MEASUREMENT AND REPRESENTATION OF SENSATIONS

Edited by

Hans Colonius
Universität Oldenburg

Ehtibar N. Dzhafarov
Purdue University

Camera ready copy for this book was provided by the editors.

Lawrence Erlbaum Associates, Inc., Publishers
10 Industrial Avenue
Mahwah, New Jersey 07430
www.erlbaum.com

Cover design by Tomai Maridou

Library of Congress Cataloging-in-Publication Data

Measurement and representation of sensations / Hans Colonius and Ehtibar N. Dzhafarov, editors.
p. cm.

Includes bibliographical references and index.

ISBN 0-8058-5353-7 (alk. paper)
1. Sensory stimulation—Measurement. 2. Sensory discrimination. 3. Senses and sensations. I. Colonius, Hans, II. Dzhafarov, Ehtibar N.

QP435.M285 2006
152.1—dc22 2005052082
CIP

Books published by Lawrence Erlbaum Associates are printed on acid-free paper, and their bindings are chosen for strength and durability.

Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

1
Regular Minimality: A Fundamental Law of Discrimination

Ehtibar N. Dzhafarov[1] and Hans Colonius[2]

[1] Purdue University
[2] Universität Oldenburg

1. INTRODUCTION

The term *discrimination* in this chapter is understood in the meaning of *telling stimuli apart*. More specifically, it refers to a process or ability by which a perceiver judges two stimuli to be different or identifies them as being the same (overall or in a specified respect). We postpone until later the discussion of the variety of meanings in which one can understand the terms *stimuli*, *perceiver*, and *same–different judgments*. For now, we can think of discrimination as pertaining to the classical psychophysical paradigm in which stimuli are being chosen from a certain set (say, of colors, auditory tones, or geometric shapes) two at a time, and presented to an observer or a group of observers who respond by saying that the two stimuli are the same, or that they are different. The response to any given pair of stimuli $(\mathbf{x}, \mathbf{y})$ in such a paradigm can be viewed as a binary random variable whose values (same–different) vary, in the case of a single observer, across the potential infinity of replications of this pair, or, in the case of a group, across the population of observers the group represents. As a result, each stimulus pair $(\mathbf{x}, \mathbf{y})$ can be assigned a certain probability, $\psi(\mathbf{x}, \mathbf{y})$, with which a randomly chosen response to $\mathbf{x}$ and $\mathbf{y}$ (paired in this order) is "the two stimuli are different,"

$$\psi(\mathbf{x}, \mathbf{y}) = \Pr[\mathbf{x} \text{ and } \mathbf{y} \text{ are judged to be different}]. \tag{1}$$

The empirical basis for considering $(\mathbf{x}, \mathbf{y})$ as an ordered pair, distinct from $(\mathbf{y}, \mathbf{x})$, is the same as for considering $(\mathbf{x}, \mathbf{x})$ as a pair of two identical stimuli rather than a single stimulus. Stimuli $\mathbf{x}$ and $\mathbf{y}$ presented to a perceiver for comparison are necessarily different in some respect, even when one refers to them as being physically identical and writes $\mathbf{x} = \mathbf{y}$: thus, $\mathbf{x}$ (say, a tone) may be presented first and followed by $\mathbf{y}$ (another tone, perhaps otherwise identical to $\mathbf{x}$); or $\mathbf{x}$ and $\mathbf{y}$ (say, aperture colors) may be presented side-by-side, one on the left, the other on the right. Dzhafarov (2002b) introduced the term *observation area* to reflect and generalize this distinction: two stimuli being compared belong to two distinct observations areas (in the examples just given, spatial locations, or ordinal positions in time). This seemingly trivial fact plays a surprisingly prominent role in the theory of perceptual discrimination. In particular, it underlies the formulation of the law of Regular Minimality, on which we focus in this chapter.

There is more to the notion of an observation area than the difference between spatiotemporal locations of stimuli, but this need not be discussed now. Formally, we refer to $\mathbf{x}$ in $(\mathbf{x}, \mathbf{y})$ as belonging to the *first observation area*, and to $\mathbf{y}$ as belonging to the *second observation area*, the adjectives "first" and "second" designating the ordinal positions of the symbols in the pair rather than the chronological order of their presentation. The difference between the two observation areas, whatever their physical meaning, is always perceptually conspicuous, and the observer is supposed to ignore it: thus, when asked to determine whether the stimulus on the left (or presented first) is identical to the stimulus on the right (presented second), the observer would normally perceive two stimuli rather than a single one, and understand that the judgment must not take into account the difference between the two spatial (or temporal) positions. In the history of psychophysics, this aspect of discrimination has not received due attention, although G. T. Fechner did emphasize its importance in his insightful discussion of the "non-removable spatiotemporal non-coincidence" of two stimuli under comparison (1887, p. 217; see also the translation in Scheerer, 1987).

It should be noted that the meaning of the term *discrimination*, as used by Fechner and by most psychophysicists after him, was different from ours. In this traditional usage, the notion of discrimination is confined to semantically unidimensional attributes (such as loudness, brightness, or attractiveness) along which two stimuli, $\mathbf{x}$ and $\mathbf{y}$, are compared in terms of which of them contains more of this attribute (greater–less judgments, as opposed to same–different ones). Denoting this semantically unidimensional attribute by $\mathcal{P}$, each ordered pair $(\mathbf{x}, \mathbf{y})$ in this paradigm is assigned probability $\gamma(\mathbf{x}, \mathbf{y})$, defined as

$$\gamma(\mathbf{x}, \mathbf{y}) = \Pr[\mathbf{y} \text{ is judged to be greater than } \mathbf{x} \text{ with respect to } \mathcal{P}]. \quad (2)$$

As a rule, although not necessarily, subjective attribute $\mathcal{P}$ is being related to its "physical correlate," a physical property representable by an axis of nonnegative reals (e.g., sound pressure, in relation to loudness). In this case, stimuli $\mathbf{x}, \mathbf{y}$ can be identified by values x, y of this physical property, and probability $\gamma(\mathbf{x}, \mathbf{y})$ can be written as $\gamma(x, y)$.[3] The physical correlate is always chosen so that $y \to \gamma(x, y)$; (i.e., function γ considered as a function of y only, for a fixed value of x) is a strictly increasing function for any value of x, as illustrated in Fig. 1, left. Clearly then, $x \to \gamma(x, y)$ is a strictly decreasing function for any value of y. Note, in Fig. 1 (left), the important notion of a *Point of Subjective Equality* (PSE). The difference between x, in the first observation area, and its PSE in the second observation area, is sometimes called the *constant error* associated with x (the term "systematic error" being preferable, because the difference between x and its PSE need not be constant in value across different values of x). The systematic error associated with y, in the second observation area, is defined analogously.

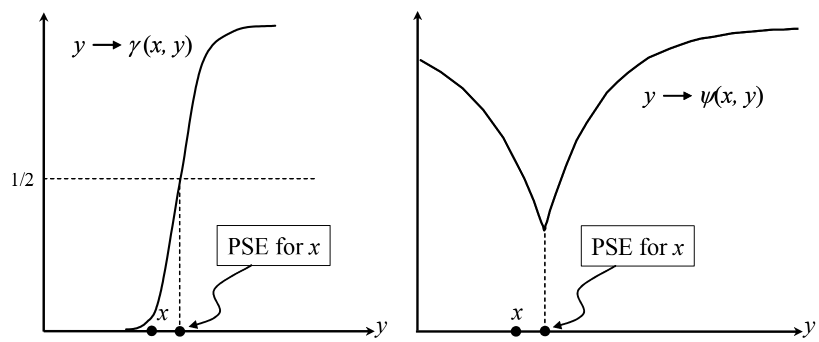


Fig. 1: Possible appearance of discrimination probability functions $\gamma(x, y) = \Pr[y \text{ is greater than } x \text{ in attribute } \mathcal{P}]$ (left) and $\psi(x, y) = \Pr[x \text{ is different from } y]$ (right), both shown for a fixed value of x, with x and y represented by real numbers (unidimensional stimuli). For $\gamma(x, y)$, the median value of y is taken as the Point of Subjective Equality (PSE) for x (with respect to $\mathcal{P}$). For $\psi(x, y)$, PSE for x is the value of y at which $\psi(x, y)$ achieves its minimum.

Same–different discrimination also may involve a semantically unidimensional attribute (e.g., "do these two tones differ in loudness?"), but it does not have to: the question can always be formulated "generically": are the two stimuli different (in anything at all, ignoring however the difference

[3] Here and throughout, we use boldface lowercase letters to denote stimuli, and lightface lowercase letters when dealing with their real-number attributes; by convenient abuse of language, however, we may refer to "stimulus x" in place of "stimulus $\mathbf{x}$ with value x."

between the observation areas). It is equally immaterial whether stimuli $\mathbf{x}, \mathbf{y}$ can be represented by real numbers, vectors of real numbers, or any other mathematical construct: physical measurements only serve as labels identifying stimuli. For convenience of graphical illustrations, however, we will assume in the beginning of our discussion that $\mathbf{x}, \mathbf{y}$ are matched in all respects except for a unidimensional physical attribute (so they can be written x, y). In such a case, discrimination probability function might look as shown in Fig. 1, right. The important notion of PSE here acquires a new meaning: for x, in the first observation area, its PSE is the stimulus in the second observation which is least discriminable from x (and analogously for PSE for y in the second observation area). That such a point exists is part of the formulation of the Regular Minimality principle.[4]

Our last introductory remark relates to a possible confusion in understanding of functions $\mathbf{y} \to \psi(\mathbf{x}, \mathbf{y})$ and $\mathbf{x} \to \psi(\mathbf{x}, \mathbf{y})$; (this remark equally applies to functions $\mathbf{y} \to \gamma(\mathbf{x}, \mathbf{y})$ and $\mathbf{x} \to \gamma(\mathbf{x}, \mathbf{y})$ for greater–less discriminations). The mathematical meaning of $\mathbf{y} \to \psi(\mathbf{x}, \mathbf{y})$, for example, is that $\mathbf{x}$ is being held constant whereas $\mathbf{y}$ varies, with ψ varying as a function of $\mathbf{y}$. It is important to keep in mind that whenever we use such a construction, the distinction between $\mathbf{x}$ and $\mathbf{y}$ is purely conceptual, and not procedural: it is *not* assumed that $\mathbf{x}$ is being held constant physically within a certain block of trials whereas $\mathbf{y}$ changes from one trial to another. To emphasize this fact, we often refer to $\mathbf{y} \to \psi(\mathbf{x}, \mathbf{y})$ and $\mathbf{x} \to \psi(\mathbf{x}, \mathbf{y})$ as *cross-sections* of function $\psi(\mathbf{x}, \mathbf{y})$, made at a fixed value of $\mathbf{x}$ or $\mathbf{y}$, respectively. The ideal procedure our analysis pertains to involves all possible pairs $(\mathbf{x}, \mathbf{y})$ being presented with equal likelihoods and with no sequential dependences. All necessary and optional deviations from this ideal procedure are only acceptable under the assumption (more easily stated than tested) that they yield discrimination probabilities $\psi(\mathbf{x}, \mathbf{y})$ which approximate those obtainable by means of the ideal procedure. Among necessary deviations from the ideal procedure, the most obvious one is that we have to use samples of $(\mathbf{x}, \mathbf{y})$ pairs with a finite number of replications per pair, rather than all possible pairs of stimuli of a certain type replicated infinite number of times each. Among optional deviations, we have various partial randomization schemes (including, as a marginal case, blocking trials with constant $\mathbf{x}$ or $\mathbf{y}$). One should contrast this understanding with Zhang's analysis (2004; see also Zhang's chapter in this volume) of the situations where $\psi(\mathbf{x}, \mathbf{y})$ critically depends on the blocking of constant-$\mathbf{x}$ or constant-$\mathbf{y}$ trials, or on which of

[4] The reason $y \to \psi(x, y)$ in Fig. 1 (right) is drawn with a "pencil-sharp" rather than rounded minimum is that the latter can be shown (Dzhafarov, 2002b, 2003a, 2003b; Dzhafarov & Colonius, 2005a) to be incompatible with the conjunction of Regular Minimality and Nonconstant Self-Dissimilarity, discussed later.

the two stimuli in a trial is semantically labeled as the "reference" to which the other stimulus is to be compared.

2. REGULAR MEDIALITY

It is useful for our discussion to stay a while longer with the greater–less discrimination probabilities, to formulate a principle which is analogous to Regular Minimality but has a simpler mathematical structure. Refer to Figs. 2 and 3. Think, for concreteness, of x, y being independently varying lengths of two otherwise identical horizontal line segments presented side-by-side, x on the left, y on the right; γ being the probability of judging y longer than x.

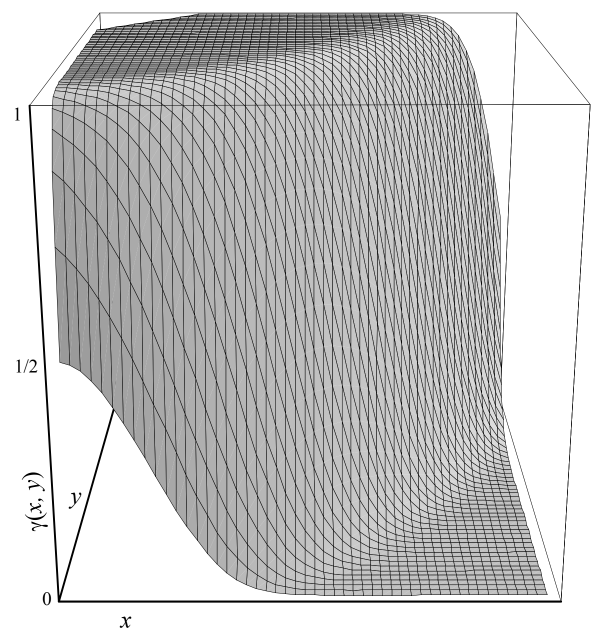


Fig. 2: Possible appearance of psychometric function $\gamma(x, y)$ for unidimensional stimuli. (This particular function was generated by a classical Thurstonian model in which x and y are mapped into independent normally distributed random variables whose means and variances change as functions of these stimuli.)

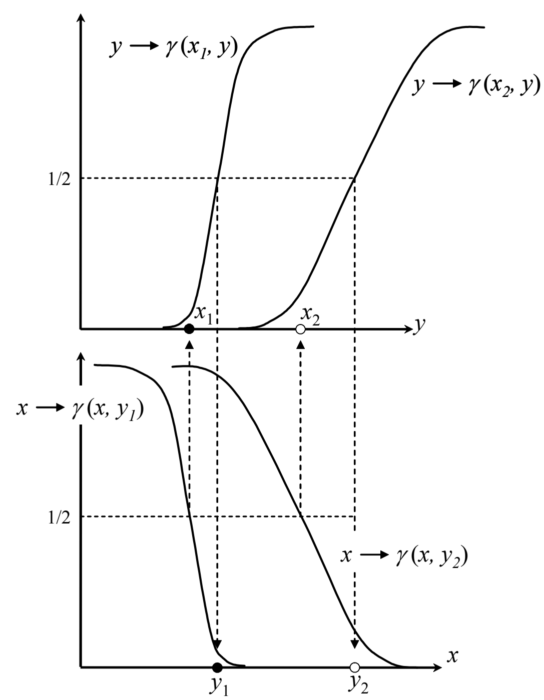


Fig. 3: Cross-sections of psychometric function $\gamma(x, y)$ shown in Fig. 2 made at two fixed values of x (upper panel) and two fixed values of y (lower panel). The figure illustrates the Regular Mediality principle for greater–less discriminations: y is the Point of Subjective Equality (PSE) for x if and only if x is the PSE for y. Thus, $\gamma(x_1, y)$ achieves level $\frac{1}{2}$ at $y = y_1$, and this is equivalent to $\gamma(x, y_1)$ achieving level $\frac{1}{2}$ at $x = x_1$ (and analogously for x_2, y_2).

We assume that, for any given x, as y changes from 0 to ∞ (or whatever the full range of presented lengths might be), function $y \to \gamma(x, y)$ increases from some value below $\frac{1}{2}$ to some value above $\frac{1}{2}$ (in Fig. 3, from 0 to 1). Because of this, the function attains $\frac{1}{2}$ at some unique value of y, by definition taken to be the PSE of x. We have therefore the following statement:

(S1) every x in $\mathcal{O}_1$ has a unique PSE y in $\mathcal{O}_2$,

where $\mathcal{O}_1, \mathcal{O}_2$ abbreviate the two observation areas. The value of y may but does not have to be equal to x. That is, we allow for a systematic error, interpretable, say, as indicating that one's perception of a given length depends on whether the segment is on the left or on the right (perceptual bias), or that the observer is predisposed to say "y is longer than x" less often or more often than to say "y is shorter than x" (response bias).

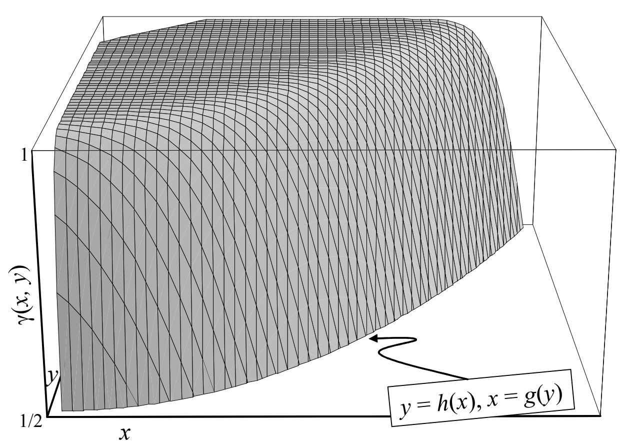


Fig. 4: The upper half of psychometric function $\gamma(x, y)$ shown in Fig. 2. The horizontal cross-section of the function at level $\frac{1}{2}$ is the PSE line, representing bijective maps h and g between the sets of all possible values for x and for y, $g \equiv h^{-1}$. By construction, $\gamma(x, h(x)) = \frac{1}{2}$ for all x; equivalently, $\gamma(g(y), y) = \frac{1}{2}$ for all y.

We further assume that, for any given y, as x changes from 0 to ∞, function $x \to \gamma(x, y)$ decreases from some value above $\frac{1}{2}$ to some value below $\frac{1}{2}$, because of which it reaches $\frac{1}{2}$ at some unique value of x, the PSE for y. We have the next statement:

(S2) every y in $\mathcal{O}_2$ has a unique PSE x in $\mathcal{O}_1$.

On a moment's reflection, we also have the third statement:

$$\text{(S3)} \quad \begin{array}{c} y \text{ in } \mathcal{O}_2 \text{ is the PSE for } x \text{ in } \mathcal{O}_1 \\ \text{if and only if} \\ x \text{ in } \mathcal{O}_1 \text{ is the PSE for } y \text{ in } \mathcal{O}_2. \end{array}$$

Indeed, $\gamma(x, y) = \frac{1}{2}$ means, by definition, that both y is a PSE for x and x is a PSE for y; and due to S1 and S2, these PSEs are unique. The seeming redundancy in the formulation of S3 serves to emphasize that the statement does not involve any switching of the physical locations of the two lines as we state their PSE relations: x remains on the left, y on the right.

The three statements just formulated, S1 to S3, constitute what can be called the *Regular Mediality* principle (Dzhafarov, 2003a). Its significance in this context is in that the formulation of Regular Minimality, as we see in the next section, is essentially identical, with the following caveats: in the Regular Minimality principle, the PSEs are defined differently, the formulations of S1 to S3 are not confined to unidimensional stimuli, and S3 is an independent statement rather than a consequence of S1 and S2.

Before we turn to Regular Minimality, however, it is useful to observe the following, in reference to Fig. 4. Statement S1 is equivalent to saying that there is a function $y = h(x)$ such that $\gamma(x, h(x)) = \frac{1}{2}$, for all x. Analogously for S2, there is a function $x = g(y)$ such that $\gamma(g(y), y) = \frac{1}{2}$, for all y. The meaning of S3 then is that g and h are inverses of each other (hence they are both bijective maps, one-to-one and onto). Geometrically, there is a single PSE line in the xy-plane, equivalently representable by $y = h(x)$ and $x = g(y)$.

3. REGULAR MINIMALITY

We give the formulation of Regular Minimality in full generality, for stimuli of arbitrary nature.

Discrimination probability function $\psi(\mathbf{x}, \mathbf{y})$ satisfies Regular Minimality if the following three statements are satisfied:

(RM1) There is a function $\mathbf{y} = \mathbf{h}(\mathbf{x})$ such that, for every $\mathbf{x}$ in $\mathcal{O}_1$, function $\mathbf{y} \to \psi(\mathbf{x}, \mathbf{y})$ achieves its minimum at $\mathbf{y} = \mathbf{h}(\mathbf{x})$ in $\mathcal{O}_2$;

(RM2) There is a function $\mathbf{x} = \mathbf{g}(\mathbf{y})$ such that, for every $\mathbf{y}$ in $\mathcal{O}_2$, function $\mathbf{x} \to \psi(\mathbf{x}, \mathbf{y})$ achieves its minimum at $\mathbf{x} = \mathbf{g}(\mathbf{y})$ in $\mathcal{O}_1$;
(RM3) $\mathbf{g} \equiv \mathbf{h}^{-1}$.

Remark 1. Strictly speaking, the formulation of Regular Minimality requires a caveat: physical labels for stimuli in the two observation areas have been assigned so that, in $\mathcal{O}_1$, $\mathbf{x}_1 = \mathbf{x}_2$ if and only if they are "psychologically indistinguishable," in the sense that $\psi(\mathbf{x}_1, \mathbf{y}) = \psi(\mathbf{x}_2, \mathbf{y})$ for all $\mathbf{y}$; and analogously for $\mathbf{y}_1, \mathbf{y}_2$ in $\mathcal{O}_2$. The notion of psychological equality (indistinguishability) is discussed later, in Section 10).

Remark 2. It follows from RM1 to RM3 that both $\mathbf{h}$ and $\mathbf{g}$ are bijective maps (one-to-one and onto), from all possible values of $\mathbf{x}$ onto all possible values of $\mathbf{y}$, and vice versa.

Remark 3. Statement RM3 can also be formulated in the form of S3 for Regular Mediality:

$$\begin{array}{c} \mathbf{y} \text{ in } \mathcal{O}_2 \text{ is the PSE for } \mathbf{x} \text{ in } \mathcal{O}_1 \\ \text{if and only if} \\ \mathbf{x} \text{ in } \mathcal{O}_1 \text{ is the PSE for } \mathbf{y} \text{ in } \mathcal{O}_2. \end{array}$$

Figures 5 and 6 provide an illustration using unidimensional stimuli. Focusing on x_1 (in $\mathcal{O}_1$) and y_1 (in $\mathcal{O}_2$), they are PSEs of each other because $y \to \psi(x_1, y)$ achieves its minimum at $y = y_1$ and $x \to \psi(x, y_1)$ achieves its minimum at $x = x_1$. Note that x_1 and y_1 need not coincide (we see later that this depends on our choice of physical labeling). Note also that the two cross-sections, $y \to \psi(x_1, y)$ and $x \to \psi(x, y_1)$, may very well have different shapes and generally cannot be reconstructed from each other. Their minima, however, are necessarily on the same level (see Fig. 7), because, due to Regular Minimality, this level is, for the first of these cross-sections, $\psi(x_1, y = y_1)$, and for the second, $\psi(x = x_1, y_1)$.

Unlike Regular Mediality, where the uniqueness of the PSE relation (statements S1 and S2) is generally lost outside the context of unidimensional stimuli, Regular Minimality applies to stimuli of arbitrary nature, including multidimensional stimuli, such as colors identified by Commission Internationale de l'Eclairage (CIE) or Munsell coordinates, discrete stimuli (such as letters of alphabet), and more complex stimuli (such as human faces or variable-trajectory variable-speed motions of a visual target), representable by one or several functions of several arguments. Figure 8 illustrates Regular Minimality for two-dimensional stimuli (the analogue of Fig. 5, being a four-dimensional hypersurface, cannot, of course, be shown graphically).

A toy example demonstrates Regular Minimality in the case of a discrete stimulus set. Symbols $\mathbf{x}_a$, $\mathbf{x}_b$, $\mathbf{x}_c$, $\mathbf{x}_d$ represent stimuli in the first

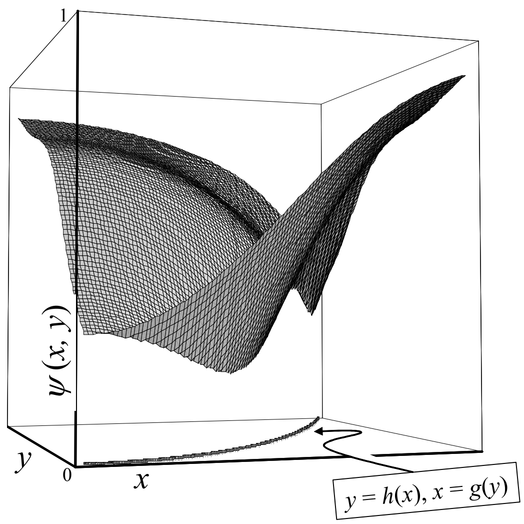


Fig. 5: Possible appearance of discrimination probability function $\psi(x, y)$ for unidimensional stimuli. (This particular function was generated by the "quadrilateral dissimilarity" model described in Section 7.2.) The function satisfies Regular Minimality. The curve in the xy-plane is the PSE line, representing bijective maps h and g between the sets of all possible values for x and for y, $g \equiv h^{-1}$. By definition of PSE, for any fixed $x, \psi(x, y)$ achieves its minimum at $y = h(x)$; and for any fixed $y, \psi(x, y)$ achieves its minimum at $x = g(y)$.

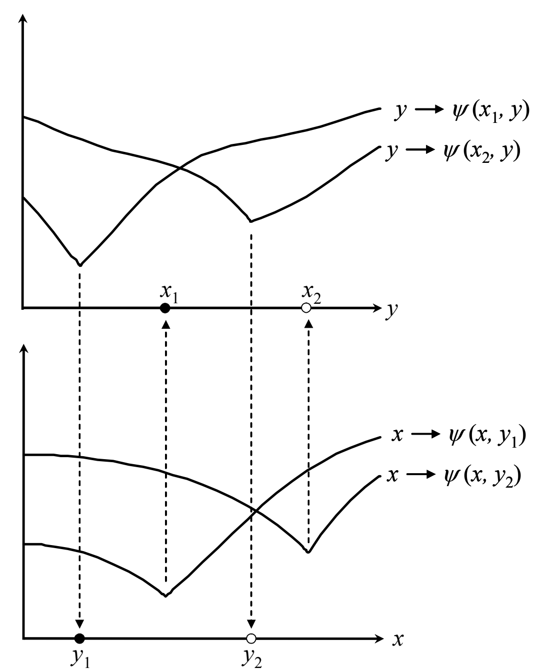


Fig. 6: Cross-sections of discrimination probability function $\psi(x, y)$ shown in Fig. 5 made at two fixed values of x (upper panel) and two fixed values of y (lower panel). The figure illustrates the Regular Minimality principle for same-different discriminations: y is the PSE for x if and only if x is the PSE for y. Thus, $\psi(x_1, y)$ achieves its minimum at $y = y_1$, while $\psi(x, y_1)$ achieves its minimum at $x = x_1$ (and analogously for x_2, y_2).

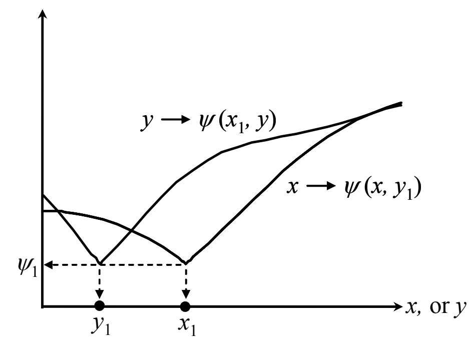


Fig. 7: The superposition of functions $\psi(x_1, y)$ and $\psi(x, y_1)$ from Fig. 6. Minimum level ψ_1 is the same in these two (generally different) functions because in both cases it equals $\psi(x_1, y_1)$.

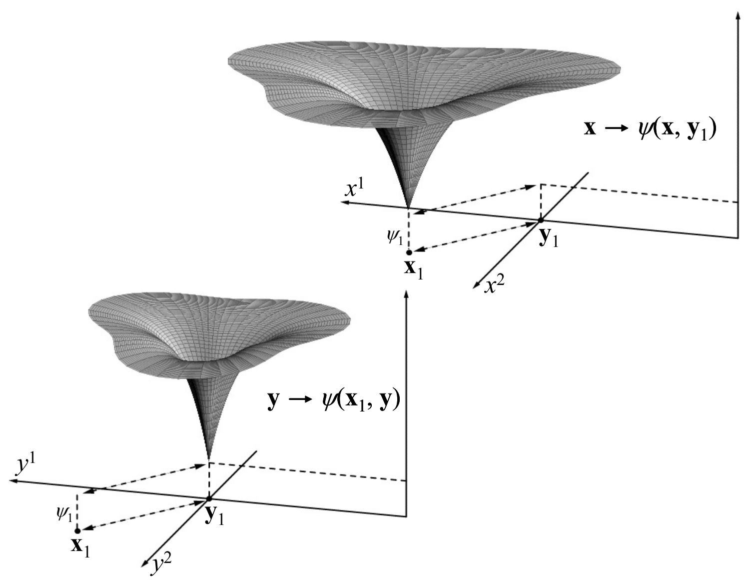


Fig. 8: Two cross-sections of a discrimination probability function, $\psi(\mathbf{x}, \mathbf{y})$, $\mathbf{x} = (x^1, x^2)$, $\mathbf{y} = (y^1, y^2)$, made at a fixed value of $\mathbf{x}$ ($\mathbf{x} = \mathbf{x}_1$, lower panel) and a fixed value of $\mathbf{y}$ ($\mathbf{y} = \mathbf{y}_1$, upper panel). The figure illustrates the Regular Minimality principle for same–different discriminations of two–dimensional stimuli: $\psi(\mathbf{x}_1, \mathbf{y})$ achieves its minimum at at $\mathbf{y} = \mathbf{y}_1$ (i.e., $\mathbf{y}_1$ is the PSE for $\mathbf{x}_1$) if and only if $\psi(\mathbf{x}, \mathbf{y}_1)$ achieves its minimum at $\mathbf{x} = \mathbf{x}_1$ (i.e., $\mathbf{x}_1$ is the PSE for $\mathbf{y}_1$). Minimum level ψ_1 is the same in the two panels because in both cases it equals $\psi(\mathbf{x}_1, \mathbf{y}_1)$. This is essentially a two-dimensional analogue of Figs. 6 and 7.

observation area, $\mathbf{y}_a$, $\mathbf{y}_b$, $\mathbf{y}_c$, $\mathbf{y}_d$ represent the same four stimuli in the second observation area. (We discuss later that, in general, stimulus sets in the two observation areas need not be the same.) The entries of the matrix represent discrimination probabilities $\psi(\mathbf{x}, \mathbf{y})$.

TOY_1	$\mathbf{y}_a$	$\mathbf{y}_b$	$\mathbf{y}_c$	$\mathbf{y}_d$
$\mathbf{x}_a$	0.6	0.6	0.1	0.8
$\mathbf{x}_b$	0.9	0.9	0.8	0.1
$\mathbf{x}_c$	1	0.5	1	0.6
$\mathbf{x}_d$	0.5	0.7	1	1

Here, Regular Minimality manifests itself in the fact that

1. every row contains a single minimal cell;
2. every column contains a single minimal cell;
3. a cell is minimal in its row if and only if it is minimal in its column.

The four PSE pairs in this example are $(\mathbf{x}_a, \mathbf{y}_c)$, $(\mathbf{x}_b, \mathbf{y}_d)$, $(\mathbf{x}_c, \mathbf{y}_b)$, and $(\mathbf{x}_d, \mathbf{y}_a)$.

4. NONCONSTANT SELF-DISSIMILARITY

Another important feature exhibited by our matrix TOY_1 is that the minima achieved by function $\psi(\mathbf{x}, \mathbf{y})$ at PSE pairs are not all on the same level:

$\mathcal{O}_1$	$\mathbf{x}_a$	$\mathbf{x}_b$	$\mathbf{x}_c$	$\mathbf{x}_d$
$\mathcal{O}_2$	$\mathbf{y}_c$	$\mathbf{y}_d$	$\mathbf{y}_b$	$\mathbf{y}_a$
ψ	0.1	0.1	0.5	0.5

The same is true for the discrimination probability function shown in Fig. 5. This is best illustrated by the "wall" erected vertically from the PSE line until it touches the surface representing $\psi(x, y)$, as shown in Fig. 9. The upper contour of the "wall" is function $\omega_1(x) = \psi(x, h(x))$ or equivalently, $\omega_2(y) = \psi(g(y), y)$, the values attained by $\psi(x, y)$ when x and y are mutual PSEs.

In general, we call the values of $\psi(\mathbf{x}, \mathbf{y})$ attained when the two arguments are each other's PSEs (i.e., $\mathbf{y} = \mathbf{h}(\mathbf{x})$, $\mathbf{x} = \mathbf{g}(\mathbf{y})$), the *self-dissimilarity values*, and we call either of functions $\omega_1(\mathbf{x}) = \psi(\mathbf{x}, \mathbf{h}(\mathbf{x}))$ and $\omega_2(\mathbf{y}) = \psi(\mathbf{y}, \mathbf{h}(\mathbf{y}))$, the *minimum level function*. Although $\omega_1(\mathbf{x})$ and $\omega_2(\mathbf{y})$ may be different functions, geometrically they describe one and the same set of

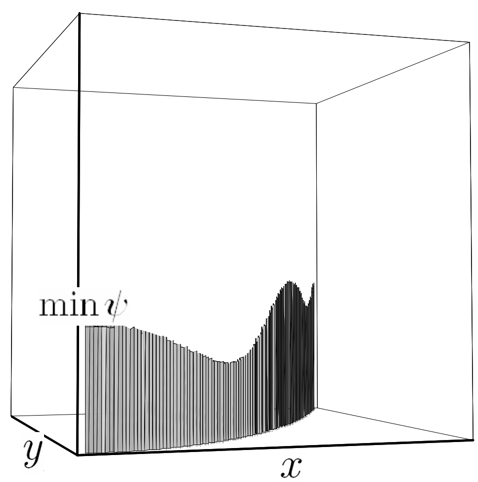


Fig. 9: The "wall" whose bottom contour is PSE line $y = h(x)$ (equivalently, $x = g(y)$) for function $\psi(x, y)$ shown in Fig. 5, and the top contour is minimum level function $\psi(x, h(x))$ (equivalently, $\psi(g(y), y)$) for the same function. The figure illustrates, in addition to Regular Minimality, the notion of Nonconstant Self-Dissimilarity: the minimum level function is not constant..

points in the $\mathbf{xy}\psi$-coordinates (in the same way $\mathbf{h}(\mathbf{x})$ and $\mathbf{g}(\mathbf{y})$ describe one and the same set of points in $\mathbf{xy}$-coordinates).

According to the principle of Nonconstant Self-Dissimilarity, $\omega_1(\mathbf{x})$ (or, equivalently, $\omega_2(\mathbf{y})$) is not *necessarily* a constant function. The modal quantifier "is not necessarily" should be understood in the following sense. For a given stimulus set presented to a given perceiver it may happen that $\omega_1(\mathbf{x})$ has a constant value across all values of $\mathbf{x}$. It may only happen, however, as a numerical coincidence rather than by virtue of a law that compels $\omega_1(\mathbf{x})$ to be constant: $\omega_1(\mathbf{x})$ considered across all possible sets of stimuli pairwise presented in all possible experiments with all possible perceivers will at least sometimes be a nonconstant function. If $\omega_1(\mathbf{x})$ is nonconstant for a particular discrimination probability function $\psi(\mathbf{x}, \mathbf{y})$, we say that Nonconstant Self-Dissimilarity is *manifest* in this function. This is the most conservative formulation of the principle. With less caution, one might hypothesize that minimum level function $\omega_1(\mathbf{x})$, at least in psychophysical applications involving same–different judgments, is never constant, provided the probabilities are measured precisely enough.

For completeness, Fig. 10 illustrates Nonconstant Self-Dissimilarity for two-dimensional stimuli, like the ones in Fig. 8. The surface that contains the minima of the cross-sections $\mathbf{y} \to \boldsymbol{\psi}(\mathbf{x}, \mathbf{y})$ is the minimum level function $\omega_2(\mathbf{y})$.

5. FUNCTIONS VIOLATING REGULAR MINIMALITY

Unlike Regular Mediality, which can be mathematically deduced from the monotonicity of cross-sections $x \to \gamma(x, y)$ and $y \to \gamma(x, y)$, Regular Minimality is not reducible to more elementary properties of $\psi(\mathbf{x}, \mathbf{y})$.

It is easy to see how Regular Minimality can be violated in discrete stimulus sets.

TOY_2	$\mathbf{y}_a$	$\mathbf{y}_b$	$\mathbf{y}_c$	$\mathbf{y}_d$
$\mathbf{x}_a$	0.1	0.6	0.1	0.8
$\mathbf{x}_b$	0.9	0.9	0.8	0.1
$\mathbf{x}_c$	1	0.5	1	0.6
$\mathbf{x}_d$	0.5	0.7	1	1

TOY_3	$\mathbf{y}_a$	$\mathbf{y}_b$	$\mathbf{y}_c$	$\mathbf{y}_d$
$\mathbf{x}_a$	0.7	0.4	0.2	0.8
$\mathbf{x}_b$	0.9	0.9	0.8	0.4
$\mathbf{x}_c$	1	0.6	0.7	0.8
$\mathbf{x}_d$	0.4	0.7	1	1

Using the same format as in matrix TOY_1, the first of the two matrices above has two equal minima in the first row, in violation of RM1. One can say here that $\mathbf{x}_a$ in $\mathcal{O}_1$ has two PSEs in $\mathcal{O}_2$ ($\mathbf{y}_a$ and $\mathbf{y}_c$), or (if the

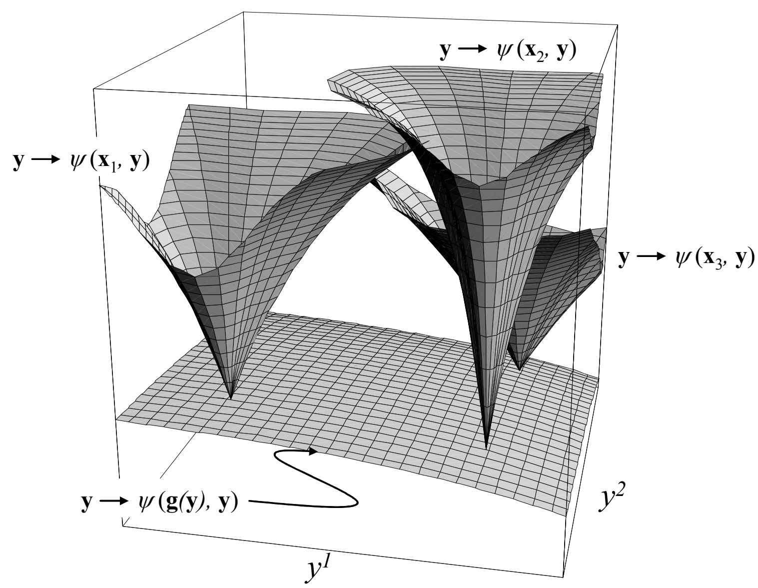


Fig. 10: An illustration of Nonconstant Self-Dissimilarity for two-dimensional stimuli. Shown are three cross-sections $\mathbf{y} \to \psi(\mathbf{x}, \mathbf{y})$, $\mathbf{x} = \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$, of discrimination probability function $\psi(\mathbf{x}, \mathbf{y})$, whose minima, $\mathbf{h}(\mathbf{x}_1)$, $\mathbf{h}(\mathbf{x}_2)$, and $\mathbf{h}(\mathbf{x}_3)$, lie on minimum level surface $\psi(\mathbf{g}(\mathbf{y}), \mathbf{y})$, where $\mathbf{g} \equiv \mathbf{h}^{-1}$. This surface is not parallel to the y^1y^2-plane, manifesting Nonconstant Self-Dissimilarity.

uniqueness of a PSE is considered part of its definition) that the PSE for $\mathbf{x}_a$ is not defined. Matrix TOY$_3$ above is of a different kind: it satisfies properties RM1 and RM2 but violates RM3. Stimulus $\mathbf{x}_c$ in $\mathcal{O}_1$ has $\mathbf{y}_b$ in $\mathcal{O}_2$ as its unique PSE; the unique PSE in $\mathcal{O}_1$ for $\mathbf{y}_b$, however, is not $\mathbf{x}_c$ but $\mathbf{x}_a$ (one could continue: and the PSE for $\mathbf{x}_a$ is not $\mathbf{y}_b$ but $\mathbf{y}_c$). In a situation like this one can say that the relation "is the PSE of" is not symmetrical, and the notion of a "*PSE pair*" is not well defined.

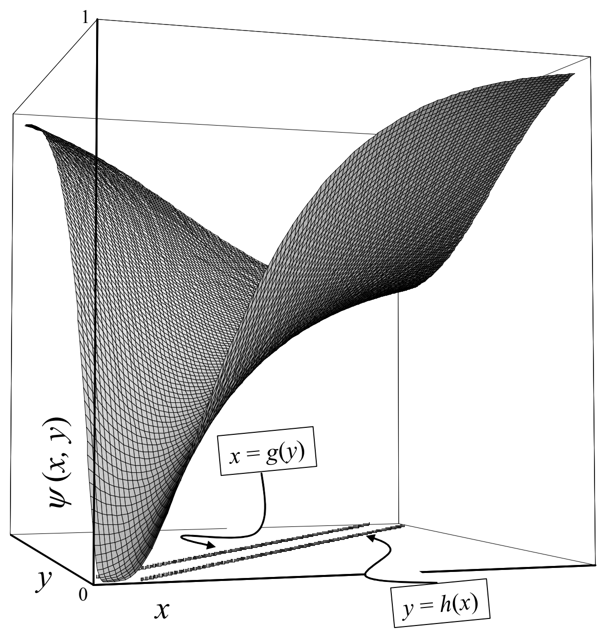


Fig. 11: An example of function $\psi(x, y)$ that violates Regular Minimality. (This particular function was generated by Luce-Galanter's Thurstonian-type model described in Section 7.1.) For a fixed value of x, $\psi(x, y)$ achieves its minimum at $y = h(x)$; for a fixed value of y, $\psi(x, y)$ achieves its minimum at $x = g(y)$. But g is not the inverse of h : the lines $y = h(x)$ and $x = g(y)$ (nearly straight lines in this example) do not coincide. Compare to Fig. 5.

Figures 11, 12, and 13 present an analogue for TOY$_3$ in a continuous (unidimensional) domain. The function depicted in these figures satisfies properties RM1 and RM2, but violates RM3: if y is the PSE for x, the latter generally will not (in this example, never) be the PSE for y, and vice versa. The notion of a PSE pair is not well defined here. Specifically,

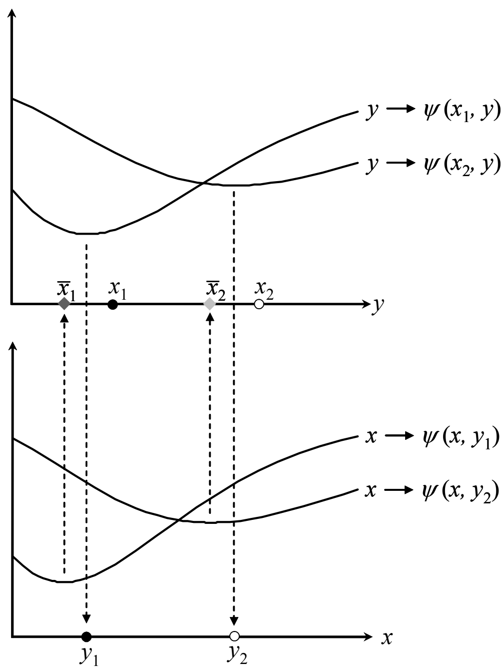


Fig. 12: Cross-sections of function $\psi\,(x, y)$ shown in Fig. 11 made at two fixed values of x (upper panel) and two fixed values of y (lower panel). The figure details violations of the Regular Minimality principle in this function: $\psi\,(x_1, y)$ achieves its minimum at $y = y_1$, but $\psi\,(x, y_1)$ achieves its minimum at a point different from $x = x_1$ (and analogously for x_2, y_2). One cannot speak of PSE pairs unambiguously in this case: for example., (x_1, y_1) and $(\bar{x}_1, y_1)$ are both "PSE pairs," with one and the same y_1 in the second observation area.

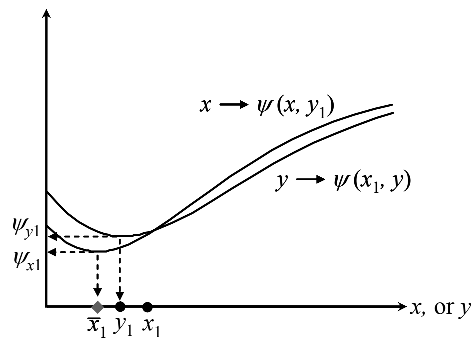
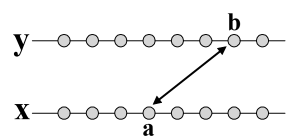


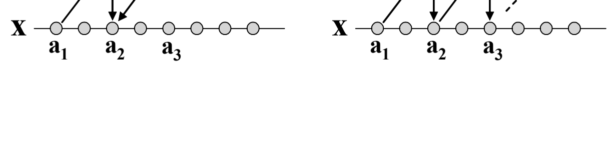
Fig. 13: The superposition of functions $\psi(x_1, y)$ and $\psi(x, y_1)$ from Fig. 12. Minimum level ψ_{x1} for the former is not the same as minimum level ψ_{y1} for the latter. Compare with Fig. 7.

one and the same stimulus (say, $x = a$ in $\mathcal{O}_1$) can be paired either with y at which $\psi(a, y)$ achieves its minimum, or with $\bar{y}$ such that $x \rightarrow \psi(x, \bar{y})$ achieves its minimum at $x = a$.

It may be useful to look at this issue more schematically. Regular Minimality can be represented by the diagram



in which the two "beaded strings" stand for stimuli in the two observation areas, and arrows stand for relation "is the PSE for." Starting at any point and traveling along the arrows, one is bound to return to this point after having visited just one other point, its PSE in the other observation area. If Regular Minimality is violated, the traveling along the arrows between the observation areas becomes more adventurous, with the potential of "wandering away" indefinitely far:



6. EMPIRICAL EVIDENCE

Discrimination probabilities of the same–different type have not been studied as intensively as those of the greater–less type. The available empirical evidence, however, seems to be in good agreement with the hypothesis that discrimination probabilities (a) satisfy Regular Minimality and (b) manifest Nonconstant Self-Dissimilarity.

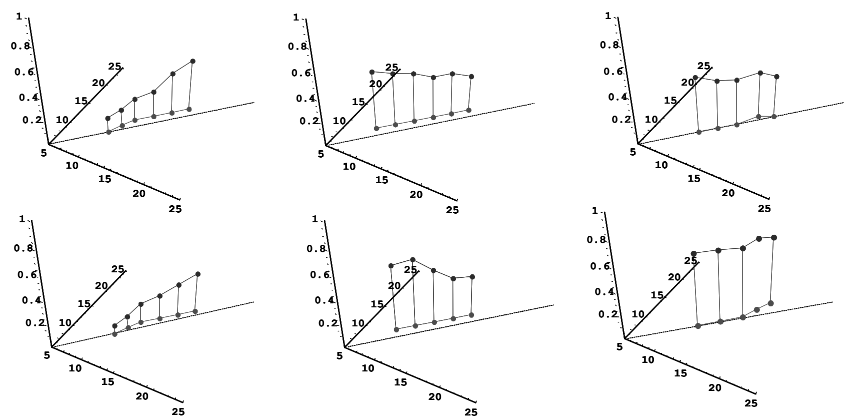


Fig. 14: An empirical version of Fig. 9, based on one of the experiments described in Dzhafarov and Colonius (2005a). x and y are lengths of two horizontal line segments, in pixels (1 pixel $\approx$ 0.86 min arc), presented side-by-side; each panel represents an experiment with a single observer. The bottom line shows estimated positions of PSEs, $y = h(x)$, the upper line shows the corresponding probabilities, $\psi(x, h(x))$ (the minimum level function). Straight lines in the xy-planes are bisectors. Each probability estimate is based on 500 to 600 replications.

In an experiment reported in Dzhafarov and Colonius (2005a), observers were asked to compare two side-by-side presented horizontal line segments (identical except for their lengths, x on the left, y on the right). The results of such an experiment are represented by a matrix of pairwise probabilities $\psi(x, y)$, with x and y values providing a dense sample of length values within a relatively small interval. Except for an occasional necessity to interpolate a minimum between two successive values, the compliance with Regular Minimality in such a matrix is verified by showing that the matrix is structured essentially like TOY_1 in Section 3 rather than TOY_2 or TOY_3 in Section 5. If (and only if) Regular Minimality is established, one can draw a single line through PSE pairs, $(x, h(x))$ or, equivalently, $(g(y), y)$, in the

xy-plane. Plotting the discrimination probability against each of these PSE pairs, we get an empirical version of the minimum level function. The results presented in Fig. 14 clearly show that Regular Minimality is satisfied, and that $\psi(x, h(x))$ is generally different for different x (i.e., Nonconstant Self-Dissimilarity is manifest). Note, in relation to the issue of canonical transformations, considered in Section 9, that x and y in a PSE pair (x, y) in these data are generally physically different, y (the length on the right) tends to be larger, indicating that the right lengths tend to be underestimated with respect to the left ones ("systematic error"). Analogous results are reported in Dzhafarov (2002b) and Dzhafarov and Colonius (2005a) for same–different discriminations of apparent motions (two-dot displays with temporal asynchrony between the dots) presented side-by-side or in a succession.

Figure 15 shows the results of an experiment by Zimmer and Colonius (2000), in which listeners made same–different judgments in response to successively presented sinusoidal tones varying in intensity (x followed by y). Regular Minimality here holds in the simplest form: x and y are mutual PSEs if (and only if) $x = y$. The minimum level function here is therefore $\psi(x, x)$ (equivalently, $\psi(y, y)$), and it clearly manifests Nonconstant Self–Dissimilarity.

Indow, Robertson, von Grunau, and Fielder (1992) and Indow (1998) reported discrimination probabilities for side-by-side presented colors varying in CIE chromaticity-luminance coordinates (a three-dimensional continuous stimulus space). With the right-hand color $\mathbf{y}$ serving as a fixed reference stimulus, function $\mathbf{x} \to \psi(\mathbf{x}, \mathbf{y})$ in this study reached its minimum at $\mathbf{x} = \mathbf{y}$,

$$\mathbf{x} \neq \mathbf{y} \Longrightarrow \psi(\mathbf{y}, \mathbf{y}) < \psi(\mathbf{x}, \mathbf{y}).$$

The experiment was not conducted with fully randomized color pairs, and it was not replicated with the left-hand color $\mathbf{x}$ used as a reference. One cannot therefore check for the compliance with Regular Minimality directly. It is reasonable to assume, however, that $\psi(\mathbf{x}, \mathbf{y})$ for side-by-side presented colors is order-balanced,

$$\psi(\mathbf{x}, \mathbf{y}) = \psi(\mathbf{y}, \mathbf{x}),$$

and under this assumption, it is easily seen, the inequality above implies Regular Minimality in the simplest form: $\mathbf{x}$ and $\mathbf{y}$ are mutual PSEs if (and only if) $\mathbf{x} = \mathbf{y}$. Nonconstant Self-Dissimilarity is a prominent feature of Indow's data: for instance, with reference color $\mathbf{y}$ changing from grey to red to yellow to green to blue, the probability $\psi(\mathbf{y}, \mathbf{y})$ for one observer increased from 0.07 to 0.33.

The conjunction of the simplest form of Regular Minimality with prominent Nonconstant Self-Dissimilarity was also obtained in two large data sets

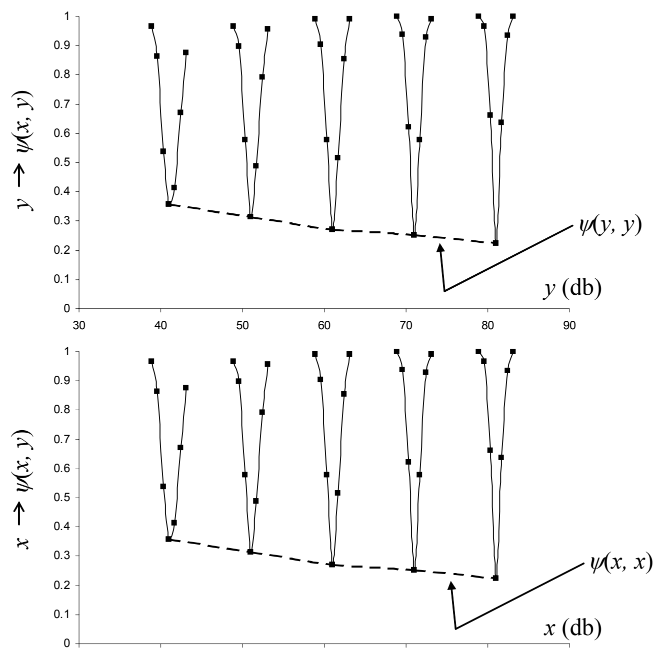


Fig. 15: An empirical version of Fig. 6, based on an experiment reported in Zimmer and Colonius (2000). x and y represent intensity of pure tones of a fixed frequency. The data are shown for a single listener. The PSEs in this case are physically identical, $h(x) = x$; that is, for any x, $\psi(x, y)$ achieves its minimum at $y = x$, and for any y, $\psi(x, y)$ achieves its minimum at $x = y$. The value of $\psi(x, x)$ decreases with increasing x.

involving discrete stimuli (36 Morse codes for letters and digits in Rothkopf, 1957, and 32 Morse code-like stimuli in Wish, 1967; sequential presentation in both cases). Below is a small fragment of Rothkopf's matrix: $\psi(\mathbf{x}, \mathbf{y})$ in each cell, for $\mathbf{x}, \mathbf{y} = D, H, K, S, W$. Each value on the main diagonal is the smallest probability in both its row and its column (Regular Minimality), and this value varies along the diagonal from 0.04 to 0.14 (Nonconstant Self-Dissimilarity). (A single deviation from this pattern found in Wish's data can be attributed to a statistical estimation error; for details, see Chapter 2 in this volume.)

RO	**D**	**H**	**K**	**S**	**W**
D	.12	.64	.19	.71	.82
H	.75	.13	.91	.63	.91
K	.27	.89	.09	.98	.67
S	.70	.41	.87	.04	. 88
W	.78	.85	.71	.88	.14

7. THE CONJUNCTION OF REGULAR MINIMALITY AND NONCONSTANT SELF-DISSIMILARITY

When dealing with stimulus sets containing finite number of elements, it is easy to construct examples of discrimination probability matrices that both satisfy Regular Minimality and manifest Nonconstant Self-Dissimilarity (as our matrix TOY$_1$ shown earlier). Here is a simple algorithm: given an n-element stimulus set, create any sequence (i_1, j_1), ..., (i_n, j_n), with $(i_1, ..., i_n)$ and $(j_1, ..., j_n)$ being two complete permutations of $(1, ..., n)$; fill in cells (i_1, j_1), ..., (i_n, j_n) with probability values $\psi_1 \geq ... \geq \psi_n$; fill in the rest of the i_1th row and j_1th column by values greater than ψ_1; fill in the rest of the i_2th row and j_2th column by values greater than ψ_2; etc. In thus a created matrix, the i_kth row stimulus (interpreted as a stimulus in $\mathcal{O}_1$) and the j_kth column stimulus (in $\mathcal{O}_2$) will be mutual PSEs ($k = 1, ..., n$), and Nonconstant Self-Dissimilarity will be manifest if at least one of the inequalities in $\psi_1 \geq ... \geq \psi_n$ is strict. It is equally easy to construct examples that do not satisfy Regular Minimality (as TOY$_2$ and TOY$_3$ matrices above) or do not manifest Nonconstant Self-Dissimilarity (set $\psi_1 = ... = \psi_n$ in the algorithm just given).

The construction of examples is less obvious in the case of continuous stimulus sets, as in our Fig. 5 and Fig. 11. It is instructive therefore to consider theoretical models which generate functions $\psi(\mathbf{x}, \mathbf{y})$ that always satisfy the conjunction of Regular Minimality and Nonconstant Self-Dissimilarity, as well as theoretical models whose generated functions $\psi(\mathbf{x}, \mathbf{y})$ always violate this conjunction of properties. We consider the latter class of models first.

7.1. Thurstonian-type models

To avoid technicalities, we confine our discussion here to the unidimensional case, with x, y taking on their values on intervals of reals, finite or infinite. The results to be mentioned, however, generalize to arbitrary continuous spaces of stimuli.

Consider the following scheme, well familiar to psychophysicists. Let any pair (x, y) presented to an observer for a same–different comparison be mapped into a pair of perceptual images, (P_x, Q_y), and let P_x and Q_y be mutually independent random entities taking on their values in some perceptual space, of arbitrary nature.[5] In any given trial, the observer experiences two realizations of these random entities, (p, q), and there is a decision rule that maps some of the possible (p, q)-pairs into response "same" and the remaining ones into response "different." The decision rule can be arbitrary, and so can be the distributions of P_x and Q_y in the perceptual space, except for the following critical constraint: we assume that P_x and Q_y are "well-behaved" in response to small changes in x and y. This means the following. The distribution of P_x is determined by the probabilities with which p falls within various measurable subsets of the perceptual space, and these probabilities generally change as x changes within an arbitrarily small interval of values. Intuitively, P_x is well-behaved if the rate of these changes cannot get arbitrarily high. The well-behavedness of Q_y is defined analogously.[6] As shown in Dzhafarov (2003a), no $\psi(x, y)$ generated by such a model can both satisfy Regular Minimality and manifest

[5] Notation conventions: P_x, Q_y, and $S_{x,y}$ designate random entities whose distributions depend on their index. Random entities are called random variables if their realizations p, q, s, are real numbers (with the Lebesgue sigma-algebra).

[6] In terminology of Dzhafarov (2003a), this is the "well-behavedness in the narrow (or absolute) sense": for any $x = a$, the right-hand and left-hand derivatives of $\Pr[P_x \in \mathfrak{p}]$ with respect to x exist and are bounded across all measurable sets $\mathfrak{p}$ and all values of x within an arbitrarily small interval $[a - \varepsilon, a + \varepsilon]$ (and analogously for y and Q_y). This requirement can be considerably weakened, with respect to both the class of (x, y)-values and the class of measurable sets for which it is supposed to hold (details in Dzhafarov, 2003a, b). The simplest and perhaps most important example of a non-well-behaved P_x is a deterministic entity, having a single possible value for every x.

Nonconstant Self-Dissimilarity. This means, in particular, that with such a model,

1. if $\psi(x,y)$ satisfies Regular Minimality, then $\psi(x,y) \equiv$ *constant* across all PSE pairs (x,y) (i.e., Regular Minimality can only coexist with *Constant* Self-Dissimilarity);
2. if $y \to \psi(x,y)$ achieves a minimum at $y = h(x)$, if $x \to \psi(x,y)$ achieves a minimum at $x = g(y)$, and if either $\psi(x, h(x))$ or$\psi(g(y), y)$ is nonconstant across, respectively, x and y values, then g cannot coincide with h^{-1} (i.e., even if RM1 and RM2 are satisfied, Nonconstant Self-Dissimilarity forces RM3 of Regular Minimality to be violated).

The class of such models has its historical origins in Thurstone's analysis of greater–less discriminations (Thurstone, 1927a, 1927b), because of which in Dzhafarov (2003a, 2003b) such models are referred to as "Thurstonian-type" (see Fig. 16). The simplest Thurstonian-type model for same–different discriminations is presented in Luce and Galanter (1963): the perceptual space is the set of reals, P_x and Q_y are normally distributed, and the decision rule is "respond 'different' if and only if $|p - q| > \varepsilon$," for some $\varepsilon > 0$. If the means and the variances of these normal distributions, $\left(\mu_P(x), \sigma_P^2(x)\right)$ and $\left(\mu_Q(y), \sigma_Q^2(y)\right)$, are piecewise smooth functions of x and y (which is sufficient although not necessary for P_x and Q_y to be well-behaved), then the resulting $\psi(x,y)$ must violate the conjunction of Regular Minimality and Nonconstant Self-Dissimilarity. Figures 11, 12, and 13 are generated by means of such a model (with x, y positive, and $\mu_P(x)$, $\sigma_P^2(x)$, $\mu_Q(y)$, $\sigma_Q^2(y)$ linear transformations of their arguments).

Most Thurstonian-type models proposed in the literature for same–different discriminations involve univariate or multivariate normal distributions for perceptual images of stimuli (Dai, Versfeld, & Green, 1996; Ennis, 1992; Ennis, Palen, & Mullen, 1988; Luce & Galanter, 1963; Sorkin, 1962; Suppes & Zinnes, 1963; Thomas, 1996; Zinnes & MacKay, 1983). With these and other distributions possessing finite density in $\mathbb{R}^n$ $(n \geq 1)$, a piecewise smooth dependence of their parameters on x or y implies their well-behavedness, hence the impossibility of generating a discrimination probability function with both Regular Minimality and Nonconstant Self-Dissimilarity. Luce (1977) called Thurstonian models the "essence of simplicity": "this conception of internal representation of signals is so simple and so intuitively compelling that no one ever really manages to escape from it. No matter how one thinks about psychophysical phenomena, one seems to come back to it" (p. 462). Luce refers here to the simplest Thurstonian models, involving unidimensional random representations and simple decision rules based on values of $p - q$. These models do work well for greater–less discriminations, generating functions like the one shown in

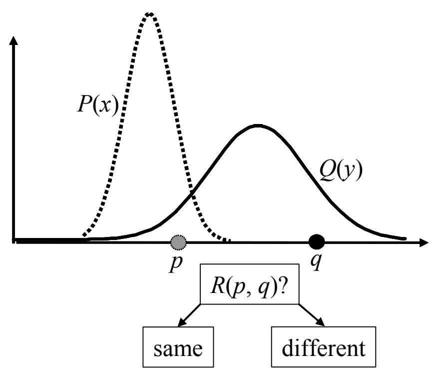


Fig. 16: Schematic representation of a Thurstonian-type model. Stimuli x and y are mapped into their "perceptual images," random variables $P(x)$ and $Q(y)$ (here, independently normally distributed on a set of reals). Response "same" or "different" is given depending on whether the realizations p, q of $P(x)$ and $Q(y)$ in a given trial stand or do not stand in a particular relation, R, to each other (e.g., $|p-q|$ exceeds or does not exceed some ε, or p, q fall or do not fall within one and the same interval in some partitioning of the set of reals). In general, p and q may be elements of an arbitrary set, the decision rule may be probabilistic (i.e., every pair p, q may lead to response "different" with some probability $\pi(p, q)$), and "perceptual images" $P(x)$ and $Q(y)$ may be stochastically interdependent, provided they are selectively attributable to x and y, respectively (in the sense of Dzhafarov, 2003c).

Figs. 2-4, subject to Regular Mediality. In the context of same–different discriminations, however, if the properties of Regular Minimality and Nonconstant Self-Dissimilarity do hold empirically, as data seem to suggest, Thurstonian-type models fail even if one allows for arbitrary decision rules and arbitrarily complex (but well-behaved) distributions for P_x and Q_y.[7] Moreover, the failure in question extends to the models in which decision rules are probabilistic rather than deterministic, that is, where each pair (p, q) can lead to both responses, "same" and "different," with certain probabilities (Dzhafarov, 2003b).

Finally, the failure in question extends to models with stochastically interdependent P_x and Q_y, provided P_x can still be considered an image of x (and not also of y) whereas Q_y is considered an image of y (and not also of x). The *selective attribution* of P_x and Q_y to x and y, respectively, is understood in the meaning explicated in Dzhafarov (2003c): one can find mutually independent random entities C, C_1, C_2, whose distributions do not depend on either x or y, such that

$$P_x = \pi\left(x, C, C_1\right), \quad Q_y = \theta\left(y, C, C_2\right), \tag{3}$$

where π, θ are some measurable functions. In other words, P_x and Q_y depend on x and y selectively, and their stochastic interdependence is due to a *common source of variability*, C. The latter may represent, for example, random fluctuations in the arousal or attention level, or in receptive fields' sensitivity profiles. P_x and Q_y then are *conditionally independent* at any fixed value c of C, because random entities $\pi\left(x, c, C_1\right)$ and $\theta\left(y, c, C_2\right)$ have independent sources of variability, C_1, C_2. As shown in Dzhafarov (2003b), if, for any c, $\pi\left(x, c, C_1\right)$ and $\theta\left(y, c, C_2\right)$ are well-behaved in the sense explained earlier (in which case we call P_x and Q_y themselves well-behaved), the resulting discrimination probability functions cannot both satisfy Regular Minimality and manifest Nonconstant Self-Dissimilarity.

The selectiveness in the attribution of P_x to x and Q_y to y is an important caveat. In Dzhafarov's (2003a, 2003b) terminology which we follow here, it is a necessary condition for calling a stochastic model Thurstonian-type. *Any* function $\psi\left(x, y\right)$ can be accounted for by a model in which x and y jointly map into a perceptual property, $S_{x,y}$, which then either maps into responses "same" and "different" probabilistically, or is a random entity itself, mapped into the responses by means of a certain decision rule (these two conceptual schemes are mathematically equivalent). For example, $S_{x,y}$ may be a nonnegative random variable interpretable as a measure

[7] The well-behavedness constraint, in some form, is critical here: as shown in Dzhafarov (2003a), *any* function $\psi\left(x, y\right)$ can be generated by a Thurstonian-type model if $P\left(x\right)$ and $Q\left(y\right)$ are allowed to have arbitrary distributions arbitrarily depending on, respectively, x and y. The well-behavedness constraint, however, is unlikely to be violated in a model designed to fit or simulate empirical data.

of "subjective dissimilarity" between x and y, and the decision rule be as in the classical signal detection theory: respond "different" if and only if the realization of s of $S_{x,y}$ exceeds some $\varepsilon > 0$. A model of the latter variety can be found, for example, in Takane and Sergent (1983). With this approach, $S_{x,y}$ can always be set up in such a way that $\psi(x, y)$ possesses both Regular Minimality and Nonconstant Self-Dissimilarity. Once this is done, Dzhafarov's (2003a, 2003b) results would indicate that $S_{x,y}$ cannot be computed from any two well-behaved random entities P_x and Q_y selectively attributable to x and y (e.g., subjective dissimilarity $S_{x,y}$ cannot be presented as $|P_x - Q_y|$ in Luce and Galanter's model mentioned earlier). In other words, $S_{x,y}$ must be an "emergent property," not reducible to the separate (and well-behaved) perceptual images of x and of y. We discuss such models next, but we prefer to do this within the conceptually more economic (but equivalent) theoretical language in which $S_{x,y}$ is treated as a deterministic quantity, $S(x, y)$, mapped into responses "same" and "different" probabilistically.

7.2. "Quadrilateral dissimilarity," "uncertainty blobs," etc.

At this point, we can switch back to stimuli $\mathbf{x}, \mathbf{y}$ of arbitrary nature, as the case of unidimensional stimuli is technically no simpler than the general case. We consider a measure of subjective dissimilarity, $S(\mathbf{x}, \mathbf{y})$, a deterministic quantity (i.e., having a fixed value for any $\mathbf{x}, \mathbf{y}$) related to discrimination probabilities by

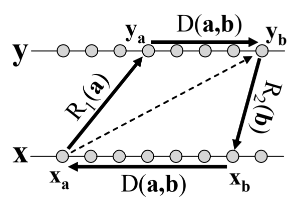
$$\psi(\mathbf{x}, \mathbf{y}) = \beta(S(\mathbf{x}, \mathbf{y})), \tag{4}$$

where β is some strictly increasing function. Such a model is distinctly non-Thurstonian as it does not involve individual random images for individual stimuli. Rather the models of this class are in the spirit of what Luce and Edwards (1958) called "the old, famous psychological rule of thumb: equally often noticed differences are equal" (p. 232), provided one keeps in mind that the "difference," understood to mean dissimilarity $S(\mathbf{x}, \mathbf{y})$, cannot be a true distance (as this would force constant minima at $\mathbf{x} = \mathbf{y}$).[8] As it turns out, for a broad class of possible definitions of $S(\mathbf{x}, \mathbf{y})$, such

[8] The Probability-Distance hypothesis, as it is termed in Dzhafarov (2002a), according to which $\psi(\mathbf{x}, \mathbf{y})$ is an increasing transformation of some distance $D(\mathbf{x}, \mathbf{y})$, is as traditional in psychophysics as is the Thurstonian-type modeling. In the context of unidimensional stimuli and greater–less discrimination probabilities $\gamma(x, y)$ this hypothesis is known as the "Fechner problem" (Falmagne, 1971; Luce & Edwards, 1958). See Dzhafarov (2002a) for history and a detailed discussion.

models only generate discrimination probability functions that are subject to both Regular Minimality and Nonconstant Self-Dissimilarity. Intuitively, the underlying idea is that the dissimilarity between stimulus $\mathbf{x}$ in $\mathcal{O}_1$ and stimulus $\mathbf{y}$ in $\mathcal{O}_2$ involves (a) the distance between $\mathbf{x}$ and the PSE $\mathbf{g}(\mathbf{y})$ of $\mathbf{y}$ (both in $\mathcal{O}_1$), (b) the distance between $\mathbf{y}$ and the PSE $\mathbf{h}(\mathbf{x})$ of $\mathbf{x}$ (both in $\mathcal{O}_2$), and (c) some slowly changing "residual" dissimilarities within the PSE pairs themselves, $(\mathbf{x}, \mathbf{h}(\mathbf{x}))$ and $(\mathbf{g}(\mathbf{y}), \mathbf{y})$.[9]

As before, the "beaded strings" in the diagram below schematically represent stimulus sets in the two observation areas, but the arrows now designate the components of a possible dissimilarity measure between $\mathbf{x_a}$ and $\mathbf{y_b}$. The PSE relation is indicated by identical index at $\mathbf{x}$ and $\mathbf{y}$: thus, $(\mathbf{x_a}, \mathbf{y_a})$ and $(\mathbf{x_b}, \mathbf{y_b})$ are PSE pairs.



We assume some distance measure D among stimuli *within* either of the observation areas: the notation $D(\mathbf{a}, \mathbf{b})$ indicates that the distance between $\mathbf{x_a}$ and $\mathbf{x_b}$ in $\mathcal{O}_1$ is the same as that between their respective PSEs, $\mathbf{y_a}$ and $\mathbf{y_b}$, in $\mathcal{O}_2$. By definition of distance, $D(\mathbf{a}, \mathbf{b}) \geq 0$, $D(\mathbf{a}, \mathbf{b}) = 0$ if and only if $\mathbf{a} = \mathbf{b}$, $D(\mathbf{a}, \mathbf{b}) = D(\mathbf{b}, \mathbf{a})$, and $D(\mathbf{a}, \mathbf{b}) + D(\mathbf{b}, \mathbf{c}) \geq D(\mathbf{a}, \mathbf{c})$.[10] We also assume the existence of the "residual" dissimilarity within the PSE pairs, *across* the two observation areas: for any PSE pair $(\mathbf{x_c}, \mathbf{y_c})$, this dissimilarity is a nonnegative number denoted $R_1(\mathbf{c})$ if computed from $\mathcal{O}_1$ to $\mathcal{O}_2$, and $R_2(\mathbf{c})$ if computed from $\mathcal{O}_2$ to $\mathcal{O}_1$. Generally, $R_1(\mathbf{c}) \neq R_2(\mathbf{c})$. The overall dissimilarity is computed as

$$S(\mathbf{x_a}, \mathbf{y_b}) = R_1(\mathbf{a}) + 2D(\mathbf{a}, \mathbf{b}) + R_2(\mathbf{b}). \tag{5}$$

Note that

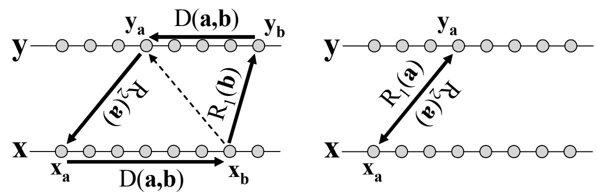
$$S(\mathbf{x_b}, \mathbf{y_a}) = R_2(\mathbf{a}) + 2D(\mathbf{a}, \mathbf{b}) + R_1(\mathbf{b})$$

[9] The choice of β is irrelevant for our discussion, because the properties of Regular Minimality and Nonconstant Self-Dissimilarity are invariant under all strictly increasing transformations of $\psi(\mathbf{x}, \mathbf{y})$. This is a fact with considerable theoretical implications, some of which is discussed in Chapter 2 of this volume (possible transformation of discrimination probabilities).

[10] Note that the first and the second $\mathbf{a}$ in $(\mathbf{a}, \mathbf{a})$, as well as in $(\mathbf{a}, \mathbf{b})$ and $(\mathbf{b}, \mathbf{a})$, generally stand for different stimuli, $\mathbf{x_a}$, and $\mathbf{y_a}$. We are essentially using here a canonical transformation of stimuli, formally introduced in Section 9.

is generally different from $S(\mathbf{x_a}, \mathbf{y_b})$, and for $\mathbf{a} = \mathbf{b}$,

$$S(\mathbf{x_a}, \mathbf{y_a}) = R_1(\mathbf{a}) + R_2(\mathbf{a}).$$



The conjunction of Regular Minimality and Nonconstant Self-Dissimilarity is ensured by positing that $R_1(\mathbf{c})$, $R_2(\mathbf{c})$ need not be the same for all $\mathbf{c}$, and that

$$|R_1(\mathbf{a}) - R_1(\mathbf{b})| < 2D(\mathbf{a}, \mathbf{b}), \quad |R_2(\mathbf{a}) - R_2(\mathbf{b})| < 2D(\mathbf{a}, \mathbf{b}).$$

These inequalities are a form of the Lipschitz condition imposed on the growth rate of R_1 and R_2. Figures 5 to 7 were generated in accordance with this "quadrilateral dissimilarity" model: we chose $\beta(s)$ in (4) as $1 - \exp(-\theta s - \eta)$, and put $D(a, b) = \gamma|a - b|$, $R_1(a) = \sin(\theta_1 a - \eta_1)$, $R_2(b) = \sin(\theta_2 b - \eta_2)$, with all Greek letters representing appropriately chosen positive constants; labels a, b in this example are related to stimuli x_a, y_b by $x_a = \sqrt{a}$ and $y_b = b$ (so that x and y are mutual PSEs if and only if $x = \sqrt{y}$).

Except for technicalities associated with R_1 and R_2 and for the fact that identically labeled $\mathbf{x}$ and $\mathbf{y}$ in (5) are generally different stimuli, the mathematical form of (5) is essentially the same as in Krumhansl's (1978) model. Somewhat more directly, the "quadrilateral dissimilarity" in (5) is related to the dissimilarity between two "uncertainty blobs," as introduced in Dzhafarov (2003b). Figure 17 provides an illustration. The "common space" in which the blobs are defined has the same meaning as the set of indices $\mathbf{a}, \mathbf{b}$ assigned to stimuli $\mathbf{x}, \mathbf{y}$ in the description of the quadrilateral dissimilarity above: that is, $\mathbf{x_a}$ and $\mathbf{y_b}$ are mapped into blobs centered at $\mathbf{a}$ and $\mathbf{b}$, respectively. The intrinsicality of metric D^* means that for a certain class of curves in the space, one can compute their lengths, and the distance between two points is defined as the length of the shortest line connecting them (a geodesic). By the assumptions made, a D^*-geodesic line connecting $\mathbf{a}$ to $\mathbf{b}$ can be produced beyond these points until it crosses the borders of the two blobs, at points $\mathbf{aa}$ and $\mathbf{bb}$. It is easy to see that no point in the first blob and no point in the second one are separated by D^*-distance exceeding $D^*(\mathbf{aa}, \mathbf{bb})$. Taking this largest possible distance for $S(\mathbf{x_a}, \mathbf{y_b})$,

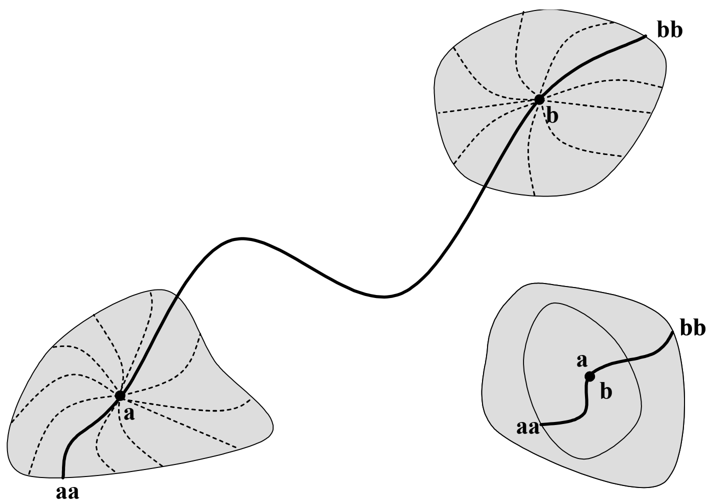


Fig. 17: Schematic representation of the "uncertainty blobs" model (Dzhafarov, 2003b). The figure plane represents a "common space" $\mathfrak{S}$ with some intrinsic metric D^* such that any two points in the space can be connected by a geodesic curve, and each geodesic curve can be produced beyond its endpoints. Each stimulus $\mathbf{x}$ in $\mathcal{O}_1$ (or $\mathbf{y}$ in $\mathcal{O}_2$) is mapped into a "blob," a D^*-circle in $\mathfrak{S}$ centered at $\mathbf{a} = \mathbf{f}_1(\mathbf{x})$ with radius $R_1(\mathbf{a})$ (respectively, centered at $\mathbf{b} = \mathbf{f}_2(\mathbf{y})$ with radius $R_2(\mathbf{b})$), such that $\mathbf{f}_1(\mathbf{x}) = \mathbf{f}_2(\mathbf{y})$ if and only if $\mathbf{x}, \mathbf{y}$ are mutual PSEs (as shown in the right lower corner). Dissimilarity $S(\mathbf{x}, \mathbf{y})$ is defined as the largest D^*-distance between the two blobs, here shown as the length of the geodesic line connecting points **aa** and **bb**.

we have then

$$S\left(\mathbf{x_a}, \mathbf{y_b}\right) = R_1\left(\mathbf{a}\right) + D^*\left(\mathbf{a}, \mathbf{b}\right) + R_2\left(\mathbf{b}\right),$$

which is identical to (5) on putting $D^*\left(\mathbf{a}, \mathbf{b}\right) = 2D\left(\mathbf{a}, \mathbf{b}\right)$. To make this identity complete, all we have to do is stipulate that the radii of the blobs change relatively slowly, in the same meaning as shown earlier,

$$\left|R_1\left(\mathbf{a}\right) - R_1\left(\mathbf{b}\right)\right| < D^*\left(\mathbf{a}, \mathbf{b}\right), \quad \left|R_2\left(\mathbf{a}\right) - R_2\left(\mathbf{b}\right)\right| < D^*\left(\mathbf{a}, \mathbf{b}\right).$$

8. RANDOM VARIABILITY IN STIMULI AND IN NEUROPHYSIOLOGICAL REPRESENTATIONS OF STIMULI

In the foregoing, we tacitly assumed that once stimulus labels have been assigned, they are always identified correctly. In a continuous stimulus set, however, stimuli are bound to be identified with only limited precision. Confining, for simplicity, the discussion to unidimensional stimuli, one and the same "apparent" physical label (i.e., the value of stimulus as known to the experimenter, say, 10 min arc, 50 cd/m^2, 30 dB) generally corresponds to at least slightly different "true" stimuli in different trials. To put this formally, apparent stimuli x, y chosen from a stimulus set correspond to random variables P_x, Q_y taking on their values in the same set of stimuli (quantities $P_x - x$, $Q_y - y$ being the measurement, or identification errors). In every trial, a pair of apparent stimuli (x, y) is probabilistically mapped into a pair of true stimuli (p, q), which in turn is mapped into the response "different" with probability $\psi\left(p, q\right)$ (about which we assume that it satisfies Regular Minimality). We have therefore

$$\psi_{app}\left(x, y\right) = \int_{q \in \mathfrak{I}} \int_{p \in \mathfrak{I}} \psi\left(p, q\right) \mathrm{d}F_x\left(p\right) \mathrm{d}F_y\left(q\right), \tag{6}$$

where $\psi_{app}\left(x, y\right)$ is discrimination probability as a function of apparent stimuli; $F_x\left(p\right)$, $F_y\left(q\right)$ are the distribution functions for true stimuli P_x, Q_y with apparent values x and y; and $\mathfrak{I}$ is the interval of all possible stimulus values.

If we assume that P_x, Q_y are stochastically independent and well-behaved (e.g., if they possess finite densities whose parameters change smoothly with the corresponding apparent stimuli, as in the classical Gaussian measurement error model), then the situation becomes formally equivalent to a Thurstonian-type model, only "perceptual space" here is replaced with the set of true stimuli. Applying the results described in Section 7.1, we come to the following conclusion: although the true discrimination probability function, $\psi(p,q)$, satisfies Regular Minimality, the apparent discrimination probability function, $\psi_{app}(x,y)$, generally does not. Indeed, it is easy to show that the minimum values of functions $y \to \psi_{app}(x,y)$ and $x \to \psi_{app}(x,y)$ computed from (6) will not generally be on a constant level (across, respectively, all possible x and all possible y); and we know that $\psi_{app}(x,y)$, being computed from a Thurstonian-type model with well-behaved random variables, cannot simultaneously exhibit the properties of Nonconstant Self-Dissimilarity and Regular Minimality. If the independent measurement errors for x and y are not negligible, therefore, one can expect apparent violations of Regular Minimality even if the principle does hold true.

This analysis, as we know, can be generalized to stochastically interdependent P_x, Q_y, provided they are selectively attributable to x and y, respectively. Stated explicitly, if P_x and Q_y are representable as in (3) (with C being a source of error common to both observation areas and C_1, C_2 being error sources specific to the first and second observation areas), and if $\pi(x,c,C_1)$ and $\theta(y,c,C_2)$ are well-behaved for any value c of C, then Regular Minimality can be violated in $\psi_{app}(x,y)$. Conversely, if $\psi_{app}(x,y)$ does not violate Regular Minimality, then the aforementioned model for measurement error cannot be correct: either measurement errors for x and y cannot be selectively attributed to x and y, or $\pi(x,c,C_1)$ and $\theta(y,c,C_2)$ are not well behaved. As an example of the latter, $\pi(x,c,C_1)$ and $\theta(y,c,C_2)$ may be deterministic quantities (see Footnote 6), or equivalently, representation (3) may have the form

$$P_x = \pi(x,C), \quad Q_y = \theta(y,C). \tag{7}$$

Clearly, when statistical error in estimating $\psi_{app}(x,y)$ is involved, all such statements should be "gradualized": thus, the aforementioned measurement error model may hold, but the variability in $\pi(x,c,C_1)$ and $\theta(y,c,C_2)$ may be too small to make the expected violations of Regular Minimality observable on a sample level.

Now, the logic of this discussion remains valid if instead of understanding P_x and Q_y as stimulus values we use these random entities to designate certain neurophysiological states, or processes evoked by stimuli x and y (which we now take as identified precisely). The mapping from stimuli to responses involves brain activity, and at least at sufficiently peripheral levels thereof we can speak of "separate" neurophysiological representations of x and y. Clearly, the response given in a given trial (same or different) depends on the values of these representations, $P_x = p$ and $Q_y = q$, irrespective of which stimuli x, y they represent. We need not decide here where the neurophysiological representation of stimuli ends and the response formation begins. Whatever the nature and complexity of P_x, Q_y, our conclusion will be the same: if $\psi(x, y)$ satisfies Regular Minimality (and manifests Nonconstant Self-Dissimilarity), then either P_x, Q_y cannot be selectively attributed to x and y, respectively (in which case they probably should not be called neurophysiological representations of x and y in the first place), or else, they are not well-behaved: for example, they covary in accordance with (7), or still simpler, are deterministic entities,

$$P_x = \pi(x), \quad Q_y = \theta(y)$$

(perhaps a kind of neurophysiological analogues of the "uncertainty blobs" depicted in Fig. 17).

A word of caution is due here: the mathematical justification for this analysis is derived from Dzhafarov (2003a, 2003b) and, strictly speaking, is confined to continuous stimulus spaces only (although not just unidimensional spaces considered here for simplicity): the definition of well-behavedness is based on the behavior of random entities P_x, Q_y in response to arbitrarily small changes in x and y. Restrictions imposed by the Regular Minimality and Nonconstant Self-Dissimilarity on possible representations of discrete stimulus sets remain to be investigated.

9. CANONICAL REPRESENTATION OF STIMULI AND DISCRIMINATION PROBABILITIES

We have seen that the conjunction of Regular Minimality and Nonconstant Self-Dissimilarity has a powerful restrictive effect on the possible theories of perceptual discrimination. In particular, it rules out two most traditional ways of modeling discrimination probabilities: by monotonically relating them to some distance measure imposed on stimulus space, and by deriving them from well-behaved random representations selectively attributable to stimuli being compared. The following characterization therefore is well worth emphasizing. Regular Minimality and Nonconstant Self-Dissimilarity are *purely psychological* properties, in the sense this term is used in Dzhafarov and Colonius (2005a, 2005b): they are completely independent of the physical measures or descriptions used to identify the individual stimuli in a stimulus space. If $\psi(\mathbf{x}, \mathbf{y})$ satisfies Regular Minimality and manifests Nonconstant Self-Dissimilarity, then the same remains true after all stimuli $\mathbf{x}$ (in $\mathcal{O}_1$) and/or all stimuli $\mathbf{y}$ (in $\mathcal{O}_2$) have been relabeled by means of arbitrary bijective transformations. In other words, insofar as the identity of a stimulus is preserved, its physical description is irrelevant. In the next section, we see that the preservation of a stimulus's identity itself has a prominent "psychological" (nonphysical) aspect.

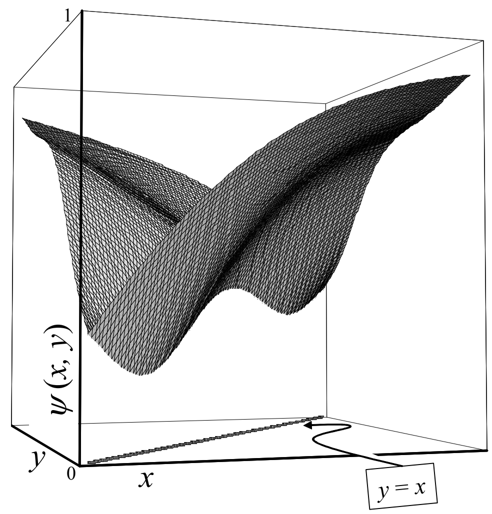


Fig. 18: A canonical form for discrimination probability function $\psi(\mathbf{x}, \mathbf{y})$ shown in Fig. 5. The PSE line $y = h(x)$ transforms into $y = x$.

In this section, we consider the identity-preserving transformations of stimuli that make the formulation of Regular Minimality especially convenient for theoretical developments. We have already used this device (*canonical transformation* of stimuli, or bringing $\psi(\mathbf{x}, \mathbf{y})$ into a *canonical form*) in the previous section. It only remains to describe it systematically.

The simplest form of Regular Minimality is observed when $\mathbf{x}$ and $\mathbf{y}$ are mutual PSEs if and only if $\mathbf{x} = \mathbf{y}$. That is,

$$\mathbf{x} \neq \mathbf{y} \Longrightarrow \psi(\mathbf{x}, \mathbf{y}) > \max\{\psi(\mathbf{x}, \mathbf{x}), \psi(\mathbf{y}, \mathbf{y})\}, \tag{8}$$

or equivalently,

$$\mathbf{x} \neq \mathbf{y} \Longrightarrow \psi(\mathbf{x}, \mathbf{x}) < \min\{\psi(\mathbf{x}, \mathbf{y}), \psi(\mathbf{y}, \mathbf{x})\}. \tag{9}$$

It is possible that in the case of discrete stimuli (such as letters of alphabet or Morse codes), Regular Minimality always holds in this form. In general, however, PSE function $\mathbf{y} = \mathbf{h}(\mathbf{x})$ may deviate from the identity function. Thinking of the situations when the stimulus sets in the two observation areas are different (see Section 11), $\mathbf{x} = \mathbf{y}$ may not even be a meaningful equality.

It is always possible, however, to relabel the stimuli in the two observation areas in such a way that (a) the stimulus sets in $\mathcal{O}_1$ and $\mathcal{O}_2$ are identical

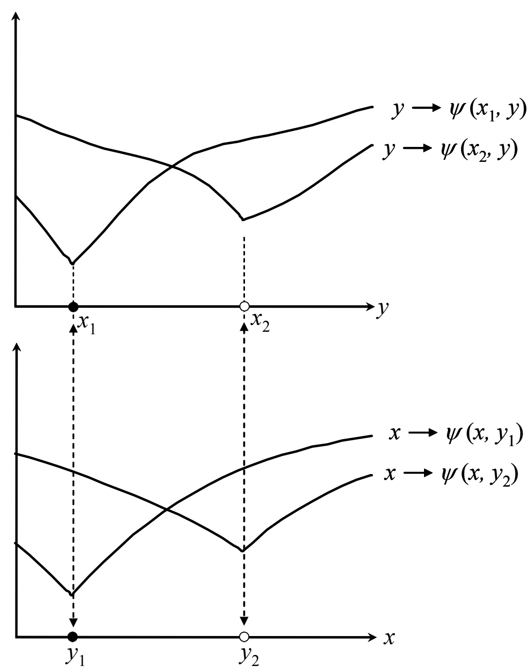


Fig. 19: Analogous to Fig. 6, but the cross-sections are those of discrimination probability function $\psi(\mathbf{x}, \mathbf{y})$ in a canonical form, as shown in Fig. 18. y is the PSE for x (equivalently, x is the PSE for y) if and only if $x = y$.

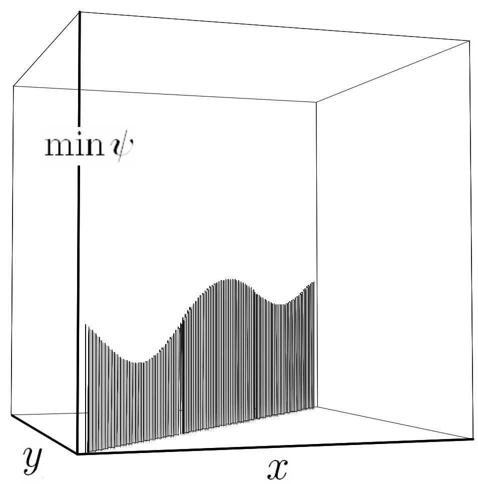


Fig. 20: Analogous to Fig. 9, but for discrimination probability function $\psi(\mathbf{x}, \mathbf{y})$ in a canonical form, as shown in Fig. 18. The transformation of the PSE line into $y = x$ does not, of course, change the contour of the minimum level function, exhibiting Nonconstant Self-Dissimilarity.

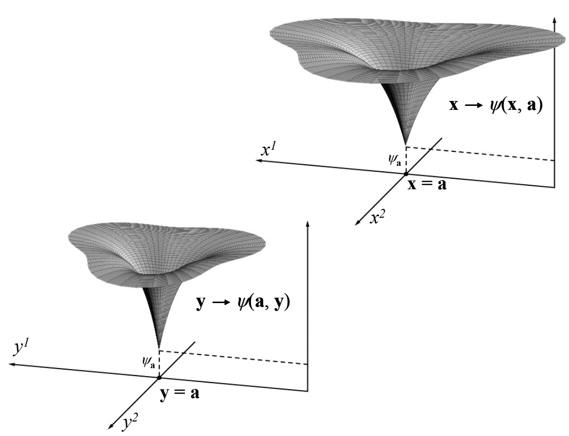


Fig. 21: Analogous to Fig. 8, but the two cross-sections are those of discrimination probability function $\psi(\mathbf{x}, \mathbf{y})$ in a canonical form. The cross-sections are made at $\mathbf{x} = \mathbf{a}$ (lower panel, with $\psi(\mathbf{a}, \mathbf{y})$ reaching its minimum at $\mathbf{y} = \mathbf{a}$) and $\mathbf{y} = \mathbf{a}$ (upper panel, with $\psi(\mathbf{x}, \mathbf{a})$ reaching its minimum at $\mathbf{x} = \mathbf{a}$).

and (b) Regular Minimality is satisfied in the simplest form, (8) to (9). We know that Regular Minimality implies a bijective correspondence between the stimulus sets in $\mathcal{O}_1$ and $\mathcal{O}_2$. It is always possible, therefore, to form a set $\mathbf{S}$ of "common stimulus labels" (or simply, "common stimuli") and to map it by means of two bijective functions, $\mathbf{f}_1^{-1}$ and $\mathbf{f}_2^{-1}$, onto the stimulus sets in $\mathcal{O}_1$ and $\mathcal{O}_2$ in such a way that, for any $\mathbf{a} \in \mathbf{S}$, $\left(\mathbf{f}_1^{-1}(\mathbf{a}), \mathbf{f}_1^{-1}(\mathbf{a})\right)$ is a pair of mutual PSEs. Equivalently, $\mathbf{f}_1(\mathbf{x}) = \mathbf{f}_2(\mathbf{y})$ if and only if $(\mathbf{x}, \mathbf{y})$ is a pair of mutual PSEs (see the legend to Fig. 17). Once this is done, one can redefine ψ by

$$\psi_{old}(\mathbf{x}, \mathbf{y}) = \psi_{new}(\mathbf{f}_1(\mathbf{x}), \mathbf{f}_2(\mathbf{y})).$$

As an example, matrix TOY_1 in Section 3 allows for the relabeling shown below,

$\mathcal{O}_1$	$\mathbf{x}_a$	$\mathbf{x}_b$	$\mathbf{x}_c$	$\mathbf{x}_d$
$\mathcal{O}_2$	$\mathbf{y}_c$	$\mathbf{y}_d$	$\mathbf{y}_b$	$\mathbf{y}_a$
common label	**A**	**B**	**C**	**D**

The following, therefore, is a canonical transformation of TOY_1:

TOY_1	$\mathbf{y}_a$	$\mathbf{y}_b$	$\mathbf{y}_c$	$\mathbf{y}_d$
$\mathbf{x}_a$	0.6	0.6	0.1	0.8
$\mathbf{x}_b$	0.9	0.9	0.8	0.1
$\mathbf{x}_c$	1	0.5	1	0.6
$\mathbf{x}_d$	0.5	0.7	1	1

$\Rightarrow$

TOY_0	**A**	**B**	**C**	**D**
A	0.1	0.8	0.6	0.6
B	0.8	0.1	0.9	0.9
C	1	0.6	0.5	1
D	1	1	0.7	0.5

For continuous stimuli, given a PSE function, $\mathbf{y} = \mathbf{h}(\mathbf{x})$, any pair of functions $\mathbf{f}_1 \equiv \mathbf{f}$, $\mathbf{f}_2 \equiv \mathbf{h} \circ \mathbf{f}$, for any (bijective) $\mathbf{f}$, provides a canonical transformation. Figures 18, 19, 20, and 21 illustrate canonical forms for our earlier examples.

10. PSYCHOLOGICAL IDENTITY OF STIMULI

Up to this point, we implicitly assumed that all stimuli in either of the observation areas are psychologically distinct, in the following sense: if $\mathbf{x}_1 \neq \mathbf{x}_2$ in $\mathcal{O}_1$, then at least for one stimulus $\mathbf{y}$ in $\mathcal{O}_2$,

$$\psi(\mathbf{x}_1, \mathbf{y}) \neq \psi(\mathbf{x}_2, \mathbf{y});$$

and analogously for any $\mathbf{y}_1 \neq \mathbf{y}_2$ in $\mathcal{O}_2$. Put differently, if $\psi(\mathbf{x}_1, \mathbf{y}) = \psi(\mathbf{x}_2, \mathbf{y})$ for all $\mathbf{y}$ in $\mathcal{O}_2$, then $\mathbf{x}_1 = \mathbf{x}_2$; and if $\psi(\mathbf{x}, \mathbf{y}_1) = \psi(\mathbf{x}, \mathbf{y}_2)$ for all $\mathbf{x}$ in $\mathcal{O}_1$, then $\mathbf{y}_1 = \mathbf{y}_2$. On a moment's reflection, this is not a real constraint. If $\mathbf{x}_1 \neq \mathbf{x}_2$, but $\psi(\mathbf{x}_1, \mathbf{y}) = \psi(\mathbf{x}_2, \mathbf{y})$ for all $\mathbf{y}$ in $\mathcal{O}_2$ (in which case, one can say that that $\mathbf{x}_1$ and $\mathbf{x}_2$ are "*psychologically equal*"), one can always relabel the stimuli so that $\mathbf{x}_1$ and $\mathbf{x}_2$ receive identical labels. For example, if aperture colors are initially labeled by their radiometric spectra (radiometric intensity as a function of wavelength), we know that there are an infinity of distinct spectra that are, for a given level of adaptation, equally distinguishable from any given spectrum (metameric). As a result, all mutually metameric colors can be merged and assigned a single label, say, a triple of CIE color coordinates. Figure 22 provides a schematic illustration.

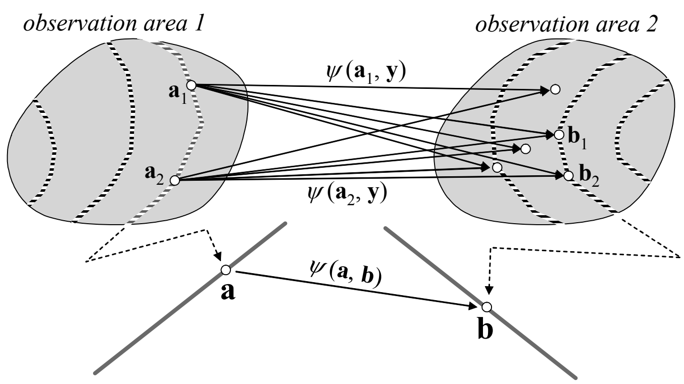


Fig. 22: Equivalence class of psychologically equal stimuli (shown by striped lines). $\mathbf{a}_1$ and $\mathbf{a}_2$ in $\mathcal{O}_1$ are psychologically equal because $\psi(\mathbf{a}_1, \mathbf{y}) = \psi(\mathbf{a}_2, \mathbf{y})$ for every $\mathbf{y}$ in $\mathcal{O}_2$; these two stimuli therefore are assigned a common label, $\mathbf{a}$. Equivalence classes $\mathbf{a}$ and $\mathbf{b}$ are treated as single stimuli in $\mathcal{O}_1$ and $\mathcal{O}_2$, respectively, with $\psi(\mathbf{a}, \mathbf{b})$ put equal to $\psi(\mathbf{x}, \mathbf{y})$ for any $\mathbf{x} \in \mathbf{a}$, $\mathbf{y} \in \mathbf{b}$. The Regular Minimality condition is assumed to hold for these "reduced" stimulus sets (sets of equivalence classes, shown by the two straight lines).

The example below shows a matrix of discrimination probabilities that, following the procedure of "lumping together" psychologically equal stimuli, yields our matrix TOY_1.

		$\mathbf{y}_a$	$\mathbf{y}_a$	$\mathbf{y}_b$	$\mathbf{y}_b$	$\mathbf{y}_c$	$\mathbf{y}_d$	$\mathbf{y}_d$
	TOY$_{11}$	$\mathbf{y}_1$	$\mathbf{y}_2$	$\mathbf{y}_3$	$\mathbf{y}_4$	$\mathbf{y}_5$	$\mathbf{y}_6$	$\mathbf{y}_7$
$\mathbf{x}_a$	$\mathbf{x}_1$	0.6	0.6	0.6	0.6	0.1	0.8	0.8
$\mathbf{x}_a$	$\mathbf{x}_2$	0.6	0.6	0.6	0.6	0.1	0.8	0.8
$\mathbf{x}_b$	$\mathbf{x}_3$	0.9	0.9	0.9	0.9	0.8	0.1	0.1
$\mathbf{x}_b$	$\mathbf{x}_4$	0.9	0.9	0.9	0.9	0.8	0.1	0.1
$\mathbf{x}_c$	$\mathbf{x}_5$	1	1	0.5	0.5	1	0.6	0.6
$\mathbf{x}_d$	$\mathbf{x}_6$	0.5	0.5	0.7	0.7	1	1	1
$\mathbf{x}_d$	$\mathbf{x}_7$	0.5	0.5	0.7	0.7	1	1	1

$\Rightarrow$

TOY$_1$	$\mathbf{y}_a$	$\mathbf{y}_b$	$\mathbf{y}_c$	$\mathbf{y}_d$
$\mathbf{x}_a$	0.6	0.6	0.1	0.8
$\mathbf{x}_b$	0.9	0.9	0.8	0.1
$\mathbf{x}_c$	1	0.5	1	0.6
$\mathbf{x}_d$	0.5	0.7	1	1

Thus, $\{\mathbf{x}_1, \mathbf{x}_2\} \to \mathbf{x}_a$, $\{\mathbf{y}_1, \mathbf{y}_2\} \to \mathbf{y}_a$, and so forth. In this example, each equivalence class of psychologically equal stimuli in $\mathcal{O}_1$ bijectively maps onto an equivalence class of psychologically equal stimuli in $\mathcal{O}_2$: $\{\mathbf{x}_1, \mathbf{x}_2\} \longleftrightarrow \{\mathbf{y}_1, \mathbf{y}_2\}$, $\{\mathbf{x}_5\} \longleftrightarrow \{\mathbf{y}_5\}$, and so forth. Although we cannot think of a realistic counterexample, on this level of abstraction there is no reason to postulate such a correspondence. The matrix below illustrates the point.

		$\mathbf{y}_a$	$\mathbf{y}_c$	$\mathbf{y}_b$	$\mathbf{y}_c$	$\mathbf{y}_b$	$\mathbf{y}_d$	$\mathbf{y}_d$
	TOY$_{12}$	$\mathbf{y}_1$	$\mathbf{y}_2$	$\mathbf{y}_3$	$\mathbf{y}_4$	$\mathbf{y}_5$	$\mathbf{y}_6$	$\mathbf{y}_7$
$\mathbf{x}_a$	$\mathbf{x}_1$	0.6	0.1	0.6	0.1	0.6	0.8	0.8
$\mathbf{x}_b$	$\mathbf{x}_2$	0.9	0.8	0.9	0.8	0.9	0.1	0.1
$\mathbf{x}_c$	$\mathbf{x}_3$	1	1	0.5	1	0.5	0.6	0.6
$\mathbf{x}_d$	$\mathbf{x}_4$	0.5	1	0.7	1	0.7	1	1
$\mathbf{x}_d$	$\mathbf{x}_5$	0.5	1	0.7	1	0.7	1	1
$\mathbf{x}_d$	$\mathbf{x}_6$	0.5	1	0.7	1	0.7	1	1
$\mathbf{x}_d$	$\mathbf{x}_7$	0.5	1	0.7	1	0.7	1	1

$\Rightarrow$

TOY$_1$	$\mathbf{y}_a$	$\mathbf{y}_b$	$\mathbf{y}_c$	$\mathbf{y}_d$
$\mathbf{x}_a$	0.6	0.6	0.1	0.8
$\mathbf{x}_b$	0.9	0.9	0.8	0.1
$\mathbf{x}_c$	1	0.5	1	0.6
$\mathbf{x}_d$	0.5	0.7	1	1

This matrix, too, following the relabeling shown, yields matrix TOY$_1$, but the equivalence classes in $\mathcal{O}_1$ cannot be paired with equinumerous equivalence classes in $\mathcal{O}_2$ (e.g., $\{\mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_6, \mathbf{x}_7\}$ does not have a four-element counterpart in $\mathcal{O}_2$). It is critical for the requirement of Regular Minimality, however, that the resulting sets of the equivalence classes themselves contain equal numbers of elements in the two observation areas: $\{\mathbf{x}_a, \mathbf{x}_b, \mathbf{x}_c, \mathbf{x}_d\}$ and $\{\mathbf{y}_a, \mathbf{y}_b, \mathbf{y}_c, \mathbf{y}_d\}$. Regular Minimality, in effect, says that one can establish a bijection between the equivalence classes in $\mathcal{O}_1$ and the equivalence classes in $\mathcal{O}_2$ in such a way that the corresponding elements (equivalence classes treated as redefined stimuli) are mutual PSEs.

11. VARIETY OF PARADIGMS

Here, we describe a variety of meanings in which one can understand same-different judgments, observation areas, and the very terms *stimuli* and *perceiver*.

It was mentioned in the introductory paragraph of this chapter that the sameness or difference of two stimuli can be judged "overall" or "in a specified respect." Expanding on that, the definition of a discrimination probability function, (1), can be generalized in two ways:

$$\psi(\mathbf{x}, \mathbf{y}) = \Pr[\mathbf{x} \text{ and } \mathbf{y} \text{ are different with respect to } \mathcal{A}], \tag{10}$$

meaning that all differences other than those in a designated property $\mathcal{A}$ (shape, size, color, etc.) should be ignored; and

$$\psi(\mathbf{x}, \mathbf{y}) = \Pr[\mathbf{x} \text{ and } \mathbf{y} \text{ are different in any respect other than } \mathcal{B}], \tag{11}$$

meaning that any differences in a designated property $\mathcal{B}$ (which again can be shape, size, color, etc.) should be ignored. As follows from our discussion of the two distinct observation areas, the "generic" definition (1) is in fact a special case of (11), with $\mathcal{B}$ designating the perceptual difference between the two observation areas.

In psychophysical experiments, the observation areas usually mean different locations in space or time, but the scope of possible meanings is much broader. Thus, $\mathcal{O}_1$ and $\mathcal{O}_2$ may be defined by the modality of stimulus, as in the grapheme-morpheme comparisons (e.g., a written syllable $\mathbf{x}$ compared with a pronounced syllable $\mathbf{y}$): in this case, the ordering of two stimuli in $(\mathbf{x}, \mathbf{y})$ is determined by which of them is written and which pronounced, irrespective of their temporal order. As another example, when a green color patch and a red color patch of variable intensities are compared in brightness, the two fixed colors serve to define the two observation areas, irrespective of the spatial positions or temporal order of the patches.

A combination of several such observation-area-defining attributes (say, colors $\times$ locations) or simply more than just two values of a given attribute (say, several locations) may lead to multiple observation areas, in which case stimulus pairs should be encoded as $((\mathbf{x},o), (\mathbf{y},o'))$, where $\mathbf{x}, \mathbf{y}$ are labels identifying the stimuli in all respects except for their observation areas, the latter being designated by o, o' (with $o \neq o'$). Although the relation among $\psi((\mathbf{x},o), (\mathbf{y},o'))$ for different pairs of distinct o, o' is beyond the scope of this chapter, our hypothesis is that Regular Minimality should be satisfied for all such pairs.

In some applications, the difference between the observation areas is known or assumed to be immaterial. Thus, when asked to compare the attractiveness of two photographs, their spatial arrangement may very well be immaterial (or even undefined, if the perceiver is allowed to move them freely). Our analysis still applies to such cases: although formally distinguishing $(\mathbf{a}, \mathbf{b})$ and $(\mathbf{b}, \mathbf{a})$, we simply impose the order-balance, or symmetry condition, $\psi(\mathbf{x}, \mathbf{y}) = \psi(\mathbf{y}, \mathbf{x})$. Counterintuitive as it may sound, the order-balancedness does not imply that Regular Minimality can only be satisfied in a canonical form. If $\psi(\mathbf{x}, \mathbf{y}) = \psi(\mathbf{y}, \mathbf{x})$, the PSE relation $\mathbf{y} = \mathbf{h}(\mathbf{x})$ is equivalent to the PSE relation $\mathbf{x} = \mathbf{h}(\mathbf{y})$. Comparing this to properties (RM1 to RM3), in Section 3, we see that $\mathbf{h} \equiv \mathbf{h}^{-1}$. The functional equation $\mathbf{h} \equiv \mathbf{h}^{-1}$ is known as Babbage's equation (see Kuczma, Choczewski, & Ger, 1990), and it has more solutions than just an identity function, although the latter often is the only realistic solution (e.g., it is the only nondecreasing solution in the case of unidimensional stimuli).

One can significantly broaden the class of paradigms which can be treated as same-different comparisons by applying the term *stimuli*, in a purely formal way, to any two sets of entities, $\mathbf{M}_1$ and $\mathbf{M}_2$ (stimuli in the first and second observation areas, respectively), that can be endowed with a probability function $\psi : \mathbf{M}_1 \times \mathbf{M}_2 \to [0, 1]$. The term *perceiver* then, may designate any device or computational procedure which, in response to any ordered pair $\mathbf{x} \in \mathbf{M}_1$, $\mathbf{y} \in \mathbf{M}_2$, produces a certain output with probability $\psi(\mathbf{x}, \mathbf{y})$. We propose that this output can be interpreted as meaning "$\mathbf{x}$ is different from $\mathbf{y}$" if and only if function $\psi(\mathbf{x}, \mathbf{y})$ satisfies Regular Minimality. In other words, Regular Minimality may serve as a criterion (necessary and sufficient condition) for the inclusion of otherwise vastly different paradigms in the category of same-different comparisons.

To give a very "nonpsychophysical" example, consider a class $\mathbf{M}$ of statistical models, and a class $\mathfrak{D}$ of possible results of some experiment. Each model from $\mathbf{M}$ can be fitted to each possible result, and rejected or retained in accordance with some statistical criterion $\mathcal{C}$. Given two models, $\mathbf{x}, \mathbf{y} \in \mathbf{M}$, and a certain experimental outcome $d_0 \in \mathfrak{D}$, consider a procedure that consists of (a) fitting $\mathbf{x}$ to d_0 and specifying thereby all free parameters of $\mathbf{x}$ (b) repeatedly generating outcomes $d \in \mathfrak{D}$ by means of thus specified $\mathbf{x}$; and (c) fitting $\mathbf{y}$ to every generated outcome d and rejecting or retaining it in accordance with criterion $\mathcal{C}$. Then the probability $\psi(\mathbf{x}, \mathbf{y})$ with which model $\mathbf{y}$ is rejected by an outcome generated by model $\mathbf{x}$ can be taken as the probability of discriminating $\mathbf{y}$ from $\mathbf{x}$, provided $\psi(\mathbf{x}, \mathbf{y})$ satisfies Regular Minimality. In this example, the "observation area" of a model is defined by the role in which this model is employed: $\mathbf{M}_1$ represents the models specified by fitting them to d_0 and used to generate outcomes d, whereas $\mathbf{M}_2$ represents the models tested by applying them to thus generated d. The "perceiver" in this case, from whose "point of view" the models are being compared, is the entire computational procedure, specified by d_0 and $\mathcal{C}$. One would normally expect that Regular Minimality for a well-defined class of models should be satisfied canonically, (8) to (9). This is, however, a secondary consideration, because the models in $\mathbf{M}_2$, as we know, can always be relabeled so that the PSE of model $\mathbf{x} \in \mathbf{M}_1$ is assigned label $\mathbf{x}$.

As another example, let $\mathbf{M}_1$ be a set of categories or sources, each of which can be exemplified by a variety of entities (e.g., lung dysfunctions exemplified by X-ray films), and let $\mathbf{M}_2$ be the same set of categories or sources when they are judged to be or not to be exemplified by a given entity ("does this X-ray film indicate this lung dysfunction?"). The probability with which an entity exemplifying category $\mathbf{x}$ is judged not to belong to category $\mathbf{y}$ then can be taken as $\psi(\mathbf{x}, \mathbf{y})$, provided ψ satisfies Regular Minimality. Again, in a well-calibrating expert system, one would expect Regular Minimality to hold canonically, but any form of Regular Minimality can be recalibrated into a canonical form.

12. CONCLUSION

The principle according to which any well-defined discrimination probability function $\psi(\mathbf{x}, \mathbf{y})$, defined by (1), (10), or (11), should satisfy Regular Minimality, seems to have all the hallmarks of a fundamental law:

(A) It cannot be derived from more elementary properties of discrimination probabilities. In this respect, it is very different from the Regular Mediality principle for greater-less judgments (Section 2).

(B) It is conceptually simple, almost obvious, yet has unexpectedly restrictive consequences for theoretical modeling of discrimination probabilities (Sections 7.1, 7.2, and 8), especially when combined with the property of Nonconstant Self-Dissimilarity (Section 4).

(C) Its conceptual plausibility allows one to use it as a criterion for classifying a paradigm into the category of same-different judgments (Section 11).

(D) It is born out by available experimental evidence (although much more work remains to be done before one can call this evidence abundant; see Section 6).

(E) It can serve as a benchmark against which to consider empirical evidence: if the latter exhibits deviations from Regular Minimality, one is warranted to look for other possible causes before discarding the principle itself (Section 8).

We conclude this chapter by a brief comment on the last characterization. Stimulus uncertainty, which we discussed in Section 8 is only one of many factors which, if Regular Minimality does in fact hold true, predictably leads to its apparent violations in data. Skipping over the relatively obvious issue of sampling errors (both in estimating probabilities and in choosing a representative subset of a stimulus space), perhaps the most important factor working against the principle of Regular Minimality in real-life experiments is the possibility of mixing together discrimination probability functions with different PSE functions. It is easy to see that if Regular Minimality is satisfied in both $\psi_1(\mathbf{x}, \mathbf{y})$ and $\psi_2(\mathbf{x}, \mathbf{y})$, defined on the same set of stimuli, and if their respective PSE functions are $\mathbf{y} = \mathbf{h}_1(\mathbf{x})$ and $\mathbf{y} = \mathbf{h}_2(\mathbf{x})$, then linear combinations $\alpha\psi_1(\mathbf{x}, \mathbf{y}) + (1-\alpha)\psi_2(\mathbf{x}, \mathbf{y})$ $(0 \leq \alpha \leq 1)$ will generally violate Regular Minimality, unless $\mathbf{h}_1 \equiv \mathbf{h}_2$. In a psychophysical experiment with continuous stimuli (like the one related to Fig. 14), it seems desirable to use very large numbers of replications per stimulus pair to increase the reliability of the statistical estimates of discrimination probabilities. In a very long experiment, however, it seems likely that the discrimination probability function would gradually change, because of which the resulting probability estimates will be those of a linear combination of functions $\psi_t(\mathbf{x}, \mathbf{y})$, with t being the time at which $(\mathbf{x}, \mathbf{y})$ was presented. If PSE functions $\mathbf{y} = \mathbf{h}_t(\mathbf{x})$ also vary in time, this mixture may very well exhibit violations of Regular Minimality. Analogous considerations apply to group experiments: there we may have to deal with heterogeneous mixtures of functions $\psi_k(\mathbf{x}, \mathbf{y})$, with k representing different members of a group.

Acknowledgment. This research was supported by National Science Foundation grant SES 0318010 to Purdue University.

References

Dai, H., Versfeld, N. J., & Green, D. M. (1996). The optimum decision rules in the same-different paradigm. *Perception and Psychophysics*, 58, 1–9.

Dzhafarov, E. N. (2002a). Multidimensional Fechnerian scaling: Probability-distance hypothesis. *Journal of Mathematical Psychology,* 46, 352–374.

Dzhafarov, E. N. (2002b). Multidimensional Fechnerian scaling: Pairwise comparisons, regular minimality, and nonconstant self-similarity. *Journal of Mathematical Psychology,* 46, 583–608.

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

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

Dzhafarov, E. N. (2003c). Selective influence through conditional independence. *Psychometrika*, 68, 7–26.

Dzhafarov, E. N., & Colonius, H. (2005a). Psychophysics without physics: A purely psychological theory of Fechnerian Scaling in continuous stimulus spaces. *Journal of Mathematical Psychology*, 49, 1–50.

Dzhafarov, E.N., & Colonius, H. (2005b). Psychophysics without physics: Extension of Fechnerian Scaling from continuous to discrete and discrete-continuous stimulus spaces. *Journal of Mathematical Psychology*, 49, 125–141.

Ennis, D. M. (1992). Modeling similarity and identification when there are momentary fluctuations in psychological magnitudes. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 279–298). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Ennis, D. M., Palen, J. J., & Mullen, K. (1988). A multidimensional stochastic theory of similarity. *Journal of Mathematical Psychology*, 32, 449–465.

Falmagne, J.-C. (1971). The generalized Fechner problem and discrimination. *Journal of Mathematical Psychology*, 8, 22–43.

Fechner, G. T. (1887). Über die psychischen Massprinzipien und das Webersche Gesetz [On the principles of mental measurement and Weber's Law]. Philosophische Studien, 4, 161–230.

Indow, T. (1998). Parallel shift of judgment-characteristic curves according to the context in cutaneous and color discrimination. In C. E. Dowling, F. S. Roberts, P. Theuns (Eds.), *Recent progress in mathematical psychology* (pp. 47–63). Mahwah, NJ: Erlbaum Associates, Inc.

Indow, T., Robertson, A. R., von Grunau, M., & Fielder, G. H. (1992). Discrimination ellipsoids of aperture and simulated surface colors by matching and paired comparison. *Color Research and Applications*, 17, 6–23.

Krumhansl, C. L. (1978). Concerning the applicability of geometric models to similarity data: The interrelationship between similarity and spatial density. *Psychological Review*, 85, 445–463.

Kuczma, M., Choczewski, B., & Ger, R. (1990). *Iterative functional equations.* Cambridge, England: Cambridge University Press.

Luce, R. D. (1977). Thurstone's discriminal processes fifty year later. *Psychometrika*, 42, 461–489.

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

Luce, R. D., & Galanter, E. (1963). Discrimination. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology*, v. 1 (pp. 103–189). New York: Wiley.

Rothkopf, E. Z. (1957). A measure of stimulus similarity and errors in some paired-associate learning tasks. *Journal of Experimental Psychology*, 53, 94–102.

Scheerer, E. (1987). The unknown Fechner. *Psychological Research*, 49, 197–202.

Sorkin, R. D. (1962). Extension of the theory of signal detectability to matching paradigms in psychoacoustics. *Journal of the Acoustical Society of America*, 34, 1745–1751.

Suppes, P., & Zinnes, J. L. (1963). Basic measurement theory. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology*, v. 1 (pp. 3–76). New York: Wiley.

Takane, Y., & Sergent, J. (1983). Multidimensional scaling models for reaction times and same-different judgments, *Psychometrika*, 48, 393–423.

Thomas, R. D. (1996). Separability and independence of dimensions within the same-different judgment task. *Journal of Mathematical Psychology*, 40, 318–341.

Thurstone, L. L. (1927a). Psychophysical analysis. *American Journal of Psychology*, 38, 368–389.

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

Wish, M. (1967). A model for the perception of Morse code-like signals. *Human Factors*, 9, 529–540.

Zhang, J. (2004). Dual scaling of comparison and reference stimuli in multidimensional psychological space. *Journal of Mathematical Psychology*, 48, 409–424.

Zimmer, K., & Colonius, H. (2000). Testing a new theory of Fechnerian scaling: The case of auditory intensity discrimination. *Journal of the Acoustical Society of America*, 108, 2596.

Zinnes, J. L., & MacKay, D. B. (1983). Probabilistic multidimensional scaling: Complete and incomplete data. *Psychometrika*, 48, 27–48.