General recognition theory and methodology for dimensional independence on simple cognitive manifolds

4 authors:

James T Townsend
Indiana University Bloomington
228 PUBLICATIONS 9,250 CITATIONS

Janet Aisbett
University of Newcastle
91 PUBLICATIONS 777 CITATIONS

Jerome R. Busemeyer
Indiana University Bloomington
218 PUBLICATIONS 10,561 CITATIONS

Amir Assadi
University of Wisconsin–Madison
118 PUBLICATIONS 667 CITATIONS

Townsend, J. T., Aisbett, J., Busemeyer, J., &
Assadi, A. (2005). General recognition theory and
Methodology for Dimensional Independence on Simple
Cognitive Manifolds. In H. Colonius & E. N. Dzhafarov
(Eds.), *Measurement and Representation of Senstions: Recent
Progress in Psychophysical Theory*. Erlbaum Associates

General Recognition Theory and Methodology for Dimensional Independence on Simple Cognitive Manifolds

James T. Townsend
Indiana University

Janet Aisbett
The University of Newcastle

Jerome Busemeyer
Indiana University

Amir Assadi
University of Wisconsin

1. INTRODUCTION

This paper concerns the issue of whether and how perceptual dimensions interact from a differential geometric standpoint. Earlier efforts in this direction initiated depiction of percepts, viewed 'in the large,' that is, where the percepts are sufficiently separated that discrimination is virtually perfect (Townsend & Spencer-Smith, 2004). Hence, the percepts can be treated in that framework as deterministic. In the present investigation, we take up the same type of question when discrimination is imperfect due to noise or closeness of the stimuli. This is accomplished as a generalization of General Recognition Theory, hereafter abbreviated as GRT (Ashby & Townsend, 1986; see also, Ashby, 1992, Maddox, 1992, and Thomas, 1999, 2003). The original GRT dealt with percepts as points lying in an orthogonally coordinated space associated with distinct densities associated with the stimulus set.

We have found that the present explorations in non-Euclidean spaces tend to bring up novel aspects of relationships between stimulus dimensions and perceptual dimensions that were not immediately evident in the usual Euclidean milieu. Thus, in addition to providing some "first-order" extensions of GRT to elementary manifolds, we view this paper as propadeutic to several potential new lines of inquiry.

Until we are deeper within the paper, it may seem that we are studying systems devoid of response properties. However, within the early deterministic framework it is assumed, as in Townsend & Thomas (1993) or Townsend, Solomon, Wenger, & Spencer-Smith (2001), that standard psychophysical responses are acquired. Later, as we enter the hard-to-discriminate, and therefore probabilistic milieu, we will be discussing an identification paradigm (readily generalizable to categorization) where an observation point in a manifold leads inexorably to a response. A word about our conception of perceptual entities seems in order. We believe that in fact most objects in physical stimulus space are things like shapes, sounds, etc. that lie in infinite dimensional spaces. We also think that people can process these as complex percepts sometimes homeomorphic or even diffeomorphic to the physical stimulus. For instance, think of perceiving or internally imaging a friend's face. Some of our previous papers have begun to deal with this aspect of perception (see, e.g., Townsend, Solomon, Wenger, & Spencer-Smith, 2001; Townsend & Thomas, 1993).

However, it is also evident that somehow perceiving organisms are able to filter out dimensions (e.g., brightness or color) and categorical entities (e.g., stripes on a tiger, or the orthographic RED, independent, say, of print color) from the original object. Further, most of psychology in general and psychophysics in particular treats perceptual stimuli as points in a relatively simple space, usually a space with orthogonal coordinates and often with a Euclidean or sometimes a Minkowski power metric. Other possibilities, such as tree metrics (e.g., Tversky, 1977) are occasionally considered too; Dzhafarov and Colonius (1999), build a theory based on Finsler and more general metrics that derive from discriminability functions.

We focus on points from two continua, assuming either elementary stimulus presentation (e.g., sound intensity and sound frequency), or that dimensional reduction through filtering (e.g., attentional) has already taken place (e.g., as in abstracting the color from an object). Thus, we shall treat the problem of stimulus continua, with a finite number of response assignments (*i.e.*, a type of category). However, many of the later statements concerning common experimental paradigms will be true for stimuli as discrete categories (e.g., letters of the alphabet, words, etc.).

The format of our study will be somewhat tutorial in form with occasional references to instructive volumes, since many readers may not be conversant with differential geometry.[1] Although there are many terms which involve several modifiers, we will give acronyms for very few of them to lessen opportunities for confusion. Also, we will drop some of the modifiers when it is transparent to which theoretical object we are referring.

Following Townsend & Spencer-Smith (2004) we take the concept of a

[1] Probability theory and stochastic processes,and, especially for psychometricians, linear algebra, remain the modal mathematical education for social scientists. However, this state of affairs is changing with increasing influences of many areas of applied and pure mathematics into social and especially cognitive science.

coordinate patch or simply "patch" in differential geometry as forming an appropriate level of description for a beginning treatment of dimensional independence on manifolds. First, we require a definition of a stimulus domain, which for simplicity we restrict to two dimensions. It is explicit that all possible pairs of stimulus dimension values be potentially available for perception.

Definition 1 *A two-dimensional stimulus domain is an open set in the plane: $D = U \times V \subseteq \Re^2$, where U and V may be finite or infinite open intervals, $\Re$ is the real line, and $\Re^2$ is the plane with the Euclidean metric.*

We should note that, $\Re^n$, $n = 1, 2$ etc., will not be armed with the Euclidean metric unless so stated, as in Definition 1. $\Re^n$ unadorned with a metric denotes the usual set of orthogonal coordinates. Note that the presence or absence of physical units is left open. The next definition adapts the notion of a proper coordinate patch that is also a diffeomorphism (that is, a map between manifolds that has an inverse such that both the map and the inverse are smooth– see, e.g., more or less in order of increasing sophistication: O'Neill, 1966; Boothby, 2003; Kobayashi & Nomizu, 1991, 1969; Lang, 1985). Our assumptions in this treatise will be rather tight and may in some instances be considerably loosened. For instance, the diffeomorphic assumption in Definition 2 below is fairly demanding. This constraint helps keep matters relatively straightforward, but these aspects should be generalizable in the future (e.g., as the case demands, to immersions, submersions, or non-smooth mappings–see Summary and Discussion).

Here, we investigate cognitive modeling using the simplest type of differential manifold, namely spaces that can be represented as orientable surfaces. One way that such spaces might be produced in the brain is by mapping, say, a physical object of n dimensions into a space of $n + k$ dimensions. For instance, a pair of fundamental frequencies w_1, w_2, as a stimulus may produce sounds involving the overtones of w_1 and w_2 through non-linearities. There may be other psychological dimensions produced by the same two stimulus dimensions. Thus, even a finite distribution of colors on a piece of art or in a room can stimulate distinct values in aesthetics scales simply by apparently minor rearrangement of the color placements.

Alternatively, it could be that any such psychological manifold is more chimerical. For example, if the psychological metric, is non-Euclidean, say Riemannian, then it could be associated with a certain kind of manifold even though the manifold as a surface might be rather a 'second-thought' construction. In any event, we shall keep the material simple by invoking only one extra dimension, assuming that the manifold is two dimensional and lies in three dimensional space. Generalization to higher dimensional embeddings (e.g., a two dimensional surface in a k-dimensional Euclidean space) is easy, if more tedious.

Definition 2 *A proper perceptual patch $X : D \to M \subset \Re^3$ is a one-to-one, diffeomorphic map of an open set D of $\Re^2$ into $\Re^3$, and onto M. Thus we specify X by a coordinate function $X(u,v) = (x(u,v), y(u,v), z(u,v))$ in $\Re^3$. Let us call the psychological coordinates $\phi \in \Phi$ and $\psi \in \Psi$, respectively, where $\Phi \times \Psi$ is a subset of $\Re^2$. In general, ϕ and ψ will be functions of (u, v) but neither is necessarily identical to any of x, y, or z. If we focus on the coordinates produced by holding v or u constant, respectively, and vary the other, the proper perceptual patch X is interpreted as a parameterization of D.*

There is an induced diffeomorphism $Y : \Pi \to M$ for $\Pi = \{\phi(u,v) : (u,v) \in D\} \times \{\psi(u,v) : (u,v) \in D\} \subset \Re^2$which provides the coordinate chart for the manifold M. Through this map and its inverse, the psychological experiences being studied can be considered as either points in M or as points in the set Π. Moreover, we will occasionally require the map $Z : D \to \Pi$ defined by $Z = Y^{-1}(X)$. In cases where a psychological coordinate is identical to a coordinate in the embedding space or a monotonic function of it (e.g., as in the patch $x = \phi(u,v), y = \psi(u,v), z = \gamma(u,v)$), the coordinate is directly 'readable' from M itself.

Figure 1 illustrates the general relationship between the stimulus domain D, the psychological space M and the psychological coordinates. In this example, the manifold M is the modified geographical upper half-sphere of radius 1. A perceptual patch might be defined as the map

$$X(u,v) = (cos(u)cos(v), sin(u)cos(v), sin(v)),$$

with $U = (-\pi/2, \pi/2)$ and $V = (-\pi/2, \pi/2)$.

Insert Figure 1 about here

As mentioned above, deterministic separability has been introduced by Townsend & Spencer-Smith (2004). We present a slightly modified form here. There are two fundamental classes of deterministic separability in their scheme, one called domain-range separability. This notion captures the traditional emphasis on whether perceptual dimensions interact or not, as functions of specified physical dimensions. For instance, suppose $M = \Re^2$, and let $X(u,v) = (\phi(u), \psi(v))$; this is a trivial example of domain range separability. In contrast, the fact that equi-loudness contours in sound perception are functions of both intensity and frequency implies that loudness is not deterministically a separable function of intensity.

The second class of separability is called range-alone separability and this concept pertains to properties that adhere to M itself, without reference to the map X. An example of a property associated primarily with M itself is local orthogonality; that is, the tangent vectors to the coordinates ϕ, ψ are always orthogonal at every point of M (see Townsend & Spencer-Smith, 2004). We shall mainly be concerned with domain-range separabilty in this study, but range-alone separability will be expanded on below.

It is worth remarking that some of our developments require a metric and some do not, rendering them more general objects. For instance, local orthogonality requires an inner product in the tangent space, which then readily produces a metric. However, the following definition, as well as Definitions 1 and 2 do not necessitate a metric.

Definition 3 *Deterministic domain-range, parameterization, perceptual separability (DDR-PPS) of two stimulus dimensions is defined by the condition that the proper perceptual patch X be a parameterization of $D = U \times V$, with the perceptual coordinates ϕ, ψ corresponding respectively to functions of u, v.*

Thus, irrespective of how the nervous system does it, we can think of the map X as a composition of maps from U, V to Φ, Ψ, and then from there onto M via Y. That is, $X(u,v) = Y(\phi(u), \psi(v)) = m \in M$. (See Figure 2.)

Insert Figure 2 about here

In Townsend & Spencer-Smith (2004), the designation of domain-range separability was given to the case where $M \subset \Re^3$ and $x = x(u) = \phi(u)$ and $y = y(v) = \psi(v)$, with z being an arbitrary function of u, v. We now think parameterization separability should, in contrast to the earlier treatment, be featured, since this earlier situation, now called "embedding coordinate perceptual separability," is obviously a special instance of parameterization separability (Definition 3). Proposition 1, whose proof is obvious, keeps this new characterization in order. Figure 3 illustrates this concept.

Insert Figure 3 about here

Proposition 1 *Deterministic domain-range embedding coordinate perceptual separability is a special case DDR-PPS of Definition 3.*

Is there any sense to thinking of any kind of separability notion when we focus entirely on the image space and the psychological coordinates therein, ignoring the $D \to M$ map? We have suggested that the answer to this query is "yes" (Townsend & Spencer-Smith, 2004). The idea is that even though the psychological dimensions may be intricate functions of the manipulated physical variables, there are some properties that are suggestive of separability. Local orthogonality, mentioned above, is one such option since whether or not the tangents to the psychological coordinates are at right angles does not depend on whether they are parameterizations of the physical coordinates, u, v. In another elemental kind of separability that depends only on the properties of the range space, Townsend & Thomas (1993) suggested that one simple type of deterministic perceptual separability might demand that the perceptual space can be expressed as $\{\phi(u); u \in U\} \times \{\psi(v); v \in V\} = \Phi \times \Psi$, that is, as the Cartesian product of the possible values of the psychological coordinates; regardless of whether ϕ, ψ are each functions only of u, v respectively. This observation leads to the following definition.

Definition 4 *Range Cartesian separability is defined by the condition that the perceptual coordinates consist of the Cartesian product $\Phi \times \Psi$.*

The upshot here is that there is a kind of deterministic non-separability when certain combinations of the psychological variables cannot occur. To see how range Cartesian separability can fail to hold even though the stimulus space is $D = U \times V$ and the specified coordinates are functions on D, consider the map $(\phi(u,v), \psi(u,v)) = (u\ln(v), ve^u)$, $u > 0$, $v > 1$. A level surface is created by solving for $v = e^{\phi_o/u}$ and then finding that $\psi = e^{\phi_o/u+u}$. We see that ψ is bounded below by $e^{2\sqrt{\phi o}}$ here, and hence the range is not equivalent to

$$\Phi \times \Psi \equiv \{\phi(u,v) : (u,v) \in D\} \times \{\psi(u,v) : (u,v) \in D\} = (0,\infty) \times (1,\infty).$$

The next proposition establishes the relationship between range-alone Cartesian separability and DDR-PPS.

Proposition 2 *(A) DDR-PPS implies that $\Pi = \Phi \times \Psi$ and therefore that range-alone Cartesian separability is in force.*

(B) Cartesian range-alone separabilty does not imply DDR-PPS.

Proof. A. Under these conditions, each value of ϕ corresponds to a specific v-trajectory in M, say $\phi(u) = X(u_\phi, V)$ and similarly each $\psi(v) = X(U, v_\psi)$. Then clearly, $\Pi = \cup_\phi\ X(u_\phi, V) \times \cup_\psi\ X(U, v_\psi) = \Phi \times \Psi$.

B. As a counterexample, consider the maps $\phi(u,v) = vu^a$ and $\psi(u,v) = uv^b$ where u, v, a, b are positive on the real line, a, b fixed parameters, say, $a, b > 1$. In this case, ϕ is primarily a function of u, for instance, loudness as a function primarily of intensity, secondarily of v = frequency; and vice versa for ψ = pitch. One can then solve for a given level surface, $\phi = \phi_o$, $v = u^{-a}\phi_o$ and find that $\psi = u\ (u^{-a}\phi_o)^b$ so that for every $\phi_o \in \Re^+$ all legitimate values of ψ are produced. Similar results obviously occur in the reverse function, and Cartesian range-alone separability occurs, since $\{(\phi(u,v),\psi(u,v))\} = \Phi \times \Psi$. But, observe that coordinate domain-range perceptual separability does *not* hold. ■

In principle, an investigator can empirically determine whether DDR-PPS holds through various psychophysical methods, such as variations on Stevens' direct scaling methods. Equi-loudness contours are an example. However, even if DDR-PPS is valid, it could be that probabilistic mechanisms, including correlation of noise in the two dimensional channels, could invalidate separability at the local, tough-to-discriminate stochastic level. We turn to this question next.

2. PROBABILISTIC DOMAIN-RANGE PERCEPTUAL SEPARABILITY AND PERCEPTUAL INDEPENDENCE

In general, we refrain at present from specifying the source of probabilism—it may be from the outer environment, internal noise, etc. Any such specification will depend on more micro-level or process-oriented interpretations. As far as our technical considerations presently are concerned, it matters little, although the technical difficulty can vary. Next, we will have a couple of evident propositions.

The earlier Ashby & Townsend (1986) work, as in the present, is descriptive to a point, in the sense that we are deliberating at a geometric level, without direct reference to underlying time-oriented dynamics. The reader is referred to MacMillan and Creelman (2005) for a survey of the current state of multivariate detection theory and methods in Cartesian coordinates. We assume that the perceptual activity and decisions in, say, a 2x2 factorial design take place in a way equivalent to some type of decision rule on M itself. For instance if, as we shall postulate, M is partitioned into a set of four mutually exclusive regions, then the probability measures over those regions will determine the stochastic structures of interest in our endeavor. On the other hand, if the decision occurs after filtering operations are performed that map the psychological point to a different space, such as the plane, then it could be that the topological or geometric aspects of M are inherited by this tertiary space, or not.

Next, we require the probabilistic structure obtained by placing probability densities on M. This approach stipulates that associated with each stimulus $(u, v) \in D$ is a probability density function on M given point-wise by $p(x, y, z|u, v)$ with the understanding that any such point (x, y, z) resides on M via the map X. Also there is a function $p(\phi, \psi|u, v)$ that depends conjointly on (u, v). We abuse the notation in using the same symbol " p" for both functions, allowing that there is a 1-1 relation between$(x, y, z) \in M$ and $(\phi, \psi) \in \Pi$. Indeed, the distribution on M can also be pulled back to D. In any event, we can think of p as being a member of a family depending on u, v, each pair (u, v) of which determines a unique probability density function. As discussed more below and as suggested by Figure 1, we can discuss the psychological coordinates as they exist in planar Π, or as they exist through a mapping Y from Π onto M. This facilitates the following definition.

Definition 5 *Every setting of u, v engenders a joint probability distribution on ϕ, ψ in Π, that is $p(\phi, \psi|u, v)$. Since a setting of u, v (say $u = u_o$, $v = v_o$) leads, in the deterministic sense, to values ϕ_o, $\psi_o$of ϕ,ψ, we can also express $p(\phi, \psi|u_o, v_o) = p(\phi, \psi|\phi_o, \psi_o)$.*

It is an interesting sidelight of this picture that even though a number of parameters dependent on u could be defined on Π, the parameter space will still be one dimensional. For instance, suppose the probability density on Π is normal with the mean and variance being functions of u. Then the mean and variance are clearly functionally dependent due to their dependence on u. Perhaps the relationships of stochastic model parameters and their dependence may be of aid in the geometrical side of matters (but see, e.g., M. Levine, 2003, for a discussion of the complexity of relationships between quantity of latent variables and dimensionality). The succeeding definition establishes a potential linkage between the stochastic and deterministic milieus but it will not be seen much in the following material.

Definition 6 *The distribution on M is said to satisfy convergent determinism if there exists a converging sequence of p as a function of a magnitude parameter I such that $p(\phi, \psi|u_o, v_o; I) \to p(\phi_o, \psi_o|u_o, v_o; I)$ as I grows large, where (ϕ_o, ψ_o) is the image of (u_o, v_o).*

Convergent determinism then ensures a connection between the stochastic situation and the high signal-to-noise ration situation when perception becomes deterministic. The parameter I can refer to any physical dimension(s) that increase discriminabiliity such as contrast, salience, etc. Yet, Definition 6 does not imply that reports, say, in magnitude estimation will be deterministic even though discrimination among the presented stimuli is high.

Next, we require an apparatus for integration in the manifold M. We have interest in the geometric aspects of spaces and hence an approach based on the exterior calculus of Grassman and Cartan (e.g., Kobayashi & Nomizu, 1969, 1991; see Lang, 1993 for a discussion of the relationship of measures to exterior forms on manifolds) seems called for. The latter immediately and elegantly provides for the institution of the metric into the integral.

Thus, we define a volume form (which is always possible if M is orientable, which we take as a postulate) or in our elementary case, an area form, which we denote by $\mathrm{d}M$.[2] The idea is that $\mathrm{d}M$ is an operator (technically, a covariant tensor of order 2) which operates on a pair of vectors (contravariant tensors) to measure an infinitesimal area. In a Riemannian manifold, this operator will be closely linked to the metric function $g_{ij}(m)$ defined at a point m of M, which determines scale.

Because $m = X(u,v)$ for some (u,v) in D, we can also write $g_{ij}(u,v)$. In addition, the prime vectors to employ in the computation are, in fact, the basis vectors of a tangent space fixed at point m. These basis vectors yield a parallelogram whose area is the local infinitesimal measure for which we are looking.

In the event that the metric is inherited from the map X, which would naturally be of interest to us, it is known as a Riemannian metric or first fundamental form. It turns out that we can then express

$$\mathrm{d}M(X_u, X_v) = \sqrt{g_{11}g_{22} - g_{12}^2}\mathrm{d}u\mathrm{d}v,$$

where X_u and X_v are the partial derivatives of X and $g_{ij} = \langle X_i, X_j \rangle$, the inner product of the partial derivatives with regard to i or j where either is equal to u or v. Note that we have successfully "pulled back" the computation from M to $D = U \times V$. The role of the factor $\sqrt{g_{11}g_{22} - g_{12}^2}$ is to expand or shrink various part of M relative to D.

In our case, though, we are more interested in ϕ, ψ as the psychological coordinates rather than the stimulus values u, v. We don't know, of course, if something like a direct map from stimulus space D to psychological co-

[2] Using forms for integration necessitates keeping track of the orientation of M, so that the oriented volume, $\mathrm{Vol}^o(\mathrm{v}_1,\mathrm{v}_2,...,\mathrm{v}_n)$ where the v's span the parallepiped, is + $\mathrm{Vol}(\mathrm{v}_1,\mathrm{v}_2,...,\mathrm{v}_n)$ if $\mathrm{DET}(\mathrm{v}_1,\mathrm{v}_2,...,\mathrm{v}_n) > 0$ and - $\mathrm{Vol}(\mathrm{v}_1,\mathrm{v}_2,...,\mathrm{v}_n)$ if $\mathrm{DET}(\mathrm{v}_1,\mathrm{v}_2,...,\mathrm{v}_n) < 0$. Because of this facet, some writers call such forms, "pseudo forms". It is assumed in our text that the orientation has been accounted for and the form properly "signed".

ordinate space Π, for instance via Z, occurs or if, more likely, the latter are 'picked off' or interpreted from M, as in $(\phi,\psi) = Y^{-1}X(u,v) = Z^{-1}$, using the notation from Figures 1-3. In any event, we can express the partial derivatives of Y as

$$Y_\phi = X_u \frac{\partial u}{\partial \phi} + X_v \frac{\partial v}{\partial \phi}$$

and

$$Y_\psi = X_u \frac{\partial u}{\partial \psi} + X_v \frac{\partial v}{\partial \psi}.$$

Having restructured the coordinates in terms of ϕ, ψ, we shall now refer the developments in the previous paragraph to the psychological coordinates. Thus, the stimulus set D for example is replaced by Π in the ϕ, ψ plane.

Regardless of the specific metric, the area of the region B on M can then be found from computing

$$\int_{B'} \mathrm{d}M = \int_{B'} \mathrm{d}M(Y_\phi, Y_\psi)$$

where B' is the area corresponding to B in the planar set $\Phi \times \Psi$. The terms Y_ϕ and Y_ψ refer to the standard coordinate basis vectors in the tangent space of M, $\mathrm{TM}_{(m)}$ at an arbitrary point $m = Y(\phi,\psi)$ of M. They correspond to (1,0) and (0,1) in $\Re^2$ in the plane and in fact are the push forward of these basis vectors to M. Finally, the density function, p, on M can simply be instated in the integrals above analogously to the situation in $\Re^n$.

The earlier Ashby & Townsend (1986) concept of *perceptual separability* can be viewed as a probabilistic realization of a 2-dimensional domain-range coordinate separability (also see Townsend, Hu and Ashby, 1981; Townsend, Hu & Evans, 1984; Townsend, Hu & Kadlec, 1988). Observe again that the fact that the results or models can be plotted in the plane does not imply flatness unless a flat metric is imposed.

How should we formulate the probabilistic version of domain-range separability on M? We can take a hint from the earlier construction in ordinary rectangular coordinated space by Ashby & Townsend (1986). They considered the 2x2 table of settings of the two stimulus dimensions; these in effect form (i.e., map to) a parameter space. It was assumed that the stimulus dimensions were associated with psychological variables, although allowing that each of the latter variables might, in effect, be functions of both stimulus variables. The psychological space was assumed to be 2-dimensional. Since typically, only a 2x2 performance matrix containing the relative frequencies of responses to the four stimuli are gathered from an experiment, the location of the probability distributions in this space have to be estimated in a sense from the data. This procedure is analogous to how one carries out Thurstone scaling, or signal detection or discrimination in 1-dimensional settings.

In any event, each cell in the stimulus table then was associated with a hypothetical joint probability distribution in $\Re^2$. Ashby & Townsend (1986) next defined probabilistic perceptual separabilty of the psychological dimension (e.g., ϕ) paired with stimulus dimension u, as the requirement that the marginal distribution on that dimension (ϕ), after integrating out the psychological dimension associated with stimulus variable v (e.g., ψ), would be constant over the other stimulus settings of variable v. Notice that this definition does not imply stochastic independence, or vice versa.

To initiate the proceedings, we may simply reiterate the Ashby & Townsend (1986) definition with the present notation.

Definition 7 *Probabilistic domain-range parameterization perceptual separability (PDR-PPS) of ϕ against ψ holds if and only if, for each physical stimulus value u, the marginal distribution for ϕ, namely $p(\phi|u, v)$, is invariant across settings of v. If this is the case we can write $p(\phi|u,v) = p(\phi|u)$. Separability of ψ against ϕ likewise holds when $p(\psi|u, v)$ is independent of u for each v and ψ.*

Identifying coordinate space with psychological space M, probabilistic domain-range perceptual separability of ϕ from ψ is seen to be equivalent to the condition that

$$\int_V p(x,y,z|u,v)\mathrm{d}M = \int_\Psi p(\phi,\psi|u,v)\mathrm{d}\psi = p(\phi|u,v) = p(\phi|u)$$

where the first term may be interpreted as the integral over $\{X^{-1}Y(\phi, \psi)$: $\psi \in \Psi$ and $(\phi, \psi) \in \Pi\}$ in the pull back to $U \times V$. Separability of ψ from ϕ on M is defined in the same manner.

As an example in the present context, Definition 7 suggests that a parameter set associated with one psychological dimension, say ϕ, is a function of only one physical dimension u and not the other, v. Hence, for instance, if the mean and variance were established to be the defining parameters for ϕ, then they would be functions only of u and not v. Integrating over ψ would then erase the mean and variance associated with v and leave a distribution on ϕ depending only on Φ (and implicitly on u and not v).

One way to think of what is happening in separability is that in perceiving the 'coordinates' in M in a separable fashion, the observer is performing in a way that is equivalent to mapping M into the $\Phi \times \Psi$ plane.

Proposition 3 *PDR-PPS holds if DDR-PPS holds and, for any stimulus value (u_s, v_s) and any fixed $u \in U$, the 'marginal' distribution obtained by* *integrating* $p(x, y, z|u_s, v_s)$ *over* $\{(x, y, z)|X^{-1}(x, y, z) \in (u, V)\}$ *is independent of* v_s.

Proof. Simply re-form Z^{-1}: $\Phi \times \Psi \to D$ as before, $Z^{-1}(\phi, \psi) = X^{-1}Y(\phi, \psi)$. Then $Z^{-1}(\phi(u), \psi(v)) = (u, v)$ and so $Z^{-1}(\phi, \psi) = X^{-1}(x, y, z)$. ■

Note that if the marginal condition in M stipulated in Proposition 3 is abrogated, then even if DDR-PPS is in force, PDR-PS can fail to hold.

To this point, we have the following set-up: We can investigate the properties of the stochastics on M via the plane $\Phi \times \Psi$. The geometric properties of M as expressed in the metric g_{ij} will be transferred there. Separability is defined as the invariance of a marginal distribution on either of the psychological dimensions when integrating out the other. These concepts are illustrated in the example in the penultimate section.

3. EVIDENCE ASSESSMENT

Somehow a decision must be made. Of course, the decision must be a function of the observations. As we learned from univariate signal detection theory, the decision plays an important role in performance. Within the literature on multivariate psychophysics and perception, there was scant consideration of this aspect of behavior. Ashby & Townsend (1986) showed that in some ways decisional structure played an even more critical role than stimulus structure in studying various types of perceptual independence. Thus, decisional structure could obscure perceptual independencies, reveal them authentically, or in rare cases, render an appearance of perceptual independence where there was none.

The observations in Π have to be mapped into the appropriate set of decisions. A natural intermediary is the *evidence space,* which may include a bias that is response oriented (e.g., Luce, 1963; Townsend, 1971) or stimulus oriented (Lappin, 1978; Nosofsky, 1992). Then a subsequent function on the biased evidence maps the latter into one of the finite set of decisions. We see more on evidence space below.

Ashby & Townsend (1986) found that an assumption of *decisional separability* is crucial in readily exposing the underlying perceptual interactions among dimensions. Decisional separability requires that an implicit or explicit decision based on a dimension be independent of the decision about the other dimension. In a space with orthogonal coordinates, this requirement implies that the decision boundaries are straight lines that are parallel to one axis and orthogonal to the other. However, orthogonality is not a necessary prerequisite as our new definition indicates.

Definition 8 *Decisional separability of ϕ from ψ is the condition that the criterion that ϕ exceed (or fail to exceed) a certain value ϕ_0, is defined by a boundary in M identified with the constant ϕ_0 trajectory (i.e., varying ψ while holding ϕ constant). So, the response variable on ϕ is a_1 if the observed ϕ is to the "left" of ϕ_0 and a_2 if it is to the "right" of ϕ_0. Decisional separability of ψ from ϕ is defined analogously.(See Figure 4.)*

Insert Figure 4 about here

Thus, each decision bound is itself a value of the psychological variable and simultaneously a trajectory of the other variable holding the first constant. Patently, just as with perceptual separability, the notion is not a symmetric relation. Of course, in Π we immediately find that decisional separability is true if and only if the decision boundaries are appropriately parallel to the ϕ, ψ axes. Next, we see some relationships of decisional separability to certain other important notions.

Proposition 4 *Decisional separability does not require: (A) Cartesian range separability; (B) Any type of domain-range separability.*

Proof. A. Merely observe that selecting an arbitrary criterion on either Φ, Ψ does not depend on what values exist on the other dimension.

B. Clearly, the notion of decisional separability is independent of what transpires with respect to the map from D to Π. That is, decisional separability is a range type of property. ■

How might such a property be generated? We must now consider how the percepts (ϕ, ψ) relate to evidence regarding the response alternative. The evidence space records the evidence in favor of each alternative and permits a decision to be made on some comparison based on this evidence. The mapping of the psychological space into the evidence space assumes the role of the so-called discriminant function in pattern recognition theory (e.g., Nilsson, 1965; Townsend & Landon, 1982). Set the evidence map $e : M \to E_p$ where the range space is a subset of $\Re^4$ so that

$$e(m) = (e_{11}(m), e_{12}(m), e_{21}(m), e_{22}(m))$$

and e_{ij} maps an observation into $\Re$ as the strength of perceptual indication for response $a_i b_j$. Assume that the evidence map and its inverse are continuous.

In general, the evidence function can include learning and motivational biases; and in fact, biases based on either the responses (most common, e.g., as in Luce's recognition choice theory, 1963) or as noted above, on stimulus bias (e.g., Lappin, 1978; Nosofsky, 1992). To constrain the theoretical development, we adopt the MAX rule of decision, that is, response $= a_i b_j$ if and only if $e_{ij} = \max_{kl}(e_{kl})$. This decision rule covers an enormous number of decision possibilities including both optimal and non-optimal strategies (for some of these, see Townsend & Landon, 1982). It does not include certain probabilistic decision strategies such as probability matching.

Consider the pairs of functions $e_{22} - e_{12} = e_1$, $e_{22} - e_{21} = e_2$, $e_{22} - e_{11} = e_3$, $e_{12} - e_{21} = e_4$, $e_{12} - e_{11} = e_5$, $e_{21} - e_{11} = e_6$. Let

$$\begin{aligned} E_{22} &= \{m \in M : e_1(m) > 0, e_2(m) > 0, e_3(m) > 0\} \\ &= \cap_{i=1,2,3}\{m \in M : e_i(m) > 0\} \end{aligned}$$

that is, E_{22} is where e_{22} dominates the other functions so that response $a_2 b_2$ will be made in this region. And so on, for the other response alternatives.

Proposition 5 *(A) If $S = \{m \in M : e_1(m) = 0, e_j(m) \neq 0$ for $j \neq 1\}$ and $\mathrm{d}e_i/\mathrm{d}p \neq 0$, then S is a 1-dimensional closed submanifold between a region E_{22} and E_{12}, and similarly for other regions E_{ij}. (B) Each region E_{ij} is connected.*

Proof. A. This result is a generalization into differential manifolds of the inverse function theorem (see, e.g., Hirsch, 1976).

B. This result follows from the topological fact that a continuous image of a connected manifold is connected (e.g., Munkres, 1975). ■

As yet, we have no means of enforcing decisional separability. The next definition opens a path to this end.

Definition 9 *Discriminant decomposability on both dimensions obtains if there exist two independent, 1-dimensional discriminant functions α_i, β_j, $i, j = 1, 2$, where α_i is a function only of ϕ and β_j is a function only of ψ, such that $e_{ij} = \max_{kl}[e_{kl}]$ holds if and only if $\alpha_i = \max_k[\alpha_k]$ and $\beta_j = \max_k[\beta_k]$.*

Proposition 6 *If discriminant decomposability holds on both dimensions and the differences $\alpha = \alpha_2 - \alpha_1$ and $\beta = \beta_2 - \beta_1$ are monotonic functions of their respective arguments ϕ, ψ, then decisional separability will hold.*

Proof. We can simply compute $\alpha_2 - \alpha_1$ and $\beta_2 - \beta_1$ and partition M into regions where these differences are negative or positive. Further, under the stipulations, either difference will be 0 on a tie between the α or β values, and the ties partition $\Phi \times \Psi$ space into two disjoint sections apiece. They can both be 0 only once where the two straight and orthogonal lines cross. The lines of separation and the intersection point, being of dimension 1 or 0, must be of measure 0 relative to the underlying densities. ■

A natural example forms when α_i is the likelihood of alternative u_i and β_j is the likelihood of v_j and α_1/α_2 is monotonic decreasing and similarly for the β ratio. These likelihoods can be taken from the marginal densities, whether or not independence of ϕ, ψ holds.

The next assignment is to formulate analogues to the Ashby & Townsend (1986) observable implications of separability. In that paper, it was shown that under decisional separability, an earlier statistic, *marginal response invariance* (earlier employed in testing for feature independence in a test then called *across-stimulus invariance*; Townsend, Hu & Ashby, 1981; Townsend & Ashby, 1982; Townsend, Hu & Evans, 1984; Townsend, Hu & Kadlec, 1988) was equivalent to perceptual separability.

Let the probability of giving a response identified with level "i" of physical dimension U and level "j" of dimension V be $P(a_i, b_j|u_i, v_j)$. The usual complete factorial identification paradigm has two levels on each dimension and four responses overall corresponding to the four stimulus combinations of the 2x2 values. For generality, let $k\prime$ be the complement to k, i.e., if there are two levels and k =1 then $k\prime$ =2 and vice versa. Then, following Ashby & Townsend (1986), we can state the definition of marginal response invariance as follows.

Definition 10 *Marginal response invariance is defined as the satisfaction of the condition*

$$\begin{aligned} P(a_i|u_k, v) &= P(a_i, b_j|u_k, v) + P(a_i, b_{j'}|u_k, v) \\ &= P(a_i, b_j|u_k, v\prime) + P(a_i, b_{j'}|u_k, v\prime) = P(a_i|u_k, v\prime) \end{aligned}$$

for any v, $v\prime \in V$, *and likewise for pairs* u, $u\prime \in U$.

Note that indeed, the marginal probability of responding that u is at level "i" depends on u, but not on the level of v. Next comes a proposition that indicates the conditions under which PDR-PPS will be revealed by marginal response invariance. Naturally, the stipulation concerns decisional variables, since marginal response invariance involves response frequencies while PDR-PPS only involves the perceptual distributions themselves.

Proposition 7 *(A) If decisional separability holds, then PDR-PPS implies marginal response invariance. (B) If marginal response invariance holds everywhere in the presence of decisional separability, then PDR-PPS is implied. Thus, under decisional separability, perceptual separability, and marginal response invariance are basically equivalent.*

Proof. A. Assume that decisional separability holds for ϕ over ψ, and that two stimulus values of v have been specified. Decisional separability entails that the decision pertaining to an observed value ϕ depends only on which side of a criterion ϕ_0 it falls. Then, one integrates over the proper range of ψ (*i.e.*, summing over the response probability for reporting either b_j or its alternative) and the "i" side of the division of the space by ϕ_0 in order to calculate the marginal probability of a_i (i.e., level "i" for dimension ϕ) under the two stimulus settings. Marginal response invariance states that for two distinct settings of v, the marginal probability of reporting a_i will be the same. But, it must be that the integrals formed under the two settings of v are indeed the same since the probability measure is identical as

$$\begin{aligned} P(a_2, b_1|u_i, v_j) + P(a_2, b_2|u_i, v_j) &= \int_{\Psi} \int_{\phi o}^{\phi} p(\phi, \psi|u_i, v_j) \mathrm{d}\phi \mathrm{d}\psi \\ &= \int_{\phi o}^{\phi} p(\phi|u_i) d\phi = \int_{\Psi} \int_{\phi o}^{\phi} p(\phi, \psi|u_i, v_{j'}) d\phi d\psi \\ &= P(a_2, b_1|u_i, v_{j'}) + P(a_2, b_2|u_i, v_{j'}). \end{aligned}$$

B. If marginal response invariance is in force for all settings of v and for any decisional boundary for u (and therefore ϕ) then because of decisional separability,

$$\begin{aligned} & \int_V \int_{u_o}^{\infty} p(x, y, z|u_i, v_j) \mathrm{d}M = \int_{\Psi} \int_{\phi_o}^{\phi} p(\phi, \psi|u_i, v_j) d\phi d\psi \\ = & \int_V \int_{u_o}^{\infty} p(x, y, z|u_i, v_{j'}) \mathrm{d}M = \int_{\Psi} \int_{\phi_o}^{\phi} p(\phi, \psi|u_i, v_{j'}) d\phi d\psi \end{aligned}$$

so that $\int_{\phi o}^{\phi} p(\phi|\ u_i, v_j)\ \mathrm{d}\phi = \int_{\phi o}^{\phi} p(\phi|\ u_i,\ v_{j'})\ \mathrm{d}\phi$ for all v_j and $v_{j'}$ and ϕ_0. Because of the uniqueness of the relation of the distribution with the parameter sets, it follows that both sides are invariant across v, for all ϕ_0, which implies PDR-PS. ■

Strictly speaking, without decisional separability, marginal response invariance and PDR-PS are logically unrelated. There is an asymmetry in what one might conclude from data. If separability is not experimentally supported, that is, marginal response invariance fails, then one cannot know from this analysis alone if perhaps the absence of decisional separabilitymight not have ruined the chances for true underlying separabilty to manifest itself in the data. But, if perceptual separability is supported by the success of marginal response invariance, then there are fewer cases where a combination of a certain kind of decisional bound and failure of perceptual separability might nevertheless lead to a misleading appearance of separability. An example of such an anomaly is mean shift integrality (e.g., Kadlec & Townsend, 1992; Maddox, 1992).

4. PERCEPTUAL INDEPENDENCE

Along with the perceptual and decisional separabilities, the other founding concept of general recognition theory is "perceptual independence". Independence is a notion that is so ubiquitous that it is probably impossible to confine it to one concept even in highly related areas of discourse. Ashby & Townsend (1986) chose to assign the term, with a modifier—"perceptual," to represent the finest grain level with which we were dealing, that of probabilistic independence of perceptual dimensions, within a single stimulus. That is, perceptual independence is a *within-stimulus* condition and as is perhaps befitting, this usage coincides with that in probability and stochastic processes theory. Observe again, that in contrast, perceptual separabilty refers to *across-stimulus* invariances (see, e.g., Townsend & Ashby, 1982). Hence, logically, perceptual independence has nothing to do with domain-range considerations. Nevertheless, later discussion will indicate how the domain-range map could indeed play a strategic role, despite the logical distinction.

It is straightforward to define perceptual independence for manifolds more general than the Euclidean. But on region Π, we can express matters analogously to Ashby & Townsend (1986).

Definition 11 *Perceptual independence will be said to hold for a given experimental condition* (u, v) *if* $p(\phi, \psi|u, v) = p(\phi|u, v)p(\psi|u, v)$.

Again, notice that it is not automatically assumed that the marginal probability density on ϕ is a function only of u for that would be tantamount to forcing perceptual separability as well, which is a logically distinct concept. Unfortunately, psychologists have frequently conflated separability and independence.

What is the commonly observable relation from experiments pertaining to the theoretical notion of perceptual independence? It is a statistic called sampling independence deriving from earlier usage in probing independence of feature processing (e.g., Townsend, Hu and Ashby, 1981; Ashby & Townsend, 1986). At the risk of confusing previous readers of GRT, we would like to alter the name of "sampling independence" to "report independence" to emphasize that this statistical property is an observable. We again use the definitions relating to experimentally observable response probabilities from above.

Definition 12 *Consider an experimental setting of the stimulus dimensions u, v. Then, report independence is defined by the property that* $P(a_i, b_j|u,v) = [P(a_i, b_j|u,v) + P(a_i, b_{j'}|u,v)][P(a_i, b_j|u,v) + P(a_{i'}, b_j|u,v)]$.

Hence, report independence requires that the joint likelihood of reporting a_i and b_j equals the product of the marginal probabilities of reporting them. This definition is in line with the notion of probabilistic independence, but of course, the report probabilities, both joint and marginal, are integrals over the underlying probability densities. Note that it is not necessary to require that the observer explicitly report each value of the two dimensions. As long as there is a 1-1 assignment of stimulus pairs to responses, report independence can be checked. Thus, in the first instance, the experimenter could require that the observer respond "This trial I saw dimension A at level 1 but dimension B at level 2." Alternatively, she/he could require that the observer respond R_i, $i = 1,2,3,4$, with R_1 corresponding to level 1 on both dimensions, R_2 to level 1 on dimension A but level 2 on dimension B and so on. In either case, it is straightforward to assess report independence.

Next, we require a proposition to link up perceptual independence and report independence. The natural proposition depends strongly on decisional separability as it did earlier. Obviously, just as in the case of separability, in M the densities are 'perturbed' by the stretching factor. However, given a stimulus (u_0, v_0), on integration the independence of, say, the ψ trajectory representing the psychological value say, ϕ' and the ϕ trajectory representing psychological value ψ', will be established, even if we were directly integrating in x, y, z on M.

Proposition 8 *If decisional separability holds for a given experimental setting, then perceptual independence and report independence are equivalent.*

Proof. Assume decisional separabilty.

A. Perceptual independence implies report independence. With decisional separability,

$$
\begin{aligned}
& P(a_2, b_2|u,v) \\
= & \int_{\psi o}^{\infty} \int_{\phi o}^{\infty} p(\phi, \psi|u,v) d\phi d\psi = \int_{\phi o}^{\infty} p(\phi|u,v) d\phi \int_{\psi o}^{\infty} p(\psi|u,v) d\psi \\
= & [\int_{-\infty}^{\psi o} \int_{\phi o}^{\infty} p(\phi, \psi|u,v) d\phi d\psi + \int_{\psi o}^{\infty} \int_{\phi o}^{\infty} p(\phi, \psi|u,v) d\phi d\psi] \\
[\int & {}_{\psi o}^{\infty} \int_{-\infty}^{\phi o} \int_{\phi o}^{\infty} p(\phi, \psi|u,v) d\phi d\psi + \int_{\psi o}^{\infty} \int_{\phi o}^{\infty} p(\phi, \psi|u,v)) d\phi d\psi] \\
= & [P(a_2, b_2|u,v) + P(a_2, b_1|u,v)][P(a_1, b_2|u,v) + P(a_2, b_2|u,v)].
\end{aligned}
$$

B. Report independence for all u, v implies perceptual independence. Because of decisional separability, just noting

$$P(a_2, b_2|u, v) = [P(a_2, b_2|u, v)+P(a_2, b_1|u, v)][P(a_1, b_2|u, v)+P(a_2, b_2|u, v)]$$

and reversing the argument of part A proves the result. ■

We have concentrated on how things look in Π space which is virtually identical to the Cartesian circumstances in Ashby & Townsend (1986). Moreover, unless DR-PPS is true, the map Y carrying $\Phi \times \Psi$ into M will not coincide with the X map. Some further remarks concerning the relationship between separability and independence will be made in the next section.

5. PROBABILISTIC RANGE-ALONE SEPARABILITY AND PERCEPTUAL INDEPENDENCE

What happens when we move from a tight linkage of each physical dimension to their respective perceptual images? Can, for instance, such notions as separability and independence still have any empirical referent? These issues are explored in this section.

We begin with the fairly minimalist assumption that there exist two psychological dimensions everywhere which possess tangents that are linearly independent. Further, we continue to assume that there exists a mapping that carries the equi-magnitude trajectories of each dimension into a psychological plane $\Phi \times \Psi$. Obviously though, these no longer need be coincident with the original physical dimensions, u and v. That is, no longer will $X(u, V) = Y(\phi, \Psi)$, where ϕ is the image of u.

First, since DDR-PS obviously does not hold, could there exist probability distributions such that PDR-PS still is true? This appears to be an open problem. Basically, a functional equation that captures this general issue is, for separability of ψ over ϕ, and when ψ depends on both u, v but ϕ only depends on u,

$$\begin{aligned}\frac{\partial}{\partial u} \int \; p(\phi, \psi|u, v)d\phi &= \int \frac{\partial p(\phi, \psi|u, v)}{\partial u} d\phi \\ &= \int \frac{\partial p(\phi, \psi|\phi(u), \psi(u, v))}{\partial \phi(u)} \frac{\partial \phi(u)}{\partial u} d\phi + \int \frac{\partial p(\phi, \psi|\phi(u), \psi(u, v))}{\partial \psi(u)} \frac{\partial \psi(u)}{\partial u} d\phi.\end{aligned}$$

The first term is always zero but the second being zero establishes the condition for PDR-PS. If perceptual independence holds, then the critical equation is

$$\int \frac{\partial p(\psi|\psi(u,v))}{\partial\psi(u,v)} \frac{\partial\psi(u,v)}{\partial u} d\phi = 0.$$

Since we assumed that p is a function of the parameter $\psi(u,v)$ and ψ is a function of v, both non-trivial, then it's clear that these two functions must trade-off in such a way over $\Phi \times \Psi$ that the integral is nil.

As noted earlier, if ϕ, ψ are dependent under p, even if the parameters are separate functions only of u, v, $\phi(u)$ and $\psi(v)$, then this dependence may ruin PDR-PS. Yet, it is also possible that PDR-PS is true although perhaps unlikely. This situation is akin to what has been defined as "marginal selective influence" (Townsend & Schweickert, 1989; see also Townsend, 1984, Townsend & Thomas, 1993) in response time theory.

Next, we turn to the issue of redefining the stimulus dimensions in order to enforce DDR-PS. If we had access to the internal sites of $\Phi \times \Psi$, in principle we might be able to ascertain if the marginal distributions of ϕ and ψ were invariant over the parameter settings, say $\phi\prime$, $\psi\prime$ respectively. However, behaviorally there is no way to make this stipulation empirical without creating a new stimulus domain (or doing something logically equivalent). Thus, we can now establish a new physical domain D^*, actually a diffeomorphism of the old one, D, such that the new physical coordinates u^*, v^* in D^* now correspond respectively to functions of ϕ, ψ, and hence are PDR-PS with regard to ϕ, ψ. Consider two settings (u,v) and $(u\prime, v\prime)$ of D that result in the respective joint densities $p(\phi,\psi|u,v)$ and $p(\phi,\psi|u\prime,v\prime)$. Suppose further that these settings are such that, according to the scales, ϕ remains the same whereas ψ is changed. This situation corresponds to that in the domain-range situation where u is held constant.

Definition 13 *Consider that two different settings of the stimulus are presented such that the ϕ sensations would be the same if no noise were present—the stimulus has been altered so that only the ψ dimension changes. Then range-alone, probabilistic perceptual separability (RA-PPS) of ϕ from ψ is defined as the requirement that the marginal distribution on ϕ, after integrating over the ψ variable, will be the same for both settings.*

Since we have stipulated that ϕ is unchanged by integrating over the other variable, we can simply write this requirement in terms of the respective densities under the two settings:$p(\phi,\psi|u,v) = p(\phi,\psi|u_1^*, v_1^*)$ for the first and $p(\phi,\psi|u\prime,v\prime) = p(\phi,\psi|u_1^*, v_2^*)$ for the second. Note that ϕ, as one of the two sensory parameters of the density, will be unchanged as is u^*. Similarly, in general an absence of PDR-PS finds that

$$\int_\Psi p(\phi,\psi|u,v)\mathrm{d}\psi = p(\phi|u,v) \neq p(\phi|u\prime,v\prime) = \int_\Psi p(\phi,\psi|u\prime,v\prime)\mathrm{d}\psi.$$

Is there any empirical content left in this situation? The answer is clearly "yes" as long as the scientist has some guide as to the possible underlying dimensions as functions of u, v. The idea would be to use macroscopic methods (e.g., magnitude estimation, cross-modality matching, etc.) first to map out the way in which two physical dimensions interact (e.g., as in equi-loudness contours) to produce the psychological dimensions (e.g., Baird, 1977; see also Townsend & Spencer-Smith, 2004). Then, in principle one can carry out tests such as those for marginal response invariance. Naturally, some aspects of marginal response invariance will have to change to reflect the fact that u does not accurately represent ϕ any longer and similarly for v and ψ.

Nevertheless, for each distinct combination of values of ϕ and ψ, there will still be a unique u, v that produce exactly that psychological pairing. Now that D^* is defined, such notions as marginal response invariance and decisional separability follow immediately, and at least in principle and perhaps in practice with effort, RA-PPS can still be tested.

The situation for perceptual independence is a bit different, in that since we are working with a single stimulus setting, the report independence test can be carried out immediately, without fretting about a remap of $D \to D^*$ first. Everything is as before in the domain-range milieu.

6. A PSYCHOLOGICAL SPACE AND TWO DISTRIBUTIONS

A strong contender for the most studied pair of dimensions, with regard to issues of independence, are the orientation and size of a line-at-an-angle, often in the context of a definite 0-angle line and an arc, producing pie-slice like figures (Shepard, 1964). Most published studies have assessed the orientation and size dimensions to be separable. Thus, Shepard (1964) and Nosofsky (1985) found MDS plots to be reasonably rectangular, which within the deterministic milieu suggests separability. MDS fits using a city block metric, again in highly discriminable stimuli have generally been more successful than a Euclidean metric (e.g., Dunn, 1983; Hyman & Well, 1967, 1968; Shepard, 1964). More recently, using GRT methodology, Kadlec & Hicks (1998) has substantiated both perceptual independence and perceptual separability as well as decisional separability. (Potts, Melara & Marks, 1998) provide a thorough parametric investigation of experimental conditions that do or do not lead to separability with the Shepard (1964) stimuli.

We present a simple space M that can capture the topological and probabilistic aspects of the data and serve as a guide to potential assay of the metric, as well as of so far unexplored issues such as curvature and existence– or not – of paths and geodesics in the perceptual space for orientation and size. Then we impose two distinct probability densities on M.

Let ϕ correspond to psychological size, presumably mainly a function of physical radius u, and ψ correspond to psychological orientation, presumably mainly a function of physical orientation v. We think of u as standing for size and contained in the interval (a, b), a>0, not necessarily bounded on the right in general, and v as angles in $(0, 2\pi)$, say. Psychological space M in $\Re^3$ is then assumed to be the surface of revolution produced by the formula

$$X(u,v) = (s(u,v), \phi(u,v)\cos\psi(u,v), \phi(u,v)\sin\psi(u,v)).$$

Here, the $s(u,v)$ term deforms the psychological parameters so as to describe distortion orthogonal to the cross section of the surface, along the x axis say. We assume s is psychological size ϕ, so that

$$X(u,v) = (\phi(u,v), \phi(u,v)\cos\psi(u,v), \phi(u,v)\sin\psi(u,v)).$$

For well-behaved functions ϕ and ψ, X is 1-1 and onto and is further a diffeomorphism. The various mappings are illustrated in Figure 5.

Figure 5 near here

Assuming ϕ, ψ are perceptually separable as suggested by Kadlec & Hicks (1988; and under certain circumstance, Potts, Melara & Marks, 1998) then we have at the deterministic level,

$$X(u,v) = (\phi(u), \phi(u)\cos(\psi(v)), \phi(u)\sin(\psi(v)))$$

and coordinate space $\Pi = \Phi \times \Psi$. By Definition 3, ϕ and φ are then DDR-PS (deterministically domain-range separable), and by Definition 8 they are decisionally separable from each other with the boundaries effectively slicing the cone through the y-z plane, or forming a sector along the x-axis. In order to go further, we need to retrieve some more differential geometry.

Ultimately, we need the metric. Obviously, we don't know that to begin with. We do know that the set of points in a space and the metric are independent. For illustration sake, we employ the metric inherited from the mapping Y, as Riemannian metric. In our example, $Y_\phi = (1, \cos\psi, \sin\psi)$ and $Y_\psi = (0, -\phi\sin\psi, \phi\cos\psi)$. Hence

$$g_{11} = \langle(1, \cos\psi, \sin\psi), (1, \cos\psi, \sin\psi)\rangle = 1 + (\cos\psi)^2 + (\sin\psi)^2 = 2.$$

Similarly, $g_{12} = g_{21} = 0$, and $g_{22} = \phi^2$. Then, as observed earlier, the map Y distorts the ϕ, ψ plane in a way captured by $\sqrt{g_{11}g_{22} - g_{12}^2} = \sqrt{2}\phi$.

We can further conclude two immediate points from the fact that $\langle Y_\phi, Y_\psi \rangle = 0$. First, Y_ϕ and Y_ψ are the basis vectors for a coordinate system in T_pM, and their inner product being 0 implies that they are orthogonal at every point m. Hence, this system also obeys the tenet of local orthogonal separability (Townsend & Spencer-Smith, 2004).

Furthermore, there is no interaction in the sense that, although the speed depends on ϕ, ψ (degenerately in ψ, to be sure), the off-diagonal is 0. This statement can be appreciated in the expression for the infinitesimal increment

$$\begin{aligned} ds &= \sqrt{g_{11}(\phi,\psi)(d\phi)^2 + 2g_{12}(\phi,\psi)(d\phi)(d\psi) + g_{22}(\phi,\psi)(d\psi)^2} \\ &= \sqrt{2(d\phi)^2 + \phi^2(d\psi)^2}. \end{aligned}$$

Thus, observe that even if one is, for instance, traveling on the diagonal in the plane, only the weighted Euclidean values determine magnitude—there is no interaction of the product of $d\phi \times d\psi$. The rate of progression along the ϕ or x axis is uniform at rate 2 whereas if we move around a circle, the rate of change is a function of how far out we are along the x-axis (*i.e.*, how big the size is) we are. This makes sense since the circumference is larger the farther out we are. Note that even if DDR-PS holds as functions of the physical stimuli values, the apparent rate of change along the x-axis may not be uniform as it is moderated by the term $(\partial\phi/\partial u)^2$.

A deficiency of the inherited metric of our map from some points of view may be that it is not the city block metric. Many studies, though far from all, have found that the city block fits in MDS scaling of perception of Shepard stimuli are superior to fits by the Euclidean metric. However, the studies only test members of the power metric class and sometimes only city block vs. Euclidean. As far as we know, no general Riemannian metrics have been attempted with these stimulus patterns. The Euclidean metric is the only power metric that is also a Riemannian metric. One property that all the power metrics share is the non-interaction across dimensions (see Beals, Krantz & Tversky, 1968; Townsend & Thomas, 1993) and Summary and Discussion below.

Distribution A

Suppose initially that probability density p on the manifold M depicted in Figure 5 is Gaussian about the determinate mapping of stimulus (u, v).

That is, we suppose that for points (x, y, z) in M,

$$p(x, y, z|u, v) = Nexp(-\alpha((x - x(u, v))^2 + (y - y(u, v))^2 + (z - z(u, v))^2))$$

with normalizing function $N = N(u, v)$. In deriving the associated probability distribution p in coordinate space Π it is necessary to account for the distortion in M. Setting $(\phi_0, \psi_0) = Y^{-1}X(u, v)$, we have

$$p(\phi, \psi|u, v) = \sqrt{2}\phi N(\phi_0, \psi_0)\exp(-\alpha((\phi - \phi_0)^2 + (\phi\cos(\psi) - \phi_0\cos(\psi_0))^2 + (\phi\sin(\psi) - \phi_0\sin(\psi_0))^2))$$

which by standard trigonometric relations reduces to the following:

$$p(\phi, \psi|u, v) = \sqrt{2}\phi N(\phi_0, \psi_0)exp(-2\alpha(\phi^2 + \phi_0^2 - \phi\phi_0 - \phi\phi_0\cos(\psi - \psi_0))).$$

Figure 6 near here

Figures 6 (a) and (b) illustrate this probability density on coordinate space $\Phi \times \Psi$ for two physical stimulus values (u, v).

This is an example when perceptual independence (Definition 11) fails to hold. If ϕ is a function only of u, however, the example demonstrates probabilistic separability (PDR-PS) of ϕ from ψ. To see this, observe firstly that the normalizing function $N(\phi_0, \psi_0)$ must be independent of ψ_0 because $N(\phi_0, \psi_0)^{-1} \equiv \int_\Phi \int_\Psi N(\phi_0, \psi_0)^{-1}p(\phi,\psi|u,v)\ d\psi\ d\phi = \sqrt{2} \int_\Phi \phi\exp(-2\alpha(\phi^2 + \phi_0^2 - \phi\phi_0)) \int_\Psi \exp(-2\alpha(\phi\phi_0\cos(\psi - \psi_0)))\ d\psi\ d\phi$ which is independent of ψ_0. Then $\int_\Psi p(\phi,\psi|u,v)\ d\psi = \sqrt{2}\ N(\phi)\phi\exp(-2\alpha(\phi^2 + \phi_0^2 - \phi\phi_0)) \int_\Psi \exp(-2\alpha(\phi\phi_0\cos(\psi - \psi_0)))\ d\psi$ is independent of ψ_0 and hence of v.

Even if DDR-PPS holds and ϕ is basically linear in u with slope 1, as the data suggest (e.g., Baird, 1997, p. 34) and we assume also that $\psi(v) = v$, then probabilistic separability of ψ from ϕ still will not hold. If psychological size varies from 0 to infinity, and ψ is separated from ψ_0 by π radians so that $\cos(\psi - \psi_0) = -1$, then $\int_\Phi p(\phi,\psi|\ u,v)\ d\phi = \sqrt{2}\ N(\phi_0)\ \exp(-2\alpha\phi_0^2) \int_\Phi \phi\exp(-2\alpha\phi^2)\ d\phi = \sqrt{2}\ N(\phi_0)\ \exp(-2\alpha\phi_0^2)\ (4\alpha)^{-1}$. For this term to be independent of ϕ and hence of u, necessarily $N(\phi_0)$ is proportional to $\exp(2\alpha\phi_0^2)$, say $N(\phi_0) = c\ \exp(2\alpha\phi_0^2)$. However, if $\psi = \psi_0$ so that $\cos(\psi - \psi_0) = 1$, then substituting for $N(\phi_0)$ gives $\int_\Phi p(\phi,\psi|\ u,v)\ d\phi = \sqrt{2}\ c \int_\Phi \phi \exp(-2\alpha\phi(\phi - 2\phi_0)))\ d\phi$ which is not independent of ϕ_0. So for any ψ we have demonstrated that the marginal distribution $\int_\Phi p(\phi,\psi|\ u,v)d\phi$ cannot be independent of u for all physical stimulus values $v = \psi_0$.

Distribution B

In contrast, consider the case when the probability density p is Gaussian on Π about the determinate mapping of stimulus (u, v). That is, we suppose that $p(\phi,\psi|u, v)$ is proportional to $\exp(-\alpha(\phi-\phi)^2 + (\psi-\psi_0)^2))$, where again we suppose that $\phi(u, v) = \phi_0$ and $\psi(u, v) = \psi_0$. (This density function is illustrated in Figures 6(c) and (d); note that scaling on the ϕ (size) axis has visually distorted the circularly symmetric functions.) With this assumed probabilistic structure, perceptual independence obviously holds. With DDR-PS, probabilistic separability also holds in both directions, since for example $\int_\Psi p(\phi,\psi|u, v)\, d\psi = \exp(-\alpha(\phi-\phi_0)^2) \int_\Psi \exp(-\alpha(\psi-\psi_0)^2)\, d\psi \,/\, \int_\Phi \int_\Psi p(\phi,\psi|u, v)\, d\psi\, d\phi = \exp(-\alpha(\phi-\phi_0)^2) \int_\Phi \exp(-\alpha(\phi-\phi_0)^2)\, d\phi$ is independent of $v = \psi_0$.

These two distributions, A, B along with the definitions and propositions illustrate several aspects of the theory. Note that Example A has the joint density as independent in x, y, z, as might happen if independent noise sources are intimately tied to the separate x, y, z dimensions. When interpreted in terms of Π, independence is severed between ϕ and ψ. Example A also points up potential relationships between independence and separability, even though the concepts are logically distinct: With non-independence, it takes a kind of balancing act to rid, say, the marginal distribution on ϕ, from contamination via the 'wrong' stimulus dimension, in this case v. This balancing act does not occur in this example.

Example B sees independence between ϕ and ψ, for instance, as we might expect if independent noise sources are connected with the perceptual coordinates ϕ, ψ. Of course, independence in ϕ, ψ will not transfer to independence among x, y, z. Furthermore, if DDR-PS holds then Distribution B leads directly to PDR-PPS.

Decisional separability then, in the case of Distribution B will eventuate in marginal response invariance and report independence. However, decisional separability will expose the lack of PDR-PS in the case of Distribution A through a failure of marginal response invariance. And, it will also expose an absence of perceptual independence via failure of report independence.

7. SUMMARY AND DISCUSSION

7.1. Roles of the Map X

We have developed the structures for GRT that are suitable for its use on simple 2-dimensional manifolds. We used coordinates but did not resort to extrinsic concepts (e.g., a normal to a surface). Thus, though we have emphasized surfaces, the developments immediately generalize to 2-dimensional Riemannian manifolds with diffeomorphic physical→sensory maps; and, with obvious complexification of the notation, from 2-dimensional physical Euclidean space to n-dimensional manifolds.

We have found it useful in the present enterprise to accentuate the macroscopic picture of the physical-to-sensory map X which would be deterministically visible when discrimination is high (e.g., the psychophysical loudness as a function of sound intensity, as opposed to the discrimination of two neighboring tones). This perspective was not so obviously central in Ashby & Townsend (1986), though it lay in the background of the separability concepts. But here, it makes clear the distinction between the map X and the stochastic structure (e.g., see Definition 6). Thus, we demonstrated in the last section on range-alone separability, that even if the physical space were re-coordinatized from D to D^* so that a new X^* will be deterministically separable (DDR-PPS), it may not be probabilistically separable (PDR-PPS).

The underlying map X plays other strategic roles as well. Consider the question of decisional separability. It is certainly a highly critical aspect of performance and is exceedingly important in observing the internal structure. On the encouraging side is the likelihood that the decision bounds might be easier for an observer to align with the perceptual axes when the coordinates are perceptually separable. In fact, in categorization research, it is has been found that observers first try to establish decision bounds that are aligned with the simple and evident psychological dimensions and only if that strategy is ineffectual or highly non-optimal to attempt other strategies (Ashby & Maddox, 1990, 1992). To be sure, it has been found in categorization experiments that observers can establish separable decision bounds with stimuli thought to be integral (Ashby & Gott, 1988) and non-separable decision bounds with perceptually separable stimuli, if in each case the bounds are optimal within the experimental setting (Ashby & Maddox, 1990, 1992). Yet, it does appear that it is easier for people to set up decisionally separable bounds if the experimental dimensions are indeed perceptually separable (e.g., Ashby, 1992, pp. 459ff.). Hence, it seems not unlikely that the assumption of decisional separability might be reasonable in a number of circumstances.

In another role for X, although perceptual independence is formally untied to separability X, in actuality they may be intertwined. As a simple example, suppose $M = \Pi = \Re^2$ and X is a linear orthogonal map of $D = \Re^2$ to Π but with a rotation so that ϕ, ψ are orthogonal. On the one hand, if the noise source is independent in u and v, then it will be correlated in ϕ and ψ. If decisional separabilty holds, then both the failure of perceptual separability as well as the failure of perceptual independence can be discovered.

If X is deterministically perceptual separable as in DDR-PPS, but the noise is correlated then again, with the presence of decisional separability, the satisfaction of probabilistic perceptual separability will be experimentally detectable as will the failure of independence.

7.2. Roles of Metrics

A few remarks are worth making concerning the role of metrics in our proceedings. We employed the inherited metric as a natural approach. First, the actual mappings from sensory organs to brain with the various distortions could be directly related to the way in which the Euclidean metric, typically employed in the physical sciences, is distorted on the way to the psychological representation. However, it could also be that a particular manifold M is simply one of an equivalence class of isometric spaces that capture the psychological metric. Of course, it also could be that the actual metric (assuming one exists) bears little or no relationship to the 'form' of the psychological manifold.

In another interesting aspect, despite the presence of a metric, the main properties of the theory above are (differentially) topological in nature. This feature is associated with the usual change-of-variables invariance of Lesbegue or Riemannian integration. Hence, diffeomorphic alterations of either the psychological or the physical coordinates that do not mix coordinates will not disturb the main results concerning separability. That is, if $\phi = f(\alpha)$ and $\psi = g(\beta)$, where α and β are independent parameters and f, g are 1-1, differentiably invertible, then the conclusions regarding perceptual and decisional separability will be unharmed. Independence will not be distorted to dependence. In contrast, coordinate transformations that mix the coordinates, for instance, $\phi = f(\alpha, \beta)$ will ordinarily affect both separability and independence.

Metric issues can also intrude when various questions arise in, for instance, combating mimicking. A case in point regards the attempt by Kadlec & Townsend (1992) to rule out a perceptually non-separable configuration by applying an empirical test based on measurement of a line in the perceptual space. However, this was shown to be inappropriate except when perceptual independence was in force by Thomas (1999, 2003). In oversimplified form, the metric needs to take the variance of the attendant distributions into account in making such measurements. The potential relationships of such questions to the 'true' underlying psychological metric appear intriguing but lie beyond the present scope.

With regard to the metric that was emitted from the map in our application to the Shepard (1964) stimuli in the previous section, some particular and general observations can be made. First, the usually preferred city block metric being a member of the class of power metrics as noted, it is only one member of a set of general structures that are in a sense, independent functions of the distinct dimensions. Following Beals, Krantz and Tversky (1968, also see Suppes, Krantz, Luce & Tversky, 1989) we can write a dissimilarity (not necessarily a metric) function $D(x, y) = \mathrm{F}(\mathrm{f}_1(x_1, y_1), \mathrm{f}_2(x_2, y_2))$ where F is strictly monotonic in the two arguments from the two dimensions, a property they called "decomposability". Note that F does not mix the coordinate contributions as it would say if

$$D(x,y) = F(f_1(x_1,y_1), f_2(x_2,y_2), f_1(x_1,y_1).f_2(x_2,y_2)).$$

Our metric follows this precept by virtue of $g_{12} = 0$. However, it can also be observed that F is invariant across the points in the space, except as the coordinates appear under the f_i. Our metric does not obey this stricture since $g_{22} = \phi^2$.

Townsend & Thomas (1993) made the point that qualitative properties put forth by Beals et al. (1968) such as decomposability (above), which is shared by all the power metrics, may be a more appropriate criterion for separability than the very strict city block metric per se. However, we believe that the relationships among the various notions of independence and separability, including what we might call metric separability (which would presumably invoke properties such as forms of decomposability) and DDR-PPS or PDR-PPS deserve more thought and study. Both their interrelationships and their direct relationships to task and response structure may be important considerations.

7.3. Other Approaches to Separability

Clearly this is not the place for a global review of this important but already vast topic. Ashby & Townsend (1986) provided a few linkages with earlier (and continuing) approaches such as Garner and colleagues' and Shepard and his group and Townsend & Thomas (1993) provide a more extensive, if not so current, survey. And, we have to neglect the massive psychometric literature that has stemmed from the innovations of Shepard (1964), Kruskal & Wish (1979) and Carroll & Arabie (1980). However, much more could be done even excluding the latter realms. For instance, in one sector, we are collaborating with the laboratory of *D.* Algom in Israel to integrate GRT, Garner's approach and our response time technology (e.g., Algom, Eidels, Kadlec & Townsend, 2005; Melara & Algom, 2003; Townsend & Nozawa, 1995; Townsend & Wenger, 2004) in investigation of the famous Stroop effect. Whenever there is indication of significant inhibition (as in the Stroop effect) or facilitation (as in, say redundant signal experiments) (Colonius & Townsend, 1997; Miller, 1982; Mordkoff & Egeth, 1993; Mordkoff & Yantis, 1991) across dimensions, there is prima facie evidence for non-separability.

In fact, we envision a taxonomic 'space' where various types of independence and separability hold sway in part of the space. Contextual influences in general will be viewed as disabling certain of the independence conditions. 'Pure' non-independence systems of interest will include configurality where we conceive that positive dependences may dominate (e.g., Townsend & Wenger, 2004; Wenger & Townsend, 2001) and highly inhibitory interactions such as is posited for the Stroop effect.

Multidimensional Fechnerian Scaling

Dzhafarov (2002) applied his and Colonius' Multidimensional Fechnerian Scaling (MDFS) to perceptual separability, within the same-different matching setting, rather than identification experiments as is the present work. His approach is also distinctive in several other more substantive respects. Notwithstanding, it does appear possible to make a few informal comparative observations in the spirit of scientific communication.

Dzhafarov's modus operandi is to place conditions on functions of probabilities associated with coordinates. And, the coordinates are stimulus coordinates alone—no perceptual coordinates. Both these aspects differ from the GRT approach which is to deal with psychological factors like dissimilarity or distance and bias within a multi-dimensional signal detection (identification) environment. In GRT, perceptual coordinates are present, and the probabilities are placed on these. These diversities are not necessarily at odds by any means but they do indicate quite distinct strategies. Perhaps an apt analogy is the relationship of functional equations on choice probabilities vs. random utility theory (whose scales represent a state space). For instance, Luce's choice axiom is an example of the former in its purest form. Later, investigators sought conditions in random utility theory that satisfied the choice strictures (e.g., Holman & Marley cited in Yellott's 1977 paper). Obviously, GRT is more akin to the random utility approach but could be associated with functional equations expressing certain regularities.

The first condition of Dzhafarov's theory, called "weak perceptual separability" assumes that increments of probability of detection of a stimulus difference is a function of each of the dimensional differences, for instance, on u, v in our terms. So it could be written something like

> Pr[increase in detecting a difference in going from (u, v) to $(u + su, v + sv)$; s positive real and small | stimulus (u, v)] =

> F(increase in detecting a difference in going from u to $u + su$, increase in detecting a difference in going from v to $v + sv$ | stimulus (u,v)).

Note that even though F separates out the stimulus dimensions, the probabilities for both dimensions can still depend on the entire stimulus point (u, v). In most interpretations in terms of GRT, it appears this would engender a failure of PDR-PPS in our machinery, except perhaps for special cases where v becomes 'ineffective' relative to the u probabilities and vice-versa.

The second condition, called "detachability," is that when all the dimensions but one are held fixed and the one varied, then the probability of increase in detecting a difference does not depend on the values of the fixed dimensions. Again, informally,

> Pr[increase in detecting a difference in going from (u, v) to $(u + su, v)$; s positive real and small | stimulus (u, v)] is the same for all values of v.

In this condition, contrary to the first, natural transitions to GRT would seem to imply PDR-PS. For instance, if the individual dimensional probabilities are interpreted as marginals as is natural, then the transference from FMS to GRT seems reasonable.

It appears of interest to investigate these relationships more deeply, but the present comments appear to intimate some intriguing hints.[3] Since MDFS is presently occupied mostly with same-different designs as opposed to identification, study of the relationships may require additional definitions and/or assumptions (and Robin Thomas's work on interpretations of GRT in same-different designs may be of assistance).

Potential Extensions of theTheory

Generalization to many-one maps X is a clear direction to take in the future, given the phenomena of perception. For instance, when sensory dimensions map into preference scales, it is typically expected that these functions are non-monotonic (e.g., Coombs, 1964).

However, in some cases, the need is merely apparent, not real. One instance is the constancy of many percepts, typically best handled within invariance of say, perceptual operators. In other cases, it is fair to first negotiate the early down-sizing map, and then to figure in the nature of

[3] Dzhafarov (2003a, b) has shown that certain conditions on discrimination probabilities rule out natural Thurstonian (and therefore perhaps GRT) models of same different discrimination. Discussion of this result exceeds the present scope.

the early sensory→perceptual map. This does not short-change the early physical→sensory function, but provides a division of labor. For instance, the laws of establishment of hues as functions of the infinite dimensional light spectrum, might act somewhat independently of the geometry of the hue-brightness-saturation manifold.

When actually requisite, the theory of immersions is likely a first consideration. Thus, these may be appropriate for non-monotonic perceptual→preference maps. They permit many-one and non-homeomorphic (even if the map *is* one-one) maps but assume local one-one structure, that is, that the differential ($dX : T_{(u,v)}D \rightarrow T_m M$) to the range space (i.e., the coordinate Jacobian) is of the same rank as the domain. Since Rank(dX) $\leq$ min(Rank D, Rank M), in our present milieu, an immersion suggests that certain aspects of the physical world, in particular the point-to-point correspondence, is maintained at least locally. The natural alternative to immersions would be to assume a submersion where dX is of the same rank as M. It follows that locally the full scope of the perceptual space is being utilized. Immersions are typically more useful than submersions. For instance, if X is a one-one homeomorphic immersion, then locally it is an embedding and therefore locally shares the resident topology of M (e.g., Boothby, 2003). Further, if the points of D that map to the same points of M are not too close, the probability densities map enjoys the sufficient properties that allow, say, separabilty to be present and measurable.

It seems inevitable that more sophisticated spaces will ultimately be required in the cognitive, social, and even biological sciences than have hitherto been the mainstays. There have been scattered examples of efforts to expand psychology's spatial purview over the past fifty years but we suspect that the well has hardly been tapped in this regard. Fifty years from now could see spatial models very different from even the most exotic offerings from today's cafeteria, including the present approach. One exceedingly critical feature that has been scarce in cognitive science, has been the invention of direct experimental tools that permit, encourage, and interact with evolving theories of psychological spaces. At present, short of specific geometric model fitting, there is little to aid us in direct implementation of the notions proffered here. We do however, hope that the kinds of issues that arise in the present development will help to move us forward toward pursuit of an expanded set of perspectives and tools.

References

Algom, D., Eidels, A., Kadlec, H., & Townsend, J. T. (2005). Dull news travels fast: A separate channel theory of the Stroop effect. Manuscript submitted for publication.

Ashby, F. G. (1992). Multidimensional models of categorization. In F. G. Ashby (Ed.). *Multidimensional models of perception and cognition*. Scientific Psychology Series, Hillsdale, NJ: Lawrence Erlbaum.

Ashby, F. G., & Gott, R.E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognitin, 14*, 33-53.

Ashby, F. G., & Maddox, W. T. (1990). Integrating information from separable psychological dimensions. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 598-612.

Ashby, F. G., & Maddox, W. T. (1992). Complex decision rules in categorization: Contrasting novice and experienced performance. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 50-71.

Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review, 93*, 154-179.

Baird, J. C. (1977). Sensation and judgment. Mahwah, NJ: Erlbaum.

Baird, J. C. (1997). *Sensation and judgment: Complementarity theory of psychophysics*. Scientific Psychology Series, Mahwah, NJ: Erlbaum.

Beals, R., Krantz, D.H., & Tversky, A. (1968). Foundations of multidimensional scaling. *Psychological Review, 75*, 127-142.

Boothby, W. M. (2003). *An introduction to differentiable manifolds and reimannian geometry*. San Diego, CA: Academic Press

Carroll, J. D. & Arabie, P. (1980). Multidimensional scaling. *Annual Review of Psychology, 31*, 607-649.

Colonius, H. & Townsend, J. T. (1997). Activation-state representation of models for the redundant-signals-effect. In A. A. J. Marley (Ed.), *Choice, decision and measurement*, volume in honor of R. Duncan Luce, Mahwah, NJ: Erlbaum Associates.

Coombs, C. H. (1964). *A theory of data*. New York, NY: Wiley.

Dunn, J. C. (1983). Spatial metrics of integral and separable dimensions. *Journal of Experimental Psychology: Human Perception & Performance, 9,* 242-257.

Dzhafarov, E.N. (2002). Multidimensional Fechnerian scaling: Perceptual separability. *Journal of Mathematical Psychology, 46,* 564-582.

Dzhafarov, E.N. (2003). Thurstonian-type representations for "same-different" discriminations: Deterministic decisions and independent images. *Journal of Mathematical Psychology, 47,* 208-228.

Dzhafarov, E.N. (2003). Thurstonian-type representations for "same-different" discriminations: Probabilistic decisions and interdependent images. *Journal of Mathematical Psychology, 47,* 229-243.

Dzhafarov, E.N., & Colonius, H. (1999). Fechnerian metrics in unidimensional and multidimensional stimulus spaces. *Psychonomic Bulletin & Review, 6,* 239-268.

Hirsch, M. W. (1976). *Differential topology.* New York, NY: Speringer-Verlag.

Hyman, R., & Well, A. (1967). Judgments of similarity and spatial models. *Perception & Psychophysics, 2,* 233-248.

Hyman, R., & Well, A. (1968). Perceptual separability and spatial models. *Perception & Psychophysics, 3,* 161-165.

Kadlec, H., & Hicks, C. L. (1998). Invariance of perceptual spaces and perceptual separability of stimulus dimensions. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 80-104.

Kadlec, H., & Townsend, J. T. (1992). Signal detection analysis of multidimensional interactions. In F. G. Ashby (Ed.), *Probabilistic multidimensional models of perception and cognition.* Scientific Psychology Series, Hillsdale, NJ: Lawrence Erlbaum Associates.

Kobayashi, S., & Nomizu, K. (1991). *Foundations of differential geometry* (Vol. 1). New York, NY: Wiley.

Kobayashi, S., & Nomizu, K. (1969). *Foundations of differential geometry* (Vol. 2). New York, NY: Wiley.

Kruskal, J. B. & Wish, M. (1979). *Multidimensional scaling.* Beverly Hills, CA: Sage

Lang, S. (1985). *Differential manifolds.* New York, NY: Springer-Verlag.

Lang, S. (1993). *Real and functional analysis*(3^{rd} Ed.). New York, NY: Springer-Verlag.

Lappin (1978). The relative of choice behavior and the effect of prior knowledge on the speed and accuracy of response time like experiments. In N. J. Castellan Jr., & F. Restle (Eds.) *Cognitive theory, Vol. 3.* Hillsdale, NJ: Erlbaum.

Levine, M.V. (2003). Dimension in latent variable models. *Journal of Mathematical Psychology, 47,* 450-466.

Luce, R.D. (1963). Detection and recognition. In R.D. Luce, R.R.Bush, &E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol 1). New York: Wiley.

MacMillan, N. A., & Creelman, C. D. (2005). Detection theory: A user's guide. Mahwah, NJ: Erlbaum.

Maddox, W. T. (1992). Perceptual and decision separability. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition.* Scientific Psychology Series, Hillsdale, NJ: Erlbaum.

Melara, R.D., & Algom, D. (2003). Driven by information: A tectonic theory of Stroop effects. *Psychological Review, 110,* 422-471.

Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology, 14,* 247-279.

Mordkoff, J. T., & Egeth, H. E. (1993). Response time and accuracy revisited: Converging support for the interactive race model. *Journal of Experimental Psychology: Human Perception & Performance, 19,* 981-991.

Mordkoff, J. T., & Yantis, S. (1991). An interactive race model of divided attention. *Journal of Experimental Psychology, 17,* 520-538.

Munkres, J. R. (1975). Topology: a first course. Englewood Cliffs,N.J., Prentice-Hall.

Nilsson, N. J. (1965). Learning machines: Foundations of trainable pattern classifying systems. New York, NY: McGraw Hill

Nosofsky, R.M. (1985). Overall similarity and the identification of separable-dimension stimuli: A choice model analysis. *Perception & Psychophysics, 38,* 415-432.

Nosofsky, R.N. (1992). Similarity, scaling and cognitive process models. *Annual Review of Psychology, 43,* 25-53.

O'Neill, B. (1966). Elementary differential geometry. New York, NY: Academic Press.

Potts, Melara, & Marks (1998). Circle size and diameter tilt: A new look at integrality and separability. *Perception & Psychophysics, 60*, 101-112.

Suppes, P., Krantz, D.M., Luce, R.D., & Tversky, A. (1989) Foundations of measurement, *Vol. 2: Geometrical, threshold, and probabilistic representations.* San Diego, CA: Academic Press.

Thomas, R.D. (1999). Assessing sensitivity in a multidimensional space: Some problems and a definition of a general d'. *Psychonomic Bulletin & Review,6,* 224-238.

Thomas, R.D. (2003). Further considerations of a general d' in multidimensional space. *Journal of Mathematical Psychology, 47,* 220-224.

Townsend, J. T. (1971a). Theoretical analysis of an alphabetic confusion matrix. *Perception & Psychophysics, 9,* 40-50.

Townsend, J. T. (1971b). A note on the identifiability of parallel and serial

Townsend, J. T. (1984). Uncovering mental processes with factorial experiments. *Journal of Mathematical Psychology, 28,* 363-400.

Townsend, J. T., & Ashby, F. G. (1982). An experimental test of contemporary mathematical models of visual letter recognition. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 834-864.

Townsend, J. T., Hu, G. G., & Ashby, F. G. (1981). Perceptual sampling of orthogonal straight line features. Psychological Research, 43, 259-275.

Townsend, J. T., Hu, G. G., & Evans, R. (1984). Modeling feature perception in brief displays with evidence for positive interdependencies. *Perception & Psychophysics, 36,* 35-49.

Townsend, J. T., Hu, G. G., & Kadlec, H. (1988). Feature sensitivity, bias, and interdependencies as a function of energy and payoffs. *Perception & Psychophysics, 43*(6), 575-591.

Townsend, J. T., & Landon, D.E. (1982). An experimental and theoretical investigation of the constant ratio rule and other models of visual letter recognition. *Journal of Mathematical Psychology, 25,* 119-163.

Townsend, J. T. & Nozawa, G. (1995). On the spatio-temporal properties of elementary perception: An investigation of parallel, serial and coactive theories. *Journal of Mathematical Psychology, 39,* 321-360.

Townsend, J. T., & Schweickert, R. (1989). Toward the trichotomy method: Laying the foundation of stochastic mental networks. *Journal of Mathematical Psychology, 33,* 309-327.

Townsend, J. T., Solomon, B., Wenger, M. J., & Spencer-Smith, J. B. (2001). The perfect Gestalt: Infinite dimensional Riemannian face spaces and other aspects of face cognition. Chapter in M. J. Wenger and J. T. Townsend (Eds.), *Comutational, geometric and process issues in facial cognition: Progress and challenges.* Scientific Psychology Series, Mahwah, NJ: Erlbaum Press.

Townsend, J. T., & Spencer-Smith, J. B. (2004). Two kinds of global perceptual separability and curvature. Chapter in C. Kaernbach, E. Schröger, and H. Müller (Eds.), *Psychophysics beyond sensation: Laws and invariants of human cognition.* Scientific Psychology Series, Mahwah, NJ: Erlbaum.

Townsend, J. T. & Thomas, R. (1993). On the need for a general quantitative theory of pattern similarity. In S. C. Masin (Ed.), *Foundations of perceptual theory.* Amsterdam: Elsevier Publishers.

Townsend, J. T., & Wenger, M. J. (2004). A theory of interactive parallel processing: new capacity measures and predictions for a response time inequality series. *Psychological Review, 111,* 1003-1035.

Treisman, A. (1993). Perception of features and objects. In A. Baddeley & *I.* Weiskrantz (Eds.), *Attention, selection, awareness, and control.* (pp. 6-34). Oxford: Clarendon Press.

Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology, 12,* 97-136.

Tversky, A. (1977). Features of similarity. Psychological Review, 84(4), 327-352.

Vergaghe, P., & Nakayama, K. (1994). Stimulus discriminability and visual search. *Vision Research, 18,* 2453-2467.

Wenger, M. J., & Gibson, B. S. (2004). Using hazard functions to assess changes in processing capacity in an attentional cuing paradigm. *Journal of Experimental Psychology: Human Perception & Performance, 30*(4), 708-719

Wenger, M. J., & Townsend, J. T. (2001). *Computational, geometric, and process issues in facial cognition: Progress and challenges.* Scientific Psychology Series, Mahwah, NJ: Erlbaum.

Yellott, J. (1977). The relationship between Luce's choice axiom, Thurstone's theory of comparative judgment, and the double experimental distribution. *Journal of Mathematical Psychology, 15,* 109-144.

Figure 1. Example of proper perceptual patch X from 2 dimensional physical stimulus space to a hemisperical manifold in psychological space, with coordinate map Y from psychological coorsinate space to the same manifold. The induced mapping (ϕ, ψ) to the coordinate space $\Phi \times \Psi$ is not neccesarily Cartesian, and is not necessarily onto $\Phi \times \Psi$.

Figure 2. X has deterministic domain-range parametrized perceptual separability (DDR-PPS) when the diagram above exists and commutes. Here, all maps are onto, the horizontal maps are the projections onto the Cartesian components and $Y\colon \Phi \times \Psi \to M$ exists such that $X = Y \circ (\phi \times \psi)$.

Figure 3.(a) X is a deterministic domain-range embedding if for some *f:* $\Re^2 \to \Re$ the diagram exists and commutes, where ϕ is defined on U and ψ is defined on V. (b) Example of an embedding in which the 2-D coordinate space is mapped to a surface in 3-space.

Figure 4. Decisional separability of ϕ from ψ exists if the above commutative diagram exists for i = 1 and 2 and $\Phi_1 = \{\phi\colon \phi \leq \phi_0\}$ and $\Phi_2 = \{\phi\colon \phi \geq \phi_0\}$and where $M_1 \cup M_2 \cup B = M$ and $M_1 \cap M_2 = \varnothing$. B is the decision boundary and response a_i is observed for stimulus *(u, v)* whenever *X(u, v)* $\in M_i$.

Figure 5. Example of patch mapping to a conic surface of revolution. Equi-sized squares in coordinate space map to different sized areas (2D volumes) in psychological space M, due to the distortion factor of $\sqrt{2}\phi$.

Figure 6. Unnormalized densities on $\Phi \times \Psi$ plane given stimulus (ϕ_0, ψ_0).(a) and (b): $p(\phi, \psi|u, v) = \phi \exp(-\alpha((\phi - \phi_0)^2 + (\phi\cos(\psi) - \phi_0\cos(\psi_0))^2 + (\phi\sin(\psi) - \phi_0\sin(\psi_0))^2))$. (c) and (d):$p(\phi, \psi|u, v) = \exp(-3\alpha((\phi - \phi_0)^2 + (\psi - \psi_0)^2))$. For (a) and (c): (ϕ_0, ψ_0) = *(0.7, 3)*; for (b) and (d): (ϕ_0, ψ_0) = *(0.3, 1)*.

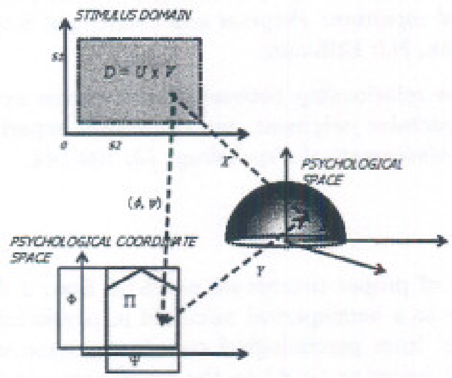

Figure 1: Example of proper perceptual patch X from 2 dimensional physical stimulus space to a hemisperical manifold in psychological space, with coordinate map Y from psychological coorsinate space to the same manifold. The induced mapping (ϕ, ψ) to the coordinate space $\Phi \times \Psi$ is not neccesarily Cartesian, and is not necessarily onto $\Phi \times \Psi$.

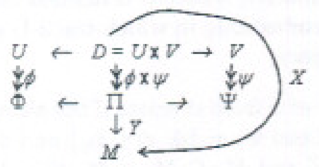

Figure 2: X has deterministic domain-range parametrized perceptual separability (DDR-PPS) when the diagram above exists and commutes. Here, all maps are onto, the horizontal maps are the projections onto the Cartesian components and $Y\colon \Phi \times \Psi \to M$ exists such that $X = Y \circ (\phi \times \psi)$.

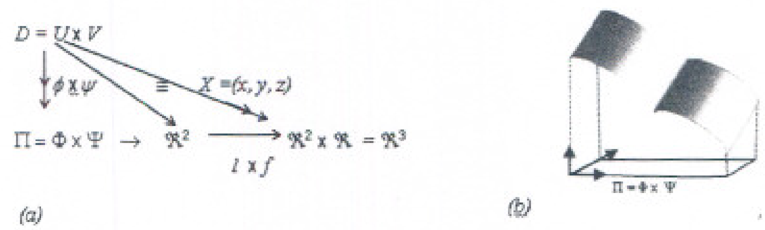

Figure 3: (a) X is a deterministic domain-range embedding if for some f: $\Re^2 \rightarrow \Re$ the diagram exists and commutes, where ϕ is defined on U and ψ is defined on V. (b) Example of an embedding in which the 2-D coordinate space is mapped to a surface in 3-space.

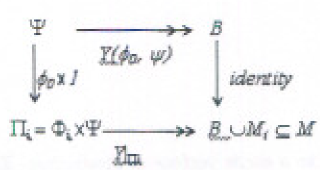

Figure 4: Decisional separability of ϕ from ψ exists if the above commutative diagram exists for i = 1 and 2 and $\Phi_1 = \{\phi\colon \phi \leq \phi_0\}$ and $\Phi_2 = \{\phi\colon \phi \geq \phi_0\}$and where $M_1 \cup M_2 \cup B = M$ and $M_1 \cap M_2 = \varnothing$. B is the decision boundary and response a_i is observed for stimulus (u, v) whenever $X(u, v) \in M_i$.

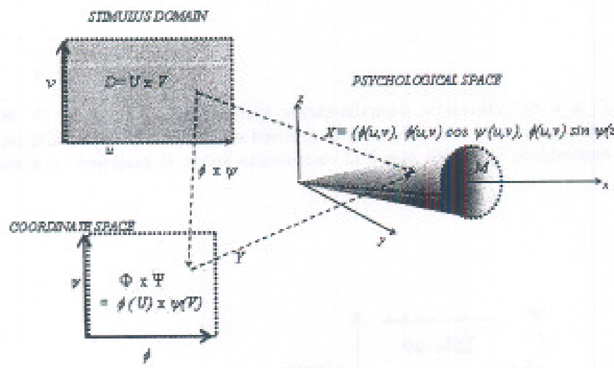

Figure 5: Example of patch mapping to a conic surface of revolution. Equi-sized squares in coordinate space map to different sized areas (2D volumes) in psychological space M, due to the distortion factor of $\sqrt{2}\phi$.

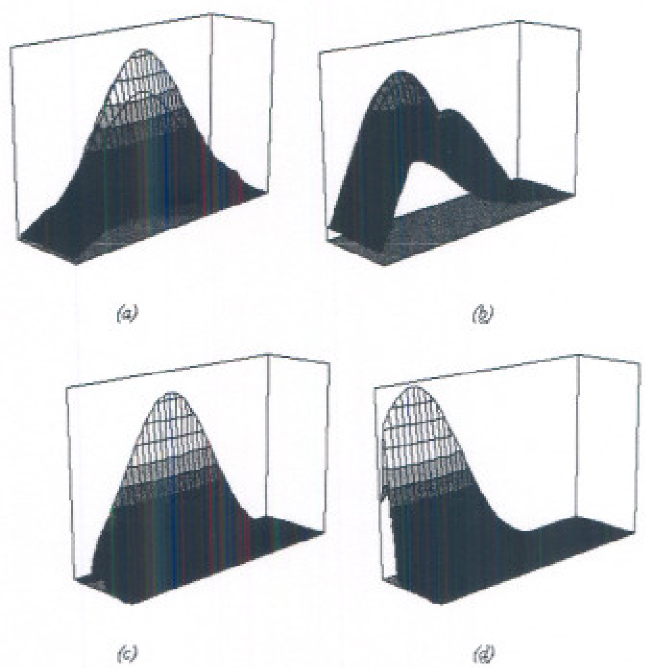

Figure 6: Unnormalized densities on $\Phi \times \Psi$ plane given stimulus (ϕ_0, ψ_0).(a) and (b): $p(\phi, \psi | u, v) = \phi \exp(-\alpha((\phi - \phi_0)^2 + (\phi \cos(\psi) - \phi_0 \cos(\psi_0))^2 + (\phi \sin(\psi) - \phi_0 \sin(\psi_0))^2))$. (c) and (d):$p(\phi, \psi | u, v) = \exp(-3\alpha((\phi - \phi_0)^2 + (\psi - \psi_0)^2))$. For (a) and (c): $(\phi_0, \psi_0) =$ *(0.7, 3)*; for (b) and (d): $(\phi_0, \psi_0) =$ *(0.3, 1)*.