Geometry, Maxwell, and Rasch:

On the Potential for Improved Measurement in the Psychosocial Sciences

William P. Fisher, Jr.

Fifth International Conference on Probabilistic Models for Measurement in Education, Psychology, Social Science and Health

23 January 2012 to 25 January 2012

Perth, Western Australia

Abstract

Geometry was held by Einstein to be the most ancient branch of physics, with all linear measurement essentially a form of practical geometry. Rasch's background in physics and mathematics informed his sense of measurement models as analogies from the structure of scientific laws, following Maxwell's method of drawing analogies from geometry as a basis for model-based reasoning. Rasch likely absorbed Maxwell's method via close and prolonged interactions with colleagues known for their use of it. Examination of the form of the relationships posited in the Pythagorean theorem, multiplicative natural laws, and Rasch models leads to a new geometrical representation of Rasch model parameter estimates. Rasch measurement conforms with a geometric model of measurement in eight ways. For their potential to be fully realized in the social sciences, Rasch's measurement ideas need to be dissociated from statistics and IRT, and instead rooted in the Maxwellian sources Rasch actually drew from. Following through on the method of analogy to geometric imagery, substantive construct models, and universally uniform unit definitions may make psychosocial measurement more intuitive and useful.

Standardized units of measurement used in science and commerce are manifestations of objective and invariant phenomena, but they are also arbitrary social conventions. Beyond the need to be consistent, convenient, memorable, and identifiable, unit sizes, ranges, and names matter little in relation to their function. But over 2,000 years passed from the origins of formalized geometry in ancient Greece to the establishment of widely uniform standards for measurement in the late eighteenth and early nineteenth centuries, and the full power of science was realized only in the wake of those standards. There is urgent need, then, to pay close attention to both the geometrically nonarbitrary and the socially arbitrary aspects of measurement and the definition of universally uniform units.

All linear measurement makes use of the geometric figure of the line. Quantitative comparisons automatically bring images of a number line to mind. Despite these associations, most statistical analysis in the social sciences does little or nothing to justify the assumption that any given numeric difference is a constant unit amount. Further, close examination of most data sets in the social sciences would show that this assumption is not justified, and is in fact contradicted by the data. So how has the geometry of lines and other figures informed measurement in the natural sciences? What aspects of geometry make it a satisfactory root metaphor for measurement? How are those features of geometry included in Rasch's approach to measurement? What motivated Rasch to formulate his models in accord with a geometrical conceptualization? What remains to be done in following through on what Rasch started? How might the role of practical geometry as source of analogies for linear measurement in the natural sciences be extended to the social sciences? In pursuing these questions, Rasch's work is seen to owe far more to physics and mathematics than to statistics and item response theory. Resituating Rasch in the history of ideas is an important step in properly framing the horizons of method and analysis in the social sciences, especially with regard to integrating those sciences with social and economic projects outside the laboratory (Hunt, 1994; Miller & O'Leary, 2007; Schaffer, 1992).

Linear Measurement as Practical Geometry

In the natural sciences, the basis for quantitative units is established, in effect, via analogies from geometry. The Pythagoreans considered tonal proportions to be the geometry of motion, encompassing sound, celestial bodies, and the human soul in a comprehensive cosmology (Isacoff, 2001, p. 38). The essential question for Copernicus was not "Does the earth move?" but, rather, "...what motions should we attribute to the earth in order to obtain the simplest and most harmonious geometry of the heavens that will accord with the facts?" (Burtt, 1954, p. 39). Both Boscovich and Legendre based their contributions to the method of least squares in geometrical formulations (Stigler, 1986, pp. 42, 46, 47, 57). Galileo "derived his rule relating time and distance using geometry" (Heilbron, 1998, p. 129). Einstein (1922) considered geometry to be "the most ancient branch of physics," according "special importance" to his view that "all linear measurement in physics is practical geometry," "because without it I should have been unable to formulate the theory of relativity" (p. 14). Pledge (1939) makes the connection in the general point that

> as the Greeks gave us the abstract ideas (point, line, etc.) with which to think of space, and the 17th century those (mass, acceleration, etc.) with which to think of mechanics, so Carnot gave us those needed in thinking of heat engines. In each case the ideas are so pervasive that we use them even to state that they never apply exactly to visible objects (p. 144).

Narens (2002) explicitly roots measurement theory in a Pythagorean sense of scientific definability focused on meaningfulness as invariance across transformations. Maxwell provides the clearest method for making any instance of linear measurement analogous with practical geometry (Black, 1962; Nersessian, 2002; Turner, 1955). Inventing the contemporary concept of mathematical modeling (Hesse, 1961, p. 206), Maxwell freed physics from the constraints of Newtonian mechanics via his concept of the abstract mathematical field (Rautio, 2005, p. 53; McMullin, 2002), and his work still stands as one of the most productive examples of the method of drawing geometric analogies of phenomena (Klein, 1974, p. 474; Rautio, 2005).

What exactly did Maxwell do? To understand his method of analogy, it is important to know that, in the eighteenth and nineteenth centuries, scientists and philosophers in many fields had been employing Newton's laws of motion as a framework for structuring investigations of a wide range of different phenomena. Newton's theory of gravitation provided the form of a Standard Model adopted across the sciences of nature, and moral philosophy, as well (Myers, 1983, pp. 65-75), as the hallmark criterion of scientific success.

> Beginning around 1770…electricity, magnetism, and heat began to yield to the sort of analysis that had ordered the motions of the planets; and just after the turn of the 19th century, the phenomena of capillarity and the behavior of light were brought into the scheme…. These achievements inspired and exemplified the program described by Laplace in 1796 and brought almost to realization (or so he thought) by Gay-Lussac in 1809: to perfect terrestrial physics by the same techniques as Newton had used to perfect celestial mechanics (Heilbron, 1993, pp. 5-6).

Nersessian (2002) concurs, saying "After Newton, the inverse-square-law model of gravitational force served as a generic model of action-at-a-distance forces for those who tried to bring all forces into the scope of Newtonian mechanics" (p. 139). Maxwell learned of the method of analogy from his colleague William Thomson (Lord Kelvin), and told him that he “intended to borrow it for a season...but applying it in a somewhat different way” (Larmor, 1937, pp. 17-18; see Nersessian, 2002, p. 144; Cropper, 2001, p. 161). The difference between Thomson’s method and Maxwell’s use of it is telling. Like Maxwell, Thomson constructed a number of analogies, such as between heat and electrostatics. But Thomson merely took existing equations describing a known physical system and substituted their parameters for the system under investigation (Nersessian, 2002, p. 144), as this seems to have been the typical way in which the Standard Model had been applied in research up to that time.

The superficiality of this method of analogy, however, would seem vulnerable to both of the errors that Maxwell (1965/1890, p. 155) sought to avoid, distraction by abstract mathematical analyses and by too-literal preconceptions of the physical phenomenon. As Maxwell put it,

> By referring everything to the purely geometrical idea of the motion of an imaginary fluid, I hope to attain generality and precision, and to avoid the dangers arising from a premature theory professing to explain the cause of the phenomena. If the results of mere speculation which I have collected are found to be of any use to experimental philosophers, in arranging and interpreting their results, they will have served their purpose, and a mature theory, in which physical facts will be physically explained, will be formed by those who by interrogating Nature herself can obtain the only true solution of the questions which the mathematical theory suggests (Maxwell, 1965/1890, p. 159).

Maxwell, then, started from simple geometric ideas and did not presume to know the relevant mathematical structure to be applied, instead constructing the source of the analogy to be mapped onto the object of investigation. In so doing, he provided “the prototype for all the great triumphs of twentieth-century physics” (Dyson, in Rautio, 2005, p. 53).

Immersed as he was for years in an intellectual milieu known for employing Maxwell’s method of analogy (Fisher, 2012b), Rasch (1960, pp. 110-115) would seem to have supposed that a basis for a similar prototype for the social sciences could be provided by deliberately structuring his models in the pattern of Maxwell’s analysis of mass, force, and acceleration. Few researchers to date, however, have found reason to note or expand upon the connection Rasch drew between his models and Maxwell’s analysis. Further, despite being focused on individual response patterns, Rasch models are routinely assumed to be statistical models of group-level aggregate patterns and, as such, to have the primary purpose of guiding data analyses. The emphasis on data analysis corresponds to de-emphasis on the construct, mapping substantive unit amounts on a number line, and the importance of defining, disseminating, and maintaining standardized units.

Thus, Rasch models are typically assumed to be data or response process models and not abstract conceptualizations of living psychological and social events, processes, and relationships. Rasch measurement practice is then almost completely defined by abstract mathematical analyses. This is exactly what Maxwell (1965/1890, p. 155) considered a distraction, saying purely mathematical simplifications are likely to cause the investigator “entirely lose sight of the phenomena to be explained; and though we may trace out the consequences of given laws, we can never obtain more extended views of the connexions of the subject.” Conversely, far from being dominated by too-literal preconceptions of the constructs studied, in practice little attention is paid to modeling constructs, though there are several significant exceptions (Dawson, Fischer, & Stein, 2006; Embretson & Daniel, 2008; Stenner, Burdick, Sanford, & Burdick, 2006; Wilson, 2005, 2008) that take up the challenge in ways roughly analogous to the approach advocated by Maxwell, as clearly psychosocial conceptions focused on psychosocial explanations of psychosocial facts.

Rasch models are more commonly employed, however, in the manner of Thomson's method of merely substituting parameter names across the different phenomena studied. The questions then arise as to if and how a shift from Thomson's method to Maxwell's might be achieved. Significant untapped potential for such a shift can be found in two sources. The first focuses on incorporating purely geometric images of constructs at the start of the modeling process. This is facilitated by the shared mathematical formalism of the Pythagorean theorem, the multiplicative structure of natural laws, and Rasch models. The second follows from closer study of Maxwell's cognitive modeling process, following Nersessian (2002, 2006, 2008), and its extension into predictive Rasch construct models.

Geometry as a Model for Measurement

Maxwell employed geometrical images as means of solving problems that stumped those employing analytic methods (Forfar, 2002, p. 8). And though the method of least squares is foundational to contemporary statistical analysis, it was originally formulated by Boscovich, who "followed in a Newtonian tradition of giving geometric descriptions rather than analytic ones," and whose work was only later expressed analytically, by Laplace (Stigler, 1986, pp. 42-43, 51). In accord with Maxwell's later use of geometry, "Boscovich's geometric approach suggested a solution to the problem that would have been far less apparent in an analytic formulation"

(Stigler, 1986, p. 47). Perhaps similar advantages for psychosocial measurement can be found in a geometric approach.

Figure 1 about here.

Figure 1 illustrates a proof of the Pythagorean theorem, where the square of the hypotenuse of a right triangle is equal to the sum of the squares of the other two sides:

$$a^2 + b^2 = c^2$$

which, for Figure 1, is

$$3^2 + 4^2 = 5^2 = 9 + 16 = 25$$

Most scientific laws are not written in this additive form (which also includes equations involving subtraction), but in a multiplicative form (which also includes equations involving division) (Crease, 2004; Taagepera, 2008; Burdick, Stone, & Stenner, 2006), like this:

$$a = f / m$$

or

$$f = m * a$$

where the acceleration of an object can be estimated by dividing the applied force by the object's mass, or the force is estimated by multiplying the mass by the acceleration. This, of course, is how Maxwell (1920/1876) presented Newton's Second Law.

Other geometric relationships do take the multiplicative form of a scientific law, such as the definition of the circle as a closed arc equidistant from a single point, with the circumference equal to pi times the radius squared. The Pythagorean theorem can also be written in the form of a multiplicative law, by means of the number *e* (2.71828...) (Maor, 1994):

$$e^9 * e^{16} = e^{25}$$

Substituting a for e^9, b for e^{16}, and c for e^{25}, this could be represented by

$$a * b = c$$

and could be solved as

$$8103 * 8{,}886{,}015 \approx 72{,}003{,}378{,}611$$

Converting back to the additive form using the natural logarithm, the equation looks like this:

$$\ln(8{,}103) = \ln(72{,}003{,}378{,}611) - \ln(8{,}886{,}015)$$

and this

$$9 = 25 - 16 .$$

Whether expressed in multiplicative or additive forms, Newton’s Second Law and the Pythagorean theorem both define the way changes in one parameter in a mathematical model result in proportionate changes in the other parameters.

Furthermore, the empirical relational structure stays the same no matter what unit characterizes the numerical relational structure. Maxwell presented Newton’s Second Law in this form:

$$A_{vj} = F_j / M_v .$$

So when catapult j's force F of 7.389 Newtons (53.445 poundals) is applied to object v's mass M of 1.6487 kilograms (3.635 pounds), the acceleration of this interaction is 4.4817 meters (14.70 feet) per second, per second. (That is, 7.389 / 1.6487 = 4.4817, or 53.445 / 3.635 ≈ 14.70). The proportional relationships are constant no matter which units are used, satisfying the criterion of meaningfulness (Mundy, 1986; Narens, 2002; Rasch, 1961). In this context, Rasch (1960, 112-113; Burdick, et al., 2006) noted that,

> If for any two objects we find a certain ratio of their accelerations produced by one instrument, then the same ratio will be found for any other of the instruments. Or, in a slightly mathematized form: The accelerations are proportional.

Conversely, it is true that if for any two instruments we find a certain ratio of the accelerations produced for one object, then the same ratio will be found for any other objects.

Citing Maxwell's presentation of Newton's Second Law as the source for the mathematical form he sought (Rasch, 1960, pp. 110-115), Rasch (1961, p. 322) then wrote his model for measuring reading ability and text reading difficulty in the multiplicative form of

$$\varepsilon_{vi} = \theta_v \sigma_i$$

and also (Rasch, 1961, p. 333) in the additive form

$$\varepsilon_{vi} = \theta_{v} + \sigma_i \, .$$

These forms of the model assert that reading comprehension ε is the product (or the sum) of person v's reading ability θ and item i's text complexity σ. The model is also often written in the equivalent forms of

$$\Pr\{X_{ni} = 1\} = e^{\beta n - \delta i} / 1 + e^{\beta n - \delta i}$$

or

$$P_{ni} = \exp(B_n - D_i) / [1 + \exp(B_n - D_i)]$$

or

$$\ln[P_{ni} / (1-P_{ni})] = B_n - D_i$$

which all effectively say that the log-odds of a correct response from person *n* on item *i* is equal to the difference between the estimate B of person *n*'s ability and the estimate D of item *i*'s difficulty. Moving the effect of *e* from one side of the equation to the other makes the response odds equal to *e* taken to the power of the difference between B and D, divided by one plus *e* to that power.

In light of the proportionality obtained in these relationships, Rasch (Rasch, 1960; also see his 1961, p. 325) formulated a separability theorem in terms that apply to both additive and multiplicative forms of the models, saying

> It is possible to arrange the observational situation in such a way that from the responses of a number of persons to the set of tests or items in question we may derive two sets of quantities, the distributions of which depend only on the test or item parameters, and only on the personal parameters, respectively. Furthermore, the conditional distribution of the whole set of data for given values of the two sets of quantities does not depend on any of the parameters (p. 122).

The separability of the parameters is evident in the proportionality of the relationships expected by the model. As any one parameter is varied relative to a second parameter, values for the third are predictable (see Table 1). For example, for a person-item interaction in which there is a 0.82 likelihood of a correct response, the odds ratio of 4.556 (0.82 / 0.18) gives a log-odds (logit) difference of 1.5 between the person ability and item difficulty estimates (see Wright and Stone, 1979, p. 16, for a table relating response probabilities to logit differences). Any ability measure that is 1.5 logits different from a difficulty calibration implies a 0.82 probability of a correct response.

Tables 1 and 2 about here.

If the 1.5 logit difference results from a comparison of a person measure of 2.0 and an item calibration of 0.5, then, to obtain the multiplicative form of the model,

$$\varepsilon_{vi} = \theta_v \sigma_i$$

we have, with the previous values entered

$$e^{2.0} = e^{1.5} * e^{0.5} ,$$

which is exactly the same equation as that previously used to illustrate Newton's Second Law:

$$7.389 = 4.4817 * 1.6487 .$$

Dividing through by 1.6487 gives

$$4.4817 = 7.389 / 1.6487 .$$

Using the natural logarithm to convert division to subtraction, we get the same equation as above:

$$\ln(4.4817) = \ln(7.389) - \ln(1.6487)$$

which reduces to the same equation in use here:

$$1.5 = 2.0 – 0.5 .$$

Taking square roots for each term allows expression of this difference in the form of the Pythagorean theorem:

$$1.2247^2 = 1.4142\ldots^2 – 0.7071^2 .$$

Seeing the relationship in this form allows visualization of the modeled difference between a person ability and an item difficulty as a right triangle with sides of 1.2247 and 0.7071, and a hypotenuse of 1.4142…, as shown in Figure 2.

Figures 2 and 3 about here.

The invariance of a given person's ability measure across items makes that measure a constant relative to the changes in the item difficulties and response probabilities. The proportions in the triangles representing the modeled expectations for each person-item interaction change, then, only with respect to the sides a and b, with the hypotenuse, c, and the 90 degree angle connecting sides a and b, remaining constant. If the item in the comparison is easier, calibrating lower on the scale, then side a becomes shorter, and the probability of a correct response increases, so side b becomes longer. For the same measure of 2 logits relative to an item calibrating at 0.01 logits and a difference of 1.99 logits, the Pythagorean expression becomes:

$$1.4142\ldots^2 – 0.1^2 = 1.4107^2$$

Figure 3, and Tables 1 and 2, show the pattern a series of triangles take as a function of a range of nonzero, positive item calibration differences from a person measure of 1.4142…. Zero differences between measures and calibrations could be assumed to be identical with the hypotenuse, omitted, or avoided by rescaling. Negative differences obtained when items calibrate higher on the scale than the measure could be illustrated via a mirror image flower- or leaf-pattern projection of the pattern in Figure 3, as shown in Figure 4, or by taking the item calibration as representing the hypotenuse.

A variety of patterns could be usefully illustrated using figures of this kind. Each item-triangle might be shown in a different shade of the same color, for instance. Off-target measures, where interpretation is not informed by the content of items near it on the scale, will have the majority of the item-triangles on only one side of the symmetry dividing line.

Figures 4 and 5 about here.

A line drawn in Figures 3 or 4 from the center of the hypotenuse to the point of any triangle's 90 degree angle is the radius of a circle formed by the points of the infinite series of all right triangles sharing that hypotenuse. This pattern is well known in the context of the curve known as Agnesi's witch.[1] Figure 5 (from Maor, 1998, p. 110) shows the circle that has the circumference of an infinite series of right triangles sharing the same hypotenuse. Triangle OAC inscribed within that circle defines a Rasch model equation similar to that of triangle 2 in Figure 3. As shown by Maor, the angle θ between sides OA and OC of the triangle OAC provides the easiest way to find the witch (point P in Figure 5), as the witch is defined by the points of the right angle corners of the triangles taking θ as the angle B opposite the initial θ. As point A is shifted along the circumference of the right side of the circle, a horizontal line drawn from A to P, keeping θ constant, results in the asymptotic curve.

[1] Thanks to Mark Stone for pointing me to Maor's book on trigonometry, from which I learned of Agnesi's witch for the first time.

It may be that the curve describes the same relation as the Rasch score-logit ogive. Support for this pure speculation may be found in the fact that, apart from the constants involved, the witch is identical with the equation for the Cauchy distribution (Maor, 1998, p. 111), the relevance of which to the Rasch model has been noted by Fischer (1995, p. 21). Maor also notes that, except for this connection with probability theory, Agnesi's witch rarely has any practical application. It has nonetheless somewhat mysteriously proven of sustained interest to mathematicians from at least the time of Fermat (1601-1665).

Further implications follow from Figure 4's representation of the modeled expectations. Departures from those expectations could be illustrated in figures evaluating the fit of the observed data to the abstract model, using variations on commonly used model fit statistics, such as the Logit Residual Index or mean square residual fit statistics. A mean square form of an individual person-item observed vs. expected index would have expected values of 1.0 and actual values ranging from near 0.0 to 2.0 and higher.

Multiplying the expected logit differences (sides b) illustrated in the triangles by these mean squares would graphically display anomalous distortions and reveal well-fitting patterns' symmetries. The distortions might take the form of pushing or pulling the expected 90-degree angle to another value, and shortening or lengthening side a as needed for fit to the observed response, however unexpected it may have been. Alternatively, the 90-degree angle and the length of side b might be retained, and the departure from expectation shown in terms of the person measure (hypotenuse) value. The anomalies might also be displayed by allowing a side b spike to project above the intersection with side a, or by having side b stop short of its expected connection with side a.

Predictive Construct Modeling

In his retirement speech and in his book, Rasch (2010/1972, 1960) explained how the structure of Newton's second law of motion (relating force, mass, and acceleration) provides a basis for objectivity in the social sciences. In his retirement speech, after describing multiple examples and elaborating the logic of the analogy in detail, as he also had in his book (Rasch, 1960, pp. 110-115), Rasch (2010/1972) concluded that,

> With all of this available to us, we will have an instrumentarium with which many kinds of problems in the social sciences can be formulated and handled with the same types of mathematical tools that physics has at its disposal—without it becoming a case of superficial analogies. (p. 1272)

But nowhere in his book, retirement lecture, or other publications does Rasch provide a theory of a substantive construct behaving in accord with the structure of a lawful regularity. As Maxwell (1965/1890, p. 155) understood would happen, the convenient analytical formulation of Rasch's models has caused us to lose sight of the phenomena to be explained, such that we "never obtain more extended views of the connexions of the subject." That is, one of physics' primary mathematical tools is the predictable conformity of abstract models and the behavior of natural phenomena. Not only can the sample studied be changed without altering the invariant parameters of the model, which is what Rasch emphasized, but cause and effect relationships can be demonstrated and controlled, as when a particular measure can be obtained by altering a sample or a new instrument configuration can be shown to calibrate to expected values.

Rasch does not describe the instrumentarium of mathematical tools in any terms except those of models and their application to data analysis. That is, he never raises the possibility that, given the establishment of laws of cognition and behavior mathematically and functionally identical to natural laws, the social sciences could calibrate instruments capitalizing on predictive theories and measuring in universally uniform reference standards, in the manner of the natural sciences (Fisher, 2009a, 2009b, 2010, 2011, 2012c; Fisher & Stenner, 2011). Rasch's focus on data analysis may be a primary reason why his models are so often treated as group-level statistical models to be fit to data, instead of as individual-level measurement models to which data are fit (Andrich, 2002). But even when the models are correctly conceived and employed, researchers are almost invariably distracted away from theorizing about the construct by the analytical subtleties afforded by the mathematical formulation.

Rasch does speak to the way in which an established law may serve as a tool for deciding whether new stimuli and objects, or items and cases, belong to the construct (Rasch, 1960, p. 124). And in his retirement speech, Rasch (2010/1972) goes a bit further and describes how the ongoing gathering of new data amounts to an implicit delimiting of the "field of validity." Exploration of the frame of reference for Newton's second law, for instance, enables the researcher to "in the end discover which physical qualities they [the relevant class of bodies and instruments] must have in contrast to those to which the law does not apply" (p. 1254). Rasch (2010/1972) points out, that

> If the frame of reference is extended, the hypothesis may no longer apply. If, for instance, you kick 1 kg butter at 20 degrees centigrade, it will stick to your shoe, and if an instrument functions not only mechanically but also magnetically, objects made from stone and iron will react in quite different ways. And if other things beside accelerations are taken for reactions—for example velocities or positions, not to mention the colour and light reflection of the bodies—then (1) [the stated law] will, of course, cease to apply. (p. 1254; also see Rasch, 1960, p. 124)

But in asserting that "Thereby you can gradually reach a clarification of the field of validity of the law," and in next taking "a closer look at the contents of the law," Rasch (2010/1972, p. 1254) does not follow Maxwell's process. Rasch does not try to explain individual-centered variation in a psychological or social phenomenon in psychological or social terms, as one would in investigations emulating Maxwell's interest in explaining a physical phenomenon in physical terms. Instead, Rasch's focus on the contents of the law is strictly mathematical. His concern is with the nature of the independence of the comparisons made in a context of infinite possibility. He shows how the frame of reference provides a means for defining all possible relevant observational situations, but he does not show, as does Maxwell for electromagnetism, what makes any given observation conform to the model in the way that it does.

But is not "clarification of the field of validity of the law" ultimately a matter of identifying the component aspects of the construct contributing to its proportionate variability? Should not an understanding of the laws governing the construct make it possible to reproduce its effects in a predictable way? If Rasch wanted to avoid superficial analogies, should not he have been compelled to arrive at a strict analogue of Maxwell's effort to physically explain physical facts, and to have then focused on psychosocial explanations of psychosocial facts?

In contrast with Rasch's and Thomson's approach to the method of analogy, Nersessian (2002) points out that Maxwell's

> kind of model-based reasoning process has the potential to lead to genuinely new representational structures, in other words, conceptual change.... Throughout his reasoning processes Maxwell abstracted from the specific mechanism to find the mathematical form of that class of mechanism, in other words, of the generic dynamical structure (p. 144)

In not seeking a substantive theory of the reading ability construct, for instance, Rasch seems to have been, contra Maxwell, “drawn aside from the subject in pursuit of analytical subtleties,” as Maxwell (1965/1890, p. 156) feared he would be if he took a purely abstract mathematical approach. Rasch (1960) is correct within the limits of his statement that “Where this law [of reading ability] can be applied it provides a principle of measurement on a ratio scale of both stimulus parameters and object parameters, the conceptual status of which is comparable to that of measuring mass and force” (p. 115). But mass and force are understood here in a limited way that makes it possible to inform applications on the basis of theory. One need not, for instance, perform a series of experiments to find out the force needed to accelerate a mass to the point that it will travel a given distance. The standardization of the metrics in which measures of mass, force, and acceleration are expressed, in conjunction with the known laws governing their relations, make it possible to know in advance of any application of a given force to a given mass what acceleration will be achieved. Rasch makes no effort to discover any kind of analogous law that might govern the difficulty of reading test items.

In his use of the method of analogy in the study of electromagnetism, Maxwell was focused on arriving at a practical system for the management of the phenomenon (Hunt, 1994; Schaffer, 1992). The properties of electrical resistance had to be grasped effectively and efficiently enough to make them commercially viable for general application in the newly emerging electrical industry. Conductors and insulators had to be understood well enough to allow manufacturing on mass scales, and that meant subjecting every meter of cable or wire to empirical tests would be impossible.

Arriving at a similar understanding of the “specific mechanisms” or operations at work in data conforming to Rasch models, “to find the mathematical form of that class of mechanism,” and "of the generic dynamical structure," requires approaches to identifying and experimentally evaluating the component elements of constructs. Such approaches must authentically test the strength of potentially real objects, respect the limits of method and remain open to questioning in ways that can accept the falsification of hypotheses. Only in the wake of these kinds of experimental tests of strength will the psychosocial sciences be able to go beyond laboratory demonstrations of real objects to their routine production in practical applications in the manner described by Ihde (1991, p. 134), Latour (1987, 2005), Wise (1995), and others. Beyond that, the expenses and inefficiencies of the psychosocial sciences' largely empiricist orientation to measurement are one of the primary reasons for the dysfunctionality of human, social, and natural capital markets (Fisher, 2009b, 2011; Fisher & Stenner, 2011).

In the wake of Rasch's work, and later large-scale studies equating high stakes reading tests (Jaeger, 1973; Rentz & Bashaw, 1977), Stenner and colleagues (Stenner, 2001; Stenner, et al., 2006) developed an effective and parsimonious predictive theory of what makes text easy or difficult to read. Others have similarly devised predictive models of other cognitive and behavioral constructs (Bunderson & Newby, 2009; Dawson, 2002, 2003, 2004; Embretson, 1998; Embretson & Daniel, 2008; Fisher, 2008; Green & Kleuver, 1992; Wilson, 2008) with the aim of achieving the degree of control over the instrumentation needed for the reliable and highly efficient automated production of assessment items (Bejar, Lawless, Morley, Wagner, Bennett, & Revuelta, 2003; Stenner & Stone, 2003).

Generalizing these accomplishments requires a systematic and methodical way of interweaving substantive qualitative content and abstract mathematical construct issues. The system for assessing constructs described by Wilson (2005; Wilson & Sloane, 2000) opens the door to a fuller realization of model-based reasoning in the psychosocial sciences in the way that it provides leverage points for theory development. That is, in the context of such a system hypotheses may be formulated and tested by iterating through a sequence of moments in a method, any one of which may serve as a point of entry or exit. Building on the way in which data, instruments, and theory have each historically served to mediate each other's interrelations (Ackermann, 1985), and focusing on the predictive control of the construct, it becomes possible to envision new horizons for quantitative social science. Of particular interest are the ways in which probabilistic measurement models may facilitate adaptive and on-the-fly practical applications for monitoring and improving the quality of psychological, social, organizational, economic, and natural ecologies (Fisher, 2012a).

Discussion

In light of the exact identity in the mathematical form of his model for measuring reading ability and Newton's Second Law, Rasch (1960, p. 115) asserted that,

> Where this law can be applied it provides a principle of measurement on a ratio scale of both stimulus parameters and object parameters, the conceptual status of which is comparable to that of measuring mass and force. Thus...the reading accuracy of a child ... can be measured with the same kind of objectivity as we may tell its weight ....

Wright (1997, p. 44), a physicist who worked with Nobelists Townes and Mulliken before turning to psychology and collaborations with Rasch, concurs, saying, "Today there is no methodological reason why social science cannot become as stable, as reproducible, and hence as useful as physics." Andrich (1988, p. 22) observes that "...when the key features of a statistical model relevant to the analysis of social science data are the same as those of the laws of physics, then those features are difficult to ignore." Empirical substantiation of the geometric metaphor's implications for spatial representation has been provided by Moulton's (1993) construction of a geographic map from distance ratings data, and by Fisher's (1988) construction of a length ruler from ordinal observations. Similar results linearly transformable into standard SI units were obtained in Choi's (1997) development of a weight measure via paired comparisons.

There are at least eight ways Rasch measurement conforms to geometry as a model for measurement, and through which a shift from Thomson's more superficial method of analogy to Maxwell's richer one might be facilitated. First, geometrically, the elements of measuring systems are defined abstractly, in the same way that a geometrical point is an indivisible line, a line is an indivisible plane, and in the same way that Newton's first law of motion posits bodies left entirely to themselves moving uniformly in straight lines. Rasch measurement defines its objects geometrically in this way, modeling the probability of a correct answer as dependent on nothing but the ability of the person responding and the difficulty of the question asked.

Second, the geometrical figure of the line provides the practical means for linear measurement's association of number with distance or length, so that substantive unit amounts may be mapped onto number lines. Graphical images, such as Wright maps of measured constructs (Wilson, 2005), support the capacity to visualize geometrically proportionate variations in the modeled parameters (Moulton, 1993), as has repeatedly proven to be crucial to scientific advances, from the Newtonian laws to Maxwell's electromagnetic theory to Einstein's theory of relativity. Rasch's (1960; Ludlow, 1985) graphical evaluations of the fit of data to his models provides an original example of this kind of thinking in the social sciences. Elucidation of the geometrical forms implied by Rasch model relationships may enable some to visually and intuitively apprehend implications and consequences previously available only in the more inconvenient and cognitively difficult analytic form.

Third, the original standard model of natural law is stated as a theorem in which the relation of two linear variables is consistently mediated by another variable, in the manner of the sides and hypotenuse of a right triangle, or the relations of mass, force, and acceleration in Newton's Second Law. These laws may be expressed in either additive or multiplicative forms, and they may be expanded to multifaceted forms (Linacre, 1989), but they must not incorporate interaction terms that destroy the possibility of invariant proportionality across the parameters in the model (Andrich, 1988, p. 67; Wright, 1999). Of course, linear scaling factors may be used to make the units for different parameters in the model commensurable (Crease, 2004; Stenner, et al., 2006). When these scaling factors are included in the formal conceptualization of unit definitions they redefine the relation of Rasch models to Item Response Theory, such that the latter becomes a special case of the former, instead of the heretofore more common reverse condition (Humphry, 2011; Humphry & Andrich, 2008).

Fourth, tools akin to the compass and straightedge enable the construction of proofs and experimental tests of hypotheses with no demands as to the existence or presence of a particular unit of measurement. The meaningful comparability of results requires patterns of invariant association across the modeled parameters, such that the same relations are shown across changes in scale (Mundy, 1986; Narens, 2002). Linear plots of item calibration estimates from different samples, or of person measures estimated from different sets of items (Figures 6-8), provide what are, in effect, geometric proofs justifying the interchangeability of independent methods of constructing the same series of right triangles.

Figures 6-8 about here.

Fifth, in the context of Rasch's probabilistic models, new conventions for taking uncertainty into account are needed. Figures 6-8 show significant differences in the correlations obtained from separate sample calibrations. All three of these data sets, however, were sampled from the same larger data matrix, and all three correlations disattenuate to 1.00, meaning that the apparent differences in the associations of the pairs of measures are due entirely to variations in error and reliability.

Sixth, substantive theory predicts and explains how changes to one facet in the overall design cause predictable changes in the relation of the other two facets. This is the capacity to intervene in the construct with new, previously unseen stimuli, instruments, questions, samples, and/or respondents, and to obtain the results expected by theory, within the error of measurement. The value of theory is realized in this context, where the expense and difficulty of obtaining the data necessary for calibrating instruments and testing the predictive power of the theory is surpassed and replaced with efficient means for reliably putting instruments to work on the basis of their designs (Bejar, et al., 2003; Bunderson & Newby, 2009; Stenner, et al., 2006; Stenner & Stone, 2003). This kind of theory-based measurement takes place on one level with Rasch's models when calibrated instruments are scored at the point of use on the basis of complete data. It takes place at another level when banks of precalibrated items are adaptively administered to produce measures interpretable in a common framework. And it takes place on yet another level when items are created on the fly and applied in computerized contexts for advanced formative applications.

Seventh, allowing nature, human and otherwise, to reveal itself by means of its exceptions is essential to the logic of discovery and the process of invention. In the same way that geometrical figures do not occur in nature and can never be drawn so as to be absolutely perfectly commensurate with their mathematical form, so, too, are there no perfectly spherical balls rolling on perfectly frictionless planes and there are no test, survey, or assessment results completely unaffected by the particular questions asked and persons answering. Despite their fundamentally opposite philosophical perspectives, Butterfield (1957, pp. 17, 25-26, 96-98) as much as Heidegger (1967, p. 89) stresses the importance of models of this kind to the progress of science. Mathematical models of geometric invariances enable the transparent representation and identification of things, and, importantly, the display of anomalous exceptions to the rule (Kuhn, 1977, p. 205; Rasch, 1960, p. 124). Rasch models are as impossible to realize in practice as the Pythagorean theorem or Newton's laws. Thus, the value of the model is not determined by its truth, but by its usefulness (Rasch, 1960, pp. 37-38; Box, 1979, p. 202). An essential reason for measuring and for employing abstract ideal models is to express expectations with such extreme clarity that exceptions will show themselves. A great many discoveries in the history of science emerged as a result of unexpected findings being recognized as answers to questions that had not yet been asked. The unexpected must be made consistently observable and reproducible before it can be discovered (Von Oech, 2001).

Eighth, decades and centuries of qualitative investigation are necessary precursors to successful quantification (Heilbron, 1993; Roche, 1998). Geometric principles informed practical surveying applications in ancient cultures predating classical Greece. The introduction of Euclidean axioms did little to change practical geometry, and though significant advances were made in the first scientific revolution before the advent of international metric standards, the standardization of measuring units in the early nineteenth century was an important factor in the explosion of productivity infusing the Industrial Revolution and the second Scientific Revolution (Alder, 2002; Roche, 1998). Rasch measurement has similarly succeeded in producing some significant advances in the absence of universally uniform units of measurement, but will likely not realize its potential until such units are widely available.

Conclusion

Linear measurement is a form of practical geometry. A basis for quantification is inferred from the meaningful invariant consistency of geometrically linear comparisons. Unit-free geometrical constructions are the original form of measurement defined as a ratio of a magnitude of a quantitative attribute to another magnitude. The existence of a unit cannot precede the existence of the magnitude identified as supporting division into ratios. As is extensively documented in the history of science, lawful regularities are identified and studied qualitatively for decades and even centuries before quantification is possible. “The road from scientific law to scientific measurement can rarely be traveled in the reverse direction” (Kuhn, 1977, p. 219) because qualitative understanding of lawful patterns is a necessary prerequisite to intuitively valid quantification.

What does this mean in practical terms in the psychosocial sciences? In the context of Rasch’s models for measurement, law-like patterns in the empirical relational structures of data from test or survey questions, ordinal observations, and response likelihoods are repeatedly exhibited when similar questions fall in similar and often highly correlated orders and relative locations across instruments and samples, as do similarly highly correlated types of examinees or respondents. That said, the replicability of various consistently reproduced orders and relative positions of location estimates across data sets is only the first phase in the process of developing predictive construct theories, and in defining and deploying a standard unit.

Continuous magnitudes of a number of important psychosocial constructs have been documented independent of samples and instruments, but little effort has yet been invested in arriving at predictive construct theories or consensus agreement on the conventions of unit size and nomenclature necessary for fully integrating mathematics and measurement in the psychosocial sciences. Though virtually everything remains to be done in defining substantively meaningful universally uniform units for the psychosocial sciences, the viability of such reference standards is supported by the presence and relevance of the same eight analogies from geometry that support the ongoing successes of the natural sciences.

References

Ackermann, J. R. (1985). *Data, instruments, and theory: A dialectical approach to understanding science*. Princeton, New Jersey: Princeton University Press.

Alder, K. (2002). *The measure of all things: The seven-year odyssey and hidden error that transformed the world*. New York: The Free Press.

Andrich, D. (1988). *Rasch models for measurement*. (Vols. series no. 07-068). Sage University Paper Series on Quantitative Applications in the Social Sciences. Beverly Hills, California: Sage Publications.

Andrich, D. (2002). Understanding resistance to the data-model relationship in Rasch's paradigm: A reflection for the next generation. *Journal of Applied Measurement*, *3*(3), 325-359.

Bejar, I., Lawless, R. R., Morley, M. E., Wagner, M. E., Bennett, R. E., & Revuelta, J. (2003). A feasibility study of on-the-fly item generation in adaptive testing. *The Journal of Technology, Learning, and Assessment*, *2*(3), 1-29; http://ejournals.bc.edu/ojs/index.php/jtla/article/view/1663.

Black, M. (1962). *Models and metaphors*. Ithaca, New York: Cornell University Press.

Box, G. E. P. (1979). Robustness in the strategy of scientific model building. In R. L. Launer & G. N. Wilkinson (Eds.), *Robustness in statistics*. New York: Academic Press, Inc.

Bunderson, C. V., & Newby, V. A. (2009). The relationships among design experiments, invariant measurement scales, and domain theories. *Journal of Applied Measurement*, *10*(2), 117-137.

Burdick, D. S., Stone, M. H., & Stenner, A. J. (2006). The Combined Gas Law and a Rasch Reading Law. *Rasch Measurement Transactions, 20*(2), 1059-60.

Burtt, E. A. (1954). *The metaphysical foundations of modern physical science* (Rev. ed.) Garden City, New York: Doubleday Anchor.

Butterfield, H. (1957). *The origins of modern science* (revised edition). New York: The Free Press.

Choi, S. E. (1997). Rasch invents "Ounces." *Rasch Measurement Transactions, 11*(2), 557.

Crease, R. (2004). The greatest equations ever. *Physics World, 17*(10), 19.

Cropper, W. H. (2001). *Great physicists: The life and times of leading physicists from Galileo to Hawking*. Oxford, England: Oxford University Press.

Dawson, T. L. (2002). New tools, new insights: Kohlberg's moral reasoning stages revisited. *International Journal of Behavioral Development, 26*(2), 154-66.

Dawson, T. L. (2003). A stage is a stage is a stage: A direct comparison of two scoring systems. *Journal of Genetic Psychology, 164*, 335-364.

Dawson, T. L. (2004). Assessing intellectual development: Three approaches, one sequence. *Journal of Adult Development, 11*(2), 71-85.

Dawson, T. L., Fischer, K. W., & Stein, Z. (2006). Reconsidering qualitative and quantitative research approaches: A cognitive developmental perspective. *New Ideas in Psychology, 24*, 229-239.

Einstein, A. (1922). Geometry and experience (G. B. &. P. Jeffery, W., Trans.). In A. Einstein, *Sidelights on relativity* (pp. 12-23). London, England: Methuen & Co., Ltd.

Embretson, S. E. (1998). A cognitive design system approach to generating valid tests: Application to abstract reasoning. *Psychological Methods, 3*(3), 380-396.

Embretson, S. E., & Daniel, R. C. (2008). Understanding and quantifying cognitive complexity level in mathematical problem solving items. *Psychology Science Quarterly, 50*, 328-344.

Fischer, G. H. (1995). Derivations of the Rasch model. In G. Fischer & I. Molenaar (Eds.), *Rasch models: Foundations, recent developments, and applications* (pp. 15-38). 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* (University of Chicago Dept. of Education, Division of the Social Sciences), 49, 0778A.

Fisher, W. P., Jr. (2008). A predictive theory for the calibration of physical functioning patient survey items. Presented at the Second Conference on Patient Reported Outcome Measurement Information Systems, Bethesda, Maryland: NIH and NIAMS, March 2-5.

Fisher, W. P., Jr. (2009a). Invariance and traceability for measures of human, social, and natural capital: Theory and application. *Measurement, 42*(9), 1278-1287.

Fisher, W. P.. Jr. (2009b). NIST Critical national need idea White Paper: metrological infrastructure for human, social, and natural capital (Tech. Rep. No. http://www.nist.gov/tip/wp/pswp/upload/202_metrological_infrastructure_for_human_social_natural.pdf). Washington, DC: National Institute for Standards and Technology.

Fisher, W. P., Jr. (2010). The standard model in the history of the natural sciences, econometrics, and the social sciences. *Journal of Physics: Conference Series, 238*(1), http://iopscience.iop.org/1742-6596/238/1/012016/pdf/1742-6596_238_1_012016.pdf.

Fisher, W. P., Jr. (2011). Bringing human, social, and natural capital to life: Practical consequences and opportunities. In N. Brown, B. Duckor, K. Draney & M. Wilson (Eds.), *Advances in Rasch Measurement, Vol. 2* (pp. 1-27). Maple Grove, MN: JAM Press.

Fisher, W. P., Jr. (2012a). Measure local, manage global: Intangible assets metric standards for sustainability. In J. Marques, S. Dhiman & S. Holt (Eds.), *Business administration education: Changes in management and leadership strategies* (p. in press). New York: Palgrave Macmillan.

Fisher, W. P., Jr. (2012b). Parameter separation theorems in econometrics and psychometrics: Some shared social history. *Historia Mathematica*, in review.

Fisher, W. P., Jr. (2012c). What the world needs now: A bold plan for new standards. *Standards Engineering*, *64*, in press.

Fisher, W. P., Jr., & Stenner, A. J. (2011). Metrology for the social, behavioral, and economic sciences (Social, Behavioral, and Economic Sciences White Paper Series). Retrieved 25 October 2011, from National Science Foundation: http://www.nsf.gov/sbe/sbe_2020/submission_detail.cfm?upld_id=36.

Forfar, J. (2002). James Clerk Maxwell: His qualities of mind and personality as judged by his contemporaries. *Mathematics Today, 38*(3), 83.

Green, K. E., & Kluever, R. C. (1992). Components of item difficulty of Raven's Matrices. *Journal of General Psychology, 119*, 189-199.

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

Heilbron, J. L. (1993). Weighing imponderables and other quantitative science around 1800. *Historical studies in the physical and biological sciences, 24* (Supplement), Part I, pp. 1-337.

Heilbron, J. L. (1998). *Geometry civilized: History, culture, and technique*. Oxford, England: Clarendon Press.

Hesse, M. (1961). *Forces and fields: A study of action at a distance in the history of physics.* London: Thomas Nelson and Sons.

Humphry, S. M. (2011). The role of the unit in physics and psychometrics. *Measurement: Interdisciplinary Research & Perspectives, 9*(1), 1-24.

Humphry, S. M., & Andrich, D. (2008). Understanding the unit in the Rasch model. *Journal of Applied Measurement, 9*(3), 249-264.

Hunt, B. J. (1994). The ohm is where the art is: British telegraph engineers and the development of electrical standards. *Osiris: A Research Journal Devoted to the History of Science and Its Cultural Influences, 9*, 48-63.

Ihde, D. (1991). *Instrumental realism: The interface between philosophy of science and philosophy of technology.* The Indiana Series in the Philosophy of Technology. Bloomington, Indiana: Indiana University Press.

Isacoff, S. M. (2001). *Temperament: The idea that solved music's greatest riddle*. New York: Alfred A. Knopf.

Jaeger, R. M. (1973). The national test equating study in reading (The Anchor Test Study). *Measurement in Education, 4*, 1-8.

Klein, H. A. (1974). *The world of measurements: Masterpieces, mysteries and muddles of metrology*. New York: Simon & Schuster.

Kuhn, T. S. (1977). The function of measurement in modern physical science. In T. S. Kuhn, *The essential tension: Selected studies in scientific tradition and change* (pp. 178-224). Chicago: University of Chicago Press. (Reprinted from Kuhn, T. S. (1961). *Isis*, *52*(168), 161-193.)

Larmor, J. (Ed.). (1937). *The origins of Clerk Maxwell's electric ideas*. [Reprinted from: *Proceedings of the Cambridge Philosophical Society*, 1936, XXXII, part v, 695-750.]). Cambridge, Massachusetts: Cambridge University Press.

Latour, B. (1987). *Science in action: How to follow scientists and engineers through society*. New York: Cambridge University Press.

Latour, B. (2005). *Reassembling the social: An introduction to Actor-Network-Theory*. Oxford, England: Oxford University Press.

Linacre, J.M. (1989). *Many-Facet Rasch Measurement*. Chicago: MESA Press.

Ludlow, L. H. (1985). A strategy for the graphical representation of Rasch model residuals. *Educational and Psychological Measurement, 45*, 851-859.

Maor, E. (1994). *e: The story of a number*. Princeton, New Jersey: Princeton University Press.

Maor, E. (1998). *Trigonometric delights*. Princeton, New Jersey: Princeton University Press.

Maxwell, J. C. (1920/1876). *Matter and motion* (J. Larmor, Ed.). New York: The Macmillan Co.

Maxwell, J. C. (1965/1890). *The scientific papers of James Clerk Maxwell* (W. D. Niven, Ed.). New York: Dover Publications.

McMullin, E. (2002). The origins of the field concept in physics. *Physics in Perspective, 4*(1), 13-39.

Miller, P., & O'Leary, T. (2007, October/November). Mediating instruments and making markets: Capital budgeting, science and the economy. *Accounting, Organizations, and Society, 32*(7-8), 701-34.

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

Mundy, B. (1986). On the general theory of meaningful representation. *Synthese, 67*(3), 391-437.

Myers, M. (1983). *The soul of modern economic man: Ideas of self-interest Thomas Hobbes to Adam Smith*. Chicago, Illinois: University of Chicago Press.

Narens, L. (2002). A meaningful justification for the representational theory of measurement. *Journal of Mathematical Psychology, 46*(6), 746-68.

Nersessian, N. J. (2002). Maxwell and "the method of physical analogy": Model-based reasoning, generic abstraction, and conceptual change. In D. Malament (Ed.), *Essays in the history and philosophy of science and mathematics* (pp. 129-166). Lasalle, Illinois: Open Court.

Nersessian, N. J. (2006). Model-based reasoning in distributed cognitive systems. *Philosophy of Science, 73*, 699-709.

Nersessian, N. J. (2008). *Creating scientific concepts*. Cambridge, Massachusetts: MIT Press.

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, Denmark: Danmarks Paedogogiske Institut.

Rasch, G. (1961). On general laws and the meaning of measurement in psychology. In J. Neyman (Ed.), *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability: Volume IV: Contributions to biology and problems of medicine* (pp. 321-333). Berkeley, California: University of California Press.

Rasch, G. (2010/1972). Retirement lecture of 9 March 1972: Objectivity in social sciences: A method problem (Cecilie Kreiner, Trans.). *Rasch Measurement Transactions, 24*(1), 1252-1272.

Rautio, J. C. (2005). Maxwell's legacy. *IEEE Microwave Magazine, 6*(2), 46-53.

Rentz, R. R., & Bashaw, W. L. (1977). The National Reference Scale for Reading: An application of the Rasch model. *Journal of Educational Measurement, 14*(2), 161-179.

Roche, J. (1998). *The mathematics of measurement: A critical history*. London: The Athlone Press.

Schaffer, S. (1992). Late Victorian metrology and its instrumentation: A manufactory of Ohms. In R. Bud & S. E. Cozzens (Eds.), *Invisible connections: Instruments, institutions, and science* (pp. 23-56). Bellingham, WA: SPIE Optical Engineering Press.

Stenner, A. J. (2001). The Lexile Framework: A common metric for matching readers and texts. *California School Library Journal, 25*(1), 41-2.

Stenner, A. J., Burdick, H., Sanford, E. E., & Burdick, D. S. (2006). How accurate are Lexile text measures? *Journal of Applied Measurement, 7*(3), 307-22.

Stenner, A. J., & Stone, M. (2003). Item specification vs. item banking. *Rasch Measurement Transactions, 17*(3), 929-30.

Stigler, S. (1986). *The history of statistics: The measurement of uncertainty before 1900.* Cambridge, Massachusetts: Harvard University Press.

Taagepera, R. (2008). *Making social sciences more scientific: The need for predictive models*. New York: Oxford University Press.

Turner, J. (1955). Maxwell on the method of physical analogy. *British Journal for the Philosophy of Science, 6*, 226-238.

Von Oech, R. (2001). *Expect the unexpected (or you won't find it): A creativity tool based on the ancient wisdom of Heraclitus*. New York: The Free Press.

Wilson, M. (2005). *Constructing measures: An item response modeling approach*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Wilson, M. (2008). Cognitive diagnosis using item response models. *Zeitschrift Für Psychologie/Journal of Psychology*, *216*(2), 74-88.

Wilson, M., & Sloane, K. (2000). From principles to practice: An embedded assessment system. *Applied Measurement in Education*, *13*(2), 181-208.

Wise, M. N. (1995). Precision: Agent of unity and product of agreement. Part III--"Today Precision Must Be Commonplace." In M. N. Wise (Ed.), *The values of precision* (pp. 352-61). Princeton, New Jersey: Princeton University Press.

Wright, B. D. (1997). A history of social science measurement. *Educational Measurement: Issues and Practice*, *16*(4), 33-45, 52.

Wright, B. D. (1999). Fundamental measurement for psychology. In S. E. Embretson & S. L. Hershberger (Eds.), *The new rules of measurement: What every educator and psychologist should know* (pp. 65-104). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Wright, B. D., Stone, M. H. (1979). *Best test design: Rasch measurement*. Chicago, Illinois: MESA Press.

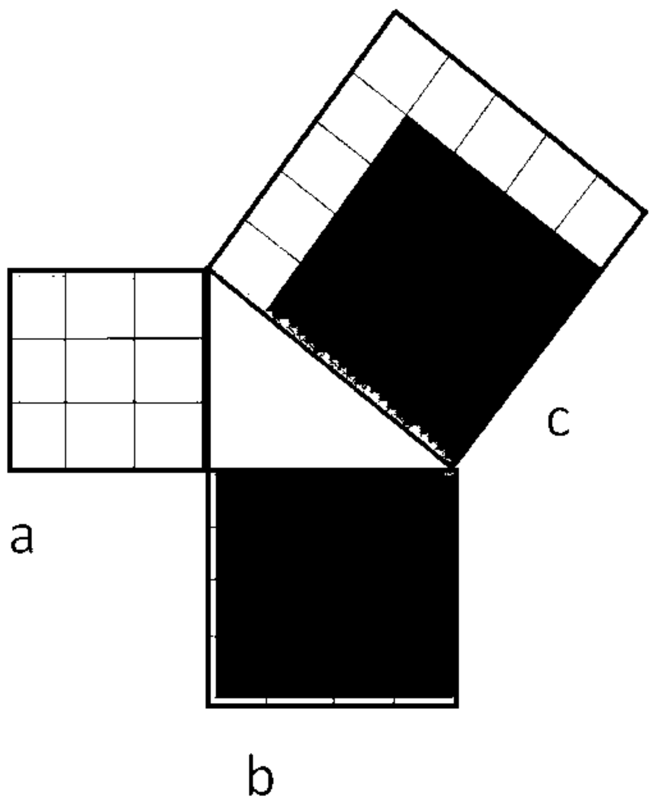

Figure 1. A proof of the Pythagorean theorem.

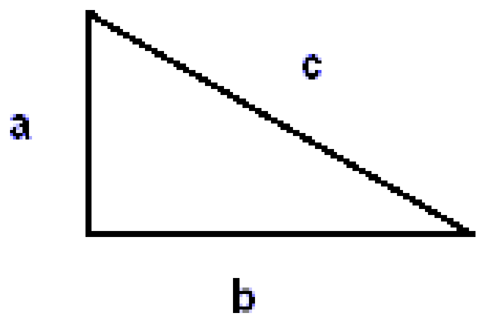

Figure 2. A right triangle with sides a (0.7071), b (1.2247), and c (1.4142…) in an arbitrary unit representing the Pythagorean equivalence of the squares 0.5 and 1.5 with 2.0..

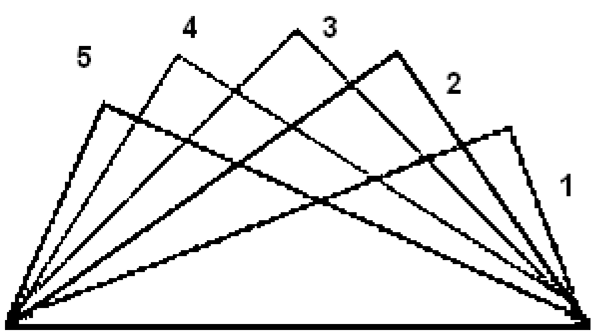

Figure 3. Right triangles sharing the same side c hypotenuse length (1.4142…) in an arbitrary unit. The values shown in Table 1 apply, with the sides to the left being side b, and the sides to the right, side a.

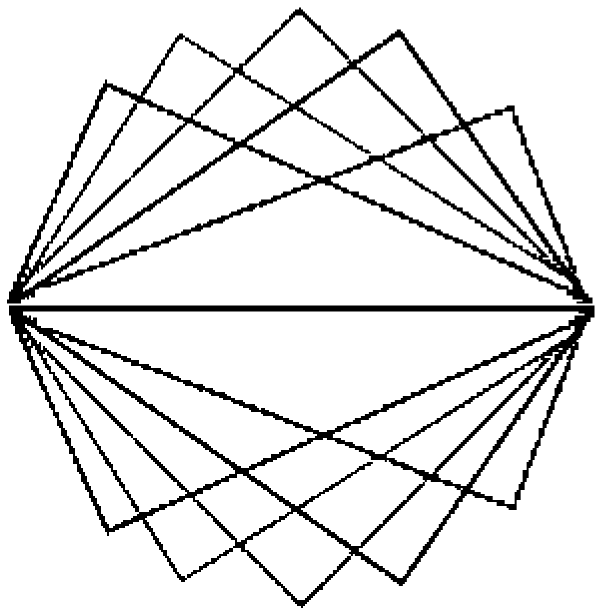

Figure 4. Symmetrical pattern showing negative measure-calibration differences as the mirror image of the positive differences.

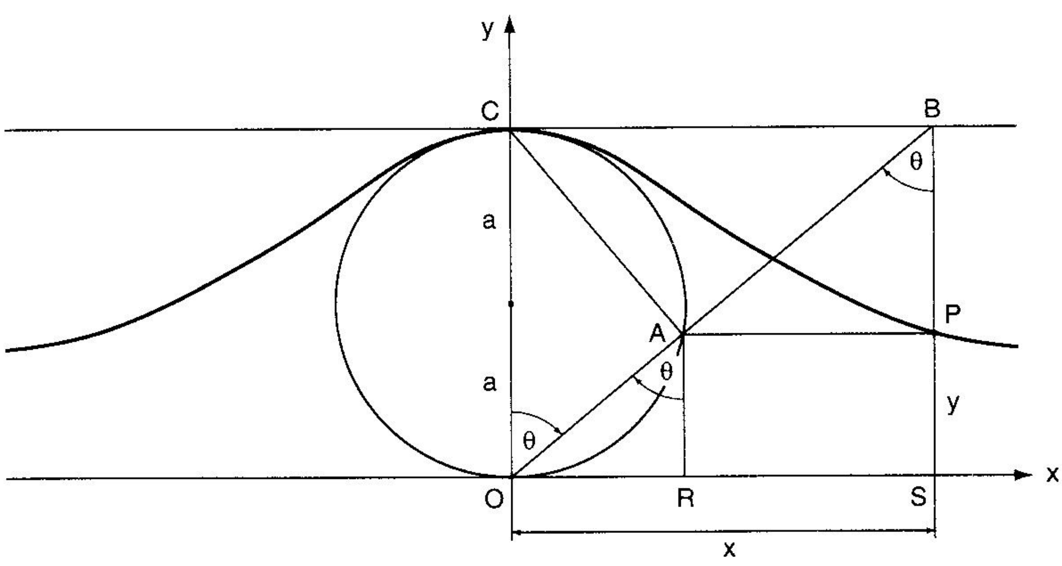

Figure 5. The Witch of Agnesi (from Maor, 1998, p. 110).

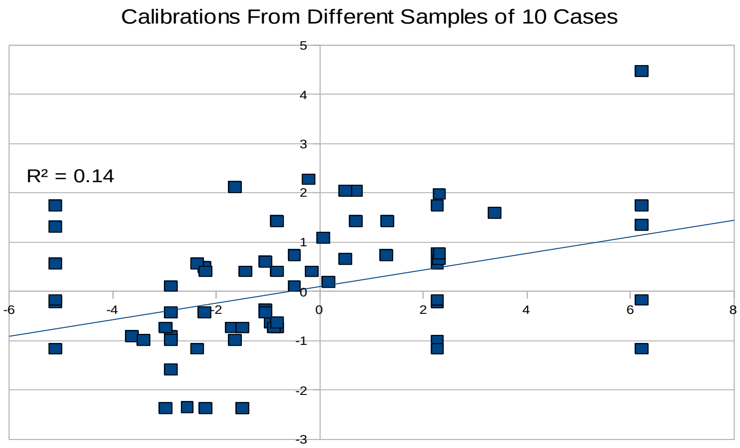

Figure 6.

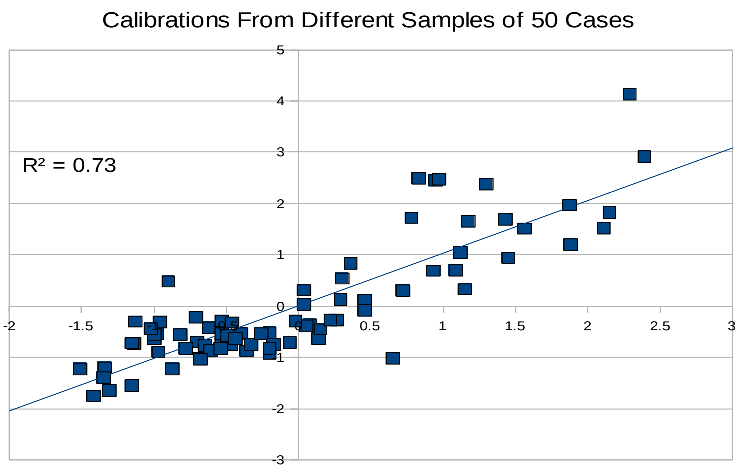

Figure 7.

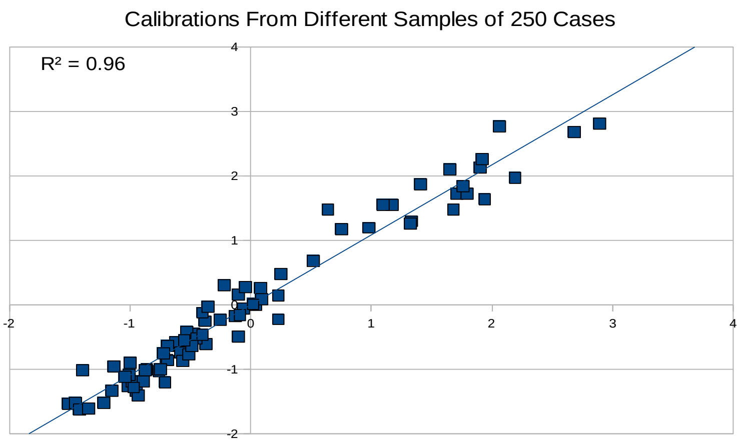

Figure 8.

Table 1. Triangle Side Lengths for Figure 3

Triangle	c	a	b
1	1.4142…	0.5	1.3229
2	1.4142…	0.8	1.17
3	1.4142…	1.0	1.0
4	1.4142…	1.2	0.75
5	1.4142…	1.3	0.56

Table 2. Triangle Sides' Square Areas for Figure 3

Triangle	c^2	a^2	b^2
1	2.0	0.25	1.75
2	2.0	0.64	1.36
3	2.0	1.00	1.00
4	2.0	1.44	0.56
5	2.0	1.69	0.31