurement in Modern Physical Science. *Isis* 52(168): 161–193. (Rpt. in *The Essential Tension: Selected Studies in Scientific Tradition and Change*. Chicago: University of Chicago Press).

Kuhn, T.S. (1977). *The Essential Tension: Selected Studies in Scientific Tradition and Change*. Chicago: University of Chicago Press.

Kuhn, T.S. (1979). Metaphor in Science. In A. Ortony (Ed.), *Metaphor and Thought*. New York: Cambridge University Press.

Latour, B. (1987). *Science in Action: How to Follow Scientists and Engineers Through Society*. New York: Cambridge University Press.

Latour, B. (1999). *Pandora's Hope: Essays on the Reality of Science Studies*. Cambridge, MA: Harvard University Press.

Linacre, J.M. (1999). Understanding Rasch measurement: Estimation methods for Rasch measures. *Journal of Outcome Measurement* 3: 382–405.

Linacre, J.M. (2003). *A User's Guide to WINSTEPS Rasch-Model Computer Program, v. 3.38*. Chicago: Winsteps.com.

Mandel, J. (1978, December). Interlaboratory testing. *ASTM Standardization News* 6: 11–12.

Martin, J. and Sugarman, J. (2001, April). Interpreting human kinds: Beginnings of a hermeneutic psychology. *Theory & Psychology* 11: 193–207.

Michell, J. (1990). *An Introduction to the Logic of Psychological Measurement*. Mahwah, NJ: LEA, Inc.

Michell, J. (2000, October). Normal Science, Pathological Science and Psychometrics. *Theory & Psychology* 10: 639–667.

Moulton, M. (1993). Probabilistic mapping. *Rasch Measurement Transactions* 7(1): 268.

Mundy, B. (1986). On the General Theory of Meaningful Representation. *Synthese* 67: 391–437.

Narens, L. (2002). *Theories of Meaningfulness*. Mahwah, NJ: LEA, Inc.

Nietzsche, F. (1967). *The Will to Power*. New York: Vintage.

O'Connell, J. (1993). Metrology: The Creation of Universality by the Circulation of Particulars. *Social Studies of Science* 23: 129–173.

Page, C. (1987, March). Axiomatics, hermeneutics, and practical rationality. *International Philosophical Quarterly* XXVII(1): 81–100.

Pledge, H.T. (1939). *Science Since 1500: A Short History of Mathematics, Physics, Chemistry, Biology*. London: His Majesty's Stationery Office.

Rasch, G. (1960). *Probabilistic Models for Some Intelligence and Attainment Tests* (reprint, with Foreword and Afterword by B.D. Wright, Chicago: University of Chicago Press, 1980). Copenhagen: Danmarks Paedogogiske Institut.

Rasch, G. (1961). On General Laws and the Meaning of Measurement in Psychology. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley: University of California Press.

Rasch, G. (1977). On Specific Objectivity: An Attempt at Formalizing the Request for Generality and Validity of Scientific Statements. *Danish Yearbook of Philosophy* 14: 58–94.

Ricoeur, P. (1974). Violence and Language. In D. Stewart and J. Bien (Eds.), *Political and Social Essays by Paul Ricoeur*. Athens: Ohio University Press.

Ricoeur, P. (1977). *The Rule of Metaphor: Multi-Disciplinary Studies of the Creation of Meaning in Language*. Trans. Robert Czerny. Toronto: University of Toronto Press.

Ricoeur, P. (1981). *Hermeneutics and the Human Sciences: Essays on Language, Action and Interpretation*. Trans. J.B. Thompson. New York: Cambridge University Press.

Risser, J. (1989). The Two Faces of Socrates: Gadamer/Derrida. In D.P. Michelfelder and R.E. Palmer (Eds.), *Dialogue & Deconstruction: The Gadamer–Derrida Encounter*. Albany: State University of New York Press.

Roche, J. (1998). *The Mathematics of Measurement: A Critical History*. London: Athlone Press.

Scott, J.F. (1960). *A History of Mathematics*. London: Taylor & Francis.

Smith, E.V., Jr. (2002). Detecting and Evaluating the Impact of Multidimensionality Using Item Fit Statistics and Principal Component Analysis of Residuals. *Journal of Applied Measurement* 3: 205–231.

Smith, M.W. (1990). Understanding of Irony in Poetry. *Rasch Measurement Transactions* 4: 89–91.

Smith, R.M. (2000). Fit Analysis in Latent Trait Measurement Models. *Journal of Applied Measurement* 1: 199–218.

Stengers, I. (2000). *Theory Out of Bounds. Vol. 19: The Invention of Modern Science*. Trans. D.W. Smith. Minneapolis: University of Minnesota Press.

Stucki, G., Daltroy, L., Katz, N., Johannesson, M. and Liang, M.H. (1996). Interpretation of Change Scores in Ordinal Clinical Scales and Health Status Measures: The Whole May Not Equal the Sum of the Parts. *Journal of Clinical Epidemiology* 49: 711–717.

Suppes, P., Krantz, D.H., Luce, R.D. and Tversky, A. (1989). *Foundations of Measurement, Volume II: Geometric and Probabilistic Representations*. New York: Academic Press.

Thurstone, L.L. (1959). *The Measurement of Values*. Chicago: University of Chicago Press.

van der Linden, W.J. and Hambleton, R.K. (1997). Item Response Theory: A Brief History. In W.J. van der Linden and R.K. Hambleton (Eds.), *Handbook of Modern Item Response Theory (IRT)*. New York: Springer-Verlag.

Waugh, R.F. (2003). Measuring Attitudes and Behaviors to Studying and Learning for University Students: A Rasch Measurement Model Analysis. *Journal of Applied Measurement* 4: 164–180.

Wernimont, G. (1978, December). Careful Intralaboratory Study Must Come First. *ASTM Standardization News* 6: 11–12.

Wilshire, B. (1982). *Role Playing and Identity: The Limits of Theatre as Metaphor*. Bloomington: Indiana University Press.

Wise, M.N. (Ed.). (1995). *The Values of Precision*. Princeton: Princeton University Press.

Wittgenstein, L. (1980). *Culture and Value*. Ed. G.H. von Wright, in collaboration with H. Nyman, Trans. B. Winch. Chicago: University of Chicago Press.

Wolfe, E.W. (2000). Equating and Item Banking With the Rasch Model. *Journal of Applied Measurement* 1: 409–434.

Wood, D. and Bernasconi, R. (1988). *Derrida and Différance*. Evanston, IL: Northwestern University Press.

Wright, B.D. (1984). Despair and Hope for Educational Measurement. *Contemporary Education Review* 3: 281–288.

Wright, B.D. (1997, Winter). A History of Social Science Measurement. *Educational Measurement: Issues and Practice* 16: 33–45, 52.

Wright, B.D. and Mok, M. (2000). Understanding Rasch Measurement: Rasch Models Overview. *Journal of Applied Measurement* 1: 83–106.

Wright, B.D. and Stone, M.H. (1979). *Best Test Design: Rasch Measurement*. Chicago: MESA Press.

Wu, M.L., Adams, R.J. and Wilson, M.R. (1998). *ConQuest: Generalized Item Response Modelling Software*. Camberwell, Victoria, Australia: Australian Council for Educational Research.