DOCUMENT RESUME

ED 331 851 TM 016 361

AUTHOR Fisher, William P., Jr.
TITLE The Hermeneutic of Additive Conjoint Measurement in Educational Research.
PUB DATE Apr 91
NOTE 45p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, April 3-7, 1991).
PUB TYPE Information Analyses (070) -- Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Adults; *Educational Research; Ethnography; *Hermeneutics; Literature Reviews; Metaphors; Models; Philosophy; Qualitative Research; *Research Problems; Social Science Research
IDENTIFIERS *Additive Models; *Conjoint Measurement

ABSTRACT

Additive conjoint measurement models for educational research are construed as practical applications of the hermeneutic theory of scientific instrumentation. It is argued that the model of the text is no less paradigmatic for quantitative method in the natural and social sciences than it is for qualitative, ethnographic methods. It is also suggested that qualitative ethnographic research limits itself if it does not follow through on its critique of quantitative thinking by demonstrating the alternatives it offers. Two studies were conducted to test the hypothesis that metaphors provide virtual meanings that can be honed into actual measures. In one study, a relatively homogenous sample of 36 Americans rated the meaningfulness of 68 entailments of the metaphor "love is a rose." In the second study, a group of Nigerians rated the meaningfulness of a local saying. Both studies support the hypothesis that the virtual measures of meaning provided by language can be sharpened into actual quantitative measures. Implications for application of conjoint measurement models are discussed. It is concluded that the methodological and epistemological dilemma of social research is more complex than merely recognizing the barriers to scientific objectivity; it is more a matter of recognizing that the model of objectivity which researchers have attempted to replicate does not exist even in the sense of a heuristic device. Six tables present study data. An 81-item list of references is included. (SLD)

The Hermeneutic of Additive Conjoint Measurement in Educational Research

William P. Fisher, Jr.
Marianjoy Rehabilitation Center
2091 Creekside Drive #1-2
Wheaton, IL 60187

Paper presented to the Philosophical Studies of Education SIG of the American Educational Research Association at the 1991 Annual Meeting in Chicago
April, 1991

CONTENTS

Objectives. The hermeneutic account of natural science (Heelan, 1972, 1983a, 1983b, 1985) offers a new and creative theory of scientific measurement and instrumentation that is remarkably similar to the theory of additive conjoint measurement frequently employed in educational research (Perline, Wright & Wainer, 1979; Brogden, 1977; Green, 1986; Wright, 1985; Rasch, 1960; Luce and Tukey, 1964). This paper takes its cue from Heelan's work and construes additive conjoint measurement models as practical applications of the hermeneutic theory of scientific instrumentation. The paper has two purposes. The first is to show that the model of the text is no less paradigmatic for quantitative method in the natural and social sciences than it is for qualitative, ethnographic methods. Second, the paper develops the idea that qualitative, ethnographic research falls short of its goals and potential as long as it does not follow through on its critique of quantitative thinking by demonstrating the meaningfulness and usefulness of the quantitative alternatives it offers.

Texts and Mathematical Objectivity. Heelan's work follows from the critique of science developed in the works of philosophers such as Husserl, Heidegger, and Gadamer. Heelan parts ways with these writers at the point where they are unwilling to undertake the project of elaborating the interpretive nature of scientific understanding. Heelan draws from recent work in the philosophy of science, where much has been done to refocus thinking on historical transformations of scientific and human understanding (Fleck, 1979; Toulmin, 1953, 1982; Kuhn, 1970; Hesse, 1970, 1972; Laudan, 1977). Heelan has focused his attention on what he calls the hermeneutic of instrumentation, a phrase that describes the way measuring instruments work to mediate the mutual implication of subject and object, of theory and data, and of measuring devices and the things measured.

Heelan's work expresses the growing awareness in the philosophy of science that the calibration of instruments is crucial to the creation of new phenomena, and that this function is the purpose of scientific research (also see Hacking, 1983, and Lynch, 1985 on this point). This awareness suggests that the social sciences, including educational research, could do well to ask more questions concerning the mediating function of instruments and the associated quality of observations. For instance, scientific instruments generally represent a single phenomenon in various manifestations of more and less along a continuum divided into equal intervals. Scientific comparison of amounts requires that the order and size of the measuring units remain constant across individual manifestations of the phenomenon measured, and it requires that the order and magnitude of differences between individuals remain constant across particular models or examples of the type of instrument involved. This sense of objectivity already bears strong resemblance to what Wright (1968) calls sample-free calibration and instrument-free measurement.

There is, however, a crucial difference between the hermeneutic and usual conceptualizations of this definition of the objective functioning of instruments. Where the usual view asserts that the unit order and size in which data are expressed is God-given, 'natural,' or otherwise determined outside of history, the hermeneutic view asserts that instrumental structure emerges from within the data-theory interaction. Heelan contrasts a sense of instruments as one-sided and data as given with a sense of instruments as two-sided and data as interpretations of a mode of being or form of life, inscribed upon technological media on one side and read off them on the other. In order for such media to perform the mediating function they must be legible, which is to say that the question and answer process, through which the theory-data interaction plays itself out, must constantly refer and defer to the matter of the text represented by the instrument; in short, the instrumental medium must separate from the data and theory of its inventor/author in the same manner that a written text separates from the particular intentions and context of its author and takes on a life of its own (Gadamer, 1989; Ricoeur, 1981).

Ackermann (1985) continues this line of thought, noting that data, in turn, mediate the relation between the theory and the instrument; theory, finally, mediates the relation of instrument and data. Scientific objectivity becomes possible, from this point of view, only when theory, data and instruments are mutually interrelated such that each supports the others. Each of the three are textual insofar as perception itself is a form of reading (Nicholson, Heelan, Lynch), but also because theory is read into the instrument and data are read off it.

The fundamental importance of the textual nature of data, instruments, and theory is made especially evident in that, even in that paragon of hard, mathematical objectivity, physics, "no elementary phenomenon is a phenomenon until it is a *registered* (observed) phenomenon" (Wheeler, 1983, pp. 184, 196, 199, 202-203, 206-207; Wheeler and Zurek, 1983, pp. xvi, xix, 769-771; Bohr, 1983, pp. 3-5, 17-18; emphasis added). There is literally nothing to say about any phenomenon until its material form is manifest in a historically, culturally, and linguistically constituted frame of reference. The process and event of enframing is how and when things come into language, and "the most important aspect of the quantum interpretation discussion is the insight it has given into the epistemological role of the conceptual framework" (Petersen, 1968, p. 185). Bohr understood complementarity as a general epistemological problem of human existence (Holton, 1988, p. 134) and stressed the importance of the conceptual enframing of questions in a manner reminiscent of Gadamer's (1989, pp. 362-379) elaboration of the hermeneutic priority of the question. "A well-known dictum of Bohr [states]: 'Every sentence I say must be understood not as an affirmation, but as a question" (Holton, 1988, p. 132). Bohr said, in effect, that any observation, scientific or otherwise, comes about only as an answer to a usually implicit question, just as Gadamer (1989, p. 362) says that "the structure of the question is implicit in all experience." Once, in a discussion at a conference, Bohr

> was forcefully stressing the primacy of language: 'Ultimately, we human beings depend on our words. We are hanging in language.' When it was objected that reality is more fundamental than language and lies beneath language, Bohr answered: 'We are suspended in language in such a way that we cannot say what is up and what is down' (Petersen, 1968, p. 188; Wheeler and Zurek, 1983, p. 5).

Bohr stressed that observations and measurements must be communicable in everyday plain language, and that these arise out of a complementary relation between subject and object, and measured and measuring, in order to counter what he saw as the complications, confusion, and solipsism arising from the interpretations of Einstein, Heisenberg, and Wigner. Bohr's hermeneutic of instrumentation, to borrow Heelan's (1983a) phrase, takes the phenomenon registered on the measuring device as its text, and recognizes with Gadamer (1989, p. 370) the importance of Collingwood's insight that "we can understand a text only when we have understood the question to which it is an answer."

Until the rise of quantum mechanics, natural science was able to ignore our immersion in language and to act as if reality actually was more fundamental than, and beneath, language. Philosophers have long been aware, however, that something was missing from our understanding of what measurement is and how it came about. The poetic beauty and subtlety of mathematics has been accompanied by such power and effectiveness that Wigner (1960) speaks of the unreasonable effectiveness of mathematics in science, and Poincare said that "the very possibility of a science of mathematics seems an insoluble contradiction" (cited in Dantzig 1930, p. 72). Husserl (1970, p. 29) remarks that Galileo "quite understandably did not feel the need to go into the manner in which the accomplishment of idealization actually arose." Understanding of this "hidden irrational element in scientific rationality" (Marcuse 1974, p. 235); of Galileo's "fateful omission" of the manner in which the mathematical idealization of nature arose (Husserl 1970, pp. 21-59, 353-354); and of "this other . . . that holds sway throughout all the sciences, but that remains hidden to the sciences themselves" (Heidegger 1977b, p. 156) is the general goal of hermeneutic phenomenology, though this is rarely recognized by those who take its general position as background to their methods in the social sciences. This goal has been called the "demythologization of science" because science "controls what is proper to it but cannot know the one whom it serves" (Gadamer 1981, p. 150; also see 1976, p. 10).

The mathematical idealization of nature that was embodied in the models and instruments of science had its determination in an ontology that was only partially explicit and became more and more neglected as the successes of science accumulated (Burtt, 1954). Thus, when writers such as Fisher (1932, p. 2) suggested that social researchers appropriate statistical models into their own work, the worst of what had been fatefully omitted from Galileo's mathematical idealization of nature in classical measurement was extended into the social sciences, and with disastrous results. Ignoring the importance of language in the construal of social reality led directly to major problems, provoking Ricoeur's (1974a, 1974b, 1981) general observation that social scientists literally do not know what they are talking about. All of the various structuralisms, poststructuralisms, and neostructuralisms in vogue over the last fifty years have been aimed at figuring out what the objects of a social scientific dialogue might be and how we can enter into such a dialogue without obliterating or domesticating individual differences. The mathematical nature of structure is sometimes noted, as in this passage from Levi-Strauss (1963, p. 283):

> Structural studies are, in the social sciences, the indirect outcome of modern developments in mathematics which have given increasing importance to the qualitative point of view in contradistinction to the quantitative point of view of traditional mathematics.

Levi-Strauss continues on to say that the qualitative perspective in mathematics allowed the development of rigorous methods for "problems that do not admit of a metrical solution." More than this, though, the qualitative perspective has revolutionized mathematics by undercutting much of what had been presumed unshakable in the foundations of the field (Kline, 1980), and by taking important steps toward the recovery of Galileo's fateful omission in the process.

By ignoring this revolution, Levi-Strauss and others who may be taken to hold that "the only acceptable definition of structure is the one provided by mathematicians" (Descombes, 1980, p. 85), such as Michel Serres, take up the qualitative point of view but without relinquishing positivism. Purporting to model observations while "forbidding oneself any hypothesis with regard to their particular 'nature'" (Bourbaki in Descombes, 1980, p. 85), the structuralist sense of qualitative mathematics "establishes between the observer and the system a relationship which is itself nonhistorical. Understanding is not seen here as the recovery of meaning....there is no 'hermeneutic circle,' there is no historicity to the relation of understanding" (Ricoeur, 1974a, pp. 33-34). The denial of any recourse to hypotheses in observation has ancient roots but reached its epitome in Newton, who sought to investigate gravity "*in a mathematical way*, to avoid all questions about the nature or quality of this force (attraction), which we would not be understood to determine by any hypothesis" (Newton in Burtt, 1954, p. 222; original emphasis). The goals of poststructuralism and neostructuralism include, of course, the historical constitution of meaning in the relation between observer and observed, recognizing that language, history and culture frame observations such that implicit hypotheses are unavoidable; the anticipatory fore-structures of understanding projected by language are just as important to the recovery of the meaning of mathematical thinking as its objectivity is (Heidegger, 1967). Ricoeur (1981, pp. 161, 218) says that the point is to

> regard structural analysis as a stage -- and a necessary one -- between a naive and a critical interpretation, between a surface and a depth interpretation, [so that we may] integrate the opposed attitudes of [scientific] explanation and [human] understanding within an overall conception of reading as the recovery of meaning.

The necessary stage of structural analysis is left undeveloped, interpretations are more naive and superficial than critical and deep, and the integration of explanation and understanding remains fundamentally incomplete as long as the mathematics of structure is ignored and unexamined; concomitantly, Galileo's fateful omission and the hidden irrational element in rationality are left unaccounted for, and the demythologization of science is not undertaken. As Ricoeur puts it,

> if there is a hermeneutics -- and here I oppose those forms of structuralism which would remain at the explanatory level -- it must be constituted across the mediation rather than against the current of structural explanation. For it is the task of understanding to bring to discourse what is initially given as structure. It is necessary to have gone *as far as possible along the route of objectification*, to the point where structural analysis discloses the *depth semantics* of a text, before one can claim to 'understand' the text in terms of the 'matter' that speaks therefrom. The *matter* of the text is not what a naive reading of the text reveals, but what the formal arrangement of the text mediates (Ricoeur, (1981, p. 93; first emphasis added).

In short, the methodologically slanted interpretations of hermeneutic phenomenology appropriated by ethnographic approaches to cultural phenomena overlook the subtle critique of quantitative thinking being made and instead use the philosophy to bluntly reject the validity of mathematical models in social science. Even when quantitative research is accorded some validity, it is done so by simply relinquishing all opposition to the one-sided and non-dialectical assignment of numbers and arrangement of structures. To go as far as possible along the route of objectification is not to accept those forms of structuralism purporting to remain strictly explanatory; it is to situate structure in history through the work of enframing.

The interpretative turns toward metaphor, performance, dialogue, and narrative have been very encouraging because they elaborate upon the conversation that we are in a manner that is much more humanly fruitful than previous approaches have been. The linguistic texture of natural science, however, is too often forgotten, even though the notion of science as the reading of the book of nature is as old as Western culture. That book is written in the language of mathematics, and

> too many philosophers have lost sight of the fact that the study of the nature of mathematics has traditionally been, and still can be, carried out in a less isolated context, that it can deal with a much broader variety of problems, and that it has a relevance to many other areas of human interest and experience (Baum, 1973, p. 14).

Heidegger (1967, pp. 66-108) points out that *ta mathemata*, the mathematical, was, for the ancient Greeks, that which is teachable and learnable, and which is, moreover, learned through a re-creative transformation of what is already known; the former aspect of mathematics is stressed to the near exclusion of the latter in contemporary discourse. The ancient relation of things that are teachable and learnable to mathematical thinking (Heidegger 1977a, pp. 247-282; 1967, pp. 66-108) and to the instrumentation of science (Heelan 1983a, 1983b, 1985) is evident in the shared etymology of our words "instrument" and "instruction." Mathematical description applies to anything that can be taught and learned because it is in the endurance and persistence of meaning that the first rudiments of scientific objectivity are manifest. *Ta mathemata* is therefore often translated as 'curriculum' (Descartes 1961, p. 17, Bochner 1966, p. 255, Bell 1931, p. 58, Heath 1931, p. 5, Dantzig 1955, p. 25, Wilder 1965, p. 284, Slavkov 1973, p. 91, Miller 1921, pp. 78, 17), but the implications of this meaning for social science are usually ignored. It is evident, however, that quantitative thinking in educational research could stand to benefit substantially from pursuit of this line of thinking.

Contrary to those senses of mathematical objectivity involving a subject outside of any involvement in what is observed, what is at issue here is the dialectic of participation and distanciation through which texts mediate the relations of observers and observed (Ricoeur, 1981). Textual mediation extends and sharpens the body's mediation of self and other. The text of what is said, written or perceived Socratically emerges from the body as a child which must be examined for its potential to take on and live a life of its own. Something said or written is instructive and instrumental to learning only to the extent that its meaning converges with the expectations of the persons involved and separates from them across particular instances of recognition and the particular signs prompting its emergence. For example, Gadamer (1980, p. 101) points out that

> 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.

The teacher erases the drawn circle, moves on to other places, times and students, draws another figure, but the figure is called the same as the previous one, the same points are made, the same rule is delineated, the same structure is built into (in-structed) the understanding of the students, and the same lesson is taught because of the capacity of symbols to be completely exhausted by the phenomena they represent. Instructional materials take on lives of their own when they become embodied in instruments, extensions of discourse practices that rule the emergence of meaning with sufficient rigor and reliability for generalized application. Clear thinking on any topic is mathematical in the sense that the manifestation of the phenomenon observed and analyzed is relatively and probabilistically independent of the particular metaphorical, numerical, and geometrical figures representing it.

Of course, under careful enough scrutiny we could come to the conclusion that the same lesson is never taught twice, just as Heraclitus said that one cannot put one's foot in the same river twice. Just as the river changes in the course of its flow, the foot changes in its cellular structure, and the meaning of the lesson may well change depending upon who is teaching it, who is learning it, or on where and when it is taught.

But if these differences are absolutely unavoidable, and are understood to encompass the whole of communication, then teaching and science are meaningless enterprises, all is difference, and meaning is dead. Not only this, but one would have to hold that the understanding of absolute difference itself disseminates into an irreconcilable plurality of incommensurable alienations because nothing could stand for anything else long enough to endure beyond the instance of its inception. Everything would be completely relative, including relativism. However, even when we observe differences in the teaching and learning of a lesson that is ostensibly the same from one situation to the next, we are observing something meaningful and communicable. As such, observations as to the limits of objectivity cannot negate the notion of objectivity; they still presume it themselves. Understanding limits provides us with ideas as to the degree of error present in any assertion of sameness. Hence, the concept of objectivity must always be contextualized by a frame of reference, and tempered by recognition of its probable and relative status.

Mathematical Models and Local Knowledge. The differences between scientific and poetic texts are no longer matters of kind; they are matters of specificity. Metaphors assert that one thing is another (the heart is a pump that circulates fluid through a system of pipes regulated by valves) in broad strokes; we live by these broad strokes (Lakoff & Johnson, 1980) in a way that lends itself to the focusing of these metaphors in measures (the heart pumps blood through the circulatory system at a particular pressure, rate and temperature) (Rothbart, 1984). What remains constant across metaphors and measures, beyond the assertion that one thing is something else (love is a rose; Mary is 5'2"), is the metaphoric twist wherein the world is reconfigured and furnished with a newly created image that moves out of the context of its birth to play a role in the dramatic game of life (Gerhart & Russell, 1984).

The metaphoric process models and leads to the clarity of mathematics and measurement by means of the dialectic. Plato "never regarded it [dialectic] as essentially different or distinct from mathematics" (Lasserre 1964, p. 28). For Plato the objects of dialogue, mathematics and learning were all embodied in the same structure of ideality. Thus, in the context of a particular metaphor, such as "love is a rose," a mathematical dialectic and dialogue that leads to the relatively invariant replication of a unit of meaning must take place. Love must manifest itself as a rose, which is to say that it must share structural features of its "system of associated commonplaces" (Black, 1962) in common with those belonging to the nature of roses. An important aspect of this commonality, this fusion of the horizons of love and roses, is the new order and organization that the conceptualization of love as a rose brings to both love and roses. Obviously, love is less a rose in the sense of being a plant, with leaves, stems, roots, flowers, scent and thorns, and that can be put in or pulled from the ground, cut, hybridized, bought, sold, and pressed, than it is in the sense of being delicate but vigorous, beautiful, inspiring, wonderful, in need of warmth, light and nurturing, potentially painful, healing (as in the medicinal use of rose hips, for instance), many-layered, etc. And a rose is not just a rose; the qualities that make a rose a rose do not exist purely in some absolute and ideal rose but emerge as an image constituted and projected by the interactions of everything historically said, written or perceived to be a rose. Besides bringing new meaning to the notion of love, our ideas as to what makes a rose a rose are reorganized through their conceptual interaction with our ideas concerning what makes love lovely.

These articulations of love in the language of roses, and vice versa, work to form a probabilistic continuum of more and less via the interaction of subject and object; the point of Heelan's (1965, 1983a, 1983b, 1985) study of measurement is that languages are virtual instruments from which actual instruments emerge. Metaphors calibrate language along a path of meaning that is relatively and probabilistically replicable across persons, times and places, and the calibration occurs through the interaction and transformation of meaning of the terms involved (Black, 1962; Ricoeur, 1977). Heelan's (1965, pp. 64-65) formula for "every well-designed measuring-process" is a model of this interaction. The formula is written

> <-> <->
> aPb where P represents the interaction characteristic of the property P, and a and b are the terms affected and so correlated by the interaction. It follows then that P founds a
> --> <->
> twofold relation: (1) aPb, which reads: 'The formal effect of P on the instrument b, enters into the definition and measurement of the property P of the object a'.
> <-- <->
> (2) aPb, which reads: 'The formal effect of P on the instrument a enters into the definition and measurement of the property P of the object b' (Heelan 1965, p. 65).

Luce and Tukey (1964) show that such a conjoint formalization of measurement parameters is essential for scientific comparisons in the behavioral sciences. Heelan goes further, asserting that all scientific and linguistic comparisons depend upon a conjoint interaction of thing and thought, object and subject, and measured and measuring that results in the emergent delineation of phenomena and the calibration of instruments. The studies of complementarity, language and measurement made by Bohr, Wheeler, and Petersen lead to the same stress on the importance of the question asked in the enframing of observations.

Rasch (1960, p. 110) remarks on the unavoidable necessity of a conjoint formalization of measurement parameters, and how this circularity is not at all vicious, but is in fact the condition of the possibility of scientific objectivity. Rasch observed that the emergence of a clearly delineated phenomenon depends upon the extent to which questions and answers share a joint order and hang together along a continuum of more and less. The conjoint model of measurement forces us to examine the questions asked for delineation of and adherence to the object of the question and answer exchange. A clear view of the thing itself is facilitated only by constantly going back to the questions and ascertaining their appropriateness and validity within the frame of reference. The circularity noted and incorporated into measurement models by Luce and Tukey, Rasch, and others therefore

> is not to be reduced to the level of a vicious circle, or even of a circle which is merely tolerated. In the circle is hidden a positive possibility of the most primordial kind of knowing. To be sure, we genuinely take hold of this possibility only when, in our interpretation, we have understood that our first, last and constant task is never to allow our fore-having, fore-sight, and fore-conception to be presented to us by fancies and popular conceptions, but rather to make the scientific theme secure by working out these fore-structures in terms of the things themselves (Heidegger, 1962, p. 195).

Wright (Wright and Masters, 1982, p. 5; Wright and Stone, 1978, pp. 10-11) interprets the necessary participation of question and answer in the creation of phenomena as mandating a focus on the frame of reference, saying that "there are no natural units[,] there are only the arbitrary units we construct and decide to use for our counting" (Wright and Masters, 1982, p. 9). Maintaining a unit of measurement is nothing if it is not a matter of making the scientific theme secure by working out the fore-structures of knowledge in terms of the things themselves, which is what takes place when scientists struggle to force data to conform to ideas, phenomena, and implicit social, economic and political orders they are passionately committed to establishing (Kuhn, 1961, 1970; Brown, 1977; Holton, 1988).

It is only sensible, then, for Luce and Tukey (1964, p. 4) to say that when an obvious order for measurement is not available, it should be arranged. The acceptance of, and willingness to work within, this state of affairs constitutes a recognition that "what is 'natural' is not 'natural' at all, here meaning self-evident for any given person that ever existed. The 'natural' is always historical" (Heidegger 1967, p. 39; translation altered). Luce and Tukey (1964, p. 4) rightly say that objections to arranging the observational situation to produce the conditions appropriate for measurement as an unethical 'tampering' with data are specious, because observations are always arranged by the effects of history and imagination, and are never self-evidently obvious.

Classical physics was able to ignore this state of affairs because of the seemingly unmediated accessibility of the data, facts, and laws of nature, resulting in what Husserl (1970) called Galileo's "fateful omission." There were very few influences that disturbed and interfered with the clear expression and unambiguous observation of natural entities. In contrast,

> The possibilities of disturbing influences which interfere with the clear expression and hence the unambiguous observation of [human] ability [and attitudes] are endless. But, if it is really the person's ability [or attitude] that we hope to measure, then it would be unreasonable not to do our best to arrange things so that it is the person's ability [or attitude] which dominates their test behavior. Indeed, isn't that what good test administration practices are for, namely, to control and minimize the intrusion of interfering influences[?] (Wright and Stone, 1978, pp. 10-11).

Of course, arranging things so that the phenomenon of interest dominates the measurement situation could hardly go so far as to remove all sources of influence, disturbance, interference and ambiguity. The stress on the frame of reference is motivated by the situation of meanings within larger contexts, which why Rasch (1960) entitled his book *Probabilistic Models...*, and is also why he wrote that "a model is not meant to be true" (Rasch 1964, pp. 24, 2, 3; 1960, pp. 37-38). Models are meant to bear meaning, to be heuristic fictions redescribing reality and guiding the conversations conducted among researchers, those contributing to research, and the researched extending the way that metaphors and analogies work in everyday language (Hesse 1970; Black 1962; Ricoeur 1977, pp. 239-246, 1981, pp. 185-190; Tracy, 1975).

The extension of language's capacity for mathematical clarity from virtual instrumentation to actual instrumentation takes place when we narrow the focus of the wider, metaphorical sense of mathematics to the more strictly quantitative sense of scientific models. Etymologically, communication and community require some common unity of meaning that can be drawn out and plotted as a story line any person competent in a language can follow. Following the direction in which an "arrow of meaning" (Ricoeur, 1981, p. 193) points is to take up a line of inquiry, a link in a chain of ideas, a train of thought. None of these linear images are actually drawn, however, without a great deal cf back and forth circularity and spiraling in directions perpendicular to, parallel to, and along the continuum. These movements take place as we try to focus our questions, to let the thing itself come into words, and to account for the potential influences, disturbances, and ambiguities that might cloud the meanings shared. Similarly, Gadamer (1989, p. 362) says that

> The essence of the question is to have sense. Now sense involves a sense of direction. Hence the sense of the question is the only direction from which the answer can be given if it is to make sense.

Some simple conditions which must be met in order to organize a conceptual framework, "place what is questioned in a particular perspective" (Gadamer, 1989, p. 362), and form a community of speakers can be abstracted from Habermas' (1979) theory of communicative competence and notion of the ideal speech situation. First, a person having greater communicative competence should always have a greater probability of communicating effectively, no matter whom this person communicates with or what particular topic is involved. Second, a topic presenting more obstacles to effective communication should have a greater probability of obstructing communication no matter who is engaged in its issues. Finally, one must be able to check on the model of discourse no matter what topic, speaker or writer is involved by investigating the extent to which what is actually said or written lives up to the two requirements for effective communication. The requirements are never met exactly, but as Ricoeur (1977, 1979, 1981, pp. 185-190, 274-296) points out under various rubrics (the rule of metaphor, the function of fiction in the shaping of reality, the theory of the heuristic fiction, and the narrative function), it is because these requirements form a frame of reference and context for the deployment of questions that meaning is able to be created and shared.

As we have already seen, something is mathematical in the wider sense of teachable, learnable and communicable only when its constitution as a phenomenon does not depend upon the particular persons or figures of speech, metaphor or geometry that happen to represent it. Habermas' theory of communicative competence enables us to refine our sense of the connection between the qualitative and quantitative senses of mathematics because of the way in which it brings the fundamental ratios of reason into view. Rationality is the proportion of communication that emerges from the interaction and complementary relation of a speaker's or writer's competence with the things communicated. It was this insight that led Bohr to emphasize the universal significance of the role of complementarity in the communication of meaning. When Bohr said that we depend upon our words, that we are hanging in language such that we cannot tell up from down, he was returning to the wider sense of mathematics for an explanation of the numbers and formulae to which the narrower quantitative sense had led. The move from the narrower to the wider senses of mathematics was prompted by the paradoxes of quantum mechanics that have made Galileo's fateful omission just as glaring in natural science as it has been in social science. Heelan's formula for "every well-designed measuring-process" given above is, then, a model for the constitution of the linguistic frame of reference on which we depend. Every speaker's communicative competence (a) interacts with the obstacles to communication presented by the topics (b) along a continuum of more and less meaning (P).

Mathematical and rational measures of meaning emerge from within this dialectic to the extent that a clear view of the amount of thing communicated is not obstructed by the particular person or topic involved. The crucial importance of additive conjoint measurement models follows from the way they model the universal complementarity of relations through which the work of enframing is accomplished. As Rasch put it,

> A person having a greater ability than another should have the greater probability of solving *any* item of the type in question, and similarly, one item being more difficult than another

one means that for *any* person the probability of solving the second item correctly is the greater one (Rasch, 1960, p. 117).

We depend upon the words organized into questions, items, and problems to communicate meaning, to say something about something to someone. This is not done blindly, but occurs through the opening up of a frame of reference and the emergence of meaningful relationships within that framework. In the same way that largely implicit questions structure and enframe observations in everyday life, science requires fairly explicit preparation of questions for the deliberate arrangement of mathematically clear and rigorous observations. Plato introduced mathematical rigor into geometry by allowing the compass and straightedge to constitute the frame of reference for geometrical questions. Within the frame of reference, the mathematically harmonious separation of figure from meaning facilitated by finely tuned instruments, represented by Rasch in his stress on the word "any" in the above passage, becomes the foremost requirement of clarity in thinking, quantitative or otherwise.

For instance, Galileo founded modern science when he based his study of gravity on an imaginary model involving a vacuum that he would never see demonstrated in practice. He was anticipated in this kind of modelling via ideality, though, by Plato, who conceived of a point as "'an indivisible line,' and a line as 'length without breadth'" (Cajori 1985: 26). These definitions make it most obvious that what is analyzed, or rather, what reveals itself through the dialectical interplay of theory and data mediated by instruments, is not the actually divisible line that is drawn out, and which is called a point, but is instead the conceptual ideality of that figurative text. In the same way, communicative competence, reading ability, consumer preferences, attitudes toward the government, and the rose-like qualities of love are all fictions that heuristically guide and dominate the way our linguistic constitution plays itself out. The ancient Greek mathematical religion of Pythagoras confused representations of things for the things themselves; the religion was thrown into crisis, with Pythagoras reportedly drowning himself, when "irrational" numbers, such as the square root of two, were discovered.

Contemporary social science is still playing out the Pythagorean crisis because it too often ignores the fundamental importance of the frame of reference. Pythagoreans were lost in unsolvable problems such as the squaring of the circle because they could not see the forest for the trees. The crisis of Pythagorean mathematics was overcome by Plato because irrational numbers live out the same conceptual existence that rational ones do. The irrationality of the square root of two, for instance, no longer threatened the heart of mathematics because the existence of this number and the line segment it represents no longer depended upon representation as a line segment of precisely drawable length or as a number that could be exactly specified. Conjoint measurement models analogously overcome the crisis of quantitative thinking in social science by first projecting mathematical fictions of what counts as a legitimate observation and following these with checks for a relatively and probabilistically credible telling of the story by the data. Attention to the epistemological importance of the work of enframing thus stands in stark contrast to the presumptuous blindness of pretending to count on observations that vary in meaning depending upon who responded, what in particular was asked, who did the asking, where the exchange took place, etc., etc.

The fact that conjoint measurement models are heuristic fictions not intended to be true does not, then, mean that they are simply one kind of statistical model among many others of roughly equivalent meaning and utility, as some in education would lead us to believe (Divgi, 1986; Wood, 1978; Whitely and Dawis, 1974). On the contrary, not just any model participates in the recovery of Galileo's fateful omission, acknowledging and extending the hermeneutically circular, dialectical, and mutually critical correlation of reader and text (Tracy, 1975) that mediates the relation of observer and observed. Models that do not implement the method of converging indices by which phenomena come into being (Ricoeur, 1976, 1981) must necessarily remain trapped in the irrelevancies of Newton's positivism (Burtt, 1954). The hidden irrational element in science that makes mathematical modeling appear so unreasonably effective may forever continue to be an unfathomable source of mystery and wonder, but this does not mean that we can relinquish ourselves to continued irresponsibility in the assignment and arithmetic manipulation of numbers. And it is just as irresponsible to abandon hope for rigor and meaning in quantitative research by focusing exclusively on non-quantitative studies as it is to ignore or downplay the mutual implication of subject and object that constitutes whatever degree of meaning we are able to create or find in numbers. Quantitative methods in social research are not going to go away simply because qualitative, ethnographic ones are more "natural," valid or human. More attention needs to be paid to the foundations of measurement, more pressure for quality data needs to be put on those who do quantitative work, and both of these tasks will be accomplished only by those who have the requisite phenomenological and hermeneutic savvy.

Metaphor as Measurement. I conducted two studies in order to test the hypothesis that metaphors provide virtual measures of meaning that can be honed into actual measures. One study asked a relatively homogenous sample of Americans to rate the meaningfulness of entailments of the metaphor "love is a rose." The other study was conducted in Niger, West Africa, in Hausa and French with a more diversified group of respondents rating the meaningfulness of the local saying "life is a mango." By pursuing the interactive complementarity of the metaphors' systems of associated commonplaces to the point of using conjoint measurement techniques to calibrate actual instrumental embodiments of those systems, both studies show that metaphorical entailments work together in such a way as to calibrate virtual instruments that do not differ in structure or function from their actual counterparts.

The first study involved the articulation of 68 entailments of the metaphor "love is a rose" and the arrangement of these along a continuum indicative of the extent to which each entailment contributes to the meaning of the metaphor. Care was taken to enter deliberately into what

Ricoeur (1981, pp. 211-212) calls the dialectic between guessing what the entailments are and how much each contributes to the structure of the metaphor's meaning, on the one hand, and the probabilistic procedures by which the guesses are validated, on the other. Items asserting that love is beautiful, delicate, healing, and in need of light, warmth, and nurturing were hypothesized to be very meaningful. Other items asserting that love is organic, colorful, fragrant, perennial, etc. were hypothesized to be moderately meaningful. Finally, other items asserting that love has a stem and leaves, and can be bought and sold were hypothesized to be least meaningful. Twenty-one items hypothesized to span the entire measurement continuum were selected as the basis for linking three different forms of a questionnaire, and were included on each form. The other 47 items were evenly distributed across the three forms, with items of similar meanings deliberately placed on separate forms so as to test specific hypotheses concerning the extent to which levels or degrees of meaning do or do not depend upon the particular turns of phrase that express them. The hypothetical order of the items is shown in Table 1.

The items were associated with a six-point rating scale, ranging from absolutely true to very true to mostly true to mostly false to very false to absolutely false. Labels of true and false were used instead of ones explicitly indicative of more and less meaning in order to present the respondents with concrete alternatives and force specific choices. The adjectives "absolutely" and "mostly" were deliberately chosen in order to provoke distinctions capable of carrying consistent weights of meaning. 36 people rated the truth or falsity of each rosy assertion about love (or lovely assertion about roses?). These data were then tested for adherence to the requirements of measurement via analysis by MSCALE (Wright, Schultz and Congdon, 1988), a computer program that performs conjoint measurement analysis (Perline, Wright, and Wainer, 1979).

The MSCALE analysis showed that three items in Group I and two in Group II did not share in and contribute toward the consistency of the data and were dropped from the analysis. The examination of the items for evidence of consistent support of the line we intend to count on to provide valid and reliable data is an important way in which respondents are given a voice in the research process. Rather than relying strictly on the researcher's expert status for the authority to say what counts, science demands a self-critical re-examination of the questions asked in the light of the responses garnered. The respondents rated these five metaphor items such that the items' positions along the continuum of more and less meaningfulness depended upon who was doing the rating. Assertions as to the delicacy, fruitfulness, and sweetness of love are apparently too ambiguous to be considered part of the structure of meanings constituting the conceptual interaction of love and roses; they are the exceptions that prove the rule, in the sense of testing it. Removing these reprobate items from the analysis by putting them on probation improves the probity of the scale but does nothing toward proving anything. These anomalous items provide a continuous proof-test of the capacity of roses and love to interact meaningfully. Perhaps these items will play key roles as entailments in another conceptual revolution in love, just as the entailments of roses did in France in the 12th century.

Love may well be delicate in some moments, like the petals of a rose, but it is also vigorous and hardy, besides also connoting some degree of roughness and indelicacy in its sexual aspect. As for being fruitful, it may not be commonly understood that the rose is a member of the same biological family as the apple, that part of the plant is technically known as the fruit, and is often referred to as an apple. Along the same lines, the fruitfulness of love may be almost entirely a matter of sexual reproduction for most people, and birth control methods have virtually destroyed any automatic association of sexuality and reproduction. Continuing in the same direction we can see that the sexual implications of the assertion that many persons can enjoy love simultaneously could easily lead some to a rating of true, and others to a rating of false, independent of the general order they find to hold across the other items.

The item asserting that love would be as sweet by any other name is a poetic reference that in fact does more to raise a question than to maintain a claim, so it may not be surprising that the respondents cannot agree on the position of this item on the scale. Finally, the item asserting that love is associated with the Virgin Mary was intended to explore the possibility of an historical and mythological connection between ancient and contemporary goddesses of love, all of whom have extensively documented associations with roses. The items concerning Venus and Aphrodite have adjacent calibration values (accounting for error), with Aphrodite having a slightly more meaningful association with love than Venus, which is perhaps indicative of some conceptual interference from the planet of the same name. Little support for the validity and reliability of the Virgin Mary item could be found in the data. Perhaps more support could be found in a culture, such as the Hispanic, that exhibits a deeper appreciation for the cult of the Virgin.

Table 2 lists the calibration values of the items. At the bottom of each page in Table 2, and in Table 3, some summary statistics are provided on the comparison of the hypothesized order and the respondents' actual order. Group I was hypothesized to be composed of the most meaningful statements on the roseness of love, Group II to be less meaningful, and Group III to be least meaningful. In fact, each pair of groups differ by a full unit of measurement (1.00). The statistical significance of differences of this size is suggested by the average errors of measurement (0.31, 0.25, and 0.35, respectively) associated with each group of items. The results of an analysis of the variance in and among the item scale values in the groups is included in Table 3. The probabilities of obtaining these variances within and between the groups by chance are minuscule ($p < .0001$), and the amount of variance accounted for by the group assignments is 60 percent. Even when individual tests are performed between each pair of the groups with the very conservative Scheffe procedure, the probability of obtaining the difference between the least meaningful and the other two groups is less than .001, and is less than .01 for the difference between the groups of items in the middle and most meaningful ranges. The statistical reliability of the scale is over 0.92, as calculated by the MSCALE program according to equations similar to those used in calculating Cronbach's alpha. Finally, the ratio of the standard deviation (1.11) to the average error (0.34) can be used to calculate the number of levels of meaning distinguished by the scale that are not due to error (Wright and Masters, 1982, p. 92); this separation index for the love is a rose scale is 4.69, meaning that there are almost five statistically distinct levels, each separated by at least three error terms from any other level, delineated by this scale.

The ranges of the three hypothetical groups or items do overlap somewhat, most noticeably between the top and middle groups. The two statements from the top group with the lowest calibration values actually come out lower on the scale than any of the items from the middle group, calibrating in the range of least meaningful assertions. These items (love dies when picked, at 0.45, and love grows in a garden, at 0.54) represent aspects of roseness that are more difficult to associate with love than was hypothesized. Conversely, two items in Group II (love has many shapes, at 2.36, and love is colorful, at 2.24) have calibrated well within the range of the most meaningful statements. If these four items are moved to their evidently more appropriate groups, the overlap in the calibration ranges of Groups I and ' is reduced from the complete encompassment of II by I to a range of 0.29 at the bottom of Group I's range and the top of Group II's. This amount of overlap is within an error of the overlap of 0.55 between Groups II and III.

Table 4 presents an analysis of the item calibration values by form of the questionnaire. The mean calibration values of the items on each form stand well within the average error of the those calibrations; this is the case both when the 21 common items are included in the calculation and when they are not. Form II ranges more toward both extremes of the scale because the one most and two least meaningful items on the scale were exclusive to this form. To say that love is beautiful is to make the most meaningful of all of the rosy assertions about love, and to say that love can be grown commercially and that dancers hold it in their teeth is to make the least meaningful assertions. Any further administration of these questionnaires would make these three items common across the forms. Form II's extreme items have associated error terms that overlap with their nearest neighbors on the scale, and the range of values produced by the items on Form

II is reduced from a high of 4.07 and low of -1.84 to a high of 3.27 and low of 0.07 when these items are ignored, a range very similar to those found on Forms I and III. Despite these differences, a one-way analysis of variance of the items grouped by form indicates a strong likelihood that the population means are not significantly different ($p > .92$): less than one percent of the variance is explained by differences in the item calibrations across the forms.

The point of creating different forms of the same instrument is to show that the mathematical structure of the variable has been delineated relatively free of interferences and disturbances introduced by particular questions or respondents. Not only does each form span very nearly the same range of the continuum, but items with similar meanings on different forms calibrate to approximately the same position on the scale, even when completely different groups of people supply the ratings. When items that would apparently seem to carry the same weight of meaning calibrate to significantly different values, opportunity for learning is created. For instance, Table 2 shows that the Form I item "love fades" carries the same weight of meaning as an item common to all of the forms, "love withers;" both scale to a value of 2.28. It is also interesting to note that the two items concerning the decorative aspect of roses applied to love (both numbered 44), which were intended to have different wordings but wound up with virtually the same wordings, have calibrated to precisely the same value (0.33), even though two completely separate groups of people have supplied the ratings. Notice, however, that saying "love grows" is more true than saying "love grows in a garden," an assertion that falls well within the range of the items rated false. On the other hand, the respondents hold that the assertion "love is tended by a gardener" belongs squarely in the middle of the moderately meaningful group. The biologically oriented items are very tightly grouped at the bottom of the scale, with items concerning the stem, petals, pressing, bush, vine, and leaves of love rating as least meaningful. Conversely, both of the items concerning the beauty of love rate as very meaningful, and the two concerning its use in perfume and fragrance rate among the least meaningful assertions.

The scale values of the 21 common items exhibit an especially marked invariance across all three of the forms and all 36 of the persons responding. The summary statistics and results of the analysis of variance shown in Table 5 indicate a very low probability ($p < .0001$) that the meaningfulness groupings have calibrated to their scale positions by chance. The Scheffe multiple comparison test, which requires larger differences between means for statistical significance than most tests, showed that all three pairings of the groups result in significant differences, with the probability of the differences occurring by chance less than .001 for the I-III pair, and less than .01 for the I-II and II-III pairs. The three groupings of the 21 common items account for virtually 80 percent of the variance. The selection of common items worked out well, with no overlap in the estimated scale values.

Table 6 provides some information on the 36 respondents. Eighteen men and 18 women from the Illinois cities of Chicago and Moline rated the metaphor's entailments on the true-false scale. The average respondent's measure on the scale was 0.24, with an error term of 0.20. The Moline respondents appear to find slightly more meaning than the Chicago respondents in the entailments of the metaphor, and women appear to find more meaning in them than men, though these differences are not statistically or substantively significant. The suggestion that such differences might exist is interesting, though, given the general impression that women and inhabitants of a small city are likely to be more romantically inclined than men and inhabitants of a major city.

The range of the estimated person measures on the scale, from 0.05 to 0.45, means that the average respondent was likely to rate the items in the most meaningful group in the category of very true, those in the less meaningful group as mostly true, and those in the least meaningful group as mostly false. The reliability of the distinctions made between the persons on the basis of their measures is indicated by the alpha coefficient of 0.89 and by the ratio of the standard deviation of the measures (0.56) to the mean error (0.20), 2.8, resulting in a value of 4.07 for Wright and Mester's (1982, p. 92) separation index. Even though the scale distinguishes four statistically distinct levels in the measures of meaning the respondents created from the entailments of the metaphor, the overall quality of the scale indicates a large degree of consensus on the meaning of the metaphor, insofar as the order of the entailments on the scale can be interpreted as meaningful evidence. This result stands in marked contrast with the assertions of those who have held that scientific precision is a result of saying exactly what is meant, and since metaphor says one thing and means another, it has no place in scientific discourse.

The results of this study suggest that research into gender and geographical differences in the meanings attributed to various metaphors and constructs would be worthwhile to undertake by means of the methods presented. This suggestion was strongly supported by another study conducted in Africa; the second study also strongly confirmed the hypothesis that the virtual measures of meaning provided by language can be focused and sharpened into actual quantitative measures. These results will be discussed at length elsewhere.

Conclusions. To date little attention has been paid to the hermeneutic that necessarily attends the logic of question and answer in educational research. However, should not a family of measurement models that allows this logic to hold sway be of special interest to fields that not only need more rigorous quantitative methods, but also, and perhaps more importantly, need to restore their ethical, political and aesthetic vision? The methodological and epistemological dilemma of social research is no longer a matter of merely recognizing that the replication of scientific objectivity in the social and human sphere is futile; it is more a matter of recognizing that the model of objectivity we were trying to replicate was mistaken, and does not exist even in the sense of an heuristic device. On the contrary, scientific understanding has a hermeneutic as thoroughly historical, cultural and linguistic as the hermeneutics of theology or literary studies do. It is just as important to introduce a softer, more relative and probabilistic form of scientific objectivity into educational research as it is to introduce a harder, less idiosyncratic and variable form of humanistic subjectivity.

When philosophy is dissociated from the pure ideality of geometrical analyses, it tends to allow its "means to assert themselves as whatever they are, and in pushing to the fore . . . they suppress what is displayed in them" (Gadamer 1980: 105), like the mechanical devices used by Sophists and Pythagoreans to manipulate and physically transform figures. The rigor of geometry "was an indispensable preliminary to the study of philosophy" (Scott 1960: 20) 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 1978: 27; original emphasis).

Husserl takes up this problem in order to begin to overcome science's "loss of meaning for life," which has come about through Galileo's "fateful omission" of the means by which nature came to be described mathematically (Husserl 1970). The "great gap which separates the new science from its classical original" was that the mathematics of modern science was seen as strictly numerical, devoid of the moral, political, aesthetic forms and ideas of Plato (Marcuse 1974: 230).

In contemporary social science, the natural fallacy and the conception of the mathematical as strictly numerical prevent the introduction of rigor and continue the alienation of human and natural by preventing recognition of the need for instruments that will allow things to communicate themselves. The instruments of social science purport to put thinking before being, and therefore succeed only in asserting themselves, suppressing the truth of whatever might have been displayed. When questions and answers cannot be arrayed into a probabilistic order of more and less, there is no support for the assertion that we can count on the numbers read off the instrument to retain a relatively fixed meaning across persons, times, and places. In such cases, the assumptions and preconceptions of the questioner go unexamined, allowing hidden or blatantly naked prejudices to masquerade as objective facts. With the arrival of conjoint measurement models, however, the only things preventing attention to what Heidegger called our first, last and constant task are our own motivations, ethics, and politics.

Scientific importance. The scientific and cultural importance of conjoint measurement in social and educational research has hardly begun to be realized. Applications in areas as diverse as paleontology, archeology, ethnography, religious history, pastoral counseling, human resources management, physical and blind rehabilitation, sports information, industrial quality control, political science, and aesthetic judgment have joined its already numerous applications in education and psychology. As the simplicity, ease of use, and practical effectiveness of conjoint measurement becomes more widely appreciated, its philosophical implications and historical consequences will be of increasing interest to educational policy makers, researchers and practitioners. By relating Heelan's hermeneutic account of natural science to the hermeneutic manifest in the application of conjoint measurement models employed in educational research, this paper joins the work of others (Andrich, 1978, 1985; Duncan, 1984; Englehard, 1989; Wright, 1984) who have also shown that conjoint measurement presents great potential for overcoming many of the problems that have traditionally plagued social and educational research.

TABLE 1

HYPOTHETICAL ORDER OF LOVE IS A ROSE QUESTIONNAIRE ITEMS

Items likely to calibrate true, with high calibration values:

From Form I:

16. Love fades.
21. Love grows.
55. Love is delicate.
57. Love grows in a garden.
61. Love is associated with Aphrodite.
65. Love inspires passion.

From Form II:

6. Love is beautiful.
11. Love produces fruit.
14. Love dies when picked.
19. Love is fragile.
23. Love needs to be fed.
52. Love can be very painful if mishandled.
60. Love is associated with the Virgin Mary.

From Form III:

18. Love is thorny.
56. Love heals.
59. Love is associated with Venus.
62. Love can be captivating in its beauty.
66. Love inspires purity.

Included on all three forms:

1. Love is a rose.
2. Love needs sunlight.
3. Love needs warmth.
15. Love blossoms.
17. Love withers.
50. Love is given to a special person.
64. Love is sweet.

TABLE 1 - Continued

Items likely to calibrate moderately true, with mid-range calibration values:

From Form I:

12. Love is organic.
29. Love has many shapes.
68. Love is thrown at stage stars.

From Form II:

31. Love is perennial.
45. Love would be as sweet by any other name.
63. Love is an ingredient in medicines.

From Form III:

5. Love is fragrant.
22. Love is tended by a gardener.
27. Love comes in different colors.
47. Love is thrown at heroes.

Included on all three forms:

4. Love is rooted in the earth.
10. Love needs water.
13. Love is more productive when pruned.
20. Love is colorful.
33. Love should be insulated from cold.
42. Love can be dug up and replanted.
43. Love can be enjoyed by many simultaneously.

TABLE 1 - Continued

Items likely to calibrate untrue, with low calibration values:

From Form I:

7. Love has a stem.
32. Love can be dissected.
36. Love is long-stemmed.
40. Love makes a pleasant bouquet.
46. Love can be spread on the floor as a carpet.
49. Love can be made into perfume.
53. Love has a biological structure.

From Form II:

8. Love has petals.
24. Love is pressed in books.
28. Love is grown commercially.
39. Love is carried in the teeth by dancers.
44. Love can be used as a decoration.*
48. Love can be put into fragrant sachets.

From Form III:

30. Love can be made into a bed.
34. Love is a bush.
35. Love is a climbing vine.
38. Love can be worn in the hair.
44. Love can be used as a decoration.*
51. Love differs from other flowers.
54. Love is a Gloire Lyonnaise.

Included on all three forms:

9. Love is leafy.
25. Love is sold in stores.
26. Love can be cross-bred.
37. Love can be arranged in a vase.
41. Love can be held in the hands.
58. Love has parts which make good tea.
68. Love attracts nectar-gathering insects.

* Item 44 was originally designed to be two different items, but one was later mistakenly edited and left nearly identical to the other; the item is counted as a common item for statistical purposes.

TABLE 2

HYPOTHETICAL ORDER OF LOVE IS A ROSE QUESTIONNAIRE ITEMS WITH CALIBRATION VALUES

Items likely to calibrate true, with high calibration values:

From Form I:	Calibration:
16. Love fades.	2.28
21. Love grows.	3.30
55. Love is delicate.	dropped
57. Love grows in a garden.	0.54
61. Love is associated with Aphrodite.	1.97
65. Love inspires passion.	2.87
From Form II:	
6. Love is beautiful.	4.07
11. Love produces fruit.	dropped
14. Love dies when picked.	0.45
19. Love is fragile.	2.16
23. Love needs to be fed.	2.85
52. Love can be very painful if mishandled.	3.27
60. Love is associated with the Virgin Mary.	dropped
From Form III:	
18. Love is thorny.	1.76
56. Love heals.	2.68
59. Love is associated with Venus.	1.40
62. Love can be captivating in its beauty.	3.07
66. Love inspires purity.	1.68
Included on all three forms:	
1. Love is a rose.	1.75
2. Love needs sunlight.	1.70
3. Love needs warmth.	2.99
15. Love blossoms.	3.14
17. Love withers.	2.28
50. Love is given to a special person.	3.52
64. Love is sweet.	2.61

Number of calibrated items: 22 Range: 4.07 to 0.45

Mean calibration value: 2.38 Sample S. D.: 0.92
Mean error: 0.31

TABLE 2 - Continued

Items likely to calibrate as moderately true, with mid-range calibration values:

From Form I:	Calibration:
12. Love is organic.	1.69
29. Love has many shapes.	2.36
68. Love is thrown at stage stars.	0.93
From Form II:	
31. Love is perennial.	1.21
45. Love would be as sweet by any other name.	dropped
63. Love is an ingredient in medicines.	1.43
From Form III:	
5. Love is fragrant.	1.43
22. Love is tended by a gardener.	1.26
27. Love comes in different colors.	1.43
47. Love is thrown at heroes.	1.08
Included on all three forms:	
4. Love is rooted in the earth.	1.02
10. Love needs water.	1.29
13. Love is more productive when pruned.	1.38
20. Love is colorful.	2.24
33. Love should be insulated from cold.	1.22
42. Love can be dug up and replanted.	1.11
43. Love can be enjoyed by many simultaneously.	dropped

Number of items calibrated: 15   Range: 2.36 to 0.93

Mean calibration value: 1.41   Sample S. D.: 0.41

Mean Error: 0.25

TABLE 2 - Continued

Items likely to calibrate as difficult to call true, with low calibration values:

From Form I:	Calibration:
7. Love has a stem.	1.21
32. Love can be dissected.	1.21
36. Love is long-stemmed.	0.28
40. Love makes a pleasant bouquet.	1.48
46. Love can be spread on the floor as a carpet.	-0.03
49. Love can be made into perfume.	0.28
53. Love has a biological structure.	1.14
From Form II:	
8. Love has petals.	0.45
24. Love is pressed in books.	0.07
28. Love is grown commercially.	-1.84
39. Love is carried in the teeth by dancers.	-0.97
44. Love can be used as decoration.*	0.33
48. Love can be put into fragrant sachets.	0.26
From Form III:	
30. Love can be made into a bed.	0.37
34. Love is a bush.	0.60
35. Love is a climbing vine.	0.90
38. Love can be worn in the hair.	-0.03
44. Love can be used as a decoration.*	0.33
51. Love differs from other flowers.	1.26
54. Love is a Gloire Lyonnaise.	0.63
Included on all three forms:	
9. Love is leafy.	0.48
25. Love is sold in stores.	-0.75
26. Love can be cross-bred.	0.35
37. Love can be arranged in a vase.	-0.05
41. Love can be held in the hands.	0.90
58. Love has parts which make good tea.	-0.09
68. Love attracts nectar-gathering insects.	0.14

Number of items calibrated: 26 Range: 1.48 to -1.84
(Item 44 counted once.)

Mean calibration value: 0.33 Sample S. D.: 0.73
Mean error: 0.35

* Item 44 was originally designed to be two different items, but one was later mistakenly edited and left nearly identical to the other; the item is counted as a common item for statistical purposes.

TABLE 3

ITEM CALIBRATIONS BY HYPOTHETICAL GROUP

	With 21 common items			Without 21 common items		
Group no.	I	II	III	I	II	III
No. items	22	15	26	15	8	18
Mean Calibration	2.38	1.41	0.33	2.29	1.42	0.42
Error	0.31	0.25	0.35	0.36	0.30	0.39
Range:						
High	4.07	2.36	1.48	4.07	2.36	1.48
Low	0.45	0.93	-1.84	0.45	0.93	-1.84

Value	Label	Sum	Mean	Std Dev	Sum of Sq	Cases
1	GROUP I: TRUE	52.3400	2.3791	.9223	17.8650	22
2	GROUP II: BORDERLINE	21.0800	1.4053	.4111	2.3660	15
3	GROUP III: FALSE	8.5800	.3300	.7349	13.5028	26
Within Groups Total		82.0000	1.3016	.7498	33.7338	63

Criterion Variable CALIBRATION VALUE

Analysis of Variance

Source	Sum of Squares	D.F.	Mean Square	F	Sig.
Between Groups	50.2473	2	25.1236	44.6858	.0000
Within Groups	33.7338	60	.5622		

Eta = .7735 Eta Squared = .5983

TABLE 4

ITEM CALIBRATIONS BY FORM

	With 21 common items			Without 21 common items		
Form number	I	II	III	I	II	III
Form length	35	33	36	15	12	15
Mean Calibration	1.38	1.24	1.29	1.43	1.12	1.30
Error	0.25	0.31	0.26	0.30	0.47	0.33
Range:						
High	3.52	4.07	3.52	3.30	4.07	3.07
Low	-0.75	-1.84	-0.75	-0.03	-1.84	-0.03

Value	Label	Sum	Mean	Std Dev	Sum of Sq	Cases
1	FORM I	21.5100	1.4340	.9827	13.5190	15
2	FORM II	13.4100	1.1175	1.7418	33.3732	12
3	FORM III	19.5200	1.3013	.8116	9.2214	15
4	COMMON ITEMS	27.5600	1.3124	1.1661	27.1954	21
Within Groups Total		82.0000	1.3016	1.1883	83.3089	63

Criterion Variable MEASURE

Analysis of Variance

Source	Sum of Squares	D.F.	Mean Square	F	Sig.
Between Groups	.6721	3	.2240	.1587	.9237
Within Groups	83.3089	59	1.4120		

Eta = .0895 Eta Squared = .0080

TABLE 5

COMMON ITEM CALIBRATIONS BY GROUP

	N	$\bar{X}$	SD	ERR	Range
Group I	7	2.57	0.70	0.20	1.70 to 3.52
Group II	6	1.38	0.44	0.17	1.02 to 1.38
Group III	8	0.16	0.49	0.22	-0.75 to 0.90
All Groups	21	1.31	0.56	0.21	-0.75 to 3.52

Value	Label	Sum	Mean	Std Dev	Sum of Sq	Cases
1	GROUP I	17.9900	2.5700	.6975	2.9188	7
2	GROUP II	8.2600	1.3767	.4418	.9757	6
3	GROUP III	1.3100	.1638	.4855	1.6500	8
Within Groups Total		27.5600	1.3124	.5550	5.5445	21

Criterion Variable MEASURE

Analysis of Variance

Source	Sum of Squares	D.F.	Mean Square	F	Sig.
Between Groups	21.6509	2	10.8254	35.1442	.0000
Within Groups	5.5445	18	.3080		

Eta = .8923 Eta Squared = .7961

TABLE 6

36 PERSON MEASURES BY SEX AND CITY

	CHICAGO	MOLINE	MEAN
WOMEN			
MEASURE	0.17	0.50	0.30
AGE	38.36	37.14	37.89
N	11	7	
MEN			
MEASURE	0.10	0.26	0.17
AGE	38.70	42.63	40.45
N	10	8	
MEAN MEASURE	0.14	0.39	0.24

References.

Ackermann, John Robert. 1985. Data, Instruments, and Theory: A Dialectical Approach to Understanding Science. Princeton: Princeton University Press.

Andrich, David. 1978. Relationships between the Thurstone and Rasch approaches to item scaling. Applied Psychological Measurement 2: 449-460.

Andrich, David. 1985. An elaboration of Guttman scaling with Rasch models for measurement. In Sociological Methodology 1985. Ed. N. B. Tuma. San Francisco: Jossey-Bass.

Baum, Robert J. 1973. Philosophy and Mathematics: From Plato to the Present. San Francisco: Freeman, Cooper & Co.

Bell, E. T. 1931. The Queen of the Sciences. Baltimore: Williams & Wilkins.

Black, Max. 1962. Models and Metaphors. Ithaca: Cornell University Press.

Bochner, Salomon. 1966. The Role of Mathematics in the Rise of Science. Princeton: Princeton University Press.

Bohr, Niels. 1983. Discussion with Einstein on epistemological problems in atomic physics. In Quantum Theory and Measurement. Ed. John A. Wheeler and W. Zurek. Princeton: Princeton University Press.

Brogden, H. E. 1977. The Rasch model, the law of comparative judgment and additive conjoint measurement. Psychometrika 42: 631-634.

Brown, Harold I. 1977. Perception, Theory and Commitment: The New Philosophy of Science. Chicago: University of Chicago Press.

Burtt, Edwin A. 1954. The Metaphysical Foundations of Modern Physical Science. New York: Harcourt, Brace & Co.

Cajori, Florian. 1985. A History of Mathematics. New York: Chelsea.

Dantzig, Tobias. 1930. Number, the Language of Science. New York: Macmillan.

Dantzig, Tobias. 1955. The Bequest of the Greeks. London: George Allen & Unwin.

Derrida, Jacques. 1978. Edmund Husserl's "Origin of Geometry": An Introduction. Trans. John P. Leavey, Jr. Stony Brook, NY: Nicholas Hayes.

Descartes, Rene. 1961. Rules for the Direction of the Mind. Trans. Laurence J. Lafleur. Indianapolis: Bobbs-Merrill.

Descombes, Vincent. 1980. Modern French Philosophy. Cambridge: Cambridge University Press.

Divgi, D. R. 1986. Does the Rasch model really work for multiple choice items? Not if you look closely. Journal of Educational Measurement 23(4): 283-296.

Duncan, Otis Dudley. 1984. Notes on Social Measurement: Historical and Critical. New York: Russell Sage Foundation.

Englehard, George, Jr. 1989. Historical views of the concept of invariance and measurement theory in the behavioral sciences. Paper presented at the Fifth International Objective Measurement Workshop, University of California, Berkeley, March.

Fisher, Ronald A. 1932. Statistical Methods for Research Workers. 4th ed. London: Oliver & Boyd.

Fleck, Ludwig. 1979. The Birth and Genesis of a Scientific Fact. Introd. Thomas Kuhn. Chicago: University of Chicago Press.

Gadamer, Hans-Georg. 1989. Truth and Method. 2d Ed. Translation revised by Joel Weinsheimer and Donald G. Marshall. New York: Crossroad.

Gadamer, Hans-Georg. 1981. Reason in the Age of Science. Trans. and intro. Frederick G. Lawrence. Studies in Contemporary German Social Thought. Ed. Thomas McCarthy, no. 2. Cambridge: MIT Press.

Gadamer, Hans-Georg. 1980. Dialogue and Dialectic: Eight Hermeneutical Studies on Plato. Trans. and intro. P. Christopher Smith. New Haven: Yale University Press.

Gadamer, Hans-Georg. 1976. Philosophical Hermeneutics. Ed. and trans. David E. Linge. Berkeley: University of California Press.

Gerhart, Mary and Allan Russell. 1984. Metaphoric Process: The [illegible] of Scientific and Religious Understanding. Foreword by Paul Ricoeur. Fort Worth: Texas [illegible] University Press.

Green, Kathy E. 1986. Fundamental measurement: A review and application of additive conjoint measurement in educational testing. Journal of Experimental Education 54(3): 141-147.

Habermas, Jurgen. 1979. Communication and the Evolution of Society. Trans. Thomas McCarthy. Boston: Beacon.

Hacking, Ian. 1983. Representing and Intervening: Introductory Topics in the Philosophy of Natural Science. Cambridge: Cambridge University Press.

Heath, Thomas L. 1931. A Manual of Greek Mathematics. Oxford at the Clarendon Press.

Heelan, Patrick. 1965. Quantum Mechanics and Objectivity: A Study in the Physical Philosophy of Werner Heisenberg. The Hague: Martinus Nijhoff.

Heelan, Patrick. 1972. Towards a hermeneutic of natural science. Journal of the British Society for Phenomenology 3: 252: 260.

Heelan, Patrick. 1983a. Natural science as a hermeneutic of instrumentation. Philosophy of Science 50(June): 181-204.

Heelan, Patrick. 1983b. Space-Perception and the Philosophy of Science. Berkeley: University of California Press.

Heelan, Patrick. 1985. Interpretation in physics: Observation and measurement. Greater Philadelphia Philosophy Consortium, March 16. Typewritten.

Heidegger, Martin. 1962. Being and Time. Trans. John Macquarrie and Edward Robinson. New York: Harper & Row.

Heidegger, Martin. 1967. What is a Thing? Trans. W. B. Barton, Jr. and Vera Deutsch. Analytic afterword by Eugene Gendlin. South Bend, IN: Regnery.

Heidegger, Martin. 1977a. Basic Writings. Ed. David F. Krell. New York: Harper & Row.

Heidegger, Martin. 1977b. The Question Concerning Technology and Other Essays. Trans. William Lovitt. New York: Harper & Row.

Hesse, Mary. 1970. Models and Analogies in Science. Notre Dame: University of Notre Dame Press.

Hesse, Mary. 1972. In defence of objectivity. Proceedings of the British Academy 58: 275-292.

Holton, Gerald. 1988. Thematic Origins of Scientific Thought. Revised ed. Cambridge: Harvard University Press.

Husserl, Edmund. 1970. The Crisis of European Sciences and Transcendental Phenomenology. Trans. David Carr. Evanston: Northwestern University Press.

Kline, Morris. 1980. Mathematics: The Loss of Certainty. New York: Oxford University Press.

Kuhn, Thomas S. 1961. The function of measurement in modern physical science. Isis 52(168): 161-193.

Kuhn, Thomas. 1970. The Structure of Scientific Revolutions. 2d ed. Chicago: University of Chicago Press.

Lakoff, George and Mark Johnson. 1980. Metaphors We Live By. Chicago: University of Chicago Press.

Lasserre, Francois. 1964. The Birth of Mathematics in the Age of Plato. London: Hutchinson.

Laudan, Larry. 1977. Progress and Its Problems. Berkeley: University of California Press.

Levi-Strauss, Claude. 1963. Structural Anthropology. New York: Basic Books.

Luce, R. Duncan and John W. Tukey. 1964. Simultaneous conjoint measurement: A new kind of fundamental measurement. Journal of Mathematical Psychology 1(1): 1-27.

Lynch, Michael. 1985. Discipline and the material form of images: An analysis of scientific visibility. Social Studies of Science 15(1): 37-66.

Marcuse, Herbert. 1974. On science and phenomenology. In Positivism and Sociology. Ed. Anthony Giddens. London: Heinemann.

Miller, G.A. 1921. Historical Introduction to Mathematical Literature. New York: The Macmillan Company.

Perline, Richard, Benjamin D. Wright and Howard Wainer. 1979. The Rasch model as additive conjoint measurement. Applied Psychological Measurement 3(2): 237-255.

Petersen, Aage. 1968. Quantum Physics and the Philosophical Tradition. Cambridge: MIT Press.

Rasch, Georg. 1960. Probabilistic Models for Some Intelligence and Attainment Tests. Copenhagen: Danmarks Paedogogiske Institut; reprint, 1980, with Foreword and Afterword by Benjamin D. Wright, Chicago: University of Chicago Press. (page references in text are to reprint edition).

Rasch, Georg. 1964. Objective comparisons. Lecture given at the UNESCO seminar, Voksenasen, Oslo, Norway. Mimeographed.

Ricoeur, Paul. 1974a. The Conflict of Interpretations: Essays in Hermeneutics. Ed. Don Ihde. Trans. Willis Domingo et al. Evanston: Northwestern University Press.

Ricoeur, Paul. 1974b. Social and Political Essays. Ed. David Stewart and Joseph Bien. Trans. Donald Siewert et al. Athens, Ohio: Ohio University Press.

Ricoeur, Paul. 1976. Interpretation Theory: Discourse and the Surplus of Meaning. Fort Worth: Texas Christian University Press.

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

Ricoeur, Paul. 1981. Hermeneutics and the Human Sciences: Essays on Language, Action and Interpretation. Ed., trans. and intro. by John B. Thompson, with a response from the author. Cambridge: Cambridge University Press.

Rothbart, Daniel. 1984. The semantics of metaphor and the structure of science. Philosophy of Science 51: 595-615.

Slavkov, Svetoslav. 1973. Mathematics, knowledge, technological advancement. In Science, Technology, Man. Sofia: Bulgarian Academy of Sciences.

Toulmin, Stephen. 1953. The Philosophy of Science: An Introduction. New York: Harper & Row.

Toulmin, Stephen. 1982. The construal of reality: Criticism in modern and postmodern science. Critical Inquiry 9(September): 93-111.

Tracy, David. 1975. Blessed Rage for Order: The New Pluralism in Theology. Minneapolis: The Winston-Seabury Press.

Wheeler, John A. 1983. Law without law. In Quantum Theory and Measurement. Ed. John A. Wheeler and W. Zurek. Princeton: Princeton University Press.

Wheeler, John A. and W. Zurek. Ed. 1983. Quantum Theory and Measurement. Princeton: Princeton University Press.

Whitely, Susan E. and Rene V. Dawis. 1974. The nature of objectivity with the Rasch model. Journal of Educational Measurement 11(2): 163-178.

Wigner, Eugene P. 1960. The unreasonable effectiveness of mathematics in the natural sciences. Communications on Pure and Applied Mathematics 13: 1-14.

Wilder, Raymond L. 1965. Introduction to the Foundations of Mathematics. New York: John Wiley & Sons.

Wood, Robert. 1978. Fitting the Rasch model: A heady tale. British Journal of Mathematical and Statistical Psychology 31: 27-32.

Wright, Benjamin D. 1968. Sample-free test calibration and person measurement. In Proceedings of the 1967 Invitational Conference on Testing Problems. Princeton: Educational Testing Service.

Wright, Benjamin D. 1984. Despair and hope for educational measurement. Contemporary Education Review 3(1): 281-288.

Wright, Benjamin D. 1985. Additivity in psychological measurement. In Measurement and Personality Assessment. Ed. Edward Roskam, 101-111. North Holland: Elsevier Science Publishers.

Wright, Benjamin D. and Geofferey Masters. 1982. Rating Scale Analysis. Chicago: MESA Press.

Wright, Benjamin D., Matthew Schultz and Richard Congdon. 1988. MSCALE: A Rasch program for ordered categories. Chicago: MESA Press.

Wright, Benjamin D. and Mark Stone. 1979. Best Test Design. Chicago: MESA Press.