TOWARD A NEW PSYCHOPHYSICS BASED ON GEOMETRY AND STOCHASTIC SYSTEMS THEORY

James T. Townsend, Joseph W. Houpt, Noah H. Silbert and Nicholas A. Altieri
Department of Psychological and Brain Sciences, Indiana University

Abstract

Our goal in this paper is to develop a global, long-term program for psychophysics. We outline an approach to based on differential geometry and stochastic systems theory. This approach unifies both the temporal aspects of perception as well as the possibly complex multidimensional spaces underlying a perceptual process. We then take a specific example of this approach, RTGRT, which is an extension of the static (geometric) General Recognition Theory model to the (noisy) matching of perceptual information to stored memory and the resultant accrual of evidence over time. The static notions of perceptual and decisional interactions are given dynamic interpretations, and associated formalisms are presented.

Ever since the brilliant innovations by our psychophysical forefathers, scientific leaders such as Fechner, Weber, Helmholtz and many others, psychophysics has been both concerned with what we might call "geometric" descriptions as well as static and dynamic mechanisms. The former is perhaps most evident in Fechner's ground-breaking set of psychophysical methods (Elemente der Psychophysik) and the latter are very much apparent in such masterpieces as Helmholtz's authoritative tomes, "On the Sensation of Tones" and "Physiological Optics".

Today, both of these approaches are alive and healthy, with abundant examples reaching out to applied areas in addition to basic research regions. In fact, many past and present members of ISP have made profound contributions to one, often both of these pathways, or indeed, to developing a static or dynamic mechanistic underpinning for descriptive measures such as psychophysical functions.

By and large, it is probably fair to say that psychophysics and its cousin psychometrics, has tended to seek more or less stable and static depictions of psychophysical constructs, especially in a geometric sense. This claim applies to most psychophysical functions, factor analysis, multi-dimensional scaling, much test theory and even the bulk of statistics, particularly multivariate statistics. Naturally, one can introduce a time-dynamic element through times-series analysis, and in fairly recent years, path or causal analysis among other temporally-based avenues. However, it appears that many of these unprincipled 'models' are too complex to be clearly identifiable in typical data analyses. Principled psychological models may provide an antidote.

Yet, it is a challenging goal to invent plausible and testable models even for elementary psychophysical functions. Is there, after almost 150 years beyond the "Elements", a satisfactory unified model for all of Fechner's various methods including how they relate to one another? It is even more daunting to provide good models for dynamic underpinnings and possibly complex, multidimensional psychophysical phenomena. In these and other cases, changing what might seem to be minor aspects of a task, anything about the stimuli, response assignments and so on, can lead to vastly different results. The terrain is strewn with obstacles so we must be resourceful and creative, but be wary of shortcuts or simplistic courses of action such as the popular but often shallow "effect discovery" stratagem.

In the following, we will outline a program for psychophysics that encompasses what we think are essential in a global and long-term sense.

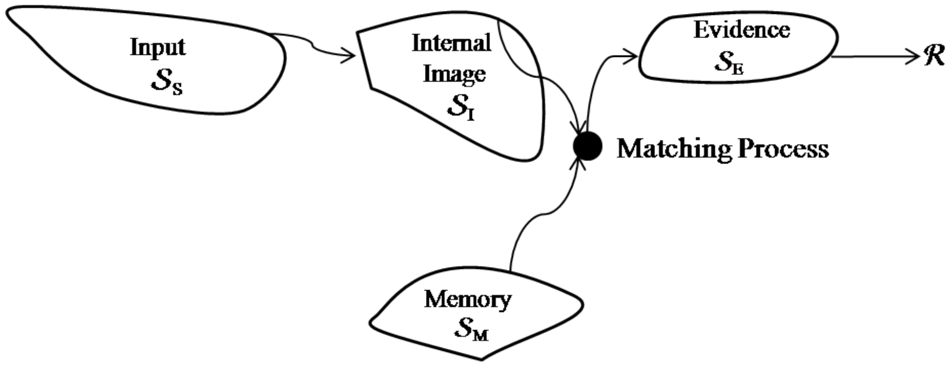

Figure 1: A geometric depiction of a stimulus presentation, which evokes internal processing and the subsequent mapping of the image to a template in memory governed by a matching process. Evidence accumulates toward one response category and a response is made.

What kind of approach can enfold both spatial and temporal characteristics? The answer lies in pure and applied mathematics that has mainly been accomplished since 1950, although as usual the roots tunnel back centuries. Basically, we wish to assess the accumulation of sensory and perceptual information over time and, especially pertinent for traditional psychophysics, temporally stable (to the extent feasible) sensations and percepts. Contemporary psychophysics, as witnessed by ISP Proceedings in recent years, is moving to embrace psychological and physiological regions once treated as 'off-limits' by conventional psychophysicists. We speak, for instance, of short-term and even long-term memory search, facets of working memory, subitizing, attentional phenomena, aesthetics, scaling of almost everything imaginable, and so on. Such adventures of course force the enlistment of higher-order processes, such as semantics.

So, we begin with some kind of trajectory, but a trajectory of what? It is our thesis that in virtually every case, the early trajectory, which undoubtedly engages the first-order stations in the nervous system, must involve high dimensional, even infinite dimensional, spaces. Consider the ecologically rare case of, say, a single-frequency tone being presented at superthreshold intensity. Even that simplest of signals is characterized by at least three potentially perceptible dimensions, phase, amplitude and frequency. Further, during the early phases, transient noise renders the signal into a function in infinite dimensional function space. Recall that white noise, for example, though continuous, possesses an infinite number of points where no derivative exists. When the set of all sounds, even nice continuous ones occurring in a finite temporal interval (t0, t1 are taken into account, we find ourselves ensconced in an infinite dimensional space of functions of that interval into the real numbers.

Perhaps an even more cogent example is the space of all faces. Probably this and other object spaces can be specified by a very large number of parameters (e.g., a set of eigenfaces as suggested by O'Toole, Wenger, & Townsend, 2001) , but just as in physics and engineering, it typically makes for more elegant science, to simply work within an infinite dimensional function space to begin with. Not only is this the approach taken by John von Neumann and Norbert Wiener when developing mathematical communications theory but Stephen Hawking (Hawking & Penrose, 1996) argues that a continuum representation of reality is still the best way to go even in the face of the well-known quantum revolution.

So, we start with a mapping of a function into some space. There will actually be a number of spaces and their exact nature will depend on whether we are depicting the physiology and anatomy or the psychological systems, although there hopefully will ultimately exist a translational map between them. Suppose for simplicity that there is a stage where an image is developing over time in the sensorium (this stage could involve several anatomical sites) and a stage where short-term or long-term memory is being accessed in order to identify or categorize the incoming input. Let the first be called $\mathcal{S}_I$ and the second $\mathcal{S}_M$. Each of these spaces can contain what is representable as an infinite number of entities, but we do assume that memory is composed of a discrete set of items, even if infinite. We will suppose that a matching process is transpiring between them in real time. We cannot get into the details of whether, for instance, this occurs in parallel (almost certainly) and whether it is holographic or a more standard kind of processing, but we do suppose that $\{\mathcal{S}_I\} \times \{\mathcal{S}_M\} \times T$, where $T = [0, \infty)$ represents the time axis.

Using this notation, we can model the internal image as a stochastic process, $\{\mathbf{i}(t), t \geq 0\}$ which develops in $\mathcal{S}_i$ based on the stimulus. The memory retrieval process is also modeled as a stochastic process, $\{\mathbf{m}(t), t \geq 0\}$. The internal image and memory are then mapped into the evidence space $\mathcal{S}_E$, based on a matching process such as a correlation. Hence, the evidence is also a stochastic process, $\{\mathbf{e}(t), t \geq 0\}$, where $\mathbf{e}(t) = f(\mathbf{i}(t), \mathbf{m}(t))$. This system is depicted in Figure 1.

In the next section we present an exemplar of this approach, RTGRT. RTGRT is a simplification of the general model outlined above since specific assumptions about the processes and geometry of the decision are included.

RTGRT

General Recognition Theory (GRT) was originally created for, and continues to be employed in, the investigation of relationships between dimensions in perception and decision making (Ashby & Townsend, 1986; Kadlec & Townsend, 1992; Maddox, 1992; Thomas, 2001). The initial theory was static, as have been most ensuing theoretical efforts and empirical applications. Nonetheless, later theoretical developments have pursued dynamic expansions of GRT. For example, Ashby (1989; 2000) developed stochastic systems extensions of a so-called "decision bound" theory GRT categorization model. We describe here a stochastic extension of a feature-complete factorial GRT identification model.

In adapting the static, deterministic GRT to a stochastic, dynamic model, there are two possible approaches. One is to simply posit distributions on finishing times. Another is to follow the approach described above, by including an information processing mechanism that predicts both the response and the time it takes to make that response. In doing so, it is difficult to avoid explicit or implicit assumptions about the processing architecture, for instance, serial, parallel, or more complex. Here we assume a parallel processing architecture, but this is not the only possibility.

The basic model, depicted in Figure 2, is formalized as follows. A stimulus S is operated on by two channels, X and Y in parallel. That part of the stimulus operated on by X is S_X, while the part of the stimulus acted upon by Y is S_Y. The internal image corresponding to X is denoted $\mathbf{i}_X(t)$ and that corresponding to Y is denoted $\mathbf{i}_Y(t)$. In the matching process, the internal image is then combined with an internal memory, or set of memories to produce evidence $\mathbf{e}_X(t)$ for the possible alternative interpretations of S_X and similarly for Y. As stated earlier, this matching process could be based on a correlation or on some type of distance metric between the internal image and the memory. At this point we leave the details of this process unspecified. For clarity, we assume that there are two alternatives, so $\mathbf{e}_X(t)$ and $\mathbf{e}_Y(t)$ are each two dimensional. Superscripts are used to denote dimension, so $\mathbf{e}_X^{(1)}(t)$ is the evidence for the first alternative.

In the general case, where non-separability (across trials) and stochastic dependence (within trials) could be present, the state space random process for evidence on the X channel

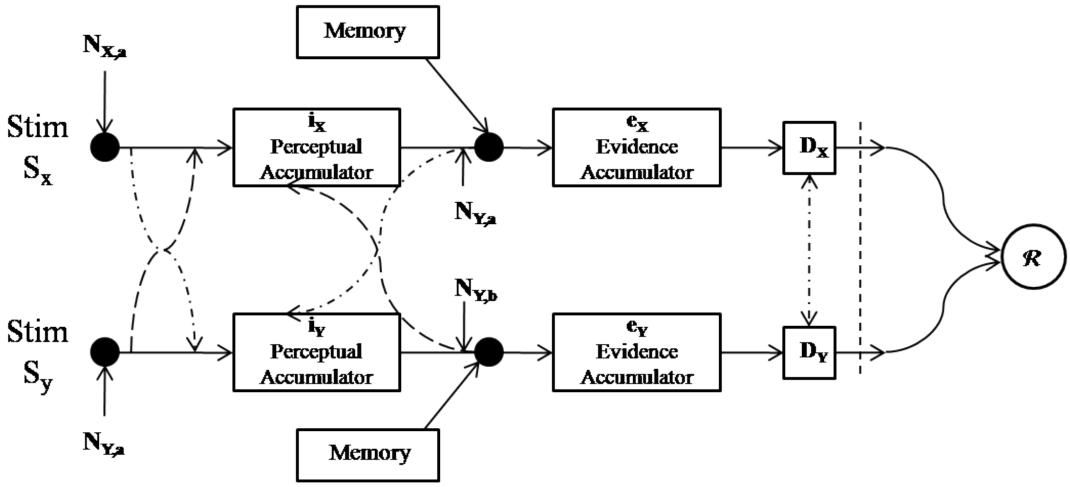

Figure 2: A linear systems instantiation of the psychological processes at work in the GRT paradigm. Each channel represents a separate stimulus dimension, which accepts different components of the input, S_X and S_Y. White noise, denoted by $\mathbf{N}_X$ and $\mathbf{N}_Y$, may also enter the system and can affect processing at different stages. Cross channel interactions allow processing in one channel to affect processing in the other and the noise may or may not be correlated. As evidence accumulates in each channel a decision is formulated prior to the response output.

can depend on both the stimulus acted upon by X and Y, the internal image in X and Y and the evidence in Y, so $\mathbf{e}_X(t) = f(S_X, S_Y, \mathbf{i}_X(t), \mathbf{i}_Y(t), \mathbf{e}_Y(t))$. The same holds true for Y, *mutatis mutandis*.

Decisions are made on each channel, based on the evidence in that channel, so are a mapping from evidence space to a response. In this model, decisions are made based on the dimension of the evidence channel that reaches its threshold first. If $C_X = \binom{C_X^{(1)}}{C_X^{(2)}}$ is the threshold and $T_X^{(1)}$ is the first passage time of the evidence on dimension 1, then $\mathbf{r}_X = 1$ if $\mathbf{e}_X^{(1)}(t) \geq C_X^{(1)}$ before $\mathbf{e}_X^{(2)}(t) \geq C_X^{(2)}$, i.e. $T_X^{(1)} < T_X^{(2)}$.

In more general circumstances, the criteria can be random functions themselves and/or depend on S, X, Y, and t and even the other criteria. However, since it seems non-causal for the decision criteria to depend on the stimulus, and given the already formidable notational complexity, we make the assumption that the "Cs" are deterministic. Thus, we can write for example, $C_X^{(1)}(\mathbf{e}_X(t), \mathbf{e}_X(t), t, C_X^{(2)}, C_Y^{(1)}, C_Y^{(2)})$.

The formula for the joint distribution of the T_X and T_Y, the completion times of X and Y, along with two (out of the four) responses is developed as follows. Let T stand for the overall processing time, aside from the base time. In the present class of models, T will designate the time required for both channels to complete processing, so $T = \max(T_X, T_Y)$. Then the joint distribution of interest is $\Pr\{(\mathbf{r}_X = r_X, \mathbf{r}_Y = r_Y) \cap T \leq t\}$.

Now, to specialize the model further, we assume for the sake of simplicity, and without loss of generality, that $(\mathbf{r}_X = 1, \mathbf{r}_Y = 1)$, and further that the total completion time is less than t. That is, $T \leq t$. Hence, we have

$$\begin{aligned}
&\Pr\{(\mathbf{r}_X = 1, \mathbf{r}_Y = 1) \cap T \leq t\} \\
&\quad = \Pr\left\{\max\left[T_X = \min\left(T_X^{(1)}, T_X^{(2)}\right) \cap T_Y = \min\left(T_Y^{(1)}, T_Y^{(2)}\right)\right] \leq t | S_X, S_Y\right\} \\
&\quad = \int_0^t \int_0^t \big[\Pr\{\mathbf{e}_X^{(1)}(t'') = C_X^{(1)} \cap \mathbf{e}_X^{(2)}(t'') < C_X^{(2)} \\
&\qquad\qquad \cap \mathbf{e}_Y^{(1)}(t') = C_Y^{(1)} \cap \mathbf{e}_Y^{(2)}(t') < C_Y^{(2)} | S_X, S_Y\big] dt' dt'' \qquad (1)
\end{aligned}$$

To produce the complete RT distribution it is necessary to convolve an independent base time distribution with the above.

Using this formalism, we can discuss the extensions of two important constructs from static GRT, perceptual separability and perceptual independence.

Perceptual separability captures the idea that the perceptual effect of one component of the stimulus does not depend on the level of the other component. Fore example, perceptual non-separability occurs in static GRT if changing the hue from red to purple affects the marginal distribution on the perceptual variable associated with stimulus shape. Importantly, perceptual separability is defined across stimuli. On the other hand, perceptual independence in static GRT is defined as stochastic independence of perceptual effects within a given stimulus. These concepts, although similar, are logically distinct in static GRT.

We define perceptual separability in information activation state-space models, of say, the X channel at, say, stimulus level $S_X = 1$ as invariance of the marginal processes $(\mathbf{i}_X^{(1)}, \mathbf{i}_X^{(2)})$ over variations of S_Y. For this model, this implies we may drop the S_Y for both $\mathbf{i}_X(t)$ and since the stimulus only affects the evidence space via the internal image, for $\mathbf{e}_X(t)$ as well. Stochastic independence of the X channel on activity in the Y channels can occur at different stages. Two possible places in which noise could enter thee system are depicted by $\mathbf{N}_a$ and $\mathbf{N}_b$ in Figure 2. Although it is not depicted in the figure, a possible source of $\mathbf{N}_b$ could be noisy memory. Stochastic dependence at any point in the system induces dependence at later stages in the system. So for example, if the only randomness is due to $\mathbf{N}_b$ ($\mathbf{N}_a = 0$), then the internal images will be stochastically independent but the evidence will not be. If there is early noise, as in $\mathbf{N}_a$, then both the internal images and the evidence will not be perceptually independent. In sum, theoretical instantiation of perceptual separability removes the S_Y influence, and stochastic independence takes away any stochastic dependence, leaving the much simplified $\mathbf{e}_X^{(1)}\left(t, \mathbf{e}_X^{(2)}, S_X, \mathbf{e}_Y^{(1)}, \mathbf{e}_Y^{(2)}, S_Y\right) = \mathbf{e}_X^{(1)}\left(t, \mathbf{e}_X^{(2)}, S_X\right)$.

These concepts can be demonstrated in terms of the model in Figure 2. First, if there is information is shared between the internal image processes $\mathbf{i}_X(t)$ and $\mathbf{i}_Y(t)$, whether it be before or after accumulation, then perceptual separability fails. These interactions are illustrated by the dashed lines between the two channels. If the only source of noise is $\mathbf{N}_b$, and $\mathbf{N}_{X,b}$ and $\mathbf{N}_{Y,b}$ are independent, then perceptual independence will hold regardless of the presence of the dashed lines. However, if there is noise at $\mathbf{N}_a$ along with the presence interactions, then perceptual separability will fail as well.

Non-separability could also come about indirectly through dependence of the X channel on the Y channel, such as if $\mathbf{N}_{X,a}$ were correlated with $\mathbf{N}_{X,b}$. In the original static GRT, if the bivariate distributions were Gaussian, then of course, the marginal distributions were invariant across differing covariances. However, in general, the dependence could be associated with non-separability of the marginal distributions, even in the static case.

Conclusion

In conclusion, our approach emphasizes a synergistic merging of the spatial (e.g., topological, geometric, etc.) facets, with the dynamic, information processing structures. Emphasis of the former without the latter too often ignores 'how we got there' whereas the latter without the former too often ignores spatial and indeed, any aspect of the information code. We anticipate that in either the static or dynamic data and theoretical depictions, the geometries and metrics currently available in our theoretical and methodological armamentarium will be insufficient, but time will tell. In addition to the potential offered in infinite dimensional Riemannian manifolds (Townsend, Solomon, Wenger, & Spencer-Smith, 2001) we suspect that general topological approaches such as that of Dzhafarov & Colonius (2007) that can assume a dissimilarity measure which doe not necessarily meet all the requirements of a metric and can do without specified dimensions, will also be increasingly useful.

References

Ashby, F. G. (1989). Stochastic general recognition theory. In D. Vickers & P. L. Smith (Eds.), *Human information processing: Measures, mechanisms, and models* (p. 435-457). Amsterdam: Elsevier.

Ashby, F. G. (2000). A stochastic version of general recognition theory. *Journal of Mathematical Psychology*, 44, 310-329.

Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review*, 91(2), 154-179.

Dzhafarov, E., & Colonius, H. (2007). Dissimilarity cumulation theory and subjective metrics. *Journal of Mathematical Psychology*, 51, 290-304.

Hawking, S., & Penrose, R. (1996). *The nature of space and time*. New Jersey: Princeton University Press.

Kadlec, H., & Townsend, J. T. (1992). Implications of marginal and conditional detection parameters for the separabilities and independence of perceptual dimensions. *Journal of Mathematical Psychology*, 36, 325-374.

Maddox, W. T. (1992). Perceptual and decisional separability. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (p. 147-180). New Jersey: Lawrence Erlbaum Associates.

OToole, A., Wenger, M., & Townsend, J. (2001). Quantitative models of perceiving and remembering faces: Precedents and possibilities. In *Computational, geometric and process issues in facial cognition: Progress and challenges* (pp. 138). Mahawah, NJ: Erlbaum Press.

Thomas, R. D. (2001a). Characterizing perceptual interactions in face identication using multidimensional signal detection theory. In M. J. Wenger & J. T. Townsend (Eds.), *Computational, geometric, and process perspectives on facial cognition* (p. 193-227). New Jersey: Lawrence Erlbaum Associates.

Thomas, R. D. (2001b). Perceptual interactions of facial dimensions in speeded classication and identication. *Perception & Psychophysics*, 63(4), 625-650.

Townsend, J. T., Solomon, B., Wenger, M. J., & Spencer-Smith, J. B. (2001). The perfect gestalt: Innite dimensional riemannian face spaces and other aspects of face cognition. In M. J. Wenger & J. T. Townsend (Eds.), *Computational, geometric and process issues in facial cognition: Progress and challenges*. Mahawah, NJ: Erlbaum Press.