Psychological Bulletin
1986, Vol. 99, No. 2, 166-180

0033-2909/86/$00.75

Measurement: The Theory of Numerical Assignments

Louis Narens
University of California, Irvine

R. Duncan Luce
Harvard University

In this article we review some generalizations of classical theories of measurement for concatenation (e.g., mass or length) and conjoint structures (e.g., momentum of mass–velocity pairs or loudness of intensity–frequency pairs). The earlier results on additive representations are briefly surveyed. Generalizations to nonadditive structures are outlined, and their more complex uniqueness results are described. The latter leads to a definition of scale type in terms of the symmetries (automorphisms) of the underlying qualitative structure. The major result is that for any measurement onto the real numbers, only three possible scale types exist that are both rich in symmetries but not too redundant: ratio, interval, and another lying between them. The possible numerical representations for concatenation structures corresponding to these scale types are completely described. The interval scale case leads to a generalization of subjective expected-utility theory that copes with some empirical violations of the classical theory. Partial attempts to axiomatize concatenation structures of these three scale types are described. Such structures are of interest because they make clear that there is a rich class of nonadditive concatenation and conjoint structures with representations of the same scale types as those used in physics.

Many scientists and philosophers are well aware of what the physicist E. P. Wigner in 1960 called "the unreasonable effectiveness of mathematics in the natural sciences." Some, like Wigner, have remarked on it; a few, like the ancient philosopher Pythagoras (c. 582–500 B.C.) have tried to explain it. Today as throughout much of history, it is still considered a mystery. There is, however, a part of applied mathematical science that is slowly chipping away at a portion of the mystery. This subfield, usually called "measurement theory," focuses on how numbers enter into science. Part of the field searches for rules—axioms—that allow one to assign numbers to entities in such a way as to capture their empirical relations numerically. Another part attempts to use such qualitative axioms to understand, to some degree, the nature and form of a variety of empirical relations among various dimensions. Such relations, when stated numerically, are commonly called "laws." In recent times, a few leading mathematicians, philosophers, physicists, statisticians, economists, and psychologists have developed new processes for measurement. This work has resulted in the detailed mathematical development of new structures, has provided scientists with a greater understanding of the range of mathematical structures they are likely to encounter and use in their science, and has generated some long-lasting controversies that are only now beginning to be resolved. For surveys that go into far more technical detail than this article, see Krantz, Luce, Suppes, and Tversky (1971, in press), Narens (1985), Pfanzagl (1968, 1971), and Roberts (1979).

This work was supported in part by National Science Foundation Grant IST-83-05819 to Harvard University.

Correspondence concerning this article should be addressed to R. Duncan Luce, Department of Psychology and Social Relations, William James Hall, Harvard University, 33 Kirkland Street, Cambridge, Massachusetts 02138.

Origins of Measurement Theory

Empirical Structures for Concatenations

The origin of modern measurement theory can be traced back at least to the investigations in the late 19th century by H. v. Helmholtz, the eclectic physician-physicist, into the formal nature of certain basic physical attributes, such as mass and length, which he recognized as having the same intrinsic mathematical structure as the positive real numbers together with addition and their natural order $\geq$. We denote this system by $\langle \mathrm{Re}^+, \geq, + \rangle$. In such cases, one can observe a natural empirical ordering relation, $\succsim$, over a set of objects, where the order reflects qualitatively the degree or amount of the to-be-measured attribute that is exhibited by the objects. One can also find a natural empirical operation, $\circ$, that combines any two objects exhibiting the attribute into a composite object that also exhibits the attribute. For example, for mass one can use an equal-arm pan balance in a vacuum to establish the order. (To be sure, it is rare now to order masses in this way, but conceptually such a procedure underlies mass measurement.) When two objects, x and y, are placed in separate pans and the balance fails to tilt, they are said to exhibit mass to the same degree, that is to be equivalent in mass, which is written $x \sim y$. Otherwise, the object in the pan that drops, say x, is said to have the greater mass, which is written $x \succ y$. Placing two objects x and y in the same pan constitutes the operation of combining, and the result is denoted $x \circ y$. In the abstract model, the combining operation goes under the generic name of *concatenation*. If we let X denote the set of all objects under consideration, including all the combinations that can be formed using $\circ$, then the potential observations from the pan balance yield the mathematical structure $\mathfrak{X} = \langle X, \succsim, \circ \rangle$. We call this a *qualitative* structure, whereas a possible representing structure such as $\langle \mathrm{Re}^+, \geq, + \rangle$ is called a *numerical (representing)* structure.

One reason for studying the abstract nature of such measurement is that the same mathematical system can apply to a wide variety of attributes. We have already mentioned that the structures under consideration serve as a basis for measuring a number of the basic physical quantities: mass, length, duration, and charge. Less obvious (see below) is that much the same structure underlies the measurement of probability.

Additive Representations

Helmholtz (1887) stated physically plausible assumptions about the structure—assumptions about $\succsim$, about $\circ$, and about their interplay—and showed that when the assumptions are true, measurement can be carried out in the following sense: There exists a mathematical mapping, φ, called a *homomorphism,* from X into positive real numbers such that for each x and y in X, (a) $x \succsim y$ if and only if $\varphi(x) \geq \varphi(y)$, and (b) $\varphi(x \circ y) = \varphi(x) + \varphi(y)$. We usually say that "under the mapping φ, the ordering relation maps $\succsim$ into $\geq$ and the qualitative operation $\circ$ into $+$." Such homomorphisms of $\mathcal{X}$ into $\langle \mathrm{Re}^+, \geq, + \rangle$ are called *additive representations.* (For a formal statement of the concept of homomorphism, see Appendix 1.) Equally important, he showed that such structure-preserving measures are relatively unique: Any two differ only by a numerical multiplicative factor, and multiplying any one measure by a positive numerical factor yields another measure. These facts are often summarized by saying the measurement is unique once a unit has been selected. The more contemporary summary statement is that the set of all such homomorphisms forms a ratio scale.[1] Such a complete description of the uniqueness of the representation is called a *uniqueness theorem.*

Axioms for Extensive Quantities

In 1901 Hölder, a mathematician, published an improved version of the theory in which, among other things, he introduced the highly important concept of an Archimedean ordered group. In this work he made significant use of an axiom, dating back to the Greek mathematician Archimedes (d. 212 BC), which captures the idea of commensurability within a physical attribute by asserting that no object is infinitely larger than another for any physical attribute. Archimedes had introduced it, in part, to provide a more rigorous basis for the notion of a continuum and, in part, to avoid some of the paradoxes described by the philosopher Zeno. In our notation this property may be formulated as follows. For the sake of concreteness, consider the measurement of length for a set X of rods for which the ordering $\succsim$ is determined by placing two rods side by side and observing which spans the other, and concatenation is determined by placing two rods end to end to form another rod. For each rod x, find another rod, say x_1, equivalent (in length) to x. Then find a rod-x_2 equivalent to $x_1 \circ x$, and another x_3 equivalent to $x_2 \circ x$, and so on. The sequence $x_1, x_2, \ldots, x_n$ is called a *standard sequence* based on x. The Archimedean axiom asserts that for any two rods x and y, there is some member x_n of the standard sequence based on x that is larger than y. Or put another way, every bounded subsequence of a standard sequence is finite.

In addition to the Archimedean axiom, Hölder assumed $\mathcal{X} = \langle X, \succsim, \circ \rangle$ satisfies five other properties closely resembling the following.

1. *Weak ordering.* The relation $\succsim$ is transitive ($x \succsim y$ and $y \succsim z$ imply $x \succsim z$ for all x, y, z in X) and connected (either $x \succsim y$ or $y \succsim x$ holds for all x, y in X).
2. *Monotonicity.* The ordering and operation interlock in such a way that the concatenation of objects preserves the ordering; that is, for all x, y, w, z in X, if $x \succsim y$ and $z \succsim w$, then $x \circ z \succsim y \circ w$.
3. *Restricted solvability.* For each x, y in X, if $x \succ y$, there exists some z in X such that $x \succ y \circ z$. (This together with the other axioms implies the existence of arbitrarily small objects.)
4. *Positivity.* All objects combine to form something larger than either of them alone; that is, for all x, y in X, both

$$x \circ y \succ x \quad \text{and} \quad x \circ y \succ y.$$

5. *Associativity.* If one is combining three or more objects, it matters not at all how the grouping by pairs occurs so long as their order is maintained; that is, for all x, y, z in X,

$$x \circ (y \circ z) \sim (x \circ y) \circ z.$$

An Archimedean structure satisfying Properties 1–5 is referred to as *extensive,* and using Hölder's method each such structure can be shown to have a representation, exactly like Helmholtz's, into the ordered set of positive real numbers with addition.

It should be noted that the axioms are of two quite distinct types. Axioms 1, 2, 4, 5, and the Archimedean property must hold if an additive representation exists, that is, they are *necessary* conditions given the representation. Axiom 3 is said to be *structural* because it limits our attention to a subset of structures possessing additive representations.

Refinements: Difference Sequences and Partial Operations

Throughout this century, Hölder's axiomatization has been refined and generalized. For example, by recasting the Archimedean axiom in terms of difference sequences satisfying the recursive relation $x_{i+1} \circ u \sim x_i \circ v$ for some $v \succ u$, Roberts and Luce (1968) formulated necessary and sufficient conditions for an additive representation, and Narens (1974) showed that the Archimedean axiom can be dropped if one is willing to permit additive representations into a generalization of the real number system called the nonstandard real numbers. A particularly important modification for measurement was the generalization to concatenations that are not necessarily defined for every pair of objects (Luce & Marley, 1969). There are at least two good reasons for modifying the theory to deal with such partially defined operations. One is that it is usually impractical to concatenate arbitrarily large objects—pan balances collapse, and rods con-

[1] The term *scale* is used loosely in the literature and with much ambiguity and imprecision. Many authors, for example, refer to the usual set of representations for length as "a ratio scale for length" and speak of "the scale type of length measurement as being ratio" while simultaneously referring to individual representations as "scales," as in "the meter scale for length." We have chosen to *disambiguate* by calling the entire set of representations a "scale" and by using the term *representation* for the other use of "scale." Within this usage, concepts like "a ratio scale for . . ." and "the scale type of . . ." are sensible and retain their usual meanings.

catenate properly only on flat platforms which necessarily are bounded. Another is that some important systems are inherently bounded from above and do not—even in theory—permit unlimited concatenation. Qualitative probability is one example. Here uncertain events are ordered by a relation of "more likely than," and the union of disjoint events is taken to be concatenation, that is, if A and B are events with $A \cap B = \varnothing$, then $\circ$ is defined to be $A \circ B = A \cup B$. Measurement in this situation consists in finding a function P from uncertain events into the closed unit interval such that P preserves the "more likely than" relation, and for disjoint A and B,

$$P(A \circ B) = P(A \cup B) = P(A) + P(B).$$

In the literature, the term *extensive* is often applied to the generalization where not all concatenations are defined as well as to closed operations, as in Hölder's original system.

Nature of Fundamental Measurement

1940 Commission Report: Only Extensive Measures

With the successful axiomatization of extensive structures and the recognition of their importance for the foundations of physics, a curious debate ensued during the 1920s and 30s about what else is measurable. Some philosophers of physics—especially Campbell (1920, 1928) but also Bridgman (1922, 1931) and later Ellis (1966)—expounded the position that measurement from first principles is necessarily extensive in character. Campbell referred to scales resulting from such measurements as "fundamental," all else being "derived." Thus, momentum, density, and all other physical measures whose units can be expressed as products of powers of the fundamental units of mass, length, time, temperature, and charge were treated as derived. Although these derived measures were clearly a crucial part of the total measurement structure of physics, especially as formulated in dimensional analysis, no very careful analysis was provided of them. They together with a basis of extensive measures form the finite dimensional vector space of physical measures that is routinely invoked in dimensional analysis. However, this vector space was not developed from entirely qualitative observations; rather it was postulated as descriptive of the way numerical physical measures interlock. Ellis, in particular, clearly understood that something more was needed, and although he hinted at the solution, he failed to work it out.

At the same time, psychologists and economists were pursuing other approaches to measurement that more or less explicitly ran afoul of the dictum that fundamental measurement rests on associative, monotonic operations of combination. The debate reached its intellectual nadir with the 1940 Final Report of a Commission of the British Association for Advancement of Science (Ferguson et al., 1940) in which a majority declared fundamental measurement in psychology to be impossible because no such empirical operations could be found. Campbell, a member of the Commission and probably a major force in its creation 8 years earlier, wrote, "Why do not psychologists accept the natural and obvious conclusion that subjective measurements of loudness in numerical terms . . . are mutually inconsistent and cannot be the basis of measurement?"

Stevens's Reply: Scale Type, Not Addition

Stevens, whose work on loudness measurement with Davis in 1938 was, in part, at issue, was independently considering the same question in a series of discussions in the late 1930s with a distinguished group of scientists and philosophers: G. D. Birkhoff, R. Carnap, H. Feigl, C. G. Hempel, and G. Bergmann. Out of this arose his now widely accepted position that a key feature of measurement is not only the empirical structure and its representation, but the degree of uniqueness of the representation as is reflected in the group of transformations that take one representation into another. In contrast to Campbell, Stevens claimed that the nature of the transformations taking one representation into another was the important feature of the representation, not the particular details of any axiomatization of it.

In his 1946 and 1951 publications Stevens singled out four groups of transformations on the real or positive real numbers as relevant to measurement: one-to-one, strictly monotonic increasing, affine, and similarity (see Table 1). And he introduced the corresponding terms of *nominal, ordinal, interval,* and *ratio* to refer to the families of homomorphisms, or scales, related by these groups. Later he added a fifth group, the power group

Table 1
Measurement Scales

Transformations of R	Scale	R
$x \to x$	Absolute	Re or Re^+
$x \to k^n$ (k fixed and positive, n ranges over integers)	Discrete ratio	Re or Re^+
$x \to rx$ (r ranges over positive reals)	Ratio	Re or Re^+
$x \to k^n x + s$ (k fixed and positive, n ranges over integers, s ranges over reals)	Discrete interval	Re
$x \to sx^{k^n}$ (k fixed and positive, n ranges over integers, s ranges over reals)	Log discrete interval	Re^+
$x \to rx + s$ (r ranges over positive reals, s ranges over reals)	Interval	Re
$x \to sx^r$ (r and s range over positive reals)	Log interval	Re^+
$x \to f(x)$ (f ranges over strictly increasing functions from R onto R	Ordinal	Re or Re^+
$x \to f(x)$ (f ranges over one-to-one functions from R onto R)	Nominal	Re or Re^+

Note. Suppose $\mathfrak{X} = \langle X, \succsim, S_1, \ldots, S_n \rangle$ and $\mathfrak{R} = \langle R, \geq, R_1, \ldots, R_n \rangle$ are relational structures, R = Re or $R = Re^+$, and $\mathcal{S}$ is the set of representations of $\mathfrak{X}$ into $\mathfrak{R}$. ($\mathcal{S}$ is called the *scale from* $\mathfrak{X}$ *into* $\mathfrak{R}$.)

($x \rightarrow sx^r$, $s > 0$, $r > 0$), applicable only to measurement in the positive reals, and he referred to the corresponding scale as *log-interval*. As late as 1959 he remarked about this latter scale that "apparently it has never been put to use," which as we shall see reflects a common misunderstanding of classical physics, which in fact is full of log-interval scales that are conventionally treated as ratio scales by making specific choices for the exponents.

Although these groups of transformations played an important role in geometry and physics and seemed to encompass much of what was then known about measurement structures, Stevens offered no argument as to why these and not others should arise, and thus his analysis was more descriptive than analytical. By the 1950s it began to be clear that there are measurement structures that do not fit the scheme. As we shall see below, considerable progress toward understanding this question has been made in the past 4 years.

Having characterized scales by the type of transformation involved, Stevens went on to emphasize that scientific propositions (he was especially concerned about statistical ones) formulated in terms of measured values must exhibit invariance of meaning under the admissible transformations characterizing the scale type. As Luce (1978) showed, this concept of a meaningful proposition was a generalization of the familiar assumption in dimensional analysis that physical laws must be dimensionally invariant under changes of units. A full understanding of the concept of meaningful scientific proposition still remains a challenge. It is by no means clear what the circumstances are for which invariance under admissible transformations is an adequate criterion for meaningfulness, nor is it known what other criteria should be used when it is not. However, these involved issues are a matter for another article (Narens & Luce, in press).

Stevens's second thrust was to devise an empirical procedure for the measurement of subjective scales in psychophysics that does not presuppose an associative operation. The method, which he dubbed "magnitude estimation," has been moderately widely used because it produces quite systematic results. Nevertheless, it has proved extremely difficult to defend his assumption that the method of magnitude estimation actually results in ratio scales. Although he recognized more than anyone else at the time the significance of scale type in contrast to the particular structures exhibiting it, he seemed not to appreciate that, in fact, the concept of scale type is a theoretical one that can only be formulated precisely in terms of an explicit axiomatic model of an empirical process. He failed to acknowledge that it takes more than one's intuitions to establish that a measurement process, such as magnitude estimation, leads to a ratio scale.

Early Alternatives to Extensive Measurement

At about the same time and continuing through the next two decades, others were working on alternative measurement axiomatizations that accorded better with Stevens's view of the scope of measurement than with those of the British philosophers and physicists. Four of these developments are worth mentioning. Beginning as early as Hölder (1901), difference measurement has been axiomatized. Here one has an ordering of pairs of elements and the representation is as numerical differences or absolute values of differences (e.g., see Krantz et al., 1971, Chap. 4). Because these structures are typified by line intervals identified by their end points, it is clear that they can readily be reduced to extensive measurement, and so they were not really considered an important departure from the dictum that fundamental measurement is equivalent to extensive measurement.

The second was the investigation into structures having an operation that is monotonic with respect to the ordering but that is neither positive nor associative. In particular, Pfanzagl (1959) axiomatized structures that satisfied the condition of bisymmetry $[(x \circ y) \circ (u \circ v) \sim (x \circ u) \circ (y \circ v)]$, which is a generalization of associativity. He showed that such structures have a linear representation of the form $\varphi(x \circ y) = a\varphi(x) + b\varphi(y) + c$, where a and b are positive. When $\circ$ is also idempotent ($x \circ x \sim x$, for all x), then $c = 0$ and $b = 1 - a$ and the model is one for any process of forming weighted means. An important physical example is the temperature that results when two gases of different temperatures are mixed in fixed proportions. In addition, of course, averaging is important throughout the social sciences.

The third development, which was conceptually closely related to this although technically quite different in detail, was the earlier axiomatization by Von Neumann and Morgenstern (1947) of expected utility. Here the operation was, in essence, a weighting with respect to probabilities of a chance event and its complement. Strictly speaking, this is a form of derived measurement because numbers (probabilities) are involved in the underlying structure; however, by the mid-1950s purely qualitative theories were developed, of which the most famous is that of Savage (1954). The resulting large literature on this topic has, almost without exception, led to interval scale representations of some form of averaged utilities.

Conjoint Measurement

Perhaps the clearest demonstration of nonextensive structures for which interval-scalable, fundamental measurement is possible was the creation in the 1960s of the theory of additive conjoint measurement. Although the earlier examples had convinced many specialists that the scope of fundamental measurement is broader than Campbell had alleged, it was only with the introduction of conjoint measurement—with its simple techniques and its possible applicability throughout the social sciences as well as physics—that this view became widely accepted. A conjoint structure simply consists of an ordered structure that can be factored in a natural way into two (or more) ordered substructures. Typical examples of such structures are: the ordering by mass of objects characterized by their volume and density; the loudness ordering provided by a person for pairs of sounds, one to each ear; and the preference ordering provided by an animal for amounts of food at certain delays.

Observe two things about the above examples. First, the factorizable orderings are very closely related to the concepts of trade-offs and indifference curves that are widely used throughout science: in each case, the equivalence part of the ordering describes the trade-off between the factors that maintains at a constant value the amount of the attribute in question, be it mass, loudness, or preference. Second, no *empirical* concatenation operation is involved in a conjoint structure. Yet, as Debreu (1960) showed by using a mix of algebraic and topological assumptions and as Luce and Tukey (1964) showed using weaker and entirely algebraic assumptions, such structures can sometimes be represented multiplicatively on the positive real numbers.

More formally, assume that there are two factors, and let A denote the set of elements forming the first one and P those forming the second one. Thus, the set $A \times P$, which is composed of all ordered pairs (a, p) with a any element in A and p any element in P and is called the Cartesian product of A and P, is the set of objects under consideration. The set $A \times P$ is assumed to be ordered by the attribute in question. Let "$\succsim$" denote this ordering. For example, if A consists of various possible amounts of a food and P consists of the various possible delays in receiving the food, then an attribute of interest is the preference of some animal (or breed) for various (amount, duration) pairs. Thus $(a, p) \succsim (b, q)$ in this case means that the amount a at duration p is preferred or indifferent to the amount b at duration q. The interesting scientific questions are: What properties do we find (or expect) $\succsim$ to satisfy, and are these such that they lead to a nice numerical representation of the data?

The two most basic assumptions often made about $\succsim$ are that it is a weak order (see *Axioms for Extensive Quantities* and Appendix 2) and that it exhibits a form of monotonicity that, in this context, is called *independence.* One important consequence of independence is that the order $\succsim$ induces a unique order on each of the factors. What independence says is that if the value of one factor is held fixed, then the ordering induced by $\succsim$ on the other factor does not depend on the value selected for the fixed one, or put more formally, for all a and b in A and p and q in P,

$$(a, p) \succsim (b, p) \quad \text{if and only if} \quad (a, q) \succsim (b, q),$$

and

$$(a, p) \succsim (a, q) \quad \text{if and only if} \quad (b, p) \succsim (b, q).$$

Note that in the first statement, the value from the second factor, P, is the same on both sides of an inequality, whereas in the second statement the fixed value is from the first factor, A. The orderings induced in this fashion on A and P are denoted, respectively, $\succsim_A$ and $\succsim_P$ and are defined by

$$a \succsim_A b \quad \text{if and only if for some (and so for all) } p,$$
$$(a, p) \succsim (b, p),$$

and

$$p \succsim_P q \quad \text{if and only if for some (and so for all) } a,$$
$$(a, p) \succsim (a, q).$$

It is easy to verify that they are weak orders if $\succsim$ is.

On the assumption of weak ordering and independence, the next question is under what additional conditions do there exist real-valued mappings ψ_A on A and ψ_P on P and a function F of two real variables that is strictly monotonic in each such that for all a and b in A and p and q in P,

$$(a, p) \succsim (b, q) \quad \text{if and only if}$$
$$F[\psi_A(a), \psi_P(q)] \geq F[\psi_A(b), \psi_P(q)].$$

The two ψ functions are, in some sense, measures of the two components of the attribute, and F is the rule that describes how these measures trade off in measuring the attribute.

The first case to be studied in detail was the one of interest in classical physics, namely, the one for which the ψs map onto the positive real numbers and F displays a multiplicative trade-off so that

$$(a, p) \succsim (b, q) \quad \text{if and only if}$$
$$\psi_A(a)\psi_P(p) \geq \psi_A(b)\psi_P(q).$$

Moreover, the representations form a log-interval scale (see Table 1), which means that for each positive α and β, $\alpha(\psi_A\psi_P)^\beta$ is an equally good representation and any two multiplicative representations are so related.

In psychology and economics a different, but equivalent, representation is more usual; it is additive rather than multiplicative and is defined on all of the real numbers. This representation is obtained simply by taking the logarithm of $\psi_A\psi_P$ in the above multiplicative representation. Because of this, in the social sciences the qualitative theory is usually referred to as "additive conjoint measurement" (even when a multiplicative representation is being used), and we follow this practice in the remainder of this article.

Recoding Conjoint Structures as Concatenation Ones

The proofs of the original conjoint measurement theorems, although correct, were not especially informative and, in particular, failed to make clear that the problem could be reduced mathematically (although not empirically) to extensive measurement. This was established first by Krantz (1964), who defined an operation on $A \times P$, and later by Holman (1971), who defined an operation on just one component. The latter construction has the advantage of generalizing to nonadditive structures. Suppose, for the moment, that the structure is sufficiently "regular" (e.g., continuous) so that any equation of the form $(a, p) \sim (b, q)$ can be solved for the fourth element when the other three are specified. This condition is called *unrestricted solvability.* Turning to Figure 1, fix a_0 in A and p_0 in P, and consider any a and b in A such that, in terms of ordering $\succsim_A$ induced by $\succsim$ on the A component, $a \succ_A a_0$ and $b \succ_A a_0$. The goal is to find a way to "add" together the "intervals" from a_0 to a and from a_0 to b. The strategy is to map the a_0 to b interval onto a comparable interval on the second component that begins at p_0, and then to map the latter interval back onto the first factor, but this time with a as its starting point. The map to the second factor is achieved by solving for the element called $\pi(b)$ in the equation $(a_0, \pi(b)) \sim (b, p_0)$. And the return mapping is achieved by solving for the element called $a \circ b$ in the equation $(a \circ b, p_0) \sim (a, \pi(b))$. What Holman discovered was that a necessary condition for the conjoint structure to be additive is for this *induced concatenation* operation $\circ$ to be associative. This, in turn, is equivalent to the following property, called the *Thomsen condition,* holding throughout the conjoint structure: whenever both $(a, r) \sim (c, q)$ and $(c, p) \sim (b, r)$ hold, then so does $(a, p) \sim (b, q)$. In essence, this says that the common terms c and r cancel out, as is true of the corresponding simple additive equations involving real numbers.

A further condition, an Archimedean one, is also needed in order to prove the existence of an additive representation. Basically that axiom simply says that the induced operation meets the usual Archimedean property of extensive measurement, although it can be stated directly in terms of $\succsim$ without reference to the operation. So, in sum, conditions that are sufficient to

construct an additive representation of a conjoint structure are: weak ordering, independence, the Thomsen condition, unrestricted solvability, and the Archimedean property. What Krantz (1964) and Holman (1971) did was to show that, despite the fact that there is no empirical operation visible in an additive conjoint structure, the trade-off formulated in that structure can be recast as an equivalent associative mathematical operation. This allowed the earlier representation theorem for extensive structures to be used to prove the existence of an additive conjoint representation. This construction is such that it can actually be mimicked in practice by constructing standard sequences and using these to approximate, within a specified error, the desired measure. In the early 1970s such constructions were carried out for loudness by Levelt, Riemersma, and Bunt (1972) and by Falmagne (1976).

Generalizations: Restricted Solvability and Nonassociativity

Since the early 1960s, many variants of extensive and additive conjoint measurement have been used by scientists in a number of fields. We are not able to go into the details here, but the contributions of J.-C. Falmagne, P. C. Fishburn, D. H. Krantz, R. D. Luce, F. S. Roberts, P. Suppes, and A. Tversky deserve special note, because they repeatedly emphasized the need to understand explicitly how measurement arises in science and clearly demonstrated its potency in a number of theoretical and experimental domains.

The original theory of additive conjoint measurement and its reduction to extensive measurement was quickly seen to be too restrictive in two senses. First, in many social science situations involving trade-offs—even ones with "continuous" factors—unrestricted solvability fails to hold. For example, the loudness of a pure tone depends both on signal intensity and frequency (which is the reason for loudness as well as gain controls on an amplifier), but the limits on human hearing are such that it is not always possible to match in loudness a given tone by adjusting the frequency of another tone of prescribed intensity. (The reasons for this have to do with the processing limits of the human ear.) What does hold, however, is a form of "restricted solvability," which says, for example, that with b in A and p and q in P given, then there is an element a in A that solves the equivalent $(a, p) \sim (b, q)$ provided that we know there exist elements a' and a'' in A such that $(a', p) \succ (b, q) \succ (a'', p)$. So, for example, letting the first component be the intensity of a tone and the second its frequency if (b, q) is a given tone, and p is a given frequency then the question is whether there is an intensity a so

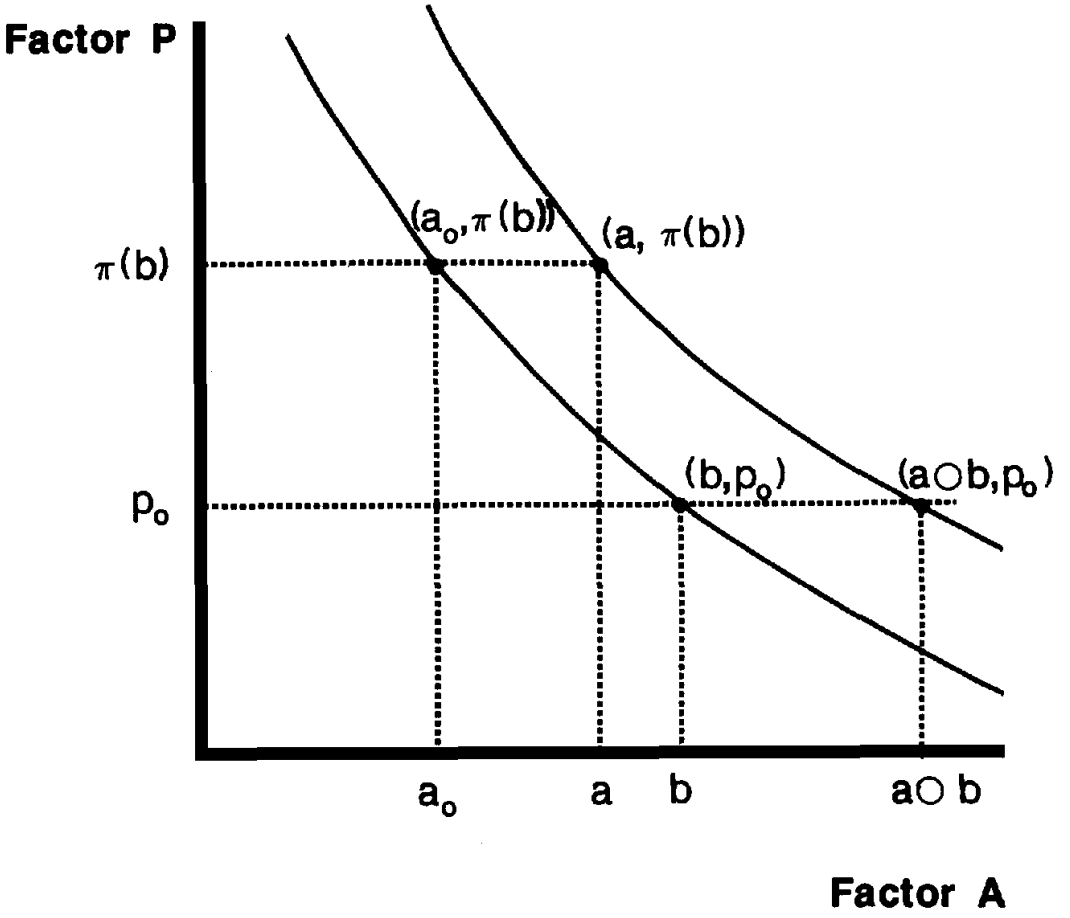

Figure 1. A graphic depiction of the solutions $\pi(b)$ and $a \cdot b$ in a conjoint structure whose components are mapped on a continuum. (The solid curves are indifference curves. Various values on factor A are denoted a_0, a, b, and $a \cdot b$, and those on factor P by P_0 and $\pi(b)$).

that the tone (a, p) is equal in loudness to (b, q). Although it is not always possible to find such an intensity, it is certainly plausible that it exists whenever there are intensities a' and a'' so that tone (a', p) is louder and the tone (a'', p) is less loud than (b, q). It turns out, although the argument is more complex, that one can still prove the existence of an additive representation with restricted solvability substituted for unrestricted solvability (see Krantz et al., 1971, Chap. 6). In terms of Holman's induced operation mentioned above, this change of assumption renders the operation a partial one, that is, one that is defined only for some pairs of elements. Because it is possible to work out a version of extensive measurement for partial operations (see *Refinements: Difference Sequences and Partial Operations*)—indeed, such is necessary to understand probability as fundamental measurement—it is still possible to carry out the construction for the conjoint structure.

Second, the property of additivity, captured in the Thomsen condition, does not always hold. Fortunately, Holman's definition of an operation, or a partial operation in the case of restricted solvability, does not in any way depend on the Thomsen condition. Thus, in general, any conjoint structure gives rise to a concatenation structure in which the induced operation satisfies monotonicity. Such induced operations are associative in exactly those cases in which the Thomsen condition holds. Moreover, one can show that a very great variety of nonassociative operations arise as induced operations of conjoint structures. This in itself was adequate reason to study nonassociative concatenation structures, which began in the mid-1970s. In trying to understand the uniqueness of nonassociative representations, a more complete theory of scale types (described in *Scale Type: General Definition*) had to be developed.

Narens and Luce (1976) showed that any concatenation structure meeting all of the conditions for an extensive structure except for associativity has a numerical representation in terms of some nonassociative numerical operation. Their proof was not constructive. Rather it rested on the classic result of the mathematician Cantor (1895), to the effect that a totally ordered set X (a weakly ordered set in which equivalence is equality) is isomorphic to a subset of real numbers under $\geq$ if and only if it includes a subset Y comparable to the rational numbers in the sense that it is countable (Y can be put in one-to-one correspondence with the integers) and order dense in X—which means that for any two distinct elements of X it is possible to find at least one element from Y that is between them. The key to the proof was to use the axioms of the nonassociative structure to show the existence of such a countable, dense subset. Since then, Krantz has developed a constructive proof (Krantz et al., in press).

At the time, Narens and Luce (1976) were much concerned by their failure to characterize fully the family of representations—the scale. They were able to show that when two homomorphisms into the same numerical system agree at a point, then under weak conditions they are identical. This result does not, however, establish how two different homomorphisms relate. That question was resolved by Cohen and Narens (1979) who showed that the group of automorphisms of this kind of concatenation structure, and so the group of transformations that describe its scale type (see Table 1), can be ordered in such a way that the Archimedean axiom holds. Thus, by what Hölder had established, the transformation group is isomorphic to a subgroup of the multiplicative group of the positive real numbers. When the subgroup is actually the entire group, we have what Stevens called a ratio scale. The other subgroups had not been previously encountered as measurement scales, but Cohen and Narens were able to give numerical examples of each type. We return to questions of scale type later (see *Scale Type: General Definition*).

Distribution of Concatenation Operations in Conjoint Structures

Once it is realized that conjoint measurement, which treats those structures Campbell spoke of as "derived," is just as free from prior measurement as is extensive measurement, a problem arises that understandably went unrecognized by the earlier investigators. An attribute, such as mass, can be fundamentally measured in more than one way. For example, the mass ordering of substances, S, and volumes, V, yields a conjoint representation $\psi_S\psi_V$, which is a measure of mass (Figure 2). At the same time, the usual extensive structure of concatenation of masses leads to the standard additive measure φ_m. Obviously, $\psi_S\psi_V$ must be an increasing function of φ_m because both measures preserve the mass ordering. Furthermore, because volumes can also be concatenated, an extensive measure of volume, φ_V, also exists, and the conjoint measure of volume, ψ_V, is a monotonic increasing function of it. From what is known about physical measurement, a particular ψ_S, call it φ_S, can be chosen so that $\varphi_m = \varphi_S\varphi_V$. This is the representation that is customarily used for this conjoint structure, and the particular substance measure φ_S is called the "density" of the substance. Note, however, that from the point of view of conjoint measurement, for each positive real α and β, $\alpha(\varphi_S)^\beta(\varphi_V)^\beta$ is an equally valid representation, and so $(\varphi_S)^\beta$ is an equally valid measure of density. Thus by selecting the exponent β to be 1 (or equivalently, by identifying ψ_V with φ_V), we have by fiat altered what is really a log-interval representation density into one that appears to be a ratio scale. (This means that in order to force density actually to be a ratio scale, more physical structure than the ordering of the density–volume pairs is needed.) As we noted earlier, Stevens (1959) failed to recognize the use of such conventions when he remarked that log-interval scales were scarce. Quite the contrary, they are exceedingly common, but are often lost sight of by the practice of making certain arbitrary choices of exponents.

The reason why the extensive and conjoint measures of the same attribute are often powers of each other is that the two structures are interlocked qualitatively by what are called "laws of distribution." In the example above, two such laws hold—one between mass and the conjoint structure and the other between volume and the conjoint structure. Such laws take the following form for an operation on a component: Let $\succsim$ be the conjoint ordering of $A \times P$, $\succsim_A$ the order induced on the first component, and $\circ_A$ a concatenation operation on A such that $\langle A, \succsim_A, \circ_A \rangle$ is an extensive structure. Following Narens and Luce (1976) and Ramsay (1976), we say that $\circ_A$ is *distributive* in the conjoint structure provided that the following condition holds for all a, b, c, d in A and p, q in P: whenever

$$(a, p) \sim (c, q) \quad \text{and} \quad (b, p) \sim (d, q),$$

then

$$(a \circ_A b, p) \sim (c \circ_A d, q).$$

When the operation is on $A \times P$, a somewhat different but equivalent formation is needed. It is not difficult to show that if the extensive and conjoint measures on the A component are related by a power function, then this qualitative distribution condition must hold. It derives from the usual numerical distribution

$$(x + y)z = xz + yz.$$

A major observation of Narens and Luce (1976) was that the converse is also true. The distributive interlock is a qualitative condition, which together with the properties of extensive and additive conjoint structures underlies the entire structure of physical dimensions. In fact, Narens (1981a) established a far more general result than the one mentioned above: One need not assume that $\circ_A$ is associative; it is sufficient that the structure involving the concatenation operation has a ratio scale representation (exactly what that entails is described in *Scale Type: General Definition* and *Possible Representations of Concatenation Structures*). Moreover, one need not assume that the conjoint structure has a multiplicative representation, as that follows from the other assumptions.

General Representation Theory

Representations and Scales

As the various examples of measurement discussed above appeared after World War II, it began to be fully appreciated that they are all special cases of a general method of measurement that has come to be referred to as "representational theory." This view, whose earliest explicit formulations were probably those of Scott and Suppes (1958) and Suppes and Zinnes (1963), holds that measurement is possible whenever the following obtains: First, the underlying empirical situation is characterized as an ordered relational structure $\mathcal{X} = \langle X, \succsim, S_1, \ldots, S_n \rangle$, where $\succsim, S_1, \ldots, S_n$ are the *primitives* of the structure (Appendix 1). These primitives are empirical relations (including possibly operations) on X that characterize the empirical situation under consideration. Second, there are restrictions—axioms—on the structure that reflect truths about the empirical situation. These are to be considered as putative empirical laws. Third, there is specified a numerically based relational structure $\mathcal{R} = \langle R, \geq, R_1, \ldots, R_2 \rangle$, where R is a subset of the real numbers and the R_i are relations and operations of comparable types to the corresponding empirical ones. Finally, the fourth feature, which accomplishes measurement, is the proof of the existence of a structure preserving mapping from $\mathcal{X}$ into $\mathcal{R}$. We refer to $\mathcal{X}$ as the *empirical* or *qualitative structure*, $\mathcal{R}$ as the *representing structure*, and the structure-preserving mapping as a *homomorphism* or a *representation*. The collection of all homomorphisms into the same representing structure is referred to as a *scale* (see Footnote 1).

The basic aim of representational theory is first, to use the axioms to show that the scale is not the empty set—this is called

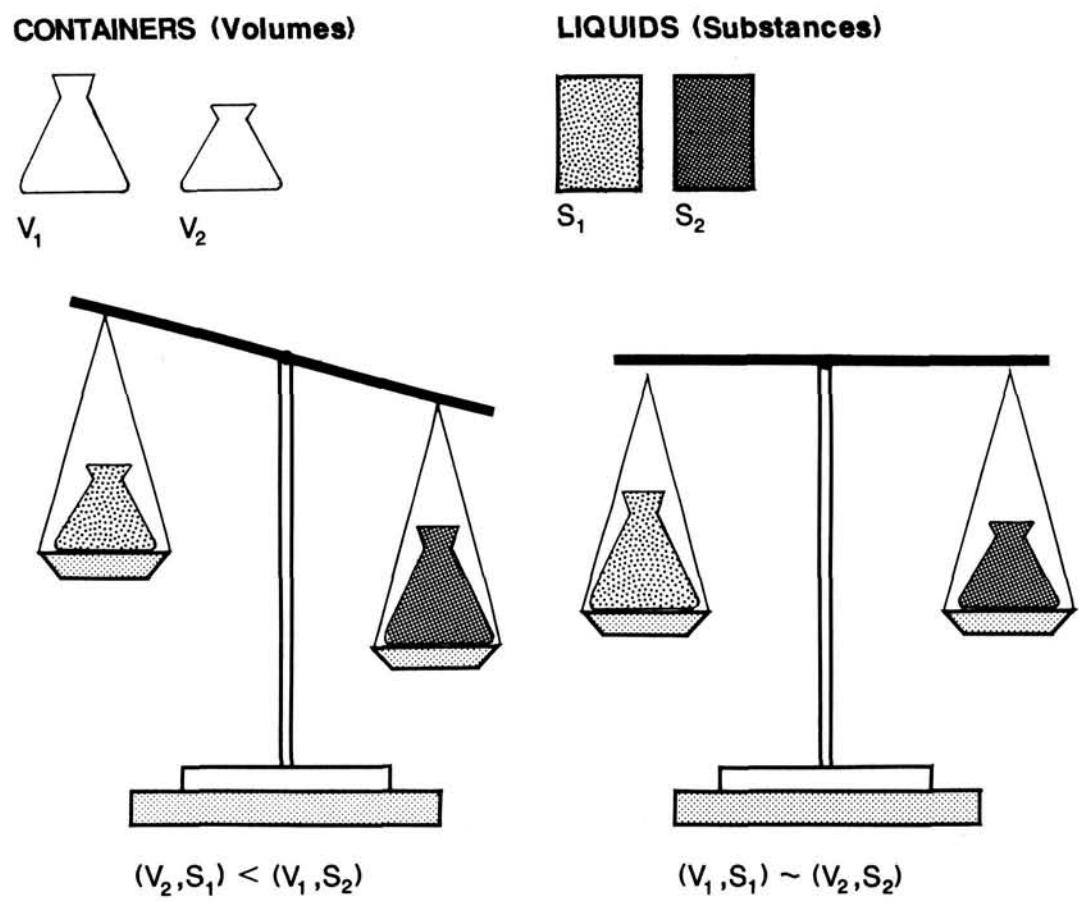

Figure 2. A pan balance determination of the mass ordering of volume-substance pairs.

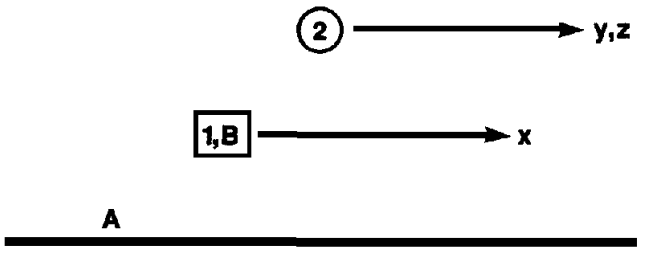

Figure 3. Observer B lives on Object 1 and perceives Object 2 as having velocity y; whereas observer A perceives Object 1 (and its resident, observer B) as having velocity x and Object 2 as having velocity z. (The concatenation operation $\circ$ is defined by $x \circ y = z$.)

the *existence* or *representation theorem*—and second, to characterize how these mappings (homomorphisms) that constitute the scale relate to one another—this is called the *uniqueness theorem*. In the classical case of extensive measurement, it is shown that a nonempty scale exists and is characterized as a ratio scale in the sense that φ and φ' are both in the scale if and only if there is a positive real constant r such that $\varphi' = r\varphi$.

It should be realized that the representing structure is not itself unique; there always are a variety of alternative ones, and different ones are used for different purposes. Velocity provides an example of this. Suppose X is a set of constant velocities in a given direction that are ordered by the distances traveled in a fixed time interval. Concatenation of velocities x and y is the velocity that is obtained by superimposing x on y. That is, $x \circ y$ is the velocity of a body that an observer on another body moving at velocity x would judge to have the velocity y. (Figure 3.) In classical physics X is taken to be all possible velocities, whereas in relativistic physics it is convenient to restrict X to velocities less than that of light. Except for that difference, the two structures are assumed to be extensive; however, in their measurement very different representations are used. In the classical case the usual additive representation is used, but in the relativistic case one selects $c > 0$ to represent the velocity of light and maps $\langle X, \succsim, \circ \rangle$ into $\langle (0, c), \geq, \oplus_c \rangle$, where $\oplus_c$ is defined as follows: for all u and v in $(0, c)$,

$$u \oplus_c v = \frac{u + v}{1 - uv/c^2}.$$

It can be shown that these two numerical representing structures are in fact isomorphic, the isomorphism being $f(u) = \tanh^{-1}(u/c)$, u in $(0, c)$. If in the relativistic case $\circ$ were represented additively, the velocity of light would be assigned the value ∞. The real reason for changing the representation from an additive one is not to avoid ∞, but rather to maintain the usual relation among velocity, distance, and duration, namely, that the former is proportional to distance traversed divided by the duration.

Homogeneity and Uniqueness

With the results about nonassociative structures as a stimulus and working within the general representational framework, Narens (1981b) proposed a method for classifying scale types which has proved useful in describing the possible representations that can arise. Although the two concepts needed, *homogeneity* and *uniqueness*, are formulated in a rather abstract way, only the former seems illusive. So we focus on it both here and in the next section.

Many of the most familiar mathematical structures used in science, such as Euclidean space, exhibit the property of being homogeneous. Like homogenized milk, each part of the space looks like each other part. This is the general intuitive concept: Every element in the domain of the structure is, from the point of view of the properties defining the structure (its primitives), just like every other element. There is no way of singling out an individual element as different from the others. To formulate this precisely and generally, two things are needed: (a) a very general concept of what we mean by a structure, and (b) the concept of an automorphism of a structure. The latter permits us to say when the structure looks the same from two points of view.

To describe the situation, a very general model of measurement is used. First, $\mathcal{X} = \langle X, \succsim, S_1, \ldots, S_n \rangle$ is a relational structure that characterizes the empirical situation in the sense that $\succsim$ is a total ordering of X (i.e., a weak ordering for which indifference, $\sim$, is actually equality) and $S_1, \ldots, S_n$ are other empirical relations. Second, $\mathcal{R} = \langle R, \geq, R_1, \ldots, R_n \rangle$ is the representing numerical structure. And third, $\mathcal{S}$ is a scale for $\mathcal{X}$. In many important scientific applications, R is either the real numbers or the positive real numbers and the elements of $\mathcal{S}$ are isomorphisms of $\mathcal{X}$ onto $\mathcal{R}$. We assume this situation throughout the rest of this article unless stated otherwise.

An automorphism is simply an isomorphism of a structure with itself, that is, a one-to-one map of the structure with itself that preserves all of the primitives. Intuitively, an automorphism corresponds to what we usually refer to as a symmetry of the structure, namely, a mapping of the structure so that things look the same before and after the mapping is completed. So, for example, if the structure is a sphere, we know that it is symmetrical in the sense that it looks exactly the same before and after any rotation about its center. Thus, for the sphere, rotations are automorphisms. The general concept applies, of course, to any relational structure. It is easy to verify that for each automorphism α of $\mathcal{X}$ and for each φ in $\mathcal{S}$, the mapping $\varphi * \alpha$, where $*$ denotes function composition [i.e., for x in X, $\varphi * \alpha(x) = \varphi(\alpha(x))$], is also in $\mathcal{S}$, and if φ' is also in $\mathcal{S}$, then $\alpha = \varphi^{-1} * \varphi'$ is an automorphism of $\mathcal{X}$. Thus, there is a one-to-one correspondence between the scale $\mathcal{S}$ and the automorphism group, and so a classification of the one is equivalent to a classification of the other. The following classification of the automorphism group in terms of its richness (called "homogeneity") and of its redundancy (called "uniqueness") has proven to be very useful. The structure is said to satisfy *1-point homogeneity* if and only if for each x, y in X, there exists an automorphism α of the structure such that $\alpha(x) = y$. This means that the structure exhibits a good deal of symmetry, because the automorphisms of a structure capture its symmetries. In geometry, this concept is equivalent to the concept of "1-transitivity," which has been extended there to apply to any M distinct points mapped by a continuous transformation to any other M distinct points of $\mathcal{X}$, in which case it is called "M-transitivity." For measurement, the generalization that is relevant is that each set of M ordered elements can be mapped by some automorphism into any other set of M comparably ordered elements. This latter condition is called *M-point*

homogeneity. For $M > 1$ it is different from the geometric concept of M-transitivity. When a structure is M-point homogeneous for every positive M, it is said to be ∞-point homogeneous.

It is convenient to abbreviate "1-point homogeneity" just to "homogeneity," but we are careful to distinguish clearly other values of M-point homogeneity.

To capture the idea of redundancy in the automorphism group, we say that the structure satisfies *N-point uniqueness* if and only if whenever two automorphisms agree at N distinct points, then they agree everywhere. If the structure is not N-point unique for any finite N, it is said to be *∞-point unique.*

Several simple observations: Suppose a structure is infinite, M-point homogeneous, and N-point unique. Then $M \leq N$; if $M' \leq M$, then the structure is M'-point homogeneous and if $N' \geq N$, then it is N'-point unique. Thus, in particular, all M-point homogeneous structures, $M \geq 1$, are 1-point homogeneous, that is, homogeneous.

Testing for Homogeneity

Although homogeneity is a concept about the structure, it is in fact usually not obvious how to recast it in terms of qualitative properties that can be readily studied empirically. In some cases, particularly when there is a primitive binary operation, such logical equivalences are known (see *Possible Representations of Concentration Structures*). (It should be mentioned that the proofs of such equivalences are usually not easy and generally require much mathematical machinery or the use of a nontrivial representation–uniqueness result.) Quite often homogeneity need not be explicitly stated because it follows as a consequence of a representation theorem. For example, in the extensive case for which there is a representation onto the positive real numbers, 1-point homogeneity easily follows from the existence and uniqueness results for additive representations. In such cases or in ones in which an empirical equivalent is known, homogeneity does not pose a serious empirical problem. Yet in many important scientific applications no such structural equivalences are known; in such cases homogeneity is simply postulated directly as a theoretical concept. Nonetheless, because of its power, it is often easy to devise simple tests to show that it does not hold even though we may not know how to test affirmatively for when it does hold. The following is one of the more useful such tests. Suppose $\mathcal{X} = \langle X, S_1, \ldots, S_n \rangle$ is a relational structure and that P is a property (one place relation) about X that is definable from the primitives $S_1, \ldots, S_n$ using ordinary first-order predicate logic. It can be shown that if $\mathcal{X}$ is 1-point homogeneous, then either $P(x)$ is false for every x in X or $P(x)$ is true for every x in X.

The following examples illustrate its use. Suppose $\mathcal{X}$ is a qualitative structure for probability (see *Refinements: Difference Sequences and Partial Operations*), and $A \succsim B$ stands for "A is at least as likely as B." Consider the predicate

$$P(A): \quad \text{for all} \quad B, A \succsim B.$$

Observe that $P(A)$ is true for A = the sure event and false for A = the null event. Thus, we know that qualitative probability is not homogeneous. This contrasts with the usual extensive models for length and mass, which are homogeneous. As a second example, consider a structure $\langle (0, 1), R \rangle$ in which (0, 1) is the open interval of real numbers between 0 and 1 and R is the ternary relation on (0, 1) defined by:

$$R(x, y, z) \quad \text{if and only if } x, y, z \text{ are in } (0, 1) \text{ and } x + y = z.$$

Consider the predicate

$$P(x): \quad \text{there exists a } z \text{ such that } R(x, x, z).$$

Because $P(1/3)$ is true and $P(2/3)$ is false, the structure $\langle (0, 1), R \rangle$ is not homogeneous.

As we shall see in the next section, the only other important case of finite point homogeneity is 2-point. Unlike 1-point homogeneity, it has proved very difficult to find qualitative equivalences to 2-point homogeneity that are empirically realizable and hold across a wide range of interesting structures. So, in practice, one either simply postulates it as a theoretical assumption or derives it (usually through a complicated mathematical argument) from the particular primitive relations under consideration. As with 1-point homogeneity, there are ways to show that it fails: (a) Because structures that are 2-point homogeneous are also 1-point homogeneous (see the end of *Homogeneity and Uniqueness*), the definitional test for 1-point homogeneity can be invoked. (b) Because a 1-point unique structure cannot be 2-point homogeneous (see the end of *Homogeneity and Uniqueness*), it suffices to show the structure is 1-point unique, and sometimes that is easy to do. (c) As we describe in *Possible Representations of Concatenation Structures,* the special case of a 2-point homogeneous structure with a primitive monotonic operation necessarily has a very restrictive form of numerical representation, and it may be possible to show by empirical tests that such a representation is simply too restrictive to model the empirical situation.

Scale Type: General Definition

Recall that in infinite structures, there is a largest value, K, of homogeneity and a smallest value, L, of uniqueness. These are referred to, respectively, as the *degree* of homogeneity and uniqueness. This pair of numbers, (K, L), is useful for classifying the type of scale exhibited by a structure; it is called the *scale type.*

It is easy to verify that if $\mathcal{S}$ is a ratio scale, then $\mathcal{S}$ is of type (1, 1); if $\mathcal{S}$ is an interval scale, then $\mathcal{S}$ is of type (2, 2); and if $\mathcal{S}$ is an ordinal scale, then $\mathcal{S}$ is of type (∞, ∞). Narens (1981a, 1981b) established the following converse of these observations. Suppose a structure has a representation *onto* the real numbers. If its scale is of type (1, 1), then a representing structure can be found such that its representations form a ratio scale; and if the scale is of type (2, 2), then it has a representing structure such that its representations form an interval scale. In addition, he showed that it is impossible for the scale to be of type (M, M) for $2 < M < \infty$. There are (∞, ∞) cases that do not have ordinal scale representations; however, this does not much matter because the ∞-point homogeneous cases—including the ordinal scalable ones—simply do not arise in empirical situations for which there is a reasonable amount of structure. Alper (1984, 1985) has shown that the only cases of structures with representations onto the positive real numbers and of scale type (K, L) with $0 < K <$

L and $1 \leq L < \infty$, are the ones in which $K = 1$ and $L = 2$, and in that case a discrete interval scale (Table 1) exists. These results give considerable insights into why so few scale types have arisen in the development of the sciences.

The whole issue of how intelligently to classify structures with either $K = 0$ or $L = \infty$ is wide open.

Possible Representations of Concatenation Structures

For the important and widely applicable case of concatenation structures of the form $\mathfrak{X} = \langle X, \succsim, \circ \rangle$, where $\succsim$ is a weak ordering on X and $\circ$ is a binary operation on X, comparable results to those given for general structures hold without the assumption that $\mathfrak{X}$ can be mapped onto the real or positive real numbers. Luce and Narens (1983, 1985) have shown that if such a concatenation structure is of scale type (K, L) with $K > 0$ and $L < \infty$, then only types (1, 1), (1, 2), and (2, 2) can occur. The latter two necessarily are idempotent, and the (1, 1) type is either idempotent, weakly positive ($x \circ x \succ x$, for all x), or weakly negative ($x \circ x \prec x$, for all x). An important sufficient condition for L to be finite is that the structure have a representation onto the positive real numbers for which the numerical operation is continuous. Continuity of an operation is usually judged to be an acceptable scientific idealization. For these three scale types, it is desirable to describe all possible candidate numerical representing structures. So, using Narens's (1981a, 1981b) results, it suffices to consider concatenation structures on the positive reals with ratio, log-interval, or log-discrete interval scales (Table 1). Suppose that $\oplus$ denotes the representing operation. Luce and Narens (1985), extending the results of Cohen and Narens (1979), have shown that in all these cases there exists a function f from the positive real numbers into itself such that f is strictly increasing, $f(x)/x$ is strictly decreasing, and the operation is given by

$$x \oplus y = yf(x/y).$$

It is worth noting that the only cases in which the above-mentioned homogeneous structure can be positive ($x \circ y \succ x$, $x \circ y \succ y$) or negative ($x \circ y \prec x$, $x \circ y \prec y$) are the (1, 1) ones with $f(1) \neq 1$. All the remaining structures are intensive in the sense that $x \circ x \sim x$ and if $x \succ y$, then $x \succ x \circ y \succ y$. (Formal properties of concatenation structures are summarized in Appendix 3.) Clearly, the above operation $\oplus$ is invariant under ratio scale transformations. The (1, 2) and (2, 2) cases simply impose additional restrictions on f. For example, consider the equation such that for all $x > 0$,

$$f(x^\rho) = f(x)^\rho.$$

The (1, 1) case is characterized by the equation holding if and only if $\rho = 1$; the (1, 2) case by its holding if and only if for some fixed $k > 0$ and variable integer n, $\rho = k^n$; and the (2, 2) case by its holding for all $\rho > 0$. In this situation, the (2, 2) case is equivalent to the existence of constants c, d, $0 < c, d < 1$ such that

$$x \oplus y = \begin{cases} x^c y^{1-c}, & \text{if } x \succ y, \\ x, & \text{if } x \sim y, \\ x^d y^{1-d} & \text{if } x \prec y. \end{cases}$$

The last representation, called the *dual bilinear representation*, shows that the (2, 2) case is highly restrictive, and that all (2, 2) operations are nothing more than two pieces of two bisymmetric operations.

So far as we know, the dual bilinear representation, except for $c = d$, has not arisen in physics, but recently Luce and Narens (1985) have used it to formulate a generalized theory of expected utility, which appears to overcome a number of the empirical disconfirmations of the classical theories of the subject. This is described in the next section.

Before turning to that, we consider two further questions: axiomatization of general concatenation structures and conditions equivalent to homogeneity. Narens and Luce (1976) showed that concatenation structures satisfying all of the axioms of extensive structures except possibly associativity had a numerical representation. Such structures, called PCSs, play an important role in measurement theory. Luce and Narens (1985) have provided a comparable axiomatization for general intensive structures. Much is known about axiomatizing homogeneity for concatenation structures. First, on the assumption of a representation onto the real numbers, certain basic algebraic properties such as associativity, bisymmetry, and right autodistributity $[(x \circ y) \circ z \sim (x \circ z) \circ (y \circ z)]$ all force homogeneity to hold.[2] Second, for a wide variety of PCSs, homogeneity is equivalent to the following structural condition: For all elements x and y and all positive integers n,

$$(x \circ y)_n = x_n \circ y_n,$$

where x_n denotes the nth element of a standard sequence based on x (see *Axioms for Extensive Quantities*). Luce (1986) has shown that a closely related, although perhaps less useful, criterion exists for homogeneity in intensive structures. The third method for establishing homogeneity is to axiomatize directly all concatenation structures of a given type. For the (2, 2) case this is equivalent to axiomatizing the dual bilinear representations, which was done in Luce (1986).

These techniques for characterizing concatenation structures by scale type can, of course, be extended to nonadditive conjoint structures, as Luce and Cohen (1983) showed; however, matters are a bit more complex than one might first anticipate. In particular, automorphisms of the conjoint structure need not always factor into automorphisms of the orderings induced on the components, and even when they do the scale types are not usually the same. We do not go into these complex details here.

Dual Bilinear Utility

A theory of preferences among gambles can be based on the idea that gambles can be "concatenated" in a special way to form other gambles, and that rationality considerations need be applied only to the simplest concatenations of gambles with gambles (Luce & Narens, 1985). When rationality considerations are more broadly invoked, even marginally, this theory reduces to the usual subjective expected utility model used throughout the social sciences.

---

[2] It is worth noting that many times when an axiom is added to a general concatenation structure that has a numerical representation, it can be formulated numerically as a functional equation. In some cases, solutions are available in the literature; a good starting point for finding such solutions is Aczél (1966).

More specifically, suppose x and y are gambles and A is an event. Then $x \circ_A y$ denotes the gamble in which x is the outcome when A occurs and y when A fails to occur. It is assumed that there is a preference ordering $\succsim$ over gambles. The Luce and Narens model ends up with a utility function U over the gambles (i.e., a real-valued function such that $g_1 \succsim g_2$ if and only if $U[g_1] \geq U[g_2]$) and two weighting functions, S^+ and S^-, defined over events such that

$$U(x \circ_A y) = \begin{cases} U(x)S^+(A) + U(y)[1 - S^+(A)], & \text{if } U(x) > U(y), \\ U(x), & \text{if } U(x) = U(y), \\ U(x)S^-(A) + U(y)[1 - S^-(A)], & \text{if } U(x) < U(y). \end{cases}$$

The weighting functions, S^+ and S^-, need not be probability functions. This model is called the *dual bilinear utility model.* The standard subjective expected utility model (SEU) arises when $S^+ = S^- =$ a (finitely additive) probability measure over the set of events (see Fishburn, 1981, for a detailed summary of SEU). The dual bilinear model may seem a little artificial at first. However, it follows from an almost universally used assumption about utility functions, namely, that the representations of an individual's utility function over gambles form an interval scale, together with some very natural and weak assumptions about $\circ_A$ and $\succsim$.

An additional reason for considering the dual bilinear model is that it is weaker than the SEU model, and there is an abundance of empirical data showing that SEU fails to describe behavior. A summary of many of the problems was given by Kahneman and Tversky (1979). Basically, the failures are concerned with three types of "rationality." The first is transitivity of preference, which has been shown to fail under some circumstances by Lichtenstein and Slovic (1971, 1973) and Grether and Plott (1979). No model, such as the present one, which associates utility with gambles can account for this. The second type of failure has to do with what Luce and Narens call "accounting equations" and Kahneman and Tversky refer to as the "framing" of gambles. An example of an accounting equation that is implied by the dual bilinear model is

$$(x \circ_A y) \circ_B y \sim (x \circ_B y) \circ_A y,$$

where A and B are independent events, such as A is an even number coming up on a roll of a die and B is a red number coming up on a turn of a roulette wheel. Observe that x is the outcome on both sides if in independent realizations of the events both A and B occur, in that order on the left and in the opposite order on the right. An example of another accounting equation is

$$(x \circ_A y) \circ_A z \sim (x \circ_A z) \circ_A (y \circ_A z),$$

where successive $\circ_A$ mean independent realizations of the event A (e.g., the first "A" refers to an even number coming up on a roll of a die, the second "A" as an even number coming up on a *different* roll of the same die, etc.). This holds in the bilinear model if and only if $S^+ = S^-$, which is true for SEU but not in general for the dual bilinear model. The earliest discussion of failures of accounting equations was by Allais (1953; see also Allais & Hagen, 1979). A third type of failure is also one of the accounting type, but it is more subtle because it involves a kind of monotonicity of events. Suppose C is an event that is disjoint from events A and B, then the assertion is

$$x \circ_A y \succsim x \circ_B y \quad \text{if and only if} \quad x \circ_{A \cup C} y \succsim x \circ_{B \cup C} y.$$

In essence, then, the pair of gambles on the right is obtainable from the pair on the left by shifting the assignment of outcomes over C and y to x. Ellsberg (1961) pointed out that this often fails for people's preferences, and this has been repeatedly confirmed. In the dual bilinear model, this equivalence holds if and only if the two weights exhibit the property that for C disjoint from A and B,

$$S^i(A) \geq S^i(B) \quad \text{if and only if}$$
$$S^i(A \cup C) \geq S^i(B \cup C), \quad i = + \text{ or } -.$$

This is true of the SEU model because of the Ss are probabilities and so $S^i(A \cup C) = S^i(A) + S^i(C)$. The basic distinction between the two types of accounting equations has to do with forcing the two weights to be identical in one case and to be probability functions in the other.

It should be noted that this model is in many ways similar to and more completely specified than the prospect theory of Kahneman and Tversky (1979), as was discussed in some detail in Luce and Narens (1985). As yet, no empirical studies have been reported that are targeted directly at the dual bilinear model.

Conclusions

In summary, a great deal is now known about the scales for inherently symmetrical, one-dimensional attributes and about how they interlock to form the systematic structure of multidimensional physical quantities. Perhaps the major milestones of the past 25 years are these: First, the development of conjoint structures, which not only provided a deep measurement analysis of the numerous nonextensive, "derived" structures of physics, but also provided a measurement approach that appears to have applications in the nonphysical sciences and has laid to rest the claim that the only possible basis for measurement is extensive structures. Second, the development of the distributive interlock between ratio scale concatenation structures and conjoint structures, which serves to explain why physical measures are all interlocked as products of powers of a few ratio scales. Third, the growing recognition of the importance of automorphism groups in classifying measurement structures, and the explicit definition of scale type in terms of degree of homogeneity and degree of uniqueness. Fourth, the application of that classification to ordered structures with a concatenation operation and to conjoint structures, thereby providing a complete catalogue for these situations of the possible representing structures for the homogeneous cases.

A number of important problems remain unresolved. For one, we do not have an adequate axiomatization of the general class of homogeneous intensive structures except for the interval scalable ones. Second, we do not have comparable results for nonhomogeneous structures, even ones with concatenation operations. This is not an esoteric question because any totally ordered structure with a partial operation—such as probability when looked at the right way—has only one automorphism, the identity map. Thus, in such cases the automorphism group fails to em-

body any structural information. Nevertheless, despite their lack of global symmetry, such structures often appear to be quite regular in other aspects, and this needs to be captured in some fashion and studied. Third, there are some important cases of interlocking concatenation and conjoint structures that are not covered by the distribution results mentioned, perhaps the most striking example being relativistic velocity as a component of the distance conjoint structure with time as the other component. Because many psychological attributes appear to be bounded, understanding this physical case may be more pertinent than it first might seem.

References

Aczél, J. (1966). *Lectures on functional equations and their applications.* New York: Academic Press.

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école Americaine [The behavior of the rational man in the face of risk: A critique of the postulates and axioms of the American school]. *Econometrica, 21,* 503–546.

Allais, M., & Hagen, O. (Eds.). (1979). *Expected utility hypothesis and the Allais paradox.* Dordrecht, The Netherlands: Reidel.

Alper, T. M. (1984). *Groups of homeomorphisms of the real line.* Unpublished Bachelor of Science thesis, Harvard University, Cambridge, Massachusetts.

Alper, T. M. (1985). A note on real measurement structures of scale type (m, m + 1). *Journal of Mathematical Psychology, 29,* 73–81.

Bridgman, P. (1922). *Dimensional Analysis,* New Haven: Yale University Press. (2nd ed. published 1931)

Campbell, N. R. (1920). *Physics: the elements.* Cambridge, England: Cambridge University Press. (Reprinted [1957] as *Foundations of Science,* New York: Dover)

Campbell, N. R. (1928). *An account of the principles of measurement and calculation.* London: Longsman, Green.

Cantor, G. (1895). Beiträge zur Begründung der transfiniten Mengenlehre [Contributions to the founding of the theory of transfinite numbers]. *Math. Ann., 46,* 481–512.

Cohen, M., & Narens, L. (1979). Fundamental unit structures: A theory of ratio scalability. *Journal of Mathematical Psychology, 20,* 193–232.

Debreu, G. (1960). Topological methods in cardinal utility theory. In K. J. Arrow, S. Karlin, & P. Suppes (Eds.), *Mathematical Methods in the Social Sciences, 1959* (pp. 16–26). Stanford, CA: Stanford University Press.

Ellis, B. (1966). *Basic concepts of measurement.* London: Cambridge University Press.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics, 75,* 643–669.

Falmagne, J.-C. (1976). Random conjoint measurement and loudness summation. *Psychological Review, 83,* 65–79.

Ferguson, A., (Chairman), Meyers, C. S. (Vice-Chairman), Bartlett, R. J., (Secretary), Banister, H., Bartlett, F. C., Brown, W., Campbell, N. R., Craik, K. J. W., Drever, J., Guild, J., Houstoun, R. A., Irwin, J. O., Kaye, G. W. C., Philpott, S. J. F., Richardson, L. F., Shaxby, J. H., Smith, T., Thouless, R. H., & Tucker, W. S. (1940). Quantitative estimates of sensory events. *The Advancement of Science. The Report of the British Association for the Advancement of Science, 2,* 331–349.

Fishburn, P. C. (1981). Subjective expected utility: A review of normative theories. *Theory and Decision, 13,* 139–199.

Grether, D., & Plott, C. (1979). Economic theory of choice and the preference reversal phenomenon. *American Economic Review, 69,* 623–638.

Helmholtz, H. von (1887). Zählen und Messen erkenntnis-theoretisch betrachet. *Philosophische Aufsätz Eduard Zeller gewidmet,* Leipzig. (Reprinted in *Wissenschaftliche Abhandlungen,* 1895, Vol. 3, pp. 356–391. C. L. Bryan (trans.), *Counting and Measuring.* Princeton, NJ: Van Nostrand, 1930)

Hölder, O. (1901). Die Axiome der Quantität und die Lehre vom Mass [The axioms of quantity and the theory of mass]. *Sächsische Akademie Wissenschaften zu Leipzig, Mathematisch-Physische Klasse, 53,* 1–64.

Holman, E. W. (1971). A note on conjoint measurement with restricted solvability. *Journal of Mathematical Psychology, 8,* 489–494.

Kahneman, D., & Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica, 47,* 263–291.

Krantz, D. H. (1964). Conjoint measurement: The Luce–Tukey axiomatization and some extensions. *Journal of Mathematical Psychology, 1,* 248–277.

Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (1971). *Foundations of Measurement* (Vol. 1). New York: Academic Press.

Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (in press). *Foundations of Measurement* (Vol. 2). New York: Academic Press.

Levelt, W. J. M., Riemersma, J. B., & Bunt, A. A. (1972). Binaural additivity of loudness. *British Journal of Mathematical and Statistical Psychology, 25,* 51–68.

Lichtenstein, S., & Slovic, P. (1971). Reversals of preferences between bids and choices in gambling decisions. *Journal of Experimental Psychology, 89,* 46–55.

Lichtenstein, S., & Slovic, P. (1973). Response-induced reversals of preferences in gambling: An extended replication in Las Vegas. *Journal of Experimental Psychology, 101,* 16–20.

Luce, R. D. (1978). Dimensionally invariant laws correspond to meaningful qualitative relations. *Philosophy of Science, 45,* 1–16.

Luce, R. D. (1986). *Uniqueness and homogeneity of real relational structures.* Manuscript submitted for publication.

Luce, R. D., & Cohen, M. (1983). Factorizable automorphisms in solvable conjoint structures I. *Journal of Pure and Applied Algebra, 27,* 225–261.

Luce, R. D., & Marley, A. A. J. (1969). Extensive measurement when concatenation is restricted and maximal elements may exist. In S. Morgenbesser, P. Suppes, & M. G. White (Eds.), *Philosophy, science, and method: Essays in honor of Ernest Nagel* (pp. 235–249). New York: St. Martin's Press.

Luce, R. D., & Narens, L. (1983). Symmetry, scale types, and generalizations of classical physical measurement. *Journal of Mathematical Psychology, 27,* 44–85.

Luce, R. D., & Narens, L. (1985). Classification of concatenation measurement structures according to scale type. *Journal of Mathematical Psychology, 29,* 1–72.

Luce, R. D., & Tukey, J. W. (1964). Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of Mathematical Psychology, 1,* 1–27.

Narens, L. (1974). Measurement without Archimedean axioms. *Philosophy of Science, 41,* 374–393.

Narens, L. (1976). Utility-uncertainty trade-off structures. *Journal of Mathematical Psychology, 13,* 296–322.

Narens, L. (1981a). A general theory of ratio scalability with remarks about the measurement-theoretic concept of meaningfulness. *Theory and Decision, 13,* 1–70.

Narens, L. (1981b). On the scales of measurement. *Journal of Mathematical Psychology, 24,* 249–275.

Narens, L. (1985). *Abstract measurement theory.* Cambridge, MA: MIT Press.

Narens, L., & Luce, R. D. (1976). The algebra of measurement. *Journal of Pure and Applied Algebra, 8,* 197–233.

Narens, L., & Luce, R. D. (in press). Meaningfulness and invariance concepts. In J. Eatwell, M. Milgate, & P. Newman (Eds.), *The new Palgrave: A dictionary of economic theory and doctrine.* New York: Macmillan.

Pfanzagl, J. (1959). A general theory of measurement—applications to utility. *Naval Research Logistics Quarterly, 6,* 283–294.

Pfanzagl, J. (1968). *Theory of Measurement.* New York: Wiley. (2nd ed. published 1971)

Ramsay, J. O. (1976). Algebraic representation in the physical and behavioral sciences. *Synthese, 33,* 419–453.

Roberts, F. S. (1979). *Measurement theory.* Reading, MA: Addison-Wesley.

Roberts, F. S., & Luce, R. D. (1968). Axiomatic thermodynamics and extensive measurement. *Synthese, 18,* 311–326.

Savage, L. J. (1954). *Foundations of statistics.* New York: Wiley. (2nd revised ed., New York: Dover, 1972)

Scott, D., & Suppes, P. (1958). Foundational aspects of theories of measurement. *Journal of Symbolic Logic, 23,* 113–128.

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

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

Stevens, S. S. (1959). Measurement, psychophysics and utility. In C. W. Churchman & P. Ratoosh (Eds.), *Measurement: Definitions and theories* (pp. 18–63). New York: Wiley.

Stevens, S. S., & Davis, H. (1938). *Hearing: Its psychology and physiology.* New York: Wiley.

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

Von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior.* Princeton, NJ: Princeton University Press.

Appendix 1

Some Structure Preserving Concepts

$\mathcal{X} = \langle X, S_0, S_1, \ldots, S_n \rangle$ is said to be a *relational structure* if and only if X is a nonempty set and $S_0, S_1, \ldots, S_n$ are relations or operations on X.

φ is said to be a *homomorphism* of the relational structure $\mathcal{X} = \langle \mathrm{X}, S_0, S_1, \ldots, S_n \rangle$ into the relational structure $\mathcal{R} = \langle R, R_0, R_1, \ldots, R_n \rangle$ if and only if φ is a function from X into R, for $k = 0, \ldots, n$, S_k and R_k have the same number, i_k, of arguments, and for all $x_1, \ldots, x_{ik}$ in X,

$$S_k(x_1, \ldots, x_{ik}) \quad \text{if and only if} \quad R_k[\varphi(x_1), \ldots, \varphi(x_{ik})],$$

if S_k is a relation, and

$$\varphi[S_k(x_1, \ldots, x_{ik})] = R_k[\varphi(x_1), \ldots, \varphi(x_{ik})]$$

if S_k is an operation. If R is a subset of reals and R_0 is the usual ordering, $\geq$, of the reals (restricted to R), then in measurement theory homomorphisms of $\mathcal{X}$ into $\mathcal{R}$ are called *representations.*

φ is said to be an *isomorphism* of $\mathcal{X}$ onto $\mathcal{R}$ if and only if φ is a homomorphism of $\mathcal{X}$ into $\mathcal{R}$, φ is onto $\mathcal{R}$, and φ is a one-to-one function.

φ is said to be an *automorphism* of $\mathcal{X}$ if and only if φ is an isomorphism of $\mathcal{X}$ onto itself.

The set of automorphisms, G, of a relational structure $\mathcal{X}$ is closed under the operation of composing functions, $*$. ($\alpha * \beta$ is defined by $\alpha * \beta(x) = \alpha[\beta(x)]$.) It is easy to show that $\langle \mathrm{G}, * \rangle$ is a group. $\langle \mathrm{G}, * \rangle$ is called the *automorphism group* of $\mathcal{X}$.

Appendix 2

Some Concepts About Conjoint Structures

Let $\succsim$ be a binary relation on the Cartesian product $\mathrm{A} \times \mathrm{P}$ and $\mathcal{C} = \langle \mathrm{A} \times \mathrm{P}, \succsim \rangle$.

$\mathcal{C}$ is said to be a *conjoint structure* if and only if the following two conditions are satisfied. (a) *Weak ordering:* $\succsim$ is transitive and connected. (b) *Independence:* For all a, b in A, if for some p in P $(a, p) \succsim (b, p)$, then for all q in P $(a, q) \succsim (b, q)$; and for all p, q in P, if for some a in A $(a, p) \succsim (a, q)$, then for all b in A $(b, p) \succsim (b, q)$.

Suppose $\mathcal{C}$ is a conjoint structure. Define $\succsim_{\mathrm{A}}$ on A as follows: For all a, b in A, $a \succsim_{\mathrm{A}} b$ if and only if for some p in P $(a, p) \succsim (a, q)$. It is easy to show that $\succsim_{\mathrm{A}}$ is a weak ordering of A. Similarly, a weak ordering $\succsim_{\mathrm{P}}$ can be defined on P.

$\mathcal{C}$ is said to satisfy (*unrestricted*) *solvability* if and only if for all a, b in A and p, q in P, there exist c in A and r in P such that $(c, p) \sim (b, q)$ and $(a, r) \sim (b, q)$.

$\mathcal{C}$ is said to satisfy *restricted solvability* if and only if for all a', a'', and b in A and p, q in P,

$$\text{if} \quad (a', p) \succsim (b, q) \succsim (a'', p),$$
$$\text{then for some } a \text{ in A } (a, p) \sim (b, q);$$

and for all a, b in A and p', p'', and q in P,

$$\text{if} \quad (a, p') \succsim (b, q) \succsim (a, p''),$$
$$\text{then for some } p \text{ in P } (a, p) \sim (b, q).$$

(Appendixes continue on next page)

Appendix 3

Some Concepts About Concatenation Structures

Let $\mathfrak{X} = \langle X, \succsim, \circ \rangle$, where $\succsim$ is a binary relation on X and $\circ$ is a binary operation on X.

$\mathfrak{X}$ is said to be a *concatenation structure* if and only if $\succsim$ is a total ordering and $\circ$ is strictly monotonic, that is, $x \succsim y$ if and only if $x \circ z \succsim y \circ z$ if and only if $z \circ x \succsim z \circ y$ for all x, y, and z in X.

$\mathfrak{X}$ is said to be a *weakly ordered concatenation structure* if and only if $\succsim$ is a weak ordering (i.e., a transitive and connected relation) and $\circ$ is strictly montonic.

$\mathfrak{X}$ is said to be *positive* if and only if $x \circ y \succ x$ and $x \circ y \succ y$ for all x, y in X.

$\mathfrak{X}$ is said to be *idempotent* if and only if $x \circ x \sim x$ for all x in X.

$\mathfrak{X}$ is said to be *intern* if and only if $x \succ y$ implies that $x \succ x \circ y \succ y$ and $x \succ y \circ x \succ x$.

$\mathfrak{X}$ is said to be *intensive* if and only if it is idempotent and intern.

$\circ$ is said to be *bisymmetric* if and only if $(x \circ y) \circ (u \circ v) \sim (x \circ u) \circ (y \circ v)$ for all x, y, u, and v in X.

Note that "averaging" operations $\oplus_\alpha$ on the real numbers of the form $r \oplus_\alpha s = \alpha r + (1 - \alpha)s$, where $0 < \alpha < 1$, are intensive and bisymmetric.

Received April 29, 1985
Revision received August 27, 1985 ■