JOURNAL OF MATHEMATICAL PSYCHOLOGY **41**, 79–87 (1997)
ARTICLE NO. MP971150

Several Unresolved Conceptual Problems of Mathematical Psychology

R. Duncan Luce

University of California, Irvine

This article is a personal commentary on the following, to my mind, unresolved issues of mathematical psychology: (1) the failure of the field to have become a fully accepted part of most departments of psychology; (2) the great difficulty we have in studying dynamic mechanisms, e.g., learning, because large samples are difficult to obtain: time samples wipe out the phenomena and subject samples are unrepresentative because of profound and ill understood individual differences; (3) the failure to unify successfully statistical and measurement theories which, I believe, are two facets of a common problem; (4) the proliferation of free parameters in many types of theories with little success in developing theories of such parameters; (5) the difficulties we have had in successfully formulating the mathematics of uncertainty and vagueness; and (6) the issues of modeling what are presumably discrete attributes and phenomena by continuous mathematics: when and how is this justified?

INTRODUCTION

For the 1995 meeting of the European Mathematical Psychology Group in Regensberg, Germany, Professor Jan Drösler urged me to assume the role of an elder statesman —a role which, aside from the elderly part, does not come very naturally. Indeed, his initial suggestion was that I propose an agenda for mathematical psychology for the coming decades. His models for my talk were Felix Klein's Erlangen program for geometry and/or David Hilbert's famous list of problems for mathematics. Being neither of the caliber of Klein or Hilbert nor concerned with pure mathematics, that tack did not strike me as viable. So, instead, I spoke on a research topic which was then occupying my attention, namely, the conditions that cause several different ways of measuring utility all to agree (Luce, 1996). Nevertheless, Drösler continued to urge that my written paper be of a more general nature, and I have succumbed to his blandishments. But rather than recommend or (mis) predict the future of the field, I have opted to speak of some of my disappointments with the mathematical psychology of my era. Of course, such disappointments do not form a particularly coherent package, and so this paper necessarily has a somewhat disjointed quality, passing from one issue to another with little by way of transition. Moreover, the topics do not seem to have a natural order, so I have organized them—and even this judgment is debatable—from what seems to me the more to the less general.

Reprint requests should be sent to R. Duncan Luce, Institute for Mathematical Behavioral Sciences, University of California, Irvine, CA 92697-5100. E-mail: rdluce@uci.edu.

LIMITED ACCEPTANCE BY PSYCHOLOGISTS

On entering the field 45 years ago I anticipated that as mathematical psychology developed, it would increasingly be incorporated into the intellectual life of departments of psychology. In the United States, that has not happened to any great extent. There are enclaves of strength such as those at Indiana, Irvine, Purdue, and Stanford plus several other departments with considerable strength in psychometrics, but most have at most one such person, usually hired to teach statistics. Departments typically make very limited mathematical demands on graduate students; a mathematical psychologist certainly is ill advised to give general colloquia that strongly emphasize mathematical theory; and our few mathematically sophisticated Ph.D.s often have difficulty finding suitable positions. Aside from those few departments with a focus on mathematical psychology or psychometrics, the only other exception to these remarks is another small set of departments that strongly emphasize vision and/or audition, where techniques of complex analysis are taken for granted.

One odd aspect of all this is that some individuals central to mathematical psychology—e.g., Atkinson, Estes, Shiffrin, Simon, Sperling, Suppes, and myself—have received considerable individual acclaim such as membership in the National Academy of Sciences. The field itself seems to be respected without being incorporated into the everyday life of psychology. One evidence for this is that the size of the Society for Mathematical Psychology has not grown over the years, certainly not in proportion to psychology as a whole or as much as certain specialty areas, such as neurosciences. This measure may, however, be misleading because a number of people developing mathematical models (including many of my colleagues) prefer to attend only meetings of substantive subfields.

What underlies this limited incorporation of mathematical psychology? I suspect that many people will argue that the field lacks many communicable findings, but I don't

0022-2496/97 $25.00

really think that position is easily defended unless one prefixes "communicable" by "easily." To my mind, three other reasons are far more significant. First, the mathematical level needed to do serious work has often turned out to be beyond and different from what can be obtained from a few lower division mathematics courses, in particular the standard calculus and linear algebra sequence. A second reason is the computer, which has made it relatively easy to simulate quite complex interactive systems. For many, it is clearly simpler and more agreeable to program than it is to study processes mathematically. Do not misunderstand; I am not a contemporary Luddite vainly trying to ban the computer, for it has clearly provided us with enormous power to augment what we can do experimentally and theoretically. Simulations, however, simply do not substitute for mathematical formulations and analysis, especially when one is trying to formulate the basic science rather than to work out the implications of well-accepted laws. One should not forget that most of the physics on which modern technology is based was developed long before modern computers. A third reason is the failure, especially in the United States, of the psychometric and mathematical modeling communities to form a strong intellectual alliance. Both groups have exhibited considerable mutual disdain.

Psychology contrasts sharply with economics, where it is routine for a substantial cadre of advanced researchers to know and to use considerable (classical) mathematics that is taught to many undergraduates. Moreover, mathematical economics and econometrics seem to coexist rather more effectively than do mathematical psychology and psychometrics. Whatever the reason, we mathematical theorists seem to be more of an endangered species in psychology.

With these political remarks said, I turn to more substantive issues.

NONSTATIONARITY AND INDIVIDUAL DIFFERENCES

When I first became aware of mathematical psychology in the late 1940s and early 1950s, an area then under active development was learning models: the operator models of Bush and Mosteller (1955), to which I contributed a nonlinear, commutative one (Luce, 1959, 1964), and the stimulus sampling/Markov chain models of Estes, Suppes, Bower, and others (see Laming, 1973, for a list of references). Because change—learning, in particular—is clearly of major interest and because a great deal of attention was paid to these dynamic models, I had anticipated that they would come to dominate the field. Although they may have helped spawn a number of things, including much work on attention, memory, and neural networks, they never did assume a role comparable to dynamics in physics. Why?

From the start we were fully aware of the inherent dilemma of studying change when there are substantial individual differences among our subjects. Ideally, one would like to study change in an individual, but it is virtually impossible to collect sufficient data to model what is going on except, possibly, in certain special situations such as lower mammals and birds in Skinnerian apparatuses. Basically, the problem is that the change is not repeatable in any simple way—one just does not know how to erase what has been learned, to return the organism to its original state, and to permit a repetition of the learning effort over and over until sufficient data are accumulated to see the details of the mechanism. Were the subjects all identical and statistically independent, as is true of classes of physical particles and as may be approximately true of relatively pure genetic strains of some mammals and insects, averaging over subjects would be fine. But whenever we take the effort to look at all carefully at individual people we see substantial differences, and so we are not justified in averaging except in the very special case where the predicted relation is linear. No other family of functions has the property that an average of several members from the family is also in the family.

These facts have not precluded most psychologists from using group data, but one must be very careful about the conclusions one attempts to draw from them and certainly one is almost never justified in naive averaging to construct learning curves, as we used to do.

Our lack of detailed knowledge about the dynamic mechanisms of behavior has the consequence that we are in absolutely no position to say anything solid about when to expect stable asymptotic behavior and when to expect chaos. The fact that some social and psychological phenomena seem rather chaotic has led some psychologists and sociologists to hope that we can use chaos theory to understand what is happening. To the best of my knowledge the asymptotic behavior, which is where one sees evidence of chaos, has been worked out only for some specific mathematical models. Little, if anything, is known about inferring the details of the dynamics from observations about the patterns of chaos. I suspect that like many other social science fads, this too will pass until we have enough well-confirmed dynamic theory to make it useful. For a sharply dissenting view, see Gregson (1988), and also note that a new journal *Nonlinear Dynamics, Psychology, & Life Sciences* has just been announced.

Is there a way to resolve the issue of dynamics in the presence of individual differences? I am not at all sure there is if the differences reflect basic differences in mechanisms among subjects. But to the degree the differences reflect distributions of parameter values in a common mechanism, then, in principle, the answer seems to be yes. However, the details certainly are not easy. Perhaps the people who have best developed strategies for this case are Jean-Paul Doignon, Jean-Claude Falmagne, and Falmagne's students Kamakshi Lakshminarayan and Michael Regenwetter

(Doignon & Falmagne, 1997; Falmagne, 1989, 1993, 1996; Falmagne & Doignon, 1997; Falmagne, Regenwetter, & Grofman, 1997; Lakshminarayan, 1996). Their underlying idea is simple enough, namely, to suppose that there is a distribution in the population of the parameter values of the dynamic process and to use group data to infer both the degree to which the model is correct and to some extent the distribution of parameter values.

The execution of this program seems to encounter two major hurdles. First, what should we suppose to be the nature of the distribution of parameters? That choice, as with most statistical models, is dictated more by mathematical convenience and tradition than by any very deep understanding of the distribution in the population. Some effort is made, using computer simulations, to show that the results are relatively robust under a range of possibilities. The other hurdle is that for even rather short learning sessions, incredibly large amounts of data must be collected and much computation is required. In work on knowledge structures, Falmagne and his associates (see Lakshminarayan, 1996) are finding that on the order of 1000 subjects are needed to study a learning chain on only five items. The growth in subjects and computer power needed tends to be exponential with the size of the learning task. These limitations strike me as probably severe.

A somewhat related issue involving statistics is my next topic.

STATISTICS AND MEASUREMENT THEORY

From my perspective, statistics and measurement theory are two faces of a common topic, namely, the recovery of (static) structure from primitive observations that are, to a degree, corrupted by sources of uncertainty or error, often lumped under the term "random noise." In statistics one typically works only with data that are already cast in numerical form and one postulates some algebraic form for the underlying structure. The two most studied are the linear model,

$$y = \sum_i \alpha_i x_i + \gamma, \tag{1a}$$

and the multiplicative one,

$$y = \prod_i x_i^{\beta_i} + \gamma. \tag{1b}$$

Others are sometimes examined by taking various conventional transformations of y, although the motive for doing so seems to be mainly to increase the validity of the statistical assumptions which usually involve a random perturbation—often assumed to be normally distributed—added to y or to the x_i.

In measurement theory the raw data are typically postulated to be ordinal in nature—some relational comparison of $\mathbf{x} = (x_1, ..., x_n)$ and $\mathbf{x}' = (x'_1, ..., x'_n)$—and the question addressed is the conditions under which the data structure exhibits numerical measures having certain properties. Until about 1985 a typical theorem took the following form: if certain properties hold on the ordinal relation, then there exist real-valued functions ϕ, ϕ_i, $i = 1, ..., n$, and F such that

$$\phi(\mathbf{x}) = F[\phi_1(x_1), ..., \phi_n(x_n)] \tag{2a}$$

is order preserving. Usually F was specified in advance, most commonly as

$$F(X_1, ..., X_2) = \sum_i X_i. \tag{2b}$$

More recently, more sophisticated questions have been addressed: Do broad classes of ordered structures have representations that are of ratio or interval scale type (Stevens, 1946, 1951) and, if so, how can their mappings be constructed (Alper, 1985, 1987; Narens, 1981a, b)? These results do not prejudge the form of the representation beyond the condition that certain symmetries of the underlying structure—-called the *translations*—take the numerical form of difference transformations. A good deal is now known about the possible representations under the totally unrealistic assumption that the data are noise free. (Much of this knowledge is summarized in the three volumes of the *Foundations of Measurement*, Krantz, Luce, Suppes & Tversky, 1971; Suppes, Krantz, Luce, & Tversky, 1989; Luce, Krantz, Suppes, & Tversky, 1990.)

The difficulty lies in putting these two faces together, neither prejudging the structural form nor pretending that the data are noise free. This has turned out not to be easy, even though the need has been recognized for decades. If one begins with noisy qualitative data, then one simply does not know how to associate statistics to the existing qualitative mathematics describing the underlying structure. And if one begins with noisy numerical data, one does not know how to do the statistical analysis of all possible monotonic transformations of the data and then to select the best fitting one. In practice, one takes one of two extreme tacks: either treat the preponderance of choices in some sort of repeated design as establishing an algebraic order or subject numerical rating data to a few numerical transformations and apply the statistical model to each, simply ignoring the fact that the usual statistical assumptions—normality, equal variances, etc.—cannot possibly be correct in all of them and hoping that it does not much matter.

Ideally, one would like to go back to basics and attempt to construct a measurement theory in which the randomness somehow is dealt with at the level of the primitives. In such a theory, a representation theorem would be rather different from the current ones. One would not construct a numerical representation on the underlying set A of objects, but rather a mapping into some family of random variables, $\{\mathbf{X}(a): a \in A\}$. One would expect the behavioral theory to characterize the family and to have the property that some statistic of central tendency, such as the mean or median, behaves algebraically like the corresponding noise-free classical theory. For example, the kind of representation I would expect for a noisy generalization of extensive measurement might be the family of gamma distributed random variables with the underlying concatenation operation, $\circ$ being represented as the sum of independent random variables and their means acting like ordinary extensive measures, i.e.,

$$E[\mathbf{X}(a \circ b)] = E[\mathbf{X}(a) + \mathbf{X}(b)] = E[\mathbf{X}(a)] + E[\mathbf{X}(b)]. \tag{3}$$

The simple fact is that we do not know how to do this. All treatments of randomness with which I am familiar presuppose a numerical representation; there simply is no *qualitative* theory of the concept. This is a limitation not just of mathematical psychology, but of probability theory in science in general. Perhaps we psychologists are more acutely aware of it than, say, classical physicists because of the great difficulty we have in making independent, repeated observations of exactly the same comparison.

A more modest proposal is to treat the data as probabilistic in nature—either as a probability distribution over algebraic models or as the primitive itself. One question raised is that of conditions under which there exists an underlying random variable representation such as was postulated early on in the discriminable dispersion models of Thurstone (1927a, b, c) and in the random utility models of Block and Marschak (1960) and Marschak (1960). For later results of this sort, see Marley (1990).

This work has recently been extended considerably by Doignon and Regenwetter (1997), Heyer and Niederée (1989, 1992), Regenwetter (1996), and Regenwetter, Marley, and Joe (1997)—but only for fundamentally ordinal structures. We still cannot really deal probabilistically with the kind of structural questions that interest many substantive scientists. For example, consider subjects choosing among gambles (or uncertain alternatives). One property with which I have been particularly concerned, arises if we include among the primitives a binary operation $\oplus$ of *joint receipt*. Let $(x, p; y)$ denote a lottery in which x is the outcome with probability p and y with probability $1-p$ and let e denote the status quo (often taken to be no exchange of money). Then the property called *segregation* asserts that for either gains, i.e., $x, y \succ e$, or for losses, i.e., $x, y \prec e$,

$$(x, p; e) \oplus y \sim (x \oplus y, p; e \oplus y) \sim (x \oplus y, p; y). \tag{4}$$

Note that from a rational perspective, this condition is fairly compelling—the two sides are completely equivalent. Luce and Fishburn (1991, 1995) and Luce (1996) have made considerable theoretical use of Eq. (4).

Using somewhat crude statistical methods, such as whether or not the choices exhibit a symmetric pattern in the differences of independently established certainty equivalents, Cho and Luce (1995) and Cho, Luce, and von Winterfeldt (1994) have concluded from their data that the evidence, although noisy, favors segregation. It is, as yet, unclear how a property like this should be studied in the context of random variable models, and until it is, I don't see much hope of carrying out really acceptable statistical analyses of substantively interesting properties and theories. At the same time, I don't think it reasonable to expect theorists to abandon their modeling and experimentalists their evaluating just because we don't know how best to do the statistics. One should never forget that the very successful physical sciences muddled along quite nicely with very primitive statistics, in part, of course, because of really large sample sizes which we rarely have.

PROLIFERATION OF PARAMETERS VERSUS THEORIES OF PARAMETERS

Most models one encounters in psychology have a fair number of unspecified parameters that are to be estimated from the data to be explained. This is true not only of the dynamic models discussed earlier, but equally for perfectly static ones such as signal detection theory (SDT) and most response time, memory, and attention models. Moreover, many of these theories are reasonably tractable only when there are very few stimuli and responses. As those numbers increase, the number of parameters often increases exceedingly rapidly. For example, in areas where SDT is used, the 2-stimulus, 2-response model, with its ubiquitous ROC curves and the well-known d' and response bias measures, is quite tractable. But go to just 3 stimuli and 3 responses and what in the $(2, 2)$ case had been a cut point on a continuum becomes an unspecified partition of 2-space into 3 (presumably connected) regions—this to account for only 6 independent conditional probabilities. Unless the class of partitions is sharply limited, the problem is totally underdetermined by the data. But on what principles does one limit the partitions? I do not sense that we have a very good understanding of how to do this throughout the world of information processing models, although considerable effort is currently underway as seen in, for example, Ashby's (1992) collection of papers.

One way to restate the problem is that we usually do not have a theory of the parameters or, put still another way, we do not understand very clearly what the parameters mean. If we did, we could then estimate them in one experiment and use those values in the model for a different experiment. But often that does not work even in the simplest of situations. For example the d' values estimated in a 2-stimulus, absolute-identification experiment do not directly predict the effective d' values in an n-stimulus, $n > 2$, experiment of the same type (Miller, 1956; Luce, Green, & Weber, 1976).

Do we have any examples of theories of parameters? To my knowledge, examples lie not in the world of information processing models, but elsewhere. In the theories of individual decision making in which the concept of utility plays a role, we find models of the following character. As mentioned earlier, the primitive data often are choices made between risky or uncertain alternatives, and the models describe constraints among these choices. These constraints do not include any parameters and they are testable in finite data samples. Such testing must be done with some care to avoid testing the conjunction of the property in question with other properties that are as empirically questionable as the one purportedly under study. This goal has not always been achieved and has probably led to some erroneous conclusions (Luce, 1992).

Well-known examples of such purely behavioral properties are transitivity of choices and monotonicity of choices with changes of a single consequence of a gamble. Less-well-known ones include, among several others that I will not go into here, the behavioral hypothesis of *segregation* mentioned above, Eq. (4). The mathematics involved in these models lies in discovering some convenient—usually numerical—representation of these data constraints. Such models involve the construction of numerical functions—a utility function over both certain and uncertain alternatives and one or more weighting functions over chance events. The earliest such model—due to von Neumann and Morgenstern (1947)—is the best known, namely, that the behavior is *as if* the person has a utility function U and is maximizing its expectation:

$$\sum_i U(x_i)\, p_i. \tag{5}$$

Here a single function is constructed from the behavior.

Gradually over the years more complex models have evolved as we have come to understand experimentally which constraints do and do not seem to hold. A major landmark was Savage's (1954) axiomatic generalization of Eq. (5) to subjective expected utility (SEU). Fundamentally, he showed that one could begin with events E_i to which no probability is attached, and if the individual's choice behavior is sufficiently systematic, it will be *as if* that person is maximizing

$$\mathrm{SEU}(g) = \sum_i U(x_i)\, S(E_i), \tag{6}$$

where $g = (x_1, E_1; ...; x_n, E_n)$ is a gamble (or uncertain alternative), U is a numerical utility function over consequences, and S is a finitely additive probability measure (unique to the decision maker) over the underlying chance events. Despite the widespread feeling that Savage's axioms are normatively compelling, much as those of elementary logic are, SEU is known to be descriptively wrong as, of course, are some axioms of logic. Ellsberg (1961) demonstrated, in a very simple situation, that people do not act as if they assign a fixed subjective probability to a truly uncertain event.

Numerous later models have attempted to understand what is happening by working with nonprobabilistic weights on events. This is not the place to detail these models except to note that in these cases we do have testable theories of these unknown functions that from another perspective seem like parameters to be estimated. Indeed, if additional (nonparametric) behavioral conditions are satisfied, one can get quite specific forms e.g., Luce and Fishburn (1991, 1995) and Miyamoto (1988) have arrived at

$$U(x) = \begin{cases} C(1 - e^{-cx}), & x \geqslant 0 \\ -K(1 - e^{kx}), & x < 0 \end{cases}, \tag{7}$$

where all constants are positive, as the only possible form for the utility function when it is concave for gains, convex for losses, and segregation and binary prospect theory (Kahneman & Tversky, 1979) hold. Others have come up with alternative forms, but in most cases behavioral conditions underlie them. It is merely a question of trying to decide empirically which, if any, of these conditions is correct.

It is perhaps worth noting that much of classical physics involved the formulation of observable "behavioral" relations (laws) and developing differential equations to summarize these relations. Examples are Newton's laws, hydrodynamic theory, Maxwell's equations for electromagnetism, classical thermodynamics, and both special and general relativity theory. Note that this type of modeling is not at all analogous to information-processing modeling. The closest physical analogies to such models of unobservable structure are the kinetic theory of gases, which came long after there was a well-developed behavioral theory of thermodynamics, and the modern theories of atomic structure.

Although it is all too easy to write down information-processing models for psychological behavior, it is an open question in my mind whether this is really the best way to proceed. Latent, not directly observable structures—the hallmark of such models—afford too many options, leading to badly underdetermined models with many free parameters. But clearly, the mainstream of psychology is passing me by.

UNCERTAINTY AND VAGUENESS

Almost all of the mathematics most of us know is grounded in set theory. So are computers. Almost none of our conversation is so grounded. We communicate, as I am trying to in this article, mostly in terms that have a certain penumbra of uncertainty or vagueness. If someone says "I'll return at dusk," you do not expect a precise set-theoretic definition of what that person means, certainly not a sharp luminosity level. This has been a well-recognized contrast between mathematics and discourse for as long as I've been thinking about science and appreciably earlier, judging by authors such as Poincaré (1929), who commented that "the physical continuum is, so to speak, a nebula not resolved; the most perfect instruments could not attain its resolution..." (p. 46).

To some degree probability theory attempts to deal with some aspects of uncertainty, but clearly that model is not really sufficient. As many, but especially Glenn Shafer (1976), have been at pains to point out, one may be so uncertain that one sometimes assigns a low likelihood both to an event and to its complement. If you do not know where Izmir is located, you probably are not very willing to assign as much as an even chance to rain there tomorrow and, if asked separately, you are also unwilling to assign as much as an even chance to no rain there tomorrow. Your uncertainly is so great that, independently assessed, your estimates won't add to 1. In other examples, increasingly fine partitions of the events lead to estimates that actually sum to far more than 1 (Fox & Tversky, 1997; Redelmeier, Koehler, Liberman, & Tversky, 1995; Tversky & Koehler, 1994). To the degree this occurs, which seems to be considerable, models such as SEU (Eq. (6)) must be wrong.

The most extensive attempts to deal directly with the thorny problem of vagueness have been multi-valued logics and Zadeh's (1975) theory of fuzzy sets (for a fairly recent summary of fuzzy set theory see Klir and Folger, 1988). The former has penetrated very little and the latter hardly at all into modern mathematics; however, fuzzy sets and relations have developed into a minor cottage industry in engineering and computer science, where they are being used in various control devices. Repeated attempts have been made to use them theoretically in psychology, but to the best of my knowledge, none has been viewed as particularly illuminating. Part of the problem, so far as I am concerned, is that the theory of fuzzy sets is, after all, classically set theoretic with an underlying membership function that is a simple numerical map that to all intents is like a distribution function. The novel aspect of the theory is the attempt to define logic-like operations of "and," "or," and "implies" in terms of creating new membership functions from old ones. So, for example, the membership function for "dusk" is almost certainly monotonic with luminosity and that for "humidity" is, holding temperature constant, monotonic with the proportion of water vapor, and so the fuzzy concept of a "humid dusk" becomes a calculation in terms of these two measures.

Perhaps; but I remain skeptical. I sense that this approach is missing the main point, but I have nothing better to offer. I had hoped to see a generalized set theory develop which, like ordinary set theory at the end of the 19th century, would prove so compelling that it would command the assent of most mathematicians as being a generalization with a wide potential for new mathematics. Among other things, it might alter the nature of what is seen as computable. That has not yet happened.

DISCRETENESS OR CONTINUITY

Under this heading I shall treat two quite different issues, one very general in the sciences that use mathematics in their formulations and the other very specific to psychology.

The general one is the philosophy-of-science problem of why continuous mathematics works at all in the sciences. If we take seriously what physicists tell us, many attributes must be discrete, although very finely so compared to the sort of sensory discreteness I will discuss below, but nonetheless continuous mathematics (in the proper hands) seems to achieve correct answers. Physicists and philosophers of physics have long been perplexed by this fact; witness the famous article of Eugene Wigner (1960) titled *The unreasonable effectiveness of mathematics in the natural sciences* (see also Narens and Luce, 1990). To my knowledge there is still no fully satisfactory answer, although the recent work of Suppes (1995; Sommers & Suppes, 1996a, b; Suppes & Chuaqui, 1995) may be a significant step forward.

Let me illustrate part of the problem in an area familiar to many of us, the representational theory of measurement. Stevens (1946, 1951) noted the somewhat surprising fact that the then-existing examples of measurement (mostly from the physical sciences) seemed to be of three types that he called ordinal, interval, and ratio. For a long time the underlying source of his apparently limited classification was not understood, but Narens (1981a,b) tackled the problem at a very general level and Alper (1985, 1987) obtained a final solution (see below) to the following effect: Suppose the domain of measurement is a continuum, the structure is homogeneous in the intuitive sense that each point is structurally indistinguishable from each other one, and such structure-preserving mappings (automorphisms) can have only a limited number of fixed points. Then there is a representation into the real numbers such that the scale type is either interval, ratio, or a subgroup between these two. Moreover, the representation can be chosen so that the translations are just that,

$$x \to x + s, \qquad s \text{ any real number,}$$

and the remaining ones, called *dilations*, are of the form

$$x \to rx + s, \qquad s \text{ any real and}$$
$$r > 0 \text{ in a multiplicative subgroup.}$$

Now, if we think of such a structure as an asymptotic idealization of the discrete world, then we are faced with trying to figure out what about the discrete cases ultimately evolves into the homogeneity and finite uniqueness of the "limiting" continuous model. This step does not seem to be understood. Of course, part of what has to be done is to make sure that a nested sequence of finite structures (which can be thought of as increasingly better finite approximations to a countable model) have suitably convergent families of numerical representations. This part seems to be less of a problem than the jump from the countable case, with very few automorphisms (symmetries), to the continuum, with its richness of automorphisms. This transition has never been formulated explicitly.

An issue of discreteness very specific to psychology is whether the internal representation of a physical stimulus drawn from a "continuum," such as intensity or frequency, is sufficiently discrete that we should not ignore that feature or whether it is so fine that a continuous representation is adequate. I find myself in a conflicted position on this issue. On the one hand, for a very long time I've been suspicious that some sensory attributes are really rather discrete but that noise tends to mask that discreteness. On the other hand, I personally prefer to work with continuous models. The discrete models, with their horrendous combinatorial aspects and typical lack of neat algebraic features, tend to leave me cold—although some of the network results of the past 10 years (White & Duquenne, 1996; Ganter & Wille, 1997) are nice and, with computers, quite feasible to use. But the real question is not personal taste, but empirical facts.

I believe that the current sweeping victory of the ideas of the continuous signal detectability theory (SDT) may provide a sobering lesson in the sociology of science. It appears to me that most psychophysicists have elected not to explore very deeply questions of evidence against continuity or near continuity. There isn't a lot of negative evidence, but what there is has been largely ignored after an initial flurry of activity. Let me mention four examples.

Early on, Stevens, Morgan, and Volkman (1941) suggested that sensory intensity might be discretely represented subjectively and that for such a simple task as detection it was sufficient to distinguish only two internal states: detect and not detect. Were that true, they argued, the psychometric function should move linearly between 0 and 1 with the intercepts standing at a 2:1 ratio. About two decades later, I (Luce, 1963) pointed out that, correspondingly, the ROC curve should have two linear limbs—which came to be called the low-threshold model. After some fairly casual examination of data suggesting some support for the 2-state model, a careful statistical analysis by Krantz (1969) demonstrated that to be wrong, but he did not reject a 3-state model. The field seemed to react to the rejection of the 2-state model as support for the continuum. Little else has been explored for these experiments. Part of the reason is that if one does not know much about the nature of the states, the estimation of their parameters is a bit of a nightmare. Still the jump one, two, infinity seems a bit precipitous.

This is especially the case when one realizes that all discrete state models predict that the ROC curve for the 2-alternative, forced-choice experiment should exhibit a flat portion with slope 1 in the region where the ROC curve crosses the negative diagonal. The continuous models do not predict such flatness. Rather, the continuous SDT model predicts that in z-score coordinates the curve is straight with slope 1 whereas the discrete models are bowed. As I have repeatedly pointed out and as has been equally studiously ignored by the field, Norman (1964) in a carefully run experiment reported 2-alternative, forced-choice ROC curves on three subjects that agree with discreteness and not with continuity. To my knowledge, no one has refuted these observations. Indeed, no one seems ever to collect 2-alternative, forced-choice ROC curves, there being a myth to the effect that this procedure, unlike the yes–no one, is unbiased. I find this an odd response of what purports to be a self-correcting science. Just why don't Norman's data reject SDT?

Two other signs of discreteness lie in the temporal area. The temporal sluggishness of the visual system led Stroud (1955) and others to wonder whether information is clustered in temporal packets of about 100 ms within which temporal order is lost or ignored. The fact that movies work quite successfully at 16 frames a second is consistent with that hypothesis, although it certainly does not entail it. Of course such crude temporal discreteness is not a general sensory phenomenon. For example, the auditory system is highly sensitive to some small—less than one millisecond—temporal differences. The major difficulty in ever testing this hypothesis was the issue of synchronization of the discreteness with the signals. How does one make sure that two temporally close signals are or are not in the same temporal quantum? This line was not further pursued.

Finally, there are the remarkable data of Kristofferson (1980, 1984) on time estimates over a wide range of times. He showed that a plot of standard deviations versus mean response is not the proportionality of Weber's law, as many of us expected, but rather more like a step function with the steps appearing at factors of 2 of the mean and with jumps in standard deviation by factors of $\sqrt{2}$. These data strongly suggest highly discrete changes in some aspect of, in this case, the combined sensory and motor systems. Unfortunately, these experiments are exceedingly difficult and tedious to conduct, and other scientists seem not to have accepted the challenge.

The simple fact seems to be that sensory psychologists just are not greatly interested in the question of whether sensation is discrete or continuous.

CONCLUDING REMARKS

Because this paper describes several, but far from all, of the unresolved issues of mathematical psychology in the second half of the 20th century, its tone has necessarily been, depending on your perspective, discouraging or challenging. To balance that a bit, let me make clear that I think the field has exhibited a good deal of cumulative progress. We know vastly more about a number of topics than we did in 1950. These include random variable models of simple choices and their close relation to geometry; the relation of dimensional analysis to theories of measurement; the entire structure of nonadditive measurement including both the very general theorems cast in terms of automorphisms as well as the general theory of binary operations and conjoint structures; psychophysical modeling of all sorts including mechanisms of spatial inference, motion detection, and rather well-developed theories of response times; theories of decision making including utility and subjective probability; and connectionist and network modeling including very sophisticated combinatorial analysis. Knowledgeable readers can add others, which I count on their continuing to do in the coming years.

REFERENCES

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

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.

Ashby, F. G. (Ed.), (1992). *Multidimensional models of perception and cognition*. Hillsdale, NJ: Erlbaum.

Block, H. D., & Marschak, J. (1960). Random orderings and stochastic theories of responses. In I. Olkin, S. Ghurye, W. Hoeffding, W. Madow, & H. Mann (Eds.), *Contributions to probability and statistics*, pp. 97–132. Stanford, CA: Stanford University Press.

Bush, R. R., & Mosteller, F. (1955). *Stochastic models for learning*. New York: Wiley.

Cho, Y., & Luce, R. D. (1995). Tests of hypotheses about certainty equivalents and joint receipt of gambles. *Organizational Behavior and Human Decision Processes*, **64**, 229–248.

Cho. Y., Luce, R. D., & von Winterfeldt (1994). Tests of assumptions about the joint receipt of gambles in rank- and sign-dependent utility theory. *Journal of Experimental Psychology*: *Human Perception and Performance*, **20**, 931–943.

Doignon, J.-P., & Falmagne, J. C. (1997). Well graded families of relations. *Discrete Mathematics*, in press.

Doignon, J.-P., & Regenwetter, M. (1997). The approval voting polytope for linear orders. *Journal of Mathematical Psychology*, in press.

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

Falmagne, J.-C. (1996). A stochastic theory for the emergence and the evolution of preference structures. *Mathematical Social Sciences*, **31**, 63–84.

Falmagne, J.-C., & Doignon, J.-P. (1997). Stochastic evolution of rationality. *Theory and Decision*, in press.

Falmagne, J.-C., Regenwetter, M., & Grofman, B. (1997). A stochastic model for the evolution of preferences. In A. A. J. Marley (Ed.), *Choice, decision, and measurement*: *Essays in honor of R. Duncan Luce*. Mahwah, NJ: Erlbaum.

Fox, C. R., & Tversky, A. (1997). A belief-based account of decision under uncertainty. *Management Sciences*, submitted.

Ganter, B., & Wille, R. (1997). *Formale Begriffsanalyse*: *Mathematische Grundlagen*. Berlin/Heidelberg: Springer-Verlag.

Gregson, R. A. M. (1988). *Nonlinear psychophysical dynamics*. Hillsdale, NJ: Erlbaum.

Heyer, D., & Niederée, R. (1989). Elements of a model-theoretic framework for probabilistic measurement. In E. E. Roskam (Ed.), *Mathematical psychology in progress*, pp. 99–112. Berlin: Springer.

Heyer, D., & Niederée, R. (1992). Generalizing the concept of binary choice systems induced by rankings: One way of probabilizing deterministic measurement structures. *Mathematical Social Sciences*, **23**, 31–44.

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

Klir, G. T., & Folger, T. A. (1988). *Fuzzy sets*, *uncertainty*, *and information*. Englewood Cliffs, NJ: Prentice Hall.

Krantz, D. H. (1969). Threshold theories of signal detection. *Psychological Review*, **76**, 308–324.

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

Kristofferson, A. B. (1980). A quantal step function in duration discrimination. *Perception & Psychophysics*, **27**, 300–306.

Kristofferson, A. B. (1984). Quantal and deterministic timing in human duration discrimination. In J. Gibbon & L. Allen (Eds.), *Timing and time perception*, pp. 3–15. Annals of the New York Academy of Sciences (Vol. 423).

Lakshminarayan, K. (1996). Stochastic learning paths in knowledge structures with exponential learning times. Institute for Mathematical Behavioral Sciences, TR-96-07. University of California at Irvine.

Laming, D. (1973). *Mathematical psychology*. London/New York: Academic Press.

Luce, R. D. (1959). *Individual choice behavior*. New York: Wiley.

Luce, R. D. (1964). Some one-parameter families of commutative learning operators. In R. C. Atkinson (Ed.), *Studies in mathematical psychology*, pp. 380–398. Stanford, CA: Stanford University Press.

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

Luce, R. D. (1996). When four distinct ways to measure utility are the same. *Journal of Mathematical Psychology*, **40**, 297–317.

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., & Fishburn, P. C. (1995). A note on deriving rank-dependent utility using additive joint receipts. *Journal of Risk and Uncertainty*, **11**, 5–16.

Luce, R. D., Green, D. M., & Weber, D. L. (1976). Attention bands in absolute identification. *Perception & Psychophysics*, **20**, 49–54.

Luce, R. D., Krantz, D. H., Suppes, P. & Tversky, A. (1990). *Foundations of measurement* (Vol. 3). San Diego: Academic Press.

Marschak, J. (1960). Binary-choice constraints and random utility indicators. In K. J. Arrow, S. Karlin, & P. Suppes (Eds.), *Mathematical methods in the social sciences*, 1959, pp. 312–329. Stanford, CA: Stanford University Press.

Miller, G. A. (1956). The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, **63**, 81–97.

Miyamoto, J. M. (1988). Generic utility theory: Measurement foundations and applications in multiattribute utility theory. *Journal of Mathematical Psychology*, **32**, 357–404.

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., & Luce, R. D. (1990). Three aspects of the effectiveness of mathematics in science. In R. Mickens (Ed.), *Mathematics and science*, pp. 122–135. Singapore: World Scientific.

Norman, D. A. (1964). Sensory thresholds, response biases, and the neural quantum theory. *Journal of Mathematical Psychology*, **1**, 88–120.

Poincaré, H. (1929). *The foundations of science* (G. B. Halsted, Trans.). New York: The Science Press.

Redelmeier, D. A., Koehler, D. J., Liberman, V., & Tversky, A. (1995). Probability judgment in medicine: Discounting unspecified possibilities. *Medical Decision Making*, **15**, 227–230.

Regenwetter, M. (1996). Random utility representation of finite *n*-ary relations. *Journal of Mathematical Psychology*, **40**, 219–234.

Regenwetter, M., Marley, A. A. J., & Joe, H. (1997). Random utility threshold models of subset choice. *Journal of Mathematical Psychology*, submitted for publication.

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

Shafer, G. (1976). *A mathematical theory of evidence*. Princeton, NJ: Princeton Univ. Press.

Sommer, R., & Suppes, P. (in press). Finite models of elementary recursive nonstandard analysis. *Proceedings of the Chilean Academy of Sciences*.

Sommer, R., & Suppes, P. (in press). Dispensing with the continuum. *Journal of Mathematical Psychology*.

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., Morgan, C. T., & Volkman, J. (1941). Theory of the neural quantum in the discrimination of loudness and pitch. *American Journal of Psychology*, **54**, 315–335.

Stroud, J. M. (1955). The fine structure of psychological time. In H. Quastler (Ed.), *Information theory in psychology*, pp. 174–207. Glencoe, IL: The Free Press.

Suppes, P. (1995). Free-variable axiomatic foundations of infinitesimal analysis: A fragment with finitary consistency proof. *Journal of Symbolic Logic*, **60**, 122–159.

Suppes, P., & Chuaqui, R. (1995). A finitarily consistent free-variable positive fragment of infinitesimal analysis. In *Proceedings of the IX Latin American Symposium on Mathematical Logic* (1991). Notas de Mathematics, Universidad Nacional del Sur, Bahia Blanca, Argentina.

Suppes, P., Krantz, D. H., Luce, R. D., & Tversky, A. (1989). *Foundations of measurement* (Vol. 2). San Diego: Academic Press.

Thurstone. L. L. (1927a). A law of comparative judgment. *Psychological Review*, **34**, 273–286.

Thurstone, L. L. (1927b). Psychophysical analysis. *American Journal of Psychology*, **38**, 68–89.

Thurstone, L. L. (1927c). Three psychophysical laws. *Psychological Review*, **34**, 424–432.

Tversky, A., & Koehler, D. J. (1994). Support theory: A nonextensional representation of subjective probability. *Psychological Review*, **101**, 547–567.

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

White, D. R., & Duquenne, V. (Eds.) (1996). Kinship networks and discrete structure analysis. *Social Networks*, **18**, 267–314.

Wigner, E. P. (1960). The unreasonable effectiveness of mathematics in the natural sciences. *Communications in Pure and Applied Mathematics*, **13**, 1–14.

Zadeh, L. A. (1975). The concept of a linguistic variable and its application to approximate reasoning. I. *Information Sciences*, **8**, 199–249.

Received: November 6, 1996