PSYCHOMETRIKA—VOL. 70, NO. 2, 227–251
JUNE 2005
DOI: 10.1007/s11336-004-1248-8

MEASUREMENT ANALOGIES: COMPARISONS OF BEHAVIORAL AND PHYSICAL MEASURES*

R. DUNCAN LUCE

UNIVERSITY OF CALIFORNIA

Two examples of behavioral measurement are explored — utility theory and a global psychophysical theory of intensity — that closely parallel the foundations of classical physical measurement in several ways. First, the qualitative attribute in question can be manipulated in two independent ways. Second, each method of manipulation is axiomatized and each leads to a measure of the attribute that, because they are order preserving, must be strictly monotonically related. Third, a law-like constraint, somewhat akin to the distribution property underlying, e.g., mass measurement, links the two types of manipulation. Fourth, given the numerical measures that result from each manipulation, the linking law between them can be recast as a functional equation that establishes the connection between the two measures of the same attribute. Fifth, a major difference from most physical measurement is that the resulting measures are themselves mathematical functions of underlying physical variables — of money and probability in the utility case and of physical intensity and numerical proportions in the psychophysical case. Axiomatizing these functions, although still problematic, appears to lead to interesting results and to limit the degrees of freedom in the representations.

For a good many years I have fretted about aspects of psychological measurement — my first papers on the topic introduced the concept of semiorders (Luce 1956) and gave a mathematical critique of Fechner's construction of a scale of subjectively equal jnds (Luce & Edwards, 1958; see also Krantz, 1971; and Iverson, Myung, & Karabatsos, 2005). My purpose here is to share with you some of my ruminations on behavioral measurement as well as to summarize some newer results. In particular, I will consider a number of parallels with and differences from the philosophical foundations of physical measurement. Note that parallel does not mean formal identity.

A feature of this work, which may not be received with favor by psychometricians, is an emphasis on algebraic structures underlying deterministic numerical representations almost to the exclusion of probabilistic or statistical features. Although in principle both structure and randomness are essential parts of the same package, so far the attempts to deal with them in a unified manner at the foundational level have not been very successful. True, the past decade has seen some significant progress that is encouraging, yet the most successful research seems, so far, to have focussed on one of them alone. The most recent developments, which have been spearheaded by Karabatsos (2005), entail Bayesian inferences applied directly to algebraic axioms. Psychometricians are highly familiar with what can be done with statistical modeling, and in part for that reason and in part because I find structural questions rather more to my taste, structure is what I address here to the great neglect of randomness.

*This article is derived from two of my talks: a keynote address presented at the annual meeting of the Psychometric Society in Pacific Grove, CA, June 14–17, 2004, and the other at the sessions in honor of Jean-Claude Falmagne's 70th birth year at the meeting in Gent, Belgium, September 2–5, 2004, of the European Mathematical Psychology Group that he founded some 31 years ago. The work described here was supported, in part, by National Science Foundation grant SES-0136431 to the University of California, Irvine, and, in part, by the University. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors(s) and do not necessarily reflect the views of the National Science Foundation. I appreciate the detailed and useful comments of my collaborators A. J. Marley and Ragnar Steingrimsson on earlier versions of the manuscript, as well as those of three referees, one of whom, Ehtibar Dzhafarov, provided very useful comments and criticisms.

Request for reprints should be sent to R. Duncan Luce, Social Science Plaza, University of California, Irvine, CA 92697-5100, USA. E-mail: rdluce@uci.edu

Because it is well understood, I first discuss some of the foundations of classical physical measurement which set the framework for foundational measurement studies in the social and behavioral sciences. I would be remiss not to mention that, very recently, attempts are being made to use quantum formalism in psychology. There was such a talk by Jerome Busemeyer about applications to psychophysics at the 2004 meetings of the Society for Mathematical Psychology and one by V. I. Danilov and A. Lambert Mogiliansky about applications to decision making at the 2004 European Mathematical Psychology Group.

1. Foundations of Physical Measurement

All measurement starts with qualitative attributes, with properties of entities that can be varied either naturally or by agents. Think of the velocity[1] of an object or a fluid under pressure, the temperature of a body, etc. Much of the foundations of classical physical measurement focuses on the trade-offs among such attributes, usually summarized as numerical variables.

Consider blobs of materials being compared using an idealized[2] equal-arm pan balance. This provides a way to study the qualitative nature of the property called mass. The resulting mass ordering, denoted $\succsim$, is defined as follows: if a and b denote two distinct blobs of material, then $a \succsim b$ means that when we place a and b on the pans of the balance, then either the balance stays level or the a–pan drops. Further, let $a \circ b$ denote the blobs a and b both being together on one pan[3]. Then one of the elementary features of physical measurement is that $\circ$ has certain testable properties — monotonicity, commutativity, and associativity — that under reasonable structural assumptions imply the existence of a real valued, non-negative measure m that preserves the order of greater mass, $\succsim$, and is additive over $\circ$, i.e.

$$a \succsim b \quad \text{iff} \qquad m(a) \geq m(b). \tag{1}$$

$$m(a \circ b) \quad = \quad m(a) + m(b). \tag{2}$$

Moreover, the representation is unique up to similarity transformations, i.e., the representations form a ratio scale. Of course, once the existence of such a measure is established, then one can use its role in certain physical laws to render measurement more convenient, e.g., a spring balance.

In addition to concatenation, one has a so-called conjoint structure relating, e.g., mass to the volume and an aspect of the particular material involved called its density. These structures are shown in Table 1 along with the usual algebraic formulas involved.

Separate axiomatizations of each of these measurement procedures $\succsim$ over $\mathcal{V} \times \mathcal{S}$ and $\succsim$ and $\circ$ over $\mathcal{M} = \mathcal{V} \times \mathcal{S}$ yield different measures of mass. Although the practice of pan-balance weight measurement evolved over many centuries in markets, the foundational theory justifying additive measures over $\circ$ was only accomplished a little over a century ago by Hölder (1901) elaborating earlier work of Helmholtz (1887). For even longer, a measure such as density was not viewed as axiomatized but as " derived" measurement, namely, $\rho = m/V$. It was formally axiomatized by Debreu (1960) in a topological setting and, independently, by Luce and Tukey (1964) in a more general algebraic one. Each axiomatization yields its own measure of mass and, so far, these are linked only by the fact that they are both order preserving. To prove that the same measure can be used for both representations, one must add some sort of linking law — a form

[1]Velocity is a vector and its length is speed. Because I confine myself to velocities in the same direction, either term can be used.

[2]Frictionless, rigid, and in a vacuum. Marc Le Menestrel at the EMPG meetings in 2004 treated the case with friction as a semiorder.

[3]A bit of caution is needed in choosing the substances involved so as to avoid explosions or fires. I mention this because similar cautions hold in some behavioral examples.

TABLE 1.
Two independent methods of manipulating the attribute of mass. The notation is m, V, and ρ for, respectively, numerical measures on ratio scales of mass, volume, and density

Set of containers	$\mathcal{V}$
Collection homogenous substances	$\mathcal{S}$
Collection of masses	$\mathcal{M} = \mathcal{V} \times \mathcal{S}$
Conjoint structure	$\langle \mathcal{V} \times \mathcal{S}, \succsim \rangle$
Representation	$m(a) = V(a)\rho(s)$
Extensive structure	$\langle \mathcal{M}, \succsim, \circ \rangle$
Representation	$m'(a \circ b) = m'(a) + m'(b)$
Linked structures	Distributivity
Representation	$m' = m$

of distribution somewhat paralleling arithmetic distribution, $(x + y)z = xz + yz$. Links of this kind are illustrated below in the behavioral examples.

Of course, there are interesting interplays between a property such as mass and many other variables such as velocity and momentum or kinetic energy. These lead to the basic physical laws that underlie the classical scheme of physical units. Most physicists know the laws but usually are not much interested in their conceptual bases.

A number of variables such as length, velocity (in classical physics), and time are formally similar to mass measurement in that they have a concatenation operation $\circ$ and a qualitative ordering $\succsim$ and that satisfy the same fundamental qualitative laws and so have an order preserving, additive representation. Voltage, amperage, and resistance are still another example.

An example that is a bit more complex than the mass one is *relativistic velocity* (in the absence of acceleration) where the conjoint and concatenation expressions are, respectively,

$$v = d/t, \tag{3}$$

$$u \circ v = \frac{u + v}{1 + \frac{uv}{c^2}}, \tag{4}$$

where d is distance traveled in time t and $u \circ v$ is velocity concatenation of frames of reference with uniform velocities u and v in the same direction as measured by an observer using the ratio of measured distance to measured time elapsed. The constant c is the speed of light measured in same units.

Note that this says at least four things: (1) The same velocity measure v is used in both the concatenation and conjoint cases. This has to be justified by some form of linkage. (2) Velocity concatenation is non-additive over $\circ$; however, there is a non-linear transformation of (4) that is additive and is called " rapidity". We will encounter something somewhat analogous in some behavioral measurement. (3) The velocity of light, c, is a (universal) constant with units of velocity. (4) The velocity and rapidity measures, despite being non-linearly related, are each ratio scales. The scale type for velocity stems from properties of conjoint measurement coupled with the law linking that structure to the concatenation one, whereas the scale type of rapidity stems entirely from properties of the concatenation by itself.

A somewhat similar example is a fluid for which variables such as pressure, density, and temperature play a role, and certain constants exist such as viscosity and the Reynolds' number that are part of the formulation of the Navier-Stokes equations governing the temporal unfolding — the dynamic properties — of a fluid. Here differential equations are essential.

Other physical laws involve combinations of variables that, in classical physics, satisfy invariance principles embodied in dimensionally invariant laws. I do not go into this interesting, but complex, topic here (see Krantz, Luce, Suppes, & Tversky, 1971, chapter 10; Luce, Krantz, Suppes, & Tversky, 1990, chapter 22; Narens, 2002).

A number of dimensional numbers, known as physical constants, have arisen in these physical formulations, and they come in at least two varieties. The first type are the constants associated with specific objects. For homogenous solids, one thinks immediately of the rest mass and the elastic (Hooke's) constant of a solid object. At the same time, a homogenous solid has a density, which is a property of the particular substance composing the solid. For fluids, density and viscosity are typical dimensional constants unique to the fluid in question and play a significant role in aerodynamics, such as in the equations for the lift and drag of an airfoil.

A second class of constants, of which the speed of light is an example, are called universal constants. They are " universal" in the sense of not being different in different contexts. In addition to light, one has Newton's gravitational constant, Planck's constant, etc. These are exceedingly important, and many philosophers of physics have speculated a good deal about them, in particular as to why their numerical values are what they are.

And there are problematic classes of constant-like attributes of which hardness is a prototypical example. Its operational definition for solids is which of two materials scratches the other one. How does this constant relate to other physical measures? Certain ad hoc formulas exist, but to my knowledge no really systematic theory of hardness has ever been developed despite its ubiquitousness and importance.

2. Are there Behavioral Analogues?

Of course, the behavioral sciences do not yet have anything as nearly as well developed as the scheme of physical measures, but we do have some forms of measurement. The question to be considered here is: In what respect do some of these behavioral examples resemble the foundational arguments for physical variables,constants, and trade-offs? A closely related question, to which I give less attention, is exactly how physical and behavioral measurement differ. I think that I am safe in saying that we do not, at least so far, have any clear analogues of universal constants. But I also think that we have putative variables and constants of behavioral entities, and I discuss two of them.

Consider people as entities — albeit complex ones who certainly cannot be thought of as composed of anything homogenous. Of course, there are natural physical measures and, at least, momentary body constants associated with individuals: weight, height, and skin, eye, and hair color, etc. But, as psychologists, we are not greatly interested in these measures (except, perhaps, personally, as when on a diet), but rather in behavioral attributes and constants. Perhaps the most immediate are measures of abilities including some concept of general or specific abilities and traits. And one likes to think that these are related to other variables such as which questions are answered correctly in a standardized test, the speed with which they are answered, etc. One underlying intuition is that a measure of ability might be a constant, at least over some extended time period, associated with the individual and that, in some sense, links question difficulty to the likelihood of a correct response. And, as psychometricians are all well aware, a number of well developed statistical theories, such as item response theory, have been proposed as governing ability comparisons among people. Our society currently is very dependent upon such measures, and they do underlie what is arguably the most successful technology to arise from psychological research, although I suspect that that distinction may soon be challenged by computer-based teaching/testing programs, e.g., those arising from the development of knowledge space theory (Doignon & Falmagne, 1999). But to my knowledge there is not yet any real structural analysis of ability, of how it relates in other than a correlational sense to other behavioral measures, such as to simple reaction times. It seems to me that the measurement of abilities and intelligence is currently far more analogous to hardness than it is to mass measurement. One can no more combine the

TABLE 2.
Physical and behavioral measurement analogues with some key features.

Physical	Behavior	Features	Scale Type
Mass, length, Voltage, etc.	Utility, Psychophysical	Laws, rapid change, and reversible	Ratio or interval
Heating	Hunger	Regularities, slow change, and reversible	Ordinal or stronger
Material fatigue	Addictiveness	Regularities, slow change, but not reversible	Unclear
Hardness, storm intensity	Abilities, traits	No real structural laws	Ordinal

ability of two people to obtain the ability of the pair than one can combine the hardness of a blob of steel and that of a blob of copper. To some degree, it may be a form of conjoint measurement.

So where are there analogies to fundamental physical measurement?

I contend that some may lie in those human attributes that have long been studied in visual and auditory psychophysics, such as brightness and loudness or chroma and pitch, and at an interface of psychology and economics that centers on preferences among uncertain alternatives and that leads to theories of utility. Typically these are variables that are fairly easily manipulated, usually in a reversible way[4]. These are very different from some other personal variables that cannot be changed rapidly in an experimental setting. Hunger is a prototypical example, and there is no really satisfactory measurement theory for or even good ad hoc measures of hunger, just crude indices such as hours of deprivation or percent of normal body weight. Hunger is not a variable that we know how to increase or decrease very rapidly — in seconds or 10s of milliseconds rather than hours — and so fairly direct comparison of states of hunger is infeasible.

Table 2 outlines my crude overall taxonomy showing some of the types of measurement with which we are faced in the physical and behavioral sciences. My topics discussed below center on the two topics in boxes. There are, of course, a good many different approaches to the manipulable variables but, as may be understandable, I will focus on some of my contributions.

3. Utility Theories

For over half a century the approach to utility theory has been based on two classes of primitives. One can be called the domain of the theory that in one way or another models our intuitions about what constitute certain, risky, and uncertain situations. Often the latter are called " gambles." Other terms that have been used are " act" (Savage, 1954) and " prospect" (Kahneman & Tversky, 1979). Gambles with known probabilities attached to the underlying chance events (risk) and with money consequences are called " lotteries." The other primitive, which is where the behavior enters, are preference patterns of an individual among sets of gambles and certain consequences. These patterns are formulated as mathematical axioms.

[4]Of course, some care is needed, now enforced in the U.S. by university IRBs, not to endanger the individual, e.g., by using overly intense signals.

Formalizing the Primitives

State spaces: For the most part, economists and statisticians have assumed that the domain of uncertainty can be modeled as a huge " state" space that enumerates all of the states of nature that can possibly occur in the time frame under consideration, and that uncertain alternatives are mappings (acts) from the state space into pure consequences[5] with the image of the mappings having only finitely many values (Savage, 1954). That is, the uncertain alternatives or acts are viewed as maps with finite support. If lotteries are involved, then they are often formulated as random variables.

Local[6] *gambles:* In contrast, psychologists (and economists when they present specific examples) take an approach that says we choose among local gambles, such as the ones that we know how to construct readily in a laboratory. A typical laboratory example of local uncertainty is an opaque urn with thoroughly mixed colored balls and an assignments of consequences to the color drawn. Thus, a random choice from the urn is an example of an " experiment" **E** underlying the uncertainty. The experimenter can vary the information given to the respondent about the color composition of the urn, anywhere from pure risk in which the number of balls having each color is provided, to pure uncertainty where no information is provided beyond the number of distinct colors involved. If we think of the partition of the balls by color which we may number $i = 1, 2, \dots, n$, then let C_i denote the event of drawing a ball of color i and let x_i be the consequence assigned to that color. We speak of the pair (x_i, C_i) as the ith branch of the entire gamble based on n different colors, and denote the whole gamble, i.e., n-tuple of branches (but dropping the parentheses) by:

$$(x_1, C_1; x_2, C_2; \dots; x_i, C_i; \dots; x_n, C_n).$$

The set $\mathbf{C}_n = \{C_1, C_2, \dots, C_i, \dots, C_n\}$ is a partition of the "universal set" $\Omega_{\mathbf{E}} = \bigcup_{i=1}^{n} C_i$ of the experiment **E**. We allow the possibly including null events in the partition.

More serious situations are also thought of locally. For example, the important issues surrounding buying insurance on a house and its contents are usually formulated as a local gamble, although few of us call them that.

Preference order and certainty equivalents: The respondent is assumed to have a preference order among the gambles and pure consequences that is revealed by their choices. If f and g are two gambles, not necessarily based on the same underlying experiment, then the statement that f is at least as preferred as g is abbreviated as $f \succsim g$. Define $\sim$ by: $f \sim g$ iff both $f \succsim g$ and $g \succsim f$.

Indeed, let us assume that for each gamble f there is a certain consequence $CE(f)$, i.e., neither risky nor uncertain, such that $CE(f) \sim f$. $CE(f)$ called a *certainty equivalent* of f.

Here I, and many others, make the sharp idealization that $CE(f)$ is a fixed, not a random, quantity. In experimental practice, as would be expected, $CE(f)$ is better described as a random "certain" consequence rather than as a unique certain consequence. In practice we often select for $CE(f)$ some measure of central tendency of the judgments.

Permutation invariance: If ρ is any permutation of $\{1, 2, \dots, n\}$, assume that

$$(x_1, C_1; \dots; x_i, C_i; \dots; x_n, C_n) \sim (x_{\rho(1)}, C_{\rho(1)}; \dots; x_{\rho(i)}, C_{\rho(i)}; \dots; x_{\rho(n)}, C_{\rho(n)}). \quad (5)$$

Ranked gambles: The indices of the branches can be arbitrary or they can be ranked in some fashion. Suppose that the indices of a gamble are ranked from best to worst consequence, i.e., so

[5]Many speak of these as " outcomes" of the gamble. I shun that usage because probabilists call the state the arises in the chance experiment underlying the gamble an outcome. Such a double use–states versus a pure consequence associated with a state–can be ambigous.

[6]Later in the psychophysical setting, the terms " local" and " global" are given a different meaning. The context should keep the meaning straight.

that $x_1 \succsim x_2 \succsim \cdots \succsim x_n$, and then $\overrightarrow{\mathbf{C}}_n = \{C_1, C_2, ..., C_n\}$ is the corresponding vector partition of the $\Omega_{\mathbf{E}}$. In this case, we sometimes denote the gamble by $f_{\overrightarrow{\mathbf{C}}_n}$. In such cases, we speak of the resulting theories as ranked. Even though we assume that (5), such rankings are sometimes very useful for having a neat formulation of the resulting numerical representation (see " Ranked weighted utility" below).

With the assumption of permutation invariance, (5), it is quite easy to become confused about statements made for ranked gambles. However, the key to keeping matters straight is that if $f_{\overrightarrow{\mathbf{C}}_n}$ is ranked, any non-trivial permutation ρ destroys the inequalities $x_1 \succsim x_2 \succsim \cdots \succsim x_n$.

Three types of trade-offs can be manipulated and studied: consequences versus consequences, consequences versus events, and events versus events. And these become the source of measuring both the utility of certain consequences and of gambles, and also lead to some variety of subjective weights, sometimes subjective probabilities, for the events.

No change from the status quo: Another important primitive, often shunned by economic theoreticians, is a special consequence that partitions the certain consequences and the gambles into perceived gains and losses (Kahneman & Tversky, 1979). The dividing line may be called "no change from the status quo," denoted by e. Some call it a "reference point." Economists try to avoid this distinction by saying that the domain of consequences is not gains and losses, but rather final total wealth; however, almost always their specific examples are described in terms of gains and losses. Psychologists take the attitude that such every-day concepts as reference points cannot be avoided, and that the role of total wealth should be incorporated much more indirectly into the theories.

Basic Underlying Assumptions

Everyone begins with some structural assumptions — that the domain of choices must, in some sense, be very rich — and with some preference assumptions. Among the latter, the following two are most important.

Transitivity: If $f \succsim g$ and $g \succsim h$, then $f \succsim h$. In an ingenious experiment, which pit small increases in consequences against small decreases in likelihood, Tversky (1969) collected data and argued that violations of transitivity were observed and that a lexicographic semiorder is a better model. Replications of the experiment and analysis confirmed that conclusion, which has been widely accepted among cognitive psychologists. His data analysis was flawed as Iverson and Falmagne (1985) pointed out. They did a better reanalysis and concluded that Tversky's conclusion of behavioral intransitivity was true for at most one of the eight respondents. Recently, however, that conclusion has also been challenged and Tversky's original position accepted. This analysis is based on a new Bayesian analysis that pits the concept of a lexicographic semiorder against a weak order. They find that except for one respondent, the lexicographic semiorder model fit the data substantially better than a weak order (Iverson, Myung, & Karabatsos, 2005; see also Myung, Karabatsos, & Iverson, 2005).

This finding is somewhat disturbing because most theories leading to numerical utility assume transitivity. For this and other reasons (see below) many of us have concluded that it may be wise to avoid direct choices between gambles in testing the theories and to use some form of certainty equivalent instead. I assume here that transitivity holds.

Co-monotonicity: Suppose that the ith branch (x_i, C_i) of a gamble f is replaced by (x_i', C_i) but everything else is unchanged, including the ranking among certain consequences, i.e., $x_{i-1} \succsim x_i' \succsim x_{i+1}$. Call the resulting gamble f'. *Co-monotonicity.* is the assumption that

$$x_i' \succsim x_i \quad \text{iff} \quad f' \succsim f.$$

For the most part, this property is sustained empirically, at least for gains, with one notable exception discovered by M. H. Birnbaum (summarized by Birnbaum, 1997, and the earlier studies cited there). The exception seems to occur when $n = 2$ and $x_2 = e$ is replaced by an x_2' that is not too large. For example, if JCE denotes some form of judged certainty equivalent (he has used several different ones)

$$JCE(\$96, .90; \$0, .10) > JCE(\$96, .90; \$15, .10).$$

Birnbaum has repeatedly found violations of monotonicity provided that the following conditions hold:

- The comparison is made using some version of certainty equivalents. The dominance is transparent when the gambles are presented as a choice.
- The value of x_1 is "large"compared to that of x_2'.
- The probability of x_1 occurring exceeds about .85.

For a Bayesian reanalysis of Birnbaum' data, see Karabatsos (2005).

One suspects that this may have something to do with a change of strategy on the part of some respondents. It is comparatively easy to estimate mentally the expected value of $(x_1, p_1; 0, 1-p_1)$, namely, $x_1 p_1$, whereas evaluating $(x_1, p_1; x_2, 1-p_1)$ may well not be based on an accurate mental estimate of expected value. See Luce (2003) for an attempt to deal axiomatically with this phenomenon.

Additive cancellation properties: The conditions called *(ranked) additive cancellation* are well known from the literature on additive conjoint measurement theory and I do not detail them. One important one is *triple cancellation* when the ranking properties are maintained (Wakker, 1991). Again, see Karabatsos (2005) for analyses favoring double and triple cancellation.

Idempotence: We assume here the property

$$e \sim (e, C_1; e, C_2; \ldots; e, C_n), \tag{6}$$

which is called *e−idempotence*. Stated another way,

$$CE(e, C_1; e, C_2; ...; e, C_n) = e.$$

$e-$idempotence can be interpreted as saying that gambling, per se, has no inherent utility. If (6) holds when e is replaced by any x, the property is simply called *idempotence.* It is a property of almost all of the major theories of utility, although judging by some recent meetings interest is growing in theories that do not suppose it. When $e-$idempotence is dropped one has the opportunity to encompass the concept of the utility of gambling, that is $U(e, C_1; e, C_2; ...; e, C_n) \neq 0$ even though $U(e) = 0$. See Luce and Marley (2000) for more on the topic including some references. Perhaps the first systematic consideration of what amount to failures of $e-$idempotence is a little known article by Meginniss (1976) (see below).

Certainty: We assume that if there is no uncertainty about the consequence of a chance experiment, then

$$(x_1, C; x_2, \varnothing; \ldots; x_n, \varnothing) \sim x_1. \tag{7}$$

Sometimes people write this simply as (x_1, C). It seems like a triviality, but it does mean that the utility over pure consequences is the same as utility over gambles.

Several Specific Representations

Ranked additive utility (RAU): I begin with a fairly general additive representation. It supposes:

1. That there exists a real valued, order preserving function U over gambles and certain consequences onto a non-negative real interval such that $U(e) = 0$, and
2. That for the ordered gambles $f_{\overrightarrow{\mathbf{C}}_n} := (x_1, C_1; \ldots; x_i, C_i; \ldots; x_n, C_n)$, there exist functions $L^{(i)}\left(z, \overrightarrow{\mathbf{C}}_n\right)$ $(i = 1, ..., n)$, where z a typical value of the function U, with $L^{(i)}\left(0, \overrightarrow{\mathbf{C}}_n\right) = 0$ that are strictly increasing in z and such that

$$U(f) = \sum_{i=1}^{n} L^{(i)}\left(U(x_i), \overrightarrow{\mathbf{C}}_n\right). \tag{8}$$

A major open issue is to axiomatize the representation (8). Given a utility function U, which will exist under the usual weak conditions, then one can invoke Wakker's (1991) axiomatization of ranked-additive conjoint measurement to get a ranked representation of the form

$$L\left(U(f_{\overrightarrow{\mathbf{C}}_n})\right) = \sum_{i=1}^{n} L_i\left(U(x_i), \overrightarrow{\mathbf{C}}_n\right).$$

For the unranked case, see Krantz, Luce, Suppes, and Tversky (1971). An important unsolved problem is to find a condition that is necessary and sufficient for $L(\cdot)$ to be the identity function.

A portion of the literature takes (8) as a starting point and considers various properties that restrict it in interesting ways. Three recent examples are due, primarily, to Luce and Marley (2005). *Ranked weighted utility (RWU):* This is the form of (8) that says that for each i, utility and weighting of events are orthogonal concepts: for some functions S_i defined over ordered event partitions and mapping into [0, 1],

$$L_i\left(U(x_i), \overrightarrow{\mathbf{C}}_n\right) = U(x_1)S_i\left(\overrightarrow{\mathbf{C}}_n\right),$$

and so

$$U\left(f_{\overrightarrow{\mathbf{C}}_n}\right) = \sum_{i=1}^{n} U(x_i)S_i\left(\overrightarrow{\mathbf{C}}_n\right). \tag{9}$$

Under RWU, note that idempotence holds iff $\sum_{i=1}^{n} S_i\left(\overrightarrow{\mathbf{C}}_n\right) = 1$.

It turns out that a simple property insures this form. Suppose $x \succ e$. Let (C_1, C_2, C_3) and (D_1, D_2) be two independently run experiments where the concept of independent is informal. In the simplest case the property states

$$((x, D_1; e, D_2), C_1; e, C_2; e, C_3) \sim ((x, C_1; e, C_2; e, C_3), D_1; e, D_2).$$

This case is illustrated graphically in Figure 1.

This is an example of what is called *status-quo event commutativity*. Luce and Marley (2005) generalize it to the case of two gains $x \succsim y \succ e$ in a fairly obvious way and then show:

Theorem 1. Over general gambles, the following statements are equivalent:

1. Both the representation RAU, (8), and status-quo event commutativity hold.
2. RWU, (9), holds.

Event commutativity is an important example of an *accounting indifference* (Luce, 1990), which means only that two formulations of an uncertain alternative are indifferent provided that

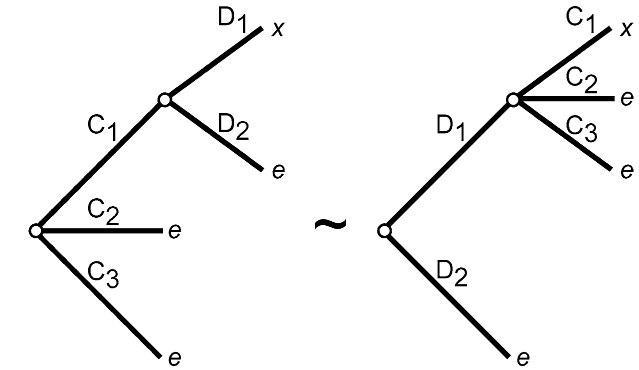

FIGURE 1.
A graphical example of status-quo event commutativity with $m = 2, n = 3$.

the same consequences arise under the same conditions, where the order of independent events occurring is immaterial. Of course, from an empirical perspective it becomes a question whether, in fact, respondents recognize the indifference of the two formulations of the same gamble. The limited data that exist tend to support event commutativity (Luce, 2000). However, in some examples, such as an airplane trip broken into two legs, such event commutativity is meaningless. Earlier we encountered three other examples of accounting indifferences: permutation invariance, e-idempotence, and certainty.

For the most part, decision theorists, especially those coming from economics and statistics, have shunned such accounting indifferences. Or, more accurately, they have automatically built them all in, without much comment, by the way that they have formulated the domain of gambles. For example, if the situation is one of risk they model the domain as a set of random variables. But compound independent random variable are nothing but convolutions, which are automatically written as first-order random variables. Savage's (1954) formulation at the level of first-order acts amounts to the same thing.

Rank dependent utility (RDU): The class of RDU models has been on center stage for 20 years for those interested in utility theory. Quiggin (1982) originated a special case of it and wrote a summary book on it for the risky context (Quiggin, 1993). Gilboa (1987) and Schmeidler (1989) generalized Quiggin's model to uncertain events, where it is often called Choquet expected utility, and Tversky and Kahneman (1992) invoked it in their extension, and improvement, of their famous prospect theory (Kahneman & Tversky, 1979) to gambles with more than two gains. This they called *cumulative prospect theory*. Luce and Fishburn (1991, 1995) and (Luce, 2000) used the term " rank-dependent utility" and studied it fairly intensively in conjunction with the existence of a binary operation; see the subsection " Joint Receipts" below.

Suppose that W is the function that arises in the idempotent binary case of RWU, i.e.,

$$U(x_1, C_1; x_2, C_2) = U(x_1)W_{C_1 \cup C_2}(C_1) + U(x_2)[1 - W_{C_1 \cup C_2}(C_1)] \quad (x_1 \succsim x_2). \tag{10}$$

If we define $C(i) := \bigcup_{j=1}^{i} C_j$ (so $\Omega_{\mathbf{E}} \equiv C(n)$), then we call the RWU representation RDU if the weights are given by:

$$S_i\left(\vec{\mathbf{C}}_n\right) = W_{\Omega_{\mathbf{E}}}\left(C(i)\right) - W_{\Omega_{\mathbf{E}}}\left(C(i-1)\right). \tag{11}$$

That is to say, the weight assigned to the ith consequence can be thought of as the incremental effect of adding C_i to the union of the prior events, $C(i-1)$. Here is a case where under a permutation the expression becomes cumbersome. Expressions of the form (11) were called capacities by Choquet (1953) but increasingly they are called Choquet weights or measures. Note that by certainty, (7), $W_C(C) = 1$. RDU is defined to be RWU with the weights given by (11).

Again, a simple accounting indifference underlies RDU, which is known — for reasons that will soon be apparent — both as *coalescing* (Luce, 1998) and as *event splitting* (Starmer & Sudgen, 1993). Consider the special case of a general gamble in which the same consequence x is assigned to both of the events C_i and C_{i+1} for some $i \in \{1, 2, ..., n\}$. The assertion is that in this case it is immaterial whether the gamble is presented with (x, C_i) and (x, C_{i+1}) as distinct branches or with $(x, C_i \cup C_{i+1})$ as a single branch:

$$\begin{aligned}&(x_1, C_1; x_2, C_2; ...; x, C_i; x, C_{i+1}; ...; x_n, C_n)\\ &\sim (x_1, C_1; x_2, C_2; ...; x, C_i \cup C_{i+1}; ...; x_n, C_n).\end{aligned} \tag{12}$$

It is obvious that this too is an accounting indifference. See Figure 2.

Theorem 2. (Luce & Marley, 2005). Over general gambles of gains and assuming certainty, (7), the following are equivalent:

1. Both RWU, (9), and coalescing, (12), hold.
2. RDU holds.

In this case idempotence holds because, by the certainty property and repeated applications of (12),

$$(x, C_1; x, C_2; ...; x, C_i; ...; x, C_n) \sim (x, C(n); x, \varnothing) \sim x$$

One sees that going from the left side to the right in (12) the reduction by coalescing is very natural and unique. But going from the right side to the left is far more ambiguous because there are very many ways to split any non-null, non-atomic event C_i into two subevents. This direction in (12) invites the term "event splitting."

Birnbaum (1999) has provided what amounts to a recipe for examples where subjects violate stochastic dominance when gambles are coalesced but not when they are appropriately partitioned. For example, consider, first, the more finely partitioned gambles. Using his mode of presentation in which probabilities lie above their corresponding consequences a pair of lotteries might be

$$g \sim \frac{.85 \quad .05 \quad .05 \quad .05}{96 \quad 96 \quad 14 \quad 12}, \quad h \sim \frac{.85 \quad .05 \quad .05 \quad .05}{96 \quad 90 \quad 12 \quad 12}.$$

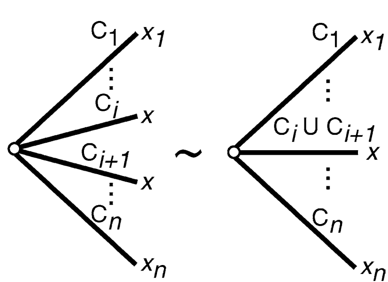

FIGURE 2.
A graphical example of coalescing.

Clearly, g dominates h. And of his sample of 100 undergraduates, 85 chose g over h. Now consider these lotteries in coalesced form

$$g' = \frac{.90 \quad .05 \quad .05}{96 \quad 14 \quad 12}, \quad h' = \frac{.85 \quad .05 \quad .10}{96 \quad 90 \quad 12}.$$

In this form 70 of 100 undergraduates chose h' over g'.

Thus, we have the (group) non-transitivity:

$$g' \sim g \succ h \sim h' \succ g'.$$

Put another way, when confronted with g' and h', most respondents apparently did not see the particular event splitting that makes the dominance transparent.

In addition, Birnbaum and others have presented a number of properties where RDU, and so cumulative prospect theory, fares very poorly. Marley and Luce (2005, in press) have given formal definitions of 23 of these properties and explored exactly how they fare under four different models: RWU, RDU, GDU (see below), and TAX, which is a particular way that Birnbaum favors writing RWU (Marley & Luce 2005, in press showed that they are equivalent). Table 3 summarizes how RDU fares on those properties where RDU either forces the property always to hold or implies that it can never hold. Of the 12 cases for which RDU predicts the property must hold, the data agree with the prediction in 3 of them, disagree in 6, and the other 3 have not been studied empirically. However, the 3 cases of agreement tell us nothing about RDU because, in fact, all RWU models predict the same thing. A 50% (or 67%, depending on how one counts these 3 successes) failure rate for predictions strikes me as rather unambiguous evidence against the theory. That coupled with the easily demonstrated failures of stochastic dominance have convinced me that RDU is just not an adequate descriptive theory. So far, few theorists have been ready to accept this conclusion, but recall that it took us quite a while to accept the earlier implications of data showing violations of subjective expected utility. But in my opinion, the descriptive defeat is just as clear cut.

Gains decomposition utility (GDU): This representation is the special case of RWU that satisfies the following accounting indifference (first stated for probabilities by Liu, 1995, although it is used implicitly by Meginniss, 1976). Again, I state the property in a simple example, and the generalizations are fairly obvious. Consider a general trinary, ranked gamble

$$g_3 := (x_1, C_1; x_2, C_2; x_3, C_3),$$

and also the subgamble that treats branches 1 and 2 as a (sub)gamble.

$$g_2 := (x_1, C_1; x_2, C_2).$$

TABLE 3.
Marley & Luce (2005) worked out in detail the predictions of models and summarized Birnbaum's empirical findings. For RDU the unambiguous predictions are

		Data		
RDU Prediction	# of Cases	# Agree	# Disagree	No Data
Must Hold	12	3[a]	6	3
Never Holds	7	3	0	4
Totals	19	6	6	7

Note: [a]These 3 must hold for any RWU model.

The property of *lower gains decomposition* (for gambles with 3 branches) is the assertion

$$g_3 \sim (g_2, C_1 \cup C_2; x_3, C_3), \tag{13}$$

which says that the original gamble, on the left, is indifferent to the compound binary pair on the right. Again, this is an accounting indifference. From this, an explicit formula can be derived for the weights of the utility representation in terms of the weights using the binary RDU form, (10), that is equivalent to both RWU and gains decomposition holding, which is called *(lower) gains decomposition utility* (GDU). That formula is complicated to state. Lower GDU implies lower gains decomposition.

Theorem 3. (Luce & Marley, 2005; Marley & Luce, 2001). Suppose that RWU, (9), holds for gains. Then any two of the following assertions implies the third:

1. RDU holds.
2. GDU holds.
3. The choice property of Luce (1959) holds: for events $C \subseteq D \subseteq E$

$$W_E(C) = W_D(C)W_E(D).$$

Liu (1995) basically used 2 and 3 to axiomatize RDU for risk and Luce (2000) did the parallel thing for uncertain gambles.

The empirical failure of RDU means that either GDU or the choice property or both are wrong descriptively.

The results in Marley and Luce (2005) show that GDU fares quite well, but not perfectly, on the various properties rejecting RDU, but no systematic attack on GDU has yet been mounted.
Subjective expected utility (SEU): The SEU representation arises from RDU when the weights are finitely additive. That is, the weight for branch i depends only on C_i, not on the rest of the event partition. So the rank dependence actually vanishes.

SEU was first formally axiomatized by Savage (1954). However, the highly innovative, posthumous book Ramsey (1931) of that youthful philosopher contains a partial axiomatization in a chapter, written in 1926.

Suppose that we generalize the concept of lower gains decomposition from partitioning on the least valued consequence to doing so on any branch (x_i, C_i). We have shown that if it holds in gambles of size $n = 3$ for $i = 1, 2, 3$, then in addition to the choice property the weights W_E are finitely additive. Thus, $S_i\left(\overrightarrow{\mathbf{C}}_n\right) = W_{\Omega_\mathbf{E}}(C_i)$, which means they are subjective probabilities, as in SEU.

Although we do not yet know an accounting indifference that when added to RDU yields SEU, we do know that coupling RDU and all forms of GDU we get SEU along with the choice property.

This fact is important in giving insight into Savage's result. By formulating the decision problem as mappings from a huge state space (Cartesian product of particular event partitions) to consequences, one builds in the entire set of accounting indifferences. Thus, once one has sufficient monotonicity and cancellation properties to get the ranked additive form (conjoint measurement) SEU follows automatically. This fact is far from apparent in Savage's proof.

Likewise, if, as commonly done, lotteries are treated as random variables, then, as mentioned above, the accounting indifferences follow automatically as convolution or, in the utility literature, as the reduction of compound gambles.

This makes clear that if one wishes to avoid the descriptively inadequate SEU or RDU, one had better shun the state space formulation. Utility theorists seem, for the most part, not yet to appreciate this point.

There is another way to describe the situation. The conceptual emphasis should be placed on the structure of the space of uncertain alternatives (gambles), not primarily on either the space of consequences or on a state space. For reasons that I have never fully understood, a number of utility theorists exhibit a form of puritanism or, if you will, conservatism that rejects using any form of compound gambles such as is involved in event commutativity and gains decomposition. *Utility of Gambling:* Although both Ramsey (1931) and von Neumann and Morgenstern (1947) explicitly mentioned this issue (as difficult to formulate within their frameworks), it has received little theoretical attention. Luce and Marley (2000) give some of the relevant references, as well as some theory, but we missed what is perhaps the most interesting paper, a newly rediscovered one that was brought to my attention by János Aczél: Meginniss (1976). He dropped the assumption of e-idempotence, (6), and he implicitly invoked (upper) gains decomposition. Because he did not impose a ranking constraint, there are no distinctions among the types of gains decomposition. Under a very restrictive form of RAU, namely that in (8), $L^{(i)}\left(U(x_i, \overrightarrow{\mathbf{C}}_n\right) = L\left(U(x_i), C_i\right)$, where L is independent of i and the only relevant event is C_i. His axiomatization resulted in two possible representations for risky gambles with a probability distribution P. One is:

$$U\left(f_P\right) = EU(f_P) + aH(P),$$

where EU is the expected utility of f_P and $H(P)$ is Shannon's entropy of the distribution P. The second representation is

$$U(f_P) = EU_{P^c}(f_P) + a\left(1 - \sum_{i=1}^{n} p_i^c\right),$$

where EU_{p^c} is the " expectation" of U relative to the weights p^c, $c \neq 1$, which do not sum to 1. The term $1 - \sum_{i=1}^{n} p_i^c$ has been called *entropy of degree c* (Aczél & Daróczy, 1975, p. 189–191). Research is under way concerning what happens when Meginniss' assumptions are weakened in natural ways and generalized to uncertain alternatives.

Joint Receipts

Given gambles f and g (including, of course, pure consequences as possible special cases) one can ask questions about their joint evaluation by a decision maker. Let $f \oplus g$ mean the receipt of both f and g. These are called *joint receipts.*

One encounters joint receipts whenever one purchases goods in a store or buys a portfolio of stocks. In traditional economics this has been modeled in the form of commodity bundles that are treated as a vector of amounts of the various goods. These vectors are, in practice, very sparse indeed. The vector corresponding to what is available in a large grocery store has many thousands of distinct item types, yet an individual shopper's usual bundle has only tens of non-zero entries. So maybe we are better off treating joint receipt as an operation, somewhat akin to placing two (or more) masses on a pan of a pan balance, rather than as a vector. And, of course, this invites a somewhat different foundational analysis: abstract algebra rather than the vector algebra of amounts.

At the level of representations, the question is: If we know, for example, $U(f)$ and $U(g)$, how does $U(f \oplus g)$ relate to them? I assume (with some loss of generality) that there is a function F such that

$$U(f \oplus g) = F\left(U(f), U(g)\right).$$

This is often called a *decomposability* assumption. The problem is to characterize F beyond its being strictly increasing in each variable.

Although the very popular RDU theory of the 1980s and 1990s has been rejected as descriptive, a sample of results of combining it with joint receipt (Luce, 2000, chapters 4–7) is illustrative. There are two classes of assumptions. One has to do with just $\oplus$ and $\succsim$ separately, much in analogy to $\circ$ and $\succsim$ in the case of mass measurement. Indeed, let us make the very same assumptions, namely, that $\oplus$ is strictly increasing in each argument, $\oplus$ is commutative and associative, and e is the identity of $\oplus$. Formally

$$\begin{aligned} x \succsim x' \quad &\text{iff} \quad x \oplus y \succsim x' \oplus y \quad (x \succ e, x' \succ e), \\ x \oplus y \quad &\sim \quad y \oplus x, \ (x \oplus y) \oplus z \sim x \oplus (y \oplus z), \ e \oplus x \sim x. \end{aligned}$$

Under well known structural assumptions, this implies the existence of an additive numerical representation V with the following properties:

$$\begin{aligned} f \succsim g \quad &\text{iff} \quad V(f) \geq V(g), \\ V(f \oplus g) \quad &= \quad V(f) + V(g), \\ V(e) \quad &= \quad 0. \end{aligned}$$

It should be mentioned that a recent article by Wu and Markle (2005) strongly suggests that in the interesting domain of mixed gains and losses, joint receipt may very well not exhibit monotonicity. Mixed consequences form a very rich and important area of research that has been neglected too much.

We have absolutely no reason at this point to suppose that the additive representation V and the representation U arising from gambles are the same measure. The only relation between them so far is that each measure preserves the same ordering $\succsim$ and that, therefore, there is some function φ from $\mathbb{R}$ to $\mathbb{R}$ such that $U = \varphi(V)$. Thus one issue is to characterize φ.

To do that, we need a link between gambles and joint receipt. Limiting ourselves to gains, one link that I have probed is called *segregation.* It was invoked informally as part of a so-called "editing" phase of evaluation by Kahneman and Tversky (1979) in their widely acclaimed[7] prospect theory. Segregation in the simplest case says

$$(x, C; e, D) \oplus y \sim (x \oplus y, C; e \oplus y, D) \sim (x \oplus y, C; y, D).$$

Note that segregation is another accounting indifference. From segregation and RDU, drawing on work of Luce (1991), Luce and Fishburn (1991, 1995), using functional equation techniques, one can show that there exists a constant δ with dimension that of $1/U$ such that

$$U(f \oplus g) = U(f) + U(g) + \delta U(f)U(g).$$

This form has come to be called *p-additive*, where p stands for " polynomial". This is because it is the polynomial form that under the transformation $V = sgn(\delta) \ln(1 + \delta U)$ yields the additive representation $V(f \oplus g) = V(f) + V(g)$. This loosely parallels the relation between relativistic velocity and the additive rapidity relation, including the fact that both U and V are ratio scales — but for different reasons, namely, the uniqueness resulting from gambles and that from joint receipts.

Birnbaum's work on the non-monotonicity of binary gambles when one of the consequences goes to e, suggest that we should by-pass e in the segregation indifference and study the case of *distribution*

[7]I say " widely acclaimed" because Kahneman received the Nobel prize in economics in 2002, and prospect theory and related empirical studies were a major factor underlying the award. Unfortunately, Tversky died before the award was made, and it is given only to the living.

$$(x, C; y, D) \oplus z \sim (x \oplus z, C; y \oplus z, D) \quad (x \succsim y \succ e, z \succsim e).$$

I did this in Luce (2003), but in the interest of giving a related but different example, I forego summarizing these results.

4. Global Psychophysics

Consider a qualitative attribute such as the loudness of pure tones as affected by signal intensity[8]. Global psychophysics is the attempt to understand the overall nature of perceived loudness throughout the full range of intensities. Other forms of perceived intensity, such as brightness, are also candidates for the theory. In contrast, local psychophysics attempts to understand very local relations of signals so close together that they tend to be confused. An auditory rule of thumb is that for a range of compared signals less than 5dB, it is local and otherwise it is global. A detection experiment is a typical local case whereas absolute identification and magnitude estimation (Stevens, 1975) studies with a wide range of intensities are typical global ones. Indeed, the vast amount of theoretical work focuses on local phenomena.

Some of the questions addressed in such global work are:

1. Is there a qualitative formulation leading to a numerical representation of loudness?
2. Assuming the answer to 1 is Yes, how does such a measure depend on various physically manipulable factors such as signal intensity and frequency.
3. How does the loudness in each of the two ears combine to produce an overall sensation of binaural loudness?
4. What signal does a respondent say has a loudness corresponding to 3 times or to $1/3$ the loudness of a given signal?

Such topics have received a good deal of empirical attention in audition and in other senses, such as brightness and heaviness. There is also a sizable theoretical literature focused on describing how the data are interrelated. Here I give a small sample based on loudness studies done jointly with my recent Ph. D. Steingrimsson (2002). Current work on brightness (based on pilot data, Steingrimsson, personal communication, April, 2005) seems to be coming up with closely parallel results.

Reinterpretation of the Primitives

Joint presentations to the two ears: The question addressed recently in Luce (2002, 2004) is whether or not the concepts of gambles and joint receipt have fairly natural reinterpretations in psychophysics, in which case the same, or closely related, mathematics can be invoked.

Suppose that an above-threshold intensity presented to the left ear is measured as follows. If x_τ denotes the physical intensity of the individual's left threshold and $x' > x_\tau$ is the actual physical intensity presented, then we treat $x := x' - x_\tau$ as the stimulus. Note that this is not a decibel difference, which corresponds to the intensity ratio x'/x_τ. For the right ear the notation is parallel with u substituted for x. Then if x is presented to the left ear and u to the right at the same time we are dealing with the stimulus pair (x, u).

[8]The theory needs to be extended to treat frequency as well as intensity both as it affects loudness and as generating a concept of pitch. Once that is done, then it should be extended further to deal with complex signals involving a spectrum of frequency-intensity pairs.

This pair resembles somewhat the joint receipt of x and u, but unlike the utility case no direct meaning is assigned to a compound symbol such as $((x, u), v)$. However, an indirect one can be given. Suppose that $z_l = z_l(x, u)$ is the signal intensity in the left ear that can in some sense be considered equivalent to (x, y), then the compound makes sense if it replaced by (z_l, v). There are three distinct ways that such matches can be established: in the left ear, in the right, or in both symmetrically:

$$(x, u) \sim (z_l, 0) \sim (0, z_r) \sim (z_s, z_s). \tag{14}$$

It is clear that z_i $(i = l, r, s)$ depends on — is a function of — both x and u. Assuming strict monotonicity in the variables, it is safe to treat this function as an operation, so by definition

$$x \oplus_l u := z_l, x \oplus_r u := z_r, x \oplus_s u := z_s. \tag{15}$$

These form three analogues to joint receipt.

At first I was optimistic that, at least among young people, we might find some for whom $\oplus_s$ is commutative and associative, in which case the mathematics would parallel utility theory. This entailed a search for people with symmetric ears in the sense that

$$(x, u) \sim (u, x). \tag{16}$$

But Steingrimsson and Luce (2005a) found only a small fraction of respondents who seemed to be symmetric. Thus, I had to revisit the theory without assuming commutativity. More on that later.

Subjective proportions: We next seek a formal analogue to gambles. The earlier notation for binary gambles $(x, C; y, D)$ can be modified in several ways leading to a psychophysical interpretation. Suppose $C \cup D = \Omega_{\mathbf{E}}$ is the universal set of experiment **E**, then it is sufficient to write the gamble as $(x, C; y)$. Next suppose that we are in a situation where $p = \Pr(C)$ exists, then this gamble can be written as an operator:

$$x \circ_p y := (x, p; y).$$

We can give this notation a psychophysical interpretation. Suppose that stimuli (x, x) and (y, y) are given, where x is more intense than y, $x > y$, and that a positive number p is also given. We may ask the respondent to choose the stimulus (z, z) that " makes the subjective interval from (y, y) to (z, z) stand in the proportion p to the subjective interval from (y, y) to (x, x)." Note that $z = z(x, y, p)$ or in operator notation

$$x \circ_p y := z(x, y; p).$$

This proportionality judgment generalizes S. S. Stevens' method of magnitude production in which $y = 0$. Stevens (1975, p.23, 157–158) discusses the relation of fractionation (in contrast to sectioning) in which the respondent is presented with two signals x and y, $x > y$, and a number p, $0 < p < 1$, and is asked for the signal that corresponds to that fraction. He suggests, although I find the text a bit opaque, that by setting $y = 0$ it becomes his method of magnitude production.

Representations of (x, u) *and* $\circ_p$

Luce (2004) arrived at behavioral properties of $\oplus_s$ and $\circ_p$ that lead to the following representation. There is a psychophysical function $\Psi(x, u)$ mapping the non-negative quadrant of the plane of intensities onto the non-negative real numbers and a distortion of numbers, W, and they are related to one another as I now describe.

First, using asymmetric matching, z_l or z_r in (14), there are two constants $\delta \geq 0, \gamma > 0$ such that the following three properties all hold for all $x \geq 0$, $u \geq 0$, $p > 0$:

$$\Psi(x,u) = \Psi(x,0) + \Psi(0,u) + \delta\Psi(x,0)\Psi(0,u) \quad (\delta \geq 0), \tag{17}$$

$$W(p) = \frac{\Psi\left[(x,x) \circ_p (y,y)\right] - \Psi(y,y)}{\Psi(x,x) - \Psi(y,y)} \quad (x > y \geq 0), \tag{18}$$

$$\Psi(x,0) = \gamma\Psi(0,x) \quad (\gamma > 0). \tag{19}$$

Note that for $\Phi = \ln(1 + \delta\Psi)$ with $\delta > 0$, Φ is strictly increasing in each variable and that $\Phi(x,u) = \Phi(x,0) + \Phi(0,u)$.

Second, using symmetric matching, z_s in (14), the theory shows that $\delta = 0$ in the summation property, (17), and that (18) also holds. However, the very restrictive constant-bias property of (19) may or may not apply under symmetric matching.

Two important special cases of (17) are: The unbiased case, where $\gamma = 1$, which means symmetry holds in the sense of (16), and the biased case with $\delta = 0$. One can show that $\delta = 0$ is satisfied if, and only if for some constant δ, $0 < \eta < 1$,

$$\Psi(x,u) = \eta\Psi(x,x) + (1-\eta)\Psi(u,u). \tag{20}$$

As this suggests, a key question to be addressed experimentally is a property that is equivalent to $\delta = 0$.

We now take up parameter-free, behavioral properties that give rise to these representations and the issue of whether or not $\delta = 0$.

Properties of the Primitives

Properties of (x,u) : The key necessary condition for $\Psi(x,u)$ to satisfy the expression (17) is the familiar *Thomsen condition* of additive conjoint measurement (Krantz et al., 1971, chapter 6):

$$\left.\begin{array}{l}(x,u') \sim (y',v)\\(y',u) \sim (y,u')\end{array}\right\} \Longrightarrow (x,u) \sim (y,v), \tag{21}$$

It is a cancellation property where one “ cancels” y' on the left ear and u' on the right one. It is not only necessary but, with other properties such as an Archimedean condition and some structural assumptions, it becomes a sufficient property for additivity.

A second property of (u,v), which I know how to state only by using the defined operation $\oplus_i$, (15), is *bisymmetry*:

$$(x \oplus_i y) \oplus_i (u \oplus_i v) = (x \oplus_i u) \oplus_i (y \oplus_i v) \quad (i = l, r, s). \tag{22}$$

This property is important because it necessary and sufficient for $\delta = 0$ in (17).

Steingrimsson and Luce (2005a) have provided experimental evidence supporting the Thomsen condition in 19 of 24 tests. Earlier studies of the Thomsen or, in context, the equivalent double cancellation property have yielded conflicting conclusions, roughly equally balanced between acceptance and rejection. We side with the studies that sustained double cancellation (Falmagne, Iverson, & Marcovici, 1979; Levelt, Riemersma, & Bunt, 1972; Schneider, 1988) but not with the others that did not (Falmagne, 1976; Gigerenzer & Strube, 1983). We do not understand with any certainty the sources for then inconsistencies although our article discusses possibilities.

Steingrimsson and Luce (2005b) provided strong evidence (6 of 6 respondents) in support of bisymmetry. This is important because the resulting $\delta = 0$ greatly simplifies some other testing.

Properties of $\circ_p$: The key necessary behavioral condition underlying the expression (18) is the analogue of binary event commutativity which, in this context, is called *(subjective) production commutativity:* For all $p > 0, q > 0$,

$$\big((x, x) \circ_p (y, y)\big) \circ_q (y, y) \sim \big((x, x) \circ_q (y, y)\big) \circ_p (y, y). \tag{23}$$

Observe that the two sides differ only in the order of applying p, q, which is the reason for the term " commutativity." This property also arose in Narens (1996) analysis of Stevens' theory. Three studies tend to support it: Ellermeier and Faulhammer (2000) for $p > 1, q > 1, y = 0$ (17/19 respondents); Zimmer (2005, in press) for $p < 1, q < 1, y = 0$ (6/7 respondents); and Steingrimsson and Luce (2005a) for $p > 1, q > 1, y > 0$, (4/4 respondents).
Links between (x, u) *and* $\circ_p$: As discussed earlier, in the measurement foundations of classical physics one must ask not only how each operation acts alone, but how they interact when one is dealing with both of them. For example, masses can be manipulated by combining them and by varying volume and density of materials. It is well known that they interact via a distribution-type law. We seek analogies of linking in this context.

One analogue is *left segregation of type* i $(i = l, r, s)$:

$$u \oplus_i (x \circ_{p,i} 0) \sim (u \oplus_i x) \circ_{p,i} (u \oplus_i 0).$$

Closely related is *right*[9] *segregation of type* i :

$$(x \circ_{p,i} 0) \oplus_i u \sim (x \oplus_i u) \circ_{p,i} (0 \oplus_i u).$$

We tested these for $i = l$ or $i = r$ except for one respondent where we tested both using $p = \frac{2}{3}$ and $p = 2$. Segregation was partially supported (8/10 tests)

Another linking condition, called *simple joint-presentation decomposition,* is:

$$(x \oplus_i u) \circ_{p,i} 0 = (x \circ_{p,i} 0) \oplus_i (u \circ_{p,i} 0) \quad (i = l, r, s). \tag{24}$$

Steingrimsson and Luce (2005b) checked this property which was partially sustained (6/8 tests). (The corresponding property that arises when $\delta \neq 0$ is more complex and would be vastly more difficult to test.)

Summary of the Theoretical Results

Luce (2004) has shown that the properties

Thomsen condition
Production commutativity
Simple joint presentation decomposition
Segregation

are necessary for the representations of (17), (19), and (18), and, under fairly mild background assumptions, they are also sufficient. Given that $\gamma \neq 1$, bisymmetry is necessary and sufficient for $\delta = 0$.

Forms of W and Ψ

Like the rank-dependent utility theory that this psychophysical theory closely resembles, it has the property that none of the behavioral properties that give rise to the representation have

[9]It must be studied separately from left segregation because commutativity of $\oplus_i$ does not hold.

any free parameters — all people are assumed to be identical at that behavioral level once one knows $\succsim$. People of course do vary in $\succsim$. Note that although testing parameter-free properties is non-trivial (but see Karabatsos, 2005), it is vastly simpler than testing any model or property that has (many) free parameters.

Despite the fact that the behavioral axioms are parameter free, the resulting representation has much freedom for individual differences, witness that the functions Ψ and W are not specified at all beyond being strictly monotonic increasing. These two statements — no parameters at the level of the axioms and vast parametric freedom at the level of the representations — seem, on the face of it, contradictory, but assuming no mathematical errors on my part, it is the case here and, of course, it has been well known in utility theory since at least Savage (1954). The source of the freedom in the representation results from the fact that the underlying order $\succsim$ over stimuli can and does differ among people. Of course, if behavioral individual differences seem to call for greater freedom, parameters will have to added to the behavioral axioms. So far, I do not think that the evidence warrants such a drastic step.

Indeed, two entirely free functions may be rather more freedom than one really wishes or needs. And so attempts have been undertaken to narrow them down. This is the topic of Steingrimsson and Luce (2005c, d, in preparation).

Three proposed forms for W : Narens (1996) arrived, in addition to subjective proportion commutativity, at the further condition

$$\big((x, x) \circ_p (y, y)\big) \circ_q (y, y) \sim (x, x) \circ_t (y, y) \; with \; t = pq \tag{25}$$

as a part of his formalization of Stevens' (1975) implicit assumptions underlying his magnitude methods. It is quite easy to show that this property together with (18) forces W to be a power function with $W(1) = 1$. For wholly different reasons, Schneider, Parker, Ostrosky, Stein, & Kanow (1974) used multidimensional scaling on similarity judgments and concluded that a power function fit the data well. In contrast, Ellermeier and Faulhammer (2000) and Steingrimsson and Luce (2005d, in preparation) have unambiguously rejected (25) empirically and so power functions for the distortion function W with $W(1) = 1$ are ruled out. But if we drop the multiplicative condition $t = pq$, then W can be a power function with $W(1) \neq 1$, which is suggested by the phenomenon of time order error in which the second of two temporally close presentations seems louder than if it is first. Data on this are currently being collected.

In utility theory data suggest that maybe W is of an inverse-S form over [0, 1], and Prelec (1998) arrived axiomatically at one such a form, namely,

$$W(p) = \begin{cases} \exp\left[-\lambda\,(-\ln p)^{\mu}\right] & (0 < p \leq 1) \\ \exp\left[\lambda'\,(\ln p)^{\mu}\right] & (1 < p) \end{cases}, \tag{26}$$

with $\mu \neq 1$. Luce (2001) provided a simpler axiomatic equivalence to (26), namely, if

$$\big((x, x) \circ_p (y, y)\big) \circ_q (y, y) \sim (x, x) \circ_t (y, y),$$

then for[10] $N = 2, 2/3$

$$\big((x, x) \circ_{p^N} (y, y)\big) \circ_{q^N} (y, y) \sim (x, x) \circ_{t^N} (y, y).$$

Zimmer and Baumann (2003a, b) reject this property where (x, u) is interpreted to be successive presentations of brief intensities x and u and Zimmer (2005) does so with two-ear presentations.

[10]Within the context of the theory, one can prove from this assumption that the property actually holds for all positive real numbers N.

Steingrimsson and Luce (2005d, in preparation) also tested it using the two ear interpretation and again reject this property. This strongly suggests that we need to consider other possible forms for W; a search is currently underway.

An earlier interesting hypothesis is the *power odds* one which came up in the context of utility (Gonzalez & Wu, 1998). Let $\Omega_p := p/(1-p)$, then the hypothesis for $p \leq 1$ is that

$$\Omega_{W(p)} = \alpha \Omega_p^{\rho} \quad (\alpha > 0, \rho > 0).$$

Questions remain: To what behavioral properties is it equivalent and what is the form when $p > 1$?

Although the Prelec function, (26), is incorrect, Steingrimsson and Luce (2005c, in preparation) show that if when there is no explicit reference the respondent using the signal and response on the preceding trial, one may develop for that function a fairly simple account for the existence of sequential effects. These effects are well known to exist. Presumably, once a more correct form for W is discovered, it too will provide a similar account for sequential effects.

A power function form for Ψ : Assuming the symmetric case of ratio production ($y = 0$), and select $z = tx$, $t > 0$. We may ask the respondent to report a ratio estimation $p_s = p_s(t, x)$. we have

$$\Psi\left((x, x) \circ_{p_s} (y, y)\right) = W(p_s)\Psi(x, x) \quad (x > 0),$$

so if we define

$$\psi_s(x) := \Psi(x, x),$$

etc., this becomes

$$W\left(p_s(t, x)\right) = \frac{\psi_s(tx)}{\psi_s(x)}. \tag{27}$$

Stevens (1975) more or less implicitly assumed that

$$p_s(t, x) = p_s(t), \tag{28}$$

in words, that the ratio production depends on the physical ratio of signals but not on the reference signal x. Hellman and Zwislocki (1961) provide supporting data for $t > 1$, although probably not for $t < 1$; whereas Beck and Shaw (1965) show the opposite pattern. The difference seems to relate to the location of the standards relative to the range.

Nonetheless, assuming (28), we have the functional equation

$$W\left(p_s(t)\right) = \frac{\psi_s(tx)}{\psi_s(x)},$$

from which it is not difficult to show that there exist constants $\alpha_s > 0$, $\beta_s > 0$ such that

$$\psi_s(x) = \alpha_s x^{\beta_s} \quad (x \geq 0), \tag{29}$$

$$p_s(t) = W^{-1}\left(t^{\beta_s}\right) \quad (t \geq 0), \tag{30}$$

Contrast this with Stevens (1975) who argued empirically that $p_s(t)$ is itself a power function. Others — I among them — have emphasized that individual data are only approximate power functions but with some distortion. Indeed, were $p(t)$ a power function, then from (30) we see that W would be a power function, which can only hold, if then, for $W(1) = 1$. Our conclusion is that it is not in general p_s but ψ_s that is a power function. It is an inferred psychophysical function, which is not directly observable, that is predicted to be a power function.

The analogous argument can be made for $i = l$ and $i = r$ which leads us to consider the two-ear psychophysical function

$$\Psi(x, u) = \alpha_l x^{\beta_l} + \alpha_r u^{\beta_r}. \tag{31}$$

One can show, under our assumptions, that this is equivalent to the following behavioral property called *multiplicative invariance*: For all signals x, u and any factor $\lambda > 0$, there exists $\sigma > 0$ such that

$$\lambda x \oplus_i \lambda^\sigma u = \lambda_i (x \oplus_i u) \quad \left(\lambda_i := \begin{cases} \lambda, & i = l \\ \lambda^\sigma, & i = r \end{cases}.\right). \tag{32}$$

It turns out that $\sigma := \beta_l/\beta_r$. Note that the property of constant bias, (19), is equivalent under this representation to $\sigma = 1$. To test (32) in general it is essential to estimate σ. One can show that under the assumption of (31),

$$z_r(x, 0)_{dB} = \sigma^2 z_l(0, x)_{dB} + c, \tag{33}$$

which provides an estimate of σ. So far, the only test of (32) was done on the assumption of constant bias, i.e., $\sigma = 1$ (Steingrimsson & Luce, 2005c, in preparation). Of their 22 respondents, three failed both of two tests, seven failed one of them, and 11 did not reject the invariance condition under constant bias. The property failed in 14 out of 44 tests. So, some people seem to satisfy the power function theory with $\sigma = 1$, but others do not. It appears the using the value of σ estimated by (33) does not suffice to makes this form work. This argues indirectly against (28). Research continues.

5. Closing Remarks

I have presented two somewhat related examples from the behavioral sciences that parallel, to a degree, some foundational formalisms of physical measurement. Such measurement is characterized by manipulating an attribute — preferences in one case and loudness (or, in principle, any intensity attribute) in the other — in at least two distinct and readily reversible ways. Both manipulations can be treated as mathematical operations: joint receipt and gambles in the utility case and joint presentation to the two ears and ratio proportions for (auditory) psychophysics. Each operator by itself yields a representation of the attribute, utility or loudness, but the two representations of the same ordering are only related by a strictly increasing function. To sharpen that connection, one must search for one or more relations linking the two operators. This is in the nature of a distribution property. Once done, that link can be converted into solving a functional equation. In this I have had the good fortune to enlist the help of some of the top experts in the world on functional equations. The guru of the field, János Aczél, has found some of the equations challenging and has enlisted the assistance of others in several countries: M. Kuczma, A. Lundberg, G. Maksa, C. T. Ng, Z. Páles. Some of the relevant functional equation papers are in the references under the first authors Aczél, Lundberg, Maksa, and Ng.

I believe that these behavioral examples are formally similar to some foundational aspects of physical measurement. For example, consider objects moving at constant velocities (no applied forces). One can concatenate velocities in the same direction to get an additive representation. Also velocity is typically defined to be the distance traversed divided by the time required to do so. In the foundations of classical physics, these two measures were linked by a simple distribution property and were shown to be proportional. In special relativity, the link is more subtle and results in a non-additive form (4) for the " summation" of velocities when velocity is measured by length/time. Although the mathematical forms differ, its role is not unlike the p-additive form that arose in the two behavioral examples. One feature of both examples is that the members of each pair of representations are ratio scales. Velocity and rapidity in the case of relativity are each ratio scales, as are U and $V = \mathrm{sgn}(\delta)\ln(1 + \delta U(x))$ in utility theory. The underlying sources of ratio-scale uniqueness, however, differ. For utility, the ratio quality of V derives from properties of $\oplus$ whereas that of U arises because of properties of $\circ_p$. For velocity, the ratio character of rapidity results from properties of $\circ$ whereas that of velocity arises from properties of the distance-time conjoint structure.

Our cases, especially in psychophysics, have the added wrinkle that the unknown functions Ψ and W relate to, respectively, physical measures of intensity and numerical proportions. That goes beyond the physical parallel and seems to be very incompletely understood, especially for W.

Of course, this work is only a drop in the bucket. Much that we would like to measure we don't know how to in fully satisfactory ways. So far as I know, modeling of this type has proved to be of little help in areas, such as ability testing, of interest to members of the psychometric community. It could become of interest were we to figure out how to treat intellectual abilities as a human " dimensional constant" relating cognitive activities in a measurable way. The fact that there are a number of results correlating performance on (basic) cognitive tests such as Raven matrices, Saul Sternberg tasks, response speed, etc. to " intelligence" (e.g., as measured by standardized tests) suggests that it might be feasible to do so were we to attend explicitly to that challenge. It would be a great accomplishment to provide improved behavioral foundations — axioms — for the ways that abilities are currently measured. I have not seen how to do that.

References

Aczél, J. (1966). *Lectures on Functional Equations and their Applications.* New York: Academic Press.

Aczél, J. (1987). *A Short Course on Functional Equations Based on Applications to the Behavioral and Social Sciences.* Dordrecht-Boston-Lancaster-Tokyo: Reidel-Kluwer.

Aczél J. (1997). Bisymmetric and consistent aggregation: Historical review and recent results. In: A.A.J. Marley (Ed.), *Choice, Decision, and Measurement: Essays in honor of R. Duncan Luce* (pp. 225–233). Mahwah, NJ: Erlbaum.

Aczél, J., & Daróczy, Z. (1975). *On Measures of Information and Their Characterizations.* New York: Academic Press.

Aczél, J., Falmagne, J.-C., & Luce, R.D. (2000). Functional equations in the behavioral sciences. *Mathematica Japonica, 52*, 469–512.

Aczél, J.,& Kuczma, M. (1991). Generalizations of a 'folk-theorem' on simple functional equations in a single variable. *Results in Mathematics, 19*, 5–21.

Aczél, J., Maksa, Gy., & Páles, Z. (1999). Solution to a functional equation arising from different ways of measuring utility. *Journal of Mathematical Analysis and Applications, 233*, 740–748.

Aczél, J., Luce, R.D., & Marley, A.A.J. (2003). A functional equation arising from simultaneous utility representations. *Results in Mathematics*, *43*, 193–197.

Beck, J., & Shaw, W.A. (1965). Magnitude of the standard, numerical value of the standard, and stimulus spacing in the estimation of loudness. *Perceptual and Motor Skills, 21*, 151–156.

Birnbaum, M.H. (1997). Violations of monotonicity in judgment and decision making. In A.A.J. Marley (Ed.), *Choice, Decision, and Measurement: Essays in Honor of R. Duncan Luce* (pp.73–100). Mahwah, NJ: Erlbaum.

Birnbaum, M.H. (1999). Paradoxes of Allais, stochastic dominance, and decision weights. In J. Shanteau, B.A. Mellers, D.A. Schum (Eds.) *Decision Science and Technology: Reflections on the Contributions of Ward Edwards* (pp. 27–52). Norwell, MA: Kluwer Academic Publishers.

Choquet, G. (1953). Theory of capacities. *Annales Inst. Fourier* 5, 131–295.

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

Doignon, J.-P., & Falmagne, J.-C. (1999). *Knowledge Spaces*. Berlin: Springer-Verlag.

Ellermeier, W., & Faulhammer, G. (2000). Empirical evaluation of axioms fundamental to Stevens' ratio-scaling approach. I. Loudness production. *Perception & Psychophysics, 62,* 1505–1511.

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

Falmagne, J.-C., Iverson, G., & Marcovici, S. (1979). Binaural " loudness" summation: Probabilistic theory and data. *Psychological Review, 86*, 25–43.

Gigerenzer, G., & Strube, G. (1983). Are there limits to binaural additivity of loudness? *Journal of Experimental Psychology: Human Perception and Performance, 9,* 126–136.

Gilboa, I. (1987). Expected utility with purely subjective non-additive probabilities. *Journal of Mathematical Economics, 16,* 65–88.

Gonzalez, R., & Wu, G. (1998). On the shape of the probability weighting function. *Cognitive Psychology, 38,* 129–166.

Hellman, R.P., & Zwislocki, J. (1961). Some factors affecting the estimation of loudness. *Journal of the Acoustical Society of America, 33,* 687–694.

Helmholtz, H.v. (1887). Zählen und Messen erkenntnis-theoretisch betrachet. *Philosophische Aufsätz Eduard Zeller gewidmet,* Leipzig. Reprinted in *Gesammelte Abhandl.,*vol. 3, 1895. English translation by C.L. Bryan, *Counting and Measuring.* Princeton, NJ: van Nostrand, 1930.

Hölder, O. (1901). Die Axiome der Quantität und die Lehre vom Mass. *Ber. Verh. Kgl. Sächsis. Ges. Wiss. Leipzig, Math.-Phys. Classe, 53*, 1–64.

Iverson, G. Myung, J.I., & Karabatsos, G. (2005). Intransitivity of preferences revisited, manuscript.

Iverson, G., & Falmagne, J.-C. (1985). Statistical issues in measurement. *Mathematical Social Sciences, 10,* 131–153.

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

Karabatsos, G. (2005). An exchangeable multinomial model for testing deterministic axioms of decision and measurement. *Journal of Mathematical Psychology, 49,* 51–69.

Krantz, D.H. (1971). Integration of just-noticeable differences. *Journal of Mathematical Psychology, 8*, 591–599.

Krantz, D.H., Luce, R.D., Suppes, P., & Tversky, A. (1971). *Foundations of Measurement,* vol. I. San Diego: Academic Press.

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

Liu, L. (1995). *A Theory of Coarse Utility and its Application to Portfolio Analysis,* University of Kansas, Ph.D. dissertation.

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

Luce, R.D. (1959). *Individual Choice Behavior: A Theoretical Analysis*, New York: Wiley.

Luce, R.D. (1990). Rational versus plausible accounting equivalences in preference judgments. *Psychological Science, 1*, 225–234.

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

Luce, R.D. (1998). Coalescing, event commutativity, and theories of utility, *Journal of Risk and Uncertainty,16*, 87–114. Errata: 1999, *18*, 99.

Luce, R.D. (2000). *Utility of Gains and Losses: Measurement-Theoretical and Experimental Approaches*. Mahwah, NJ: Erlbaum. Errata: see Luce's web page at http:www.socsci.uci.edu.

Luce, R.D. (2001). Reduction invariance and Prelec's weighting functions. *Journal of Mathematical Psychology, 45,* 167–179.

Luce, R.D. (2002). A psychophysical theory of intensity proportions, joint presentations, and matches. *Psychological Review,109*, 520–532.

Luce, R.D. (2003). Increasing increment generalizations of rank-dependent theories, *Theory and Decision, 55,* 87–146.

Luce, R.D. (2004). Symmetric and asymmetric matching of joint presentations. *Psychological Review, 111,* 446–454.

Luce, R.D., & Edwards, W. (1958). The derivation of subjective scales from just noticeable differences. *Psychological Review, 65*, 222–237.

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., Krantz, D.H., Suppes, P., & Tversky, A. (1990). *Foundations of Measurement,* Vol. III. San Diego: Academic Press.

Luce, R.D., & Marley, A.A.J. (2000). Elements of chance. *Theory and Decision, 49*, 97–126.

Luce, R.D., & Marley, A.A.J. (2005). Ranked additive utility representations of gambles: Old and new axiomatizations. *Journal of Risk and Uncertainty, 30*, 21–62.

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

Lundberg, A. (2004). Variants of the distributivity equation arising in theories of utility and psychophysics. *Aequationes Mathematicae,* in press.

Maksa, G. (1999). Solution of generalized bisymmetry type equations without surjectivity assumptions. *Aequationes Mathematicae, 57*, 50–74.

Marley, A.A.J., & Luce, R.D. (2001). Ranked-weighted utility and qualitative convolution. *Journal of Risk and Uncertainty, 23*, 135–163.

Marley, A.A.J., & Luce, R.D. (2005). Independence properties vis-à-vis several utility representations, *Theory and Decision*, in press.

Meginniss, J.R. (1976). A new class of symmetric utility rules for gambles, subjective marginal probability functions, and a generalized Bayes' rule. In *Proceedings of the American Statistical Association, Business and Economic Statistics Section* (pp. 471–476).

Myung, J.I., Karabatsos, G., & Iverson, G.J. (2005). A Bayesian approach to testing decision making axioms. *Journal of Mathematical Psychology,* in press.

Narens, L. (1996). A theory of ratio magnitude estimation. *Journal of Mathematical Psychology,* 40, 109–129.

Narens, L. (2002). *Theories of Meaningfulness.* Mahwah, NJ: Erlbaum.

Ng, C.T., Luce, R.D., & Aczél, J. (2002). Functional characterizations of basic properties of utility representations. *Monatshefte für Mathematik, 135,* 305–319.

Prelec, D. (1998). The probability weighting function. *Econometrica, 66*, 497–527.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organization,3*, 323–343.

Quiggin, J. (1993).*Generalized Expected Utility Theory: The Rank-Dependent Model.* Boston: Kluwer Academic Publishers.

Ramsey, F.P. (1931). *The Foundations of Mathematics and Other Logical Essays.* New York: Harcourt, Brace & Co. See Ch. VII. Reprinted in H.E. Kyburg, & H.E. Smokler (Eds.) (1964). *Studies in Subjective Probability* (pp. 61–92.) New York: Wiley.

Savage, L.J. (1954). *The Foundations of Statistics.* New York: Wiley.

Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica, 57*, 571–587.

Schneider, B. (1988) The additivity of loudness across critical bands: a conjoint measurement approach. *Perception & Psychophysics, 43,* 211–222.

Schneider, B., Parker, S., Ostrosky, D., Stein, D., & Kanow, G. (1974). A scale for the psychological magnitude of numbers. *Perception & Psychophysics*, *16*, 43–46.

Starmer, C., & Sugden, R. (1993). Testing for juxtaposition and event-splitting effects. *Journal of Risk and Uncertainty, 6,* 235–254.

Steingrimsson, R. (2002).*Contributions to Measuring Three Psychophysical Attributes: Testing Behavioral Axioms for Loudness, Response Time as an Independent Variable, and Attentional Intensity.* Psychology Ph.D. Dissertation, U.C. Irvine.

Steingrimsson, R., & Luce, R.D. (2005a). Empirical evaluation of a model of global psychophysical judgments: I. Behavioral properties of summations and productions. *Journal of Mathematical Psychology*, in press.

Steingrimsson, R., & Luce, R.D. (2005b). Empirical evaluation of a model of global psychophysical judgments: I. Behavioral properties linking summations and productions. *Journal of Mathematical Psychology*, in press.

Steingrimsson, R., & Luce, R.D. (2005c). Empirical evaluation of a model of global psychophysical judgments: III. Form for the Psychophysical Function and Perceptual Filtering. Submitted.

Steingrimsson, R., & Luce, R.D. (2005d). Empirical evaluation of a model of global psychophysical judgments: IV. Forms for the Weighting Function. In preparation.

Stevens, S.S. (1975). *Psychophysics: Introduction to its Perceptual, Neural, and Social Prospects,* New York: Wiley.

Tversky, A. (1969). Intransitivity of preferences. *Psychological Review, 76,* 31–48.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty,5,* 297–323.

von Neumann, J., & Morgenstern, O. (1947). *Theory of Games and Economic Behavior*, (2nd Ed). Princeton, NJ: Princeton University Press.

Wakker, P.P. (1991). Additive representations on rank-ordered sets. I. The algebraic approach. *Journal of Mathematical Psychology,35*, 501–531.

Wu, G., & Markle, A.B. (2005). An empirical test of gain-loss separability in prospect theory, under review.

Zimmer, K. (2005). Examining the validity of numerical ratios in loudness fractionation. *Perception & Psychophysics*, in press.

Zimmer, K., & Baumann, O. (2003a). Direct scaling of sensations: Are subjects able to produce consistent, and meaningful, numerical ratios? In *Proceedings of the First ISCA Tutorial & Research Workshop on Auditory Quality of Systems,* (pp. 85–90).

Zimmer, K., & Baumann, O. (2003b). Validity of Luce's transformation function relating numerals and numbers in loudness fractionation. Abstracts for Fechner Day 2003 (pp. 343–348).

Manuscript received 13 SEP 2004
Final version received 1 MAR 2005