3 A Path Taken: Aspects of Modern Measurement Theory

R. Duncan Luce
University of California, Irvine

Let me begin by welcoming William Estes to the equivalence class of emeritus Harvard professors. I can reassure him from personal experience that it does not mean a life of idleness; of course, he has assured that himself by becoming the founding editor of *Psychological Science*, the flagship journal of the American Psychological Society.

SOME PATHS OF MATHEMATICAL PSYCHOLOGY

Although Bill Estes' and my careers have run in many parallel strands, interweaving in a number of ways, we have in fact pursued quite different intellectual paths. Figure 3.1 indicates six of the main paths (plus a small illustrative sample of names for each) that have evolved in mathematical psychology during this time. A major distinction among these paths is whether one treats the topic of study as a black box whose properties are to be captured by the mathematical model or whether one attempts to model some of the structure within the black box, sometimes to the point of resembling what neural scientists tell us about neural networks.[1] Within each are further splits that to a great extent reflect different mathematical methodologies. Some scientists, like Patrick Suppes, have pursued several of these paths; others, like me, two plus a little of a third; and still others, like Bill, have been true to one.

[1]A similar split is found in other sciences—quantum mechanics and relativity theory are the classic physical examples from the first part of this century. A closely related pair at the two levels is electromagnetic theory and the theory electrons. In general, each approach seems to have its role.

As a consequence of our different paths, my remarks, unlike those of many other contributors to this symposium, do not bear directly on anything that Bill has done. Rather, I describe some of the progress achieved in one of the other paths, measurement theory.

Before doing that, a comment is warranted on the relation of measurement theory to stochastic modeling and, in particular, statistics. The distinction is this: Measurement focuses on structural relations among variables, which, to use Norman Anderson's (1981) term, constitutes a cognitive algebra. Statistics, by contrast, focuses on the chance or error aspects of our measurements. In an ideal science, they would be one subject, but so far that goal has eluded us. When we don a statistical cap, we explore the adequacy of very simple algebraic models—usually purely ordinal, additive, or multiplicative ones—and we do little to understand the qualitative conditions that underlie such a representation. When we don a measurement cap, we explore the representation, but with considerable discomfort over errors of measurement because we do not really know how simultaneously to model the algebra and the errors. Achieving unity of these approaches appears to be a difficult and probably a deep problem.

ORIGINS OF MODERN DEVELOPMENTS IN MEASUREMENT

The Campbell–Stevens debate (late 1930s and early 1940s)

Psychology, and more generally the social sciences, have found it impossible to mimic directly physical measurement that, in part at least, was based on the *counting of units*. This is familiar from length and mass measurement and had been embodied axiomatically in the model now called *extensive measurement* (Campbell, 1920, 1928; Helmholtz, 1887; Hölder, 1901; for a later treatment see chapter 3 of Krantz, Luce, Suppes, & Tversky, 1971). Coupling this failure with the powerful intuition that some psychological variables should be measured in a continuous fashion, attempts at measurement were being made outside the counting framework.

The legitimacy of these efforts was challenged by some physical scientists, notably the physicist-turned-philosopher-of-science N. R. Campbell. At his instigation, the British Association for the Advancement of Science constituted a committee to look into the issue. A number of psychologists testified, but the major counterposition was ultimately formulated by the Harvard psychophysicist S. S. Stevens (1946, 1951). The final report of the commission stated both positions (Ferguson et al., 1940). To a man, the physicists held that measurement means just one thing: It must be possible to combine entities exhibiting the attribute to be measured into an entity that also exhibits the same attribute, and

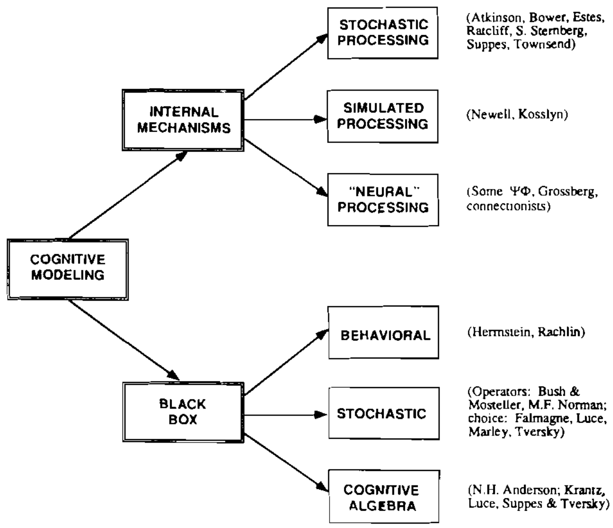

FIG. 3.1. Some paths in cognitive model building.

this combining operation is used to count the number of units approximately equal to the entity to be measured. Exactly how averaging representations, which certainly play a role in physics, were to be encompassed in this formulation was not made clear. In contrast, Stevens (and other psychologists) took the position that no one method is necessary in order for something to be classed as measurement, but rather it is a question of the degree of uniqueness achieved in the numerical representation of some body of empirical information. The key was the *scale type* as embodied in the degree of scale invariance, such as up to the specification of a unit, as in length and mass. Stevens (1946, 1951) cited five types: nominal, ordinal, interval, ratio, and absolute. The former differs from all of the others in that no ordering is involved, and today it is no longer treated as a significant part of the measurement taxonomy. Stevens (1959) later added another scale type, log-interval, but from the modern point of view that type is not really distinct from interval scales.

So far as I can tell, Stevens' taxonomy grew out of an examination of what had arisen in physics.[2] For example, on p. 34 of Stevens (1959) we find "The five scales listed . . . would seem to exhaust the possibilities, at least those of scientific interest, except perhaps for the class of scales on which no transformations would be possible. These would be ratio scales having in some sense or

[2]This examination first took place during an interdisciplinary faculty seminar at Harvard during the late 1930s. The mathematician George Birkhoff appears to have been highly influential in bringing out the significance of groups of transformations in this classification.

other a natural unit. . ." The Moh scale of hardness was a practical example of ordinal measurement; temperature (before the discovery of absolute zero) and time were examples of interval measurement; length, mass, duration (= time interval), and charge, of ratio measurement; and probability, perhaps, of absolute measurement.

Developing Evidence that Campbell Was Too Narrow (1950–1960s)

Although physics offered numerous examples of measurement that did not fall under Campbell's narrow filter, and he was obliged to devise his concept of derived measurement to accommodate them, it was not until somewhat later that it became very clear (at least to behavioral scientists) that his dictum was simply too narrow. Perhaps the single most telling example was the development of *additive conjoint measurement* (Debreu, 1960; Luce & Tukey, 1964) in which no explicit combining operation was evident in the primitives.[3] Rather, there were tradeoffs between two (or more) variables, each of which affect the attribute to be measured, and these tradeoffs can, under appropriate circumstances, be used to construct an indirect scheme for counting units. The situation described by this model was not only of considerable interest in the behavioral and social sciences, where such tradeoffs had been widely studied in learning, motivation, and perception, but it was also the actual basis of much physical measurement. Indeed, a careful examination of the entire structure of physical quantities showed that it consisted of a mix of conjointly and extensively measured variables that are interconnected by means of certain qualitative distribution laws, with neither seeming to have primacy over the other (see Krantz et al., 1971, ch. 10; and Luce, Krantz, Suppes, & Tversky, 1990, ch. 22).

In terms of Stevens' scale types, such conjoint measurement on continuous domains resulted in interval measurement when the additive representation (using all of the real numbers) is employed or log-interval when the multiplicative representation (on positive reals) is employed. The latter arises from the former by an exponential transformation.

Typically, physicists use multiplicative representations on the positive reals, as in

$$KE = \tfrac{1}{2}mv^2,$$

but there are exceptions such as dB measures and entropy. Economists and psychologists more often use an additive representation on all of the reals; for

[3]The general ideas formalized in conjoint measurement were explicitly discussed by Archimedes, but they were largely ignored during the formal development of modern measurement theory. In particular, no mention of them is found in Campbell's books.

example, utility models and many ANOVA models, but again there are exceptions such as the current attention in statistics to log-linear models.

I mention several other developments more briefly. Closely related to conjoint measurement was that of *difference measurement* (Debreu, 1958; Suppes & Winet, 1955) in which the qualitative data are judgments about the ordering of stimulus pairs according to some attribute, such as similarity. Again, in continuous domains this results in interval scales, and it serves as a building block for some models leading to geometric representations.

A different direction continues to involve an operation of combining, but instead of being positive in the sense that x combined with y is larger than both x and y, as is true of extensive measurement, the combined value lies between them. This is known as *intern* (or *intensive*) *measurement* (Aczél, 1948; Pfanzagl, 1959; von Neuman & Morgenstern, 1947). It leads to a representation in terms of some form of averaging, and in sufficiently rich contexts it usually results in interval scales but sometimes in ratio scales (see the last section).

So far, the examples have all fallen within Stevens' taxonomy, but there were even exceptions to that. Two of the most notable were Coombs' (1950) *unfolding* model for preferences and my (Luce, 1956) attempt to weaken the order axiom so that the strict order remains transitive but indifference need not be. These *semiorders* captured the qualitative notion of a threshold and under certain circumstances could be represented that way, but the scale type was never fully characterized. For recent treatments, see Fishburn (1985) and Suppes, Krantz, Luce, & Tversky, (1989).

Perplexities

So, by the mid1960s we found ourselves with a number of examples, most, but by no means all of which fell within Stevens' scheme. Moreover, despite the fact that direct counting of units was not involved, all of the examples involved addition one way or another in the representation. We clearly did not understand very well the full extent of Stevens' claim. For example, at that time the following questions were unresolved, and most psychologists suspected they would not soon be solved.

- What scale types are there other than Stevens' ordinal, interval, ratio, and absolute?
- Among all of them, whatever they are, why do Stevens' four types seem so pervasive?
- Just how crucial is addition in measurement representations?
- In particular, are nonadditive representations of any theoretical or scientific significance?

MODERN DEVELOPMENTS IN MEASUREMENT

Nonadditive Structures (1976–1985)

Mathematically, it is easy to envisage potential nonadditive conjoint representations. Consider, for example,

$$F(x,y) = x + y + \begin{cases} Cxy, & \text{if } x \geq 0,\ y \geq 0 \\ 0, & \text{if } xy < 0 \\ Kxy, & \text{if } x \leq 0,\ y \leq 0 \end{cases} \tag{1}$$

and order stimuli by:

$$(a,x) \gtrsim (b,y) \text{ if and only if } F(a,x) \geq F(b,y).$$

The question raised by Narens and Luce (1976) was: What properties need an ordering satisfy in order for it have some sort of nonadditive representation, for example, the one just described? It turns out that this question concerning conjoint structures, where it arises very naturally, can be reduced, when the domain of the conjoint structure is suitably rich, to a comparable question about an ordering and an operation of combining. And we gave suitable axioms on an operation to have a non-additive numerical representation. Basically all we did was drop associativity[4] from the axioms of extensive measurement.

But in that paper we failed to characterize the scale type. We were able to show that, like extensive measurement, if one value of the representation is prescribed, then the representation is uniquely determined. But we were not able to state the family of transformations that took one representation into another on the same numerical structure. This was somewhat perplexing, and only later did the reasons for our difficulty become apparent.

The breakthrough came when Michael Cohen, in his second-year graduate research paper at Harvard, proved that the scale type was either ratio or a proper subgroup of the similarity transformations of the ratio case. That striking result is included in Cohen & Narens (1979), where they also showed that the ratio case has a particularly simple representation. Later Luce & Narens (1985) extended this result to include additional scale types, including the interval scale case of intensive measurement. The basic form of the representation is that there exists a function f from the positive reals onto itself such that f(x) is strictly increasing and f(x)/x is strictly decreasing in x, and the numerical operation $\otimes$ is given by:

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

---

[4]If $\bigcirc$ denotes the operation, associativity says that groupings of the binary operation over three or more variables does not matter, that is, for all elements, x, y, and z, $x\bigcirc(y\bigcirc z)$ is equivalent to $(x\bigcirc y)\bigcirc z$.

Note that this form is invariant under ratio scale transformations because:

$$rx \otimes ry = ryf(rx/ry) = r(x \otimes y).$$

Detailed summaries of these developments can be found in chapter 19 of Luce *et al.* (1990) and Narens (1985).

Subconclusions

The examples that were developed during this period made clear that formal theories of measurement—formal in the sense of axiomatizing an ordering over a structure and finding a numerical representation of that information—need not end up with addition or multiplication playing any role whatsoever in the representation. Yet, addition and multiplication continued to play an explicit role in two of the stronger Stevens scale types, namely, interval and ratio. In the interval case the admissible transformations of a representation ϕ are of the form:

$$\phi' = r\phi + s,$$

and obviously both addition and multiplication are involved.

So the question arose as to the degree to which this is a deep aspect of measurement and the degree to which it is superficial.

GENERAL THEORY OF SCALE TYPES (1980s)

Symmetries

Suppose a structure $\mathcal{S}$ has a ratio scale representation. This means that if ϕ is a numerical representation (isomorphism) of the structure $\mathcal{S}$ with a numerical structure $\mathcal{N}$ then $r\phi$ is an equally good representation in the same structure. This is diagrammed in Fig. 3.2. Note that the mapping $\alpha = \phi^{-1}r\phi$ is an isomorphism between the structure $\mathcal{S}$ and itself. Physicists refer to such an isomorphism as a *symmetry* of the structure; mathematicians speak of an *automorphism* (= self-isomorphism).

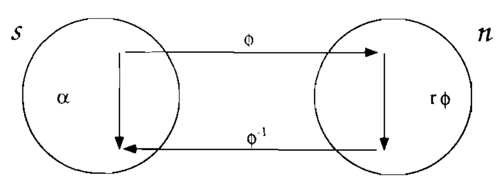

FIG. 3.2. The relations among numerical isomorphisms, admissible transformations, and symmetries for the ratio case.

Properties of Symmetries

The new focus that arose in the 1980s, initiated by Narens (1981a, b), is on the symmetries of a structure rather than on the structure itself and its representations. In a sense, the latter comes for free provided that the symmetries are sufficiently rich.

Narens posed the problem of classifying the symmetries and he introduced the following two key ideas. A structure is *homogeneous* if for each pair of points in the structure, some symmetry maps the one point into the other.[5] Ratio, interval, and ordinal scale structures whose ranges are onto an infinite interval of the real numbers are all homogeneous. Homogeneity excludes structures having an element that is structurally distinct from the others. Examples are structures with an identity (or zero) element, a maximum element, or a minimum element.

The other concept is that of a structure being *finitely unique* in the sense that for some integer N, whenever two symmetries agree at N points then they are, in fact, the same transformation. Another way of looking at the concept is that whenever the values of N distinct points are specified, the representation is uniquely determined. Both ratio and interval structures are finitely unique (with N=1 and 2, respectively). Ordinal structures are not finitely unique because it takes countably many values to specify a strictly increasing function completely.

Types of Scale Types on the Continuum

Consider structures whose ordering forms a continuum.[6] Then in terms of the concepts of homogeneity and finite uniqueness we may distinguish three quite distinct possibilities:

1. The structure is not finitely unique. This category includes ordinal cases as well as others for which the classification is not yet complete.
2. The structure is both homogeneous and finitely unique. Three distinct cases exist: ratio, interval, and ones lying between these two; e.g., $\phi' = k^n\phi + t$, where $k > 0$ is fixed, n ranges over all integers, and t over all reals.
3. The structure is not homogeneous. Here the classification, although formally complete, is not yet well understood except for some important special cases (see following).

[5]In group theory this is usually called "transitivity," but in the measurement context that usage is all-too-likely to generate confusion. The term "homogeneity" is reasonable in the sense that it means structurally that points are indistinguishable, which is pretty much the everyday meaning of the term. It is also the term used in a similar way in geometry.

[6]This means that the ordered domain is order isomorphic to the real numbers together with their natural order.

Thus, among the class of homogeneous and finitely unique structures nothing is possible other than ratio, interval, and cases in which the admissible transformation are a proper subset of the affine transformations of the interval case and properly include all of the similarity transformations of the ratio case. So far, no concrete example of this in-between scale type seems to have arisen in measurement. The reasons underlying Stevens' classification are now clear.

As I mentioned, Narens initiated the problem and achieved a partial solution; T. M. Alper, then a mathematics undergraduate at Harvard, solved it completely as his thesis, which was published as Alper (1987).[7]

Continued Role of Additivity

A symmetry is called a *translation* if either it is the identity, in which case every point is fixed in the sense of mapping into itself, or no point whatsoever is fixed. Among the affine transformations of interval scale measurement, the translations are of the form $\tau(x) = x + t$.

The key to the Alper-Narens result is in showing that:

(i) The set of all translations are homogeneous in the sense that given any two points, some translation maps the one point into the other.

(ii) The translations look like real numbers under addition, for example, $\sigma(x) = x + s$ and $\tau(x) = x + t$ imply $\tau(\sigma(x)) = x + s + t$.

The proof works equally well if (ii) is replaced by:

(ii′) The translations are isomorphic to a subgroup of the additive numbers.

Indeed, as Luce (1986, 1987) showed:

- In the homogeneous case, (i) and (ii′) provide a general recipe for measurement in the sense that given an axiom system one need only establish that (i) and (ii′) are true.
- When that is true, then it is easy to construct an explicit numerical representation in which the translations appear as multiplication by a positive constant. These representations are referred to as *unit representations*.
- These structures are exactly those that can be appended to the existing measures of physics without disrupting the nice patterns that are exploited in dimensional analysis, namely, the relations among the representations as products of powers, as is evidenced in the pattern of physical units such as gram-centimeters/second^2.

[7]He received the 1989 Young Investigator Award of the Society for Mathematical Psychology for this paper.

Some Directions Currently Being Followed

1. The small number of homogeneous, finitely unique scale types invites an attempt to classify fully the structures of each type. Detailed results are known for binary conjoint structures, that is, orderings of two factors, and for structures with a binary operation (Luce & Narens, 1985). Some of this is illustrated later in the applications.
2. Results exist about nonhomogeneous structures with an operation and with singular points (zero, maxima, or minima) but that are homogeneous between adjacent singular points (Luce, 1991b). Examples of nonhomogeneous structures are:
    - Relativistic velocity, for which both a maximum (the velocity of light) and a minimum (which is also a zero, namely no relative motion) exist.
    - Prospect theory (Kahneman & Tversky, 1979; Tversky & Kahnneman, 1990), which is a generalization of expected utility theory in which there is a distinctive zero, the status quo, that divides consequences into gains and losses.
    - Continuous probability, with both a maximum (the universal event) and a minimum (the null event), fails to be homogeneous between them.
3. A comprehensive theory of meaningful propositions within a measurement structure has been under development for some years. Stevens (1951) argued that invariance of statistical propositions under admissible scale transformations was required for these propositions to be meaningful.[8] Luce (1978) showed that the physically important principle of dimensional invariance, which underlies the method of dimensional analysis that is used in physics, engineering, and to a lesser degree in biology, is merely a special case of demanding that physical laws be invariant under the symmetries of the structure of physical quantities. More recently, Narens (in preparation) has been exploring the concept from a deep logical perspective, showing precisely the interplay of meaningfulness and invariance ideas that have arisen in geometry, physics, mathematics, and measurement.

APPLICATIONS

Philosophically the results I have described are satisfying—one feels that a previously unexplored portion of the world (or, more apt today, the solar system) has been partially mapped. But aside from increased clarity, are the results good

[8]Unfortunately, he cast his point mainly in terms of statistics per se rather than the propositions in which they occur. Doing so fueled a confused controversy that probably is still not settled for everyone. See Luce et al. (1990), Section 22.6.

for anything? I think so, and to illustrate in what way I describe briefly two areas of application, namely, psychophysical matching and the theory of choices among uncertain alternatives.

Psychophysical Matching

Consider the class of procedures where a person "matches" stimuli from one physical modality to those from another—light to sound, length to weight, and so on—but does so under instructions to maintain stimulus ratios. The most obvious example is the cross-modal matching of Stevens (1957), but there are others that are less obvious. For example, one can "match" within a single domain as follows: Define M to be the match between intensity I and the intensity increment $\Delta(I)$ that makes I and $I+\Delta(I)$ "just noticeably different."

Luce (1959) stated a theory of matching that postulated that an admissible change in the representation of the stimulus being matched should reappear as an admissible change in the representation of the matching stimulus. Indeed my claim was that this can be considered a principle of theory construction akin to dimensional invariance, but this contention was challenged by Roozeboom (1962a, b; Luce, 1962). One issue is: Why should the scientists' method of representing physical information about the stimuli be of any relevance to the person doing the matching? The major conclusion of that theory was that such matching should be represented by power functions of the stimulus representations. This was of interest because it accorded (approximately) with the empirical results of magnitude estimation and cross-modal matching.

An alternative, more substantive theory was proposed by Krantz (1972) that was based, in part, on then unpublished ideas of Shepard (published later in 1978 and 1981). In his "relation theory" it is assumed that matches are always ratio judgments, implicit if not explicit. This also led to power function interconnections which, however, were somewhat obscured in this theory by an undetermined monotonic transformation. So relation theory seems to me both too strong in its supposition that there are implicit standards and, possibly, too weak in the conclusions reached.

Another approach, one that escapes the weaknesses of each of the earlier attempts, is based on exploring the hypothesis that the psychology embodied in a matching relation M is compatible with the physics embodied in the domains being matched. This can be formulated precisely in terms of the translations of the physical structures. A matching relation M is said to be *translation consistent* if associated with each translation τ of the independent variable, there is a translation σ_τ of the dependent variable that maintains the match, that is,

$$xMu \Leftrightarrow \tau(x)M\sigma_\tau(u).$$

Suppose the structures being matched have unit representations φ and ϕ. Then Luce (1990) showed that for some positive constants k and β,

$$xMu \Leftrightarrow \varphi(u) = k\phi(x)^{\beta}.$$

For the jnd example, it says that $\Delta(I) = kI^{\beta}$, which includes Weber's law ($\beta = 1$) as a special case as well as the "near-miss to Weber's law" in pure tones (Jesteadt, Weir, & Green, 1977; McGill & Goldberg, 1968).

Other results are given concerning the class of relations having the same exponent and the relation of matching to ratio judgments.

In summary, then, this theory says that if the psychology is nicely compatible with the physics, then matching corresponds to a power function relation in terms of the unit structure representations. The compatibility is not forced, however, and it is an empirical matter, not one of units of representations, to decide if it obtains.

Utility Theory

Choices among uncertain alternatives provide a rich structure to which these measurement ideas can also be applied. The major primitives are consequences, events, finite gambles, and preference orderings. In particular, let $\mathcal{C}$ denote a (dense) set of *pure* consequences, such as sums of money received or lost. Let $\mathcal{E}$ denote a (dense) set of *chance events. A finite gamble* is simply an assignment of a pure consequence to each member of a finite partition of an event. For example, suppose the underlying event is the throw of a die, and the partition is into the subevents {1,2}, {3,4,5}, and {6}, then assigning the monetary consequences −\$5, \$2, and \$10, respectively, to these subevents is a finite gamble. Further, let the *domain of gambles,* $\mathcal{G}$, consist of all finite gambles and all those that can be formed inductively by substituting finite gambles for pure consequences, and let $\gtrsim$ denote a *preference ordering* over $\mathcal{G}$.

Classical Subjective Expected Utility (SEU) Theory

Since von Neuman and Morgenstern (1947) and Savage (1954), the major normative theory is *subjective expected utility* or, more briefly, SEU, which is based on four major tenants of rational behavior for choosing among uncertain alternatives (see Fishburn, 1982):

1. Preferences are *transitive*— $f \gtrsim g$ and $g \gtrsim h$ imply $f \gtrsim h$ —and *connected* — either $f \gtrsim g$ or $g \gtrsim f$.
2. Preferences are *monotonic:* If g' is generated from g by replacing a consequence of g by a more preferred consequence, then $g' \gtrsim g$.
3. *Independence of a common consequence:* Suppose g and h are binary gambles with consequences $x > y$. From g' from g and h' from h by changing y to x over a common subevent. Then $g \gtrsim h$ iff $g' \gtrsim h'$.
4. *Universal accounting equivalences:* For any two gambles, no matter how

> structured into successive subgambles, if they reduce to the same normal form, that is, yield the same outcomes under the same chance events (ignoring the order of occurrence of sequences of independent events), then they are judged indifferent in preference.

With suitable assumptions about the richness of the domain, these properties are equivalent to the existence of a preference order-preserving function U over $\mathcal{G}$ and a probability measure $\mathcal{P}$ over $\mathcal{E}$ such that U(g) equals the expectation of U calculated for the several consequences of g's relative to the probability measure P. (The term "subjective" of SEU reflects the fact that the expectation is calculated in terms of a decision-induced probability P, rather than an objective one, as was assumed in the original von Neumann and Morgenstern presentation.) Moreover, each of the rationality axioms is implied by the SEU representation.

Troubles With SEU

There are real and pseudo problems with SEU. One of the pseudo problems that tends to afflict cognitive psychologists is the claim that the average person surely cannot mentally calculate expectations, and so a priori the theory must be wrong. This has about the same merit as saying that the theory of physical dynamics must be wrong because a chunk of matter does not embody enough computing power to solve the differential equation describing its motion. Such assertions confuse the scientist's description of the behavior with the source of that behavior.

Turning to the real empirical issues, I mention four.

Universal accounting seems unlikely to be descriptive; however, there are surprisingly few data devoted to showing that it is wrong. Models that attempt to limit its scope can be viewed as an aspect of bounded rationality, a term introduced by Simon (1955, 1978).

Some experiments raise doubts about transitivity and monotonicity. In my opinion their significance has probably been either exaggerated and/or is in some measure a misinterpretation of the data. For more detailed discussion, see Luce (1992). Certainly both properties are normatively compelling to many scientists and decision makers.

Common consequence is highly suspect for uncertain events—those for which there is no known objective probability—as was first pointed out by Ellsberg (1961). Despite its apparent reasonableness, people who violate it do not generally alter their choices to conform to it when the violation is pointed out. This contrasts sharply with their desire to resolve violations of monotonicity and, especially, transitivity.

SEU makes no distinction between gains and losses, which is universally recognized to be behaviorally significant for decision makers. Attempts have been made to deal with this in terms of the shape of the utility function, but since in SEU utility is an interval scale measure, such discussions do not really make any sense.

So, I turn to recent attempts to modify the theory, of which there are two major types.

Rank-Dependent (RD) Utility

The rank-dependent theories, which originated with Quiggin (1982) and by now come in a number of variants (see Luce, 1988 for a listing), entail a representation that, like SEU, involves a weighted average of the U-values of each consequence where, however, the weights depend both upon the event giving rise to the consequence, as in SEU, and the rank position of that consequence relative to the other consequences in the gamble.

As an example of such a representation, consider the binary gamble (x,E;y, ¬E) where ¬E means "not E." Then the *rank-dependent* (RD) representation takes the form:

$$U(x,E;y,\neg E) = \begin{cases} U(x)S_{>}(E) + U(y)[1 - S_{>}(E)], & \text{if } x > y, \\ U(x), & \text{if } x \sim y, \\ U(x)S_{<}(E) + U(y)[1 - S_{<}(E)], & \text{if } x < y. \end{cases}$$

where the weights, S, depend not only on the event E but also on the preference order between the two consequences.

The rank-dependent representation implies monotonicity and transitivity of preference; it need not imply independence of common consequences because the weights, unlike probabilities, need not be additive over disjoint events; and most importantly it does not imply most of the accounting equivalences. The key one that it does imply is called *event commutativity:* Suppose a gamble is conducted in two stages, each of which rests on running an independent chance experiment. Suppose, further, that x is the consequence when events E and F both occur, one in the first experiment and the other in the second independent experiment, and y is the consequence otherwise. The assertion is that the decision maker is indifferent as to the order in which these events are carried out. Because event commutativity is a major implication of rank dependent theories, Alan Brothers (1990), a recent PhD of the University of California-Irvine, has studied it empirically. It appears to be satisfied by, perhaps, a quarter of student subjects, but is systematically violated by others who seem to be using simplifying heuristic rules.

In the context of binary gambles, Luce and Narens (1985) proved that rank dependence is the only possible generalization of SEU that retains transitivity and monotonicity of preference, event commutativity, and the existence of an interval scale representation.

Although rank dependence accommodates many of the empirical anomalies of SEU, it fails to take into account the special status of the status quo (or more generally of some aspiration level) and of the asymmetries of gains and losses relative to that point. To deal with that one has to abandon purely homogeneous theories and to develop those that have at least one highly distinctive point.

Rank- and Sign-Dependent (RSD) Utility

So far three related efforts in this direction have appeared. Edwards (1962) commented on the possibility of ratio scale measurement of utility when the weights are not additive and discussed the importance of the status quo in defining gains and losses. Kahneman and Tversky (1979) gave a *rank- and sign-dependent* (RSD) utility representation that they called *prospect theory*. Its domain of application was restricted to gambles with at most one positive, one negative, and one null consequence. Tversky and Kahneman (1990) generalized it to finite gambles. Recently I have been working on a different axiomatic, measurement theoretic generalization of prospect theory leading to a similar representation (Luce, 1991a; Luce & Fishburn, 1991). What sharply distinguishes our theories from the earlier ones (an exception is Pfanzagl [1959]) is not only an explicit focus on gains and losses, but the introduction of a binary operation $\oplus$ of joint receipt: If f and g are gambles, $f \oplus g$ denotes the receipt of both f and g. Having this allows one to formalize some of the editing processes discussed informally by Kahneman and Tversky.

In the simplest version of the theory, this operation is assumed to be extensive in character and so there is a utility function U that is additive[9] over $\oplus$, i.e.,

$$U(x \oplus y) = U(x) + U(y).$$

To deal with the asymmetries of gains and losses, note that any gamble g is formally equivalent to a composite gamble consisting of three subgambles:

a gamble, g^+, of only gains, which is conditional on a subevent $E(+)$ occurring,

a gamble of no change, e, conditional on a subevent $E(0)$ occurring,

a gamble, g^-, of only losses, which is conditional on a subevent $E(-)$ occurring,

where $E(+) \cup E(0) \cup E(-) = E$ is the event on which g is conditional.

A major axiom of the theory is the (rational) accounting assertion that people perceive a gamble as indifferent to its formally equivalent composite gamble of gains and losses, that is,

$$g \sim (g^+, E(+); e, E(0); g^-, E(-)).$$

[9] A somewhat weaker assumption, like Equation 1, is used in Luce and Fishburn (1991), but it is simpler here to exposit just the additive case.

A second major axiom of the theory is that this composite gamble can be recast as the joint receipt of the gamble of gains pitted against no change and the gamble of losses pitted against no change. The assumption is that decision makers (irrationally) perceive the gamble g as equivalent to this joint receipt, that is,

$$g \sim (g^+,E(+);e,E(0)\cup E(-))\oplus(g^-,E(-);e,E(0)\cup E(+))$$

And a third axiom is an editing one that strikes me as being rational in character, although it is not usually listed among the axioms of rationality because $\oplus$ is not usually among the primitives.[10] The axiom is: A finite gamble that is composed entirely of gains (losses) is perceived as indifferent to the joint receipt of the smallest gain (loss) and the modified gamble in which each consequence is reduced by that smallest gain (loss).

RSD Representation

The assumptions just described, along with transitivity and monotonicity, lead to a representation U involving two sets of weights, W^+ and W^-, both assuming values in the unit interval [0,1] and defined over all the events. This representation is both rank and sign dependent in the following sense: Let $g \sim (g^+,E(+);e,E(0);g^-,E(-))$. Then

(i) $U(g) = U(g^+)W^+[E(+)] + U(g^-)W^-[E(-)]$.
(ii) $U(g^+)$ has an ordinary RD representation with weights that also depend on +.
(iii) $U(g^-)$ has an ordinary RD representation with weights that also depend on −.
(iv) U is a ratio scale.

The ratio scale feature arises from two facts: that U(e) = 0 and that in property (i) the weights do not, in general, add to 1, that is,

$$W^+[E(+)] + W^-[E(-)] \neq 1.$$

This nonadditivity of the weights arises both because U(e) = 0 wipes out any weight assigned to E(0) and because W^+ is generally not closely related to W^-.

Kahneman and Tversky's (1979) prospect theory is the special case of RSD applied to gambles with only one gain and one loss and some added constraints

[10]As noted earlier, Pfanzagl (1959) is the only exception that I know of. He studied the axiom stated here, which he called the *consistency principle*, in the binary case but without distinguishing between gains and losses, as Fishburn and I do.

among the weights. Prospect theory has accounted for a good deal of anomalous data, and Tversky and Kahneman (unpublished) have tested, with apparent success, the more general representation. Among the issues to be studied further empirically, three stand out:

- In the domains of all gains and all losses, does event commutativity hold, as the theory says it must because the representation is rank dependent in these domains?
- Rank dependence arises primarily from the editing axiom, and so that axiom needs to be studied directly.
- The connection between gains and losses is achieved into two ways, both of which need careful examination. One is the additive representation of the joint receipt operation $\oplus$, although that can be weakened somewhat to the form of Equation 1. The other, and key, connection is the nonrational decomposition of a gamble of gains and losses into the joint receipt of two gambles, the one of gains pitted against the null consequence and the other of losses pitted against the null consequence. One study, Slovic and Lichtenstein (1968), supports the latter assumption. But both assumptions require considerably more empirical study and possible modification.

I consider the work done to date to be just a beginning of research on ratio scale theories of utility. We have explored only those ratio theories that involve a heavy dose of averaging, but as I have shown for general n-ary operations (Luce, unpublished) a large family of nonadditive, nonaveraging possibilities exists. Little is known about axiomatizing specific members of that family, but the need to restrict possibilities certainly invites the formulation of new behavioral axioms. Additional general theory about nonhomogeneous outcomes, especially those that are homogeneous between singular outcomes, is needed as input to these more psychological applications.

ON PATHS NOT TAKEN

In reflecting on the different paths taken by Bill and me and others, I was reminded of the closing lines of Robert Frost's poem *The Road Not Taken.* The "I" in the poem could be either of us as, in the 1950s, I'm not sure which path would then have been deemed the one less traveled.

> Two roads diverged in a wood, and I—
> I took the one less traveled by,
> And that has made all the difference.

ACKNOWLEDGMENTS

Preparation of this chapter was supported in part by National Science Foundation Grant IRI-8996149 to the University of California, Irvine. Louis Narens, who has been a major figure in the developments reported here, provided useful comments on this paper as did Patrick Suppes, who served as a referee.

REFERENCES

Aczél, J. (1948). On mean values. *Bulletin of the American Mathematical Society, 54,* 392–400.

Alper, T. M. (1987). A classification of all order-preserving homeomorphism groups of the reals that satisfy finite uniqueness. *Journal of Mathematical Psychology, 31,* 135–154.

Anderson, N. H. (1981). *Foundations of information integration theory.* New York: Academic Press.

Brothers, A. J. (1990). *An empirical investigation of some properties relevant to generalized expected utility.* Doctoral dissertation, University of California, Irvine, CA.

Campbell, N. R. (1920). *Physics: The elements.* Cambridge: Cambridge University Press. (Reprinted as *Foundations of science: The philosophy of theory and experiment.* New York: Dover, 1957.)

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

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

Coombs, C. H. (1950). Psychological scaling without a unit of measurement. *Psychological Review, 57,* 145–158.

Debreu, G. (1958). Stochastic choice and cardinal utility. *Econometrica, 26,* 440–444.

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.

Edwards, W. (1962). Subjective probabilities inferred from decisions. *Psychological Review, 69,* 109–135.

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

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. (1982). *The Foundations of Expected Utility.* Dordrecht: Reidel.

Fishburn, P. C. (1985). *Interval Orders and Interval Graphs.* New York: Wiley.

Helmholtz, H. von (1887). Zählen und Messen erkenntnis-thoretisch betrachet. *Philosophische Aufsätse Eduard Zeller gewidmet,* Leipzig. English translation by C. L. Bryan, (1930). *Counting and measuring.* Princeton, NJ: Princeton University Press.

Hölder, O. (1901). Die Axiome der Quantität und die Lehre vom Mass. Sächische Akademie Wissenschaften zu Leipzig, Mathematisch-Physische Klasse, 53, 1–64.

Jesteadt, W., Wier, C. C., & Green, D. M. (1977). Intensity discrimination as a function of frequency and sensation level. *Journal of the Acoustical Society of America, 61,* 169–177.

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

Krantz, D. H. (1972). A theory of magnitude estimation and cross-modality matching. *Journal of Mathematical Psychology, 9,* 168–199.

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

Luce, R. D. (1956). Semiorders and a theory of utility discrimination. *Econometrica, 24,* 178–191.

Luce, R. D. (1959). On the possible psychophysical laws. *Psychological Review, 66,* 81–95.

Luce, R. D. (1962). Comments on Rozeboom's criticisms of 'On the possible psychophysical laws.' *Psychological Review, 69,* 548–551.

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

Luce, R. D. (1986). Uniqueness and homogeneity of ordered relational structures. *Journal of Mathematical Psychology, 30,* 391–415.

Luce, R. D. (1987). Measurement structures with Archimedean ordered translation groups. *Order, 4,* 165-189.

Luce, R. D. (1988). Rank-dependent, subjective expected-utility representations. *Journal of Risk and Uncertainty, 1,* 305–332.

Luce, R. D. (1990). "On the possible psychophysical laws" revisited: Remarks on cross-modal matching. *Psychological Review, 97,* 66–77.

Luce, R. D. (1991). Rank- and sign-dependent linear utility models for binary gambles. *Journal of Economic Theory, 53,* 75–100.

Luce, R. D. (1992). Where does subjective expected utility fail descriptively? *Journal of Risk and Uncertainty, 5,* 5–27.

Luce, R. D. (unpublished). Generalized concatenation structures that are homogeneous between singular points. Manuscript submitted for publication.

Luce, R. D., & Fishburn, P. C. (1991). Rank- and sign-dependent linear utility models for finite first-order gambles, *Journal of Risk and Uncertainty, 4,* 29–59.

Luce, R. D., Krantz, D. H., Suppes, P., & Tversky, A. (1990). *Foundations of Measurement* (Vol. III). New York: Academic Press.

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.

McGill, W. J., & Goldberg, J. P. (1968). A study of the near-miss involving Weber's law and pure-tone intensity discrimination. *Perception & Psychophysics, 4,* 105–109.

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. (in preparation). *A theory of meaningfulness.*

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

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

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organization, 3,* 324–343.

Rozeboom, W. W. (1962a). The untenability of Luce's principle. *Psychological Review, 69,* 542–547.

Rozeboom, W. W. (1962b). Comment. *Psychological Review, 69,* 552.

Savage, L. J. (1954). *The foundations of statistics*. New York: Wiley.

Shepard, R. N. (1978). On the status of 'direct' psychophysical measurement. In C. W. Savage (Ed.), *Minnesota studies in the philosophy of science*, Vol. IX (pp. 441–490). Minneapolis, MN: University of Minnesota Press.

Shepard, R. N. (1981). Psychological relations and psychophysical scales: On the status of "direct" psychophysical measurement. *Journal of Mathematical Psychology, 24,* 21–57.

Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics, 69,* 99–118.

Simon, H. A. (1978). Rationality as process and as product of thought. *The American Economic Review: Papers and Proceedings, 68,* 1–16.

Slovic, P., & Lichtenstein, S. (1968). Importance of variance preferences in gambling decisions. *Journal of Experimental Psychology, 78,* 646–654.

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. (1957). On the psychophysical law. *Psychological Review, 64,* 153–181.

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.

Suppes, P., Krantz, D. H., Luce, R. D,. & Tversky, A. (1989). *Foundations of measurement* (Vol. II). New York: Academic Press.

Suppes, P., & Winet, M. (1955). An axiomatization of utility based on the notion of utility differences. *Management Science, 1,* 259–270.

Tversky, A., & Kahneman, D. (unpublished). *Cumulative prospect theory: An analysis of decision under uncertainty,* manuscript.

von Neumann, J., & Morgenstern, O. (1947). *The theory of games and economic behavior* (2nd ed.). Princeton, NJ: Princeton University Press.