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

2
Reconstructing Distances Among Objects from Their Discriminability

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

[1] Purdue University
[2] Universität Oldenburg

1. INTRODUCTION

The problem of reconstructing distances among stimuli from some empirical measures of pairwise dissimilarity is old. The measures of dissimilarity are numerous, including numerical ratings of (dis)similarity, classifications of stimuli, correlations among response variables, errors of substitution, and many others (Everitt & Rabe-Hesketh, 1997; Suppes, Krantz, Luce, & Tversky, 1989; Sankoff & Kruskal, 1999; Semple & Steele, 2003). Formal representations of proximity data, like Multidimensional Scaling (MDS; Borg & Groenen, 1997; Kruskal & Wish, 1978) or Cluster Analysis (Corter, 1996; Hartigan, 1975), serve to describe and display data structures by embedding them in low-dimensional spatial or graph-theoretical configurations, respectively. In MDS, one embeds data points in a low-dimensional Minkowskian (usually, Euclidean) space so that distances are monotonically (in the metric version, proportionally) related to pairwise dissimilarities. In Cluster Analysis, one typically represents proximity relations by a series of partitions of the set of stimuli resulting in a graph-theoretic tree structure with ultrametric or additive-tree metric distances.

Discrimination probabilities,

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

which we discussed in Chapter 1, occupy a special place among available measures of pairwise dissimilarity. The ability of telling two objects apart or identifying them as being the same (in some respect or overall) is arguably the most basic cognitive ability in biological perceivers and the most basic requirement of intelligent technical systems. At least this seems to be a plausible view, granting it is not self-evident. It is therefore a plausible position that a metric appropriately computed from the values of $\psi(\mathbf{x}, \mathbf{y})$ may be viewed as the "subjective metric," a network of distances "from the point of view" of a perceiver.

As discussed in Chapter 1, the notion of a perceiver has a variety of possible meanings, including even cases of "paper-and-pencil" perceivers, abstract computational procedures assigning to every pair $\mathbf{x}, \mathbf{y}$ the probability $\psi(\mathbf{x}, \mathbf{y})$ (subject to Regular Minimality). The example given in Chapter 1 was that of $\psi(\mathbf{x}, \mathbf{y})$ being the probability with which a data set (in a particular format) generated by a statistical model, $\mathbf{x}$, rejects (in accordance with some criterion) a generally different statistical model, $\mathbf{y}$. The pairwise determinations of sameness/difference in this example (meaning, model $\mathbf{y}$ is retained/rejected when applied to a data set generated by model $\mathbf{x}$) are usually readily available and simple. It is an attractive possibility, therefore, to have a general algorithm in which one can use these pairwise determinations to compute distances among conceptual objects (here, statistical models). The alternative, an a priori choice of a distance measure between two statistical models, may be less obvious and more difficult to justify.

This chapter provides an informal introduction to Fechnerian Scaling, a metric-from-discriminability theory which has been gradually developed by the present authors in the recent years (Dzhafarov, 2002a, 2002b, 2002c, 2002d; 2003a, 2003b; Dzhafarov & Colonius, 1999, 2001, 2005a, 2005b). Its historical roots, however, can be traced back to the work of G. T. Fechner (1801–1887). To keep the presentation on a nontechnical level, we provide details for only the mathematically simplest case of Fechnerian Scaling, the case of discrete stimulus sets (such as letters of alphabets or Morse codes); only a simplified and abridged account of the application of Fechnerian Scaling to continuous stimulus spaces is given. Notation conventions are the same as in Chapter 1.

1.1. Example

Consider the toy matrix used in Chapter 1, presented in a canonical form,

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

This matrix is used throughout to illustrate various points. We describe a computational procedure, *Fechnerian Scaling,* which, when applied to such matrices, produces a matrix of distances we call *Fechnerian*. Intuitively, they reflect the degree of subjective dissimilarity among the stimuli, "from the point of view" of the perceiver (organism, group, technical device, or a computational procedure) to whom stimuli $\mathbf{x}, \mathbf{y} \in \{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}\}$ were presented pairwise and whose responses (interpretable as "same" and "different") were used to compute the probabilities $\psi(\mathbf{x}, \mathbf{y})$ shown as the matrix entries. In addition, when the set of stimuli is finite, Fechnerian Scaling produces a set of what we call *geodesic loops*, the shortest (in some well-defined sense) chains of stimuli leading from one given object to another given object and back. Thus, when applied to our matrix TOY_0, Fechnerian Scaling yields the following two matrices:

L_0	A	B	C	D
A	A	ACBA	ACA	ADA
B	BACB	B	BCB	BDCB
C	CAC	CBC	C	CDC
D	DAD	DCBD	DCD	D

G_0	A	B	C	D
A	0	1.3	1	1
B	1.3	0	0.9	1.1
C	1	0.9	0	0.7
D	1	1.1	0.7	0

We can see in matrix L_0, for instance, that the shortest (geodesic) loop connecting $\mathbf{A}$ and $\mathbf{B}$ within the four-element space $\{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}\}$ is

$$\mathbf{A} \rightarrow \mathbf{C} \rightarrow \mathbf{B} \rightarrow \mathbf{A},$$

whereas the geodesic loop connecting $\mathbf{A}$ and $\mathbf{C}$ in the same space is

$$\mathbf{A} \rightarrow \mathbf{C} \rightarrow \mathbf{A}.$$

The lengths of these geodesic loops (whose computation will be explained later) are taken to be the Fechnerian distances between $\mathbf{A}$ and $\mathbf{B}$ and between $\mathbf{A}$ and $\mathbf{C}$, respectively. As we see in matrix G_0, the Fechnerian distance between $\mathbf{A}$ and $\mathbf{B}$ is 1.3 times the Fechnerian distance between $\mathbf{A}$ and $\mathbf{C}$.

We should recall some basic facts from Chapter 1:

(1) The row stimuli and the column stimuli in TOY_0 belong to two *distinct observation areas* (say, row stimuli are those presented on the left, or chronologically first, the column stimuli are presented on the right, or second).

(2) $\{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}\}$ are psychologically distinct, that is, no two rows or two columns in the matrix are identical (if they were, they would be merged into a single one).

(3) TOY_0 may be the result of a canonical relabeling of a matrix in which the minima lie outside the main diagonal, such as

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

The physical identity of the $\{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}\}$ in TOY_0 may therefore be different for the row stimuli and the column stimuli.

1.2. Features of Fechnerian Scaling

(A) Regular Minimality is the cornerstone of Fechnerian Scaling, and in the case of discrete stimulus sets, it is essentially the only prerequisite for Fechnerian Scaling. Due to Regular Minimality, we can assume throughout most of this chapter that our stimulus sets are canonically (re)labeled (as in TOY_0), so that

$$\mathbf{x} \neq \mathbf{y} \Longrightarrow \psi(\mathbf{x}, \mathbf{y}) > \max\{\psi(\mathbf{x}, \mathbf{x}), \psi(\mathbf{y}, \mathbf{y})\}, \tag{2}$$

or equivalently,

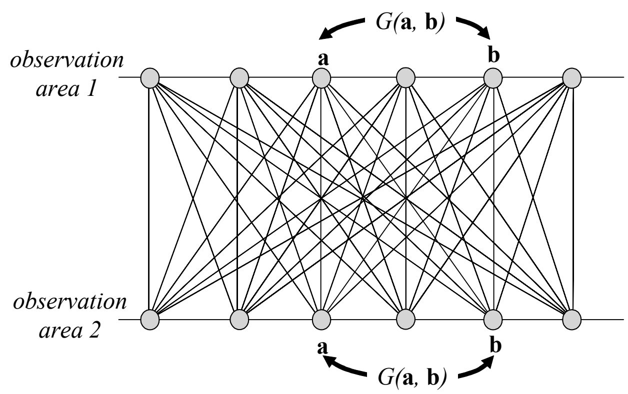
$$\mathbf{x} \neq \mathbf{y} \Longrightarrow \psi(\mathbf{x}, \mathbf{x}) < \min\{\psi(\mathbf{x}, \mathbf{y}), \psi(\mathbf{y}, \mathbf{x})\}. \tag{3}$$

In accordance with the discussion of the fundamental properties of discrimination probabilities (Chapter 1), Fechnerian Scaling does not presuppose that $\psi(\mathbf{x}, \mathbf{x})$ is the same for all $\mathbf{x}$ (Nonconstant Self-Dissimilarity), or that $\psi(\mathbf{x}, \mathbf{y}) = \psi(\mathbf{y}, \mathbf{x})$ (Asymmetry).

(B) The logic of Fechnerian Scaling is very different from the existing techniques of metrizing stimulus spaces (such as MDS) in the following respect: Fechnerian distances are computed *within* rather than *across* the two observation areas. In other words, the Fechnerian distance between $\mathbf{a}$ and $\mathbf{b}$ does not mean a distance between $\mathbf{a}$ presented first (or on the left) and $\mathbf{b}$ presented second (or on the right). Rather, we should logically distinguish $G^{(1)}(\mathbf{a}, \mathbf{b})$, the distance between $\mathbf{a}$ and $\mathbf{b}$ in the first observation area, from $G^{(2)}(\mathbf{a}, \mathbf{b})$, the distance between $\mathbf{a}$ and $\mathbf{b}$ in the second observation area. This must not come as a surprise if one keeps in mind that $\mathbf{a}$ and $\mathbf{b}$ in the first observation area are generally perceived differently from $\mathbf{a}$ and $\mathbf{b}$ in the second observation area. As it turns out, however, if Regular Minimality is satisfied and the stimulus set is put in a canonical form, then it follows from the general theory that

$$G^{(1)}(\mathbf{a}, \mathbf{b}) = G^{(2)}(\mathbf{a}, \mathbf{b}) = G(\mathbf{a}, \mathbf{b}).$$

This is illustrated in the diagram below, where the line connecting a stimulus in $\mathcal{O}1$ with a stimulus in $\mathcal{O}1$ ($\mathcal{O}$ standing for observation area) represents the probability ψ of their discrimination. Note that, for a given $\mathbf{a}, \mathbf{b}$, distance $G(\mathbf{a}, \mathbf{b})$ is computed, in general, from $\psi(\mathbf{x}, \mathbf{y})$ for all $\mathbf{x}, \mathbf{y}$, and not just from $\psi(\mathbf{a}, \mathbf{b})$. Later all of this is explained in detail.



(C) In TOY_0, a geodesic loop containing two given stimuli is defined uniquely. In general, however, this need not be the case: there may be more than one loop of the shortest possible length. Moreover, when the set of stimuli is infinitely large, whether discrete or continuous, geodesic loops may not exist at all, and the Fechnerian distance between two stimuli is then defined as the greatest lower bound (rather than minimum) of lengths of all loops that include these two stimuli.

1.3. Fechnerian Scaling and Multidimensional Scaling

MDS, when applied to discrimination probabilities, serves as a convenient reference against which to consider the procedure of Fechnerian Scaling. Assuming that discrimination probabilities $\psi(\mathbf{x}, \mathbf{y})$ are known precisely, the classical MDS is based on the assumption that for some metric $d(\mathbf{x}, \mathbf{y})$ (distance function) and some increasing transformation β,

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

This is a prominent instance of what is called the *probability-distance hypothesis* in Dzhafarov (2002b). Recall that the defining properties of a metric d are as follows: (A) $d(\mathbf{a}, \mathbf{b}) \geq 0$; (B) $d(\mathbf{a}, \mathbf{b}) = 0$ if and only if $\mathbf{a} = \mathbf{b}$; (C) $d(\mathbf{a}, \mathbf{c}) \leq d(\mathbf{a}, \mathbf{b}) + d(\mathbf{b}, \mathbf{c})$; (D) $d(\mathbf{a}, \mathbf{b}) = d(\mathbf{b}, \mathbf{a})$. In addition, one assumes in MDS that metric d belongs to a predefined class, usually the class of Minkowski metrics with exponents between 1 and 2.

It immediately follows from (A), (B), (D), and the monotonicity of β that for any distinct $\mathbf{x}, \mathbf{y}$,

$$\begin{array}{ll} \psi(\mathbf{x},\mathbf{y}) = \psi(\mathbf{y},\mathbf{x}) & \text{(Symmetry)} \\ \psi(\mathbf{x},\mathbf{x}) = \psi(\mathbf{y},\mathbf{y}) & \text{(Constant Self-Dissimilarity)} \\ \psi(\mathbf{x},\mathbf{x}) < \begin{cases} \psi(\mathbf{x},\mathbf{y}) \\ \psi(\mathbf{y},\mathbf{x}) \end{cases} & \text{(Regular Minimality)} \end{array} \tag{5}$$

We know from Chapter 1 that although the property of Regular Minimality is indeed satisfied in all available experimental data, the property of Constant Self-Dissimilarity is not. The latter can clearly be seen in the table below, a 10×10 excerpt from Rothkopf's (1957) well-known study of Morse code discriminations. In his experiment, a large number of respondents made same–different judgments in response to 36×36 auditorily presented pairs of Morse codes for letters of the alphabet and digits.[3]

Ro	B	0	1	2	3	4	5	6	7	8	9
B	16	88	83	86	60	68	26	57	83	96	96
0	95	16	37	87	92	90	92	81	68	43	45
1	86	38	11	46	80	95	86	80	79	84	89
2	92	82	36	14	69	77	59	84	83	92	90
3	81	95	74	56	11	58	56	68	90	97	97
4	55	86	90	70	31	10	58	76	90	94	95
5	20	85	86	74	76	83	14	31	86	95	86
6	67	78	71	82	85	88	39	15	30	80	87
7	77	58	71	84	84	91	40	40	11	39	74
8	86	43	61	91	88	96	89	58	44	9	22
9	97	50	74	91	89	95	78	83	48	19	6

Regular Minimality here is satisfied in the canonical form, and one can see, for example, that the Morse code for digit 6 was judged different from itself by 15% of respondents, but only by 6% for digit 9. Symmetry is clearly violated as well: thus, digits 4 and 5 were discriminated from each other in 83% of cases when 5 was presented first in the two-code sequence, but in only 58% when 5 was presented second. Nonconstant Self-similarity and

[3] This particular 10-code subset is chosen so that it forms a self-contained subspace of the 36 codes: a geodesic loop (as explained later) for any two of its elements is contained within the subset.

Asymmetry are also manifest in the 10 × 10 excerpt below from a similar study of Morse-code-like signals by Wish (1967).[4]

Wi	S	U	W	X	0	1	2	3	4	5
S	6	16	38	45	35	73	81	70	89	97
U	28	6	44	24	59	56	49	51	71	69
W	44	42	4	11	78	40	79	55	48	83
X	64	71	26	3	86	51	73	27	31	44
0	34	55	56	46	6	52	39	69	39	95
1	84	75	22	33	70	3	69	17	40	97
2	81	44	62	31	45	50	7	41	35	26
3	94	85	44	17	85	19	84	2	63	47
4	89	73	26	20	65	38	67	45	3	49
5	100	94	74	11	83	95	58	67	25	3

We can conclude, therefore, that MDS, or any other data-analytic technique based on the probability-distance hypothesis, is not supported by discrimination probability data. By contrast, Fechnerian Scaling, in the case of discrete stimulus sets, is only based on Regular Minimality, which is supported by data. Although prior to Dzhafarov (2002d), Regular Minimality has not been formulated as a basic property of discrimination, independent of its other properties (such as Constant Self-Dissimilarity), the violations of Symmetry and Constant Self-Dissimilarity have long since been noted. Tversky's (1977) contrast model and Krumhansl's (1978) distance-and-density scheme are two best known theoretical schemes dealing with these issues.

2. Multidimensional Fechnerian Scaling

MDFS (*Multidimensional Fechnerian Scaling*) is Fechnerian Scaling performed on a stimulus set whose physical description can be represented by an open connected region $\mathbf{E}$ of n-dimensional ($n \geq 1$) real-valued vectors, such that $\psi(\mathbf{x}, \mathbf{y})$ is continuous with respect to its Euclidean topology. This simply means that as $(\mathbf{x}_k, \mathbf{y}_k) \to (\mathbf{x}, \mathbf{y})$, in the conventional sense,

[4] 32 stimuli in this study were five-element sequences $T_1P_1T_2P_2T_3$, where T stands for a tone (short or long) and P stands for a pause (1 or 3 units long). We arbitrarily labeled the stimuli $A, B, ..., Z, 0, 1, ..., 5$, in the order they are presented in Wish's (1967) article. The criterion for choosing this particular subset of 10 stimuli is the same as for matrix *Ro*.

$\psi(\mathbf{x}_k, \mathbf{y}_k) \rightarrow \psi(\mathbf{x}, \mathbf{y})$. The theory of Fechnerian Scaling has been developed for continuous (arcwise connected) spaces of a much more general structure (Dzhafarov & Colonius, 2005a), but a brief overview of MDFS should suffice for understanding the main ideas underlying Fechnerian Scaling. Throughout the entire discussion, we tacitly assume that Regular Minimality is satisfied in a canonical form.

2.1. Oriented Fechnerian Distances in Continuous Spaces

Any $\mathbf{a}, \mathbf{b} \in \mathbf{E}$ can be connected by a smooth arc $\mathbf{x}(t)$, a piecewise continuously differentiable mapping of an interval $[\alpha, \beta]$ of reals into $\mathbf{E}$, such that $\mathbf{x}(\alpha) = \mathbf{a}$, $\mathbf{x}(\beta) = \mathbf{b}$ Refer to Fig. 1. The main intuitive idea underlying Fechnerian Scaling is that

(a) Any point $\mathbf{x}(t)$ on this arc, $t \in [\alpha, \beta]$, can be assigned a local measure of its difference from its "immediate neighbors," $\mathbf{x}(t + dt)$.

(b) By integrating this local difference measure along the arc, from α to β, one can obtain the "psychometric length" of this arc.

(c) By taking the infimum (the greatest lower bound) of psychometric lengths across all possible smooth arcs connecting $\mathbf{a}$ to $\mathbf{b}$, one obtains the distance from $\mathbf{a}$ to $\mathbf{b}$ in space $\mathbf{E}$.

As argued in Dzhafarov and Colonius (1999), this intuitive scheme can be viewed as the essence of Fechner's original theory for unidimensional stimulus continua (Fechner, 1860). The implementation of this idea in MDFS is as follows (see Fig. 2).

As t for a smooth arc $\mathbf{x}(t)$ moves from α to β, the value of self-discriminability $\psi(\mathbf{x}(t), \mathbf{x}(t))$ may vary (Nonconstant Self-Dissimilarity property). Therefore, to see how distinct $\mathbf{x}(t)$ is from $\mathbf{x}(t + dt)$ it would not suffice to look at $\psi(\mathbf{x}(t), \mathbf{x}(t + dt))$, or $\psi(\mathbf{x}(t + dt), \mathbf{x}(t))$; one should compute instead the *increments* in discriminability

$$\begin{aligned} \phi^{(1)}(\mathbf{x}(t), \mathbf{x}(t + dt)) &= \psi(\mathbf{x}(t), \mathbf{x}(t + dt)) - \psi(\mathbf{x}(t), \mathbf{x}(t)), \\ \phi^{(2)}(\mathbf{x}(t), \mathbf{x}(t + dt)) &= \psi(\mathbf{x}(t + dt), \mathbf{x}(t)) - \psi(\mathbf{x}(t), \mathbf{x}(t)). \end{aligned} \tag{6}$$

Both $\phi^{(1)}$ and $\phi^{(2)}$are positive due to the Regular Minimality property (in a canonical form). They are referred to as *psychometric differentials* of the *first kind* (or in the first observation area) and *second kind* (in the second observation area), respectively.

The assumptions of MDFS guarantee that the cumulation of $\phi^{(1)}(\mathbf{x}(t), \mathbf{x}(t + dt))$ (i.e., integration of $\phi^{(1)}(\mathbf{x}(t), \mathbf{x}(t + dt))/dt+$) from $t = \alpha$ to $t = \beta$ always yields a positive quantity.[5] We call this quan-

[5]Aside from Regular Minimality and continuity of $\psi(\mathbf{x}, \mathbf{y})$, the only other essential assumption of MDFS is that of the existence of a "global psychometric

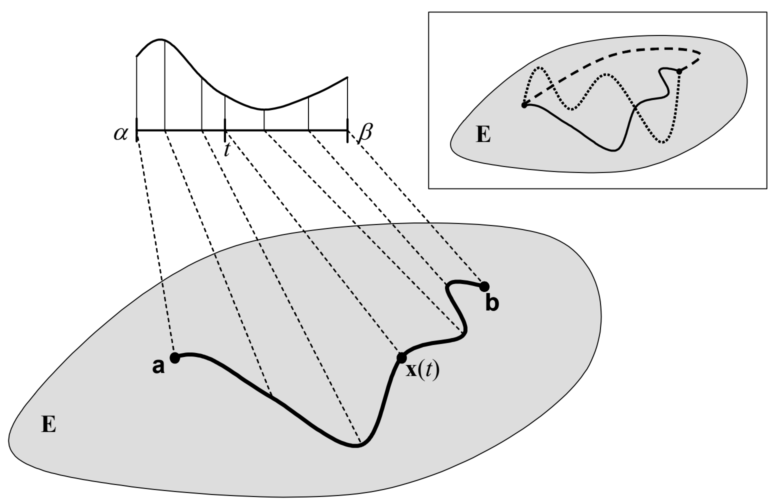


Fig. 1: The underlying idea of MDFS. $[\alpha, \beta]$ is a real interval, $\mathbf{a} \to \mathbf{x}(t) \to \mathbf{b}$ a smooth arc. The psychometric length of this arc is the integral of "local difference" of $\mathbf{x}(t)$ from $\mathbf{x}(t + dt)$, shown by vertical spikes along $[\alpha, \beta]$. The inset shows that one should compute the psychometric lengths for all possible smooth arcs leading from $\mathbf{a}$ to $\mathbf{b}$. Their infimum is the oriented Fechnerian distance from $\mathbf{a}$ to $\mathbf{b}$.

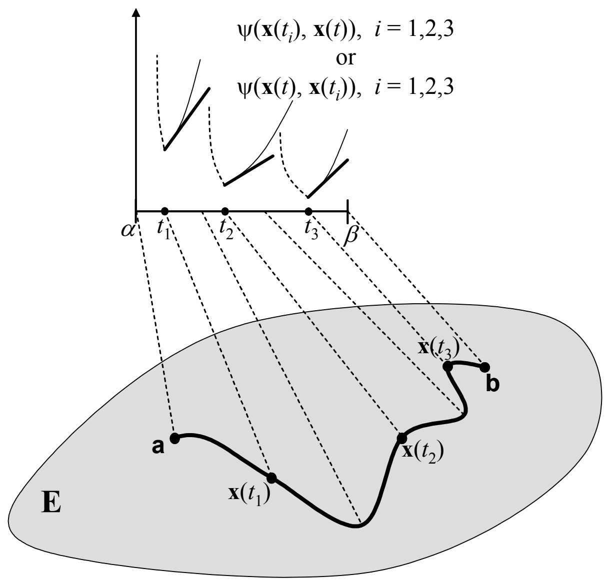


Fig. 2: The "local difference" of $\mathbf{x}(t)$ from $\mathbf{x}(t+dt)$ (as $dt \to 0+$) at a given point, $t = t_i$, is the slope of the tangent line drawn to $\psi(\mathbf{x}(t_i), \mathbf{x}(t))$, or to $\psi(\mathbf{x}(t), \mathbf{x}(t_i))$, at $t = t_i+$. Using $\psi(\mathbf{x}(t_i), \mathbf{x}(t))$ yields derivatives of the first kind, using $\psi(\mathbf{x}(t), \mathbf{x}(t_i))$ yields derivatives of the second kind. Their integration from α to β yields oriented Fechnerian distances (from $\mathbf{a}$ to $\mathbf{b}$) of, respectively, first and second kind.

tity the *psychometric length* of arc $\mathbf{x}(t)$ of the first kind, and denote it $L^{(1)}[\mathbf{a} \to \mathbf{x} \to \mathbf{b}]$, where we use the suggestive notation for arc $\mathbf{x}$ connecting $\mathbf{a}$ to $\mathbf{b}$: this notation is justified by the fact that the choice of the function $\mathbf{x}$:$[\alpha, \beta] \to \mathbf{E}$ is irrelevant insofar as the graph of the function (the curve connecting $\mathbf{a}$ to $\mathbf{b}$ in $\mathbf{E}$) remains invariant. It can further be shown that the infimum of all such psychometric lengths $L^{(1)}[\mathbf{a} \to \mathbf{x} \to \mathbf{b}]$, across all possible smooth arcs connecting $\mathbf{a}$ to $\mathbf{b}$, satisfies all properties of a distance except for symmetry. Denoting this infimum by $G_1(\mathbf{a}, \mathbf{b})$, we have (A) $G_1(\mathbf{a}, \mathbf{b}) \geq 0$; (B) $G_1(\mathbf{a}, \mathbf{b}) = 0$ if and only if $\mathbf{a} = \mathbf{b}$; (C) $G_1(\mathbf{a}, c) \leq G_1(\mathbf{a}, \mathbf{b}) + G_1(\mathbf{b}, c)$; but it is not necessarily true that $G_1(\mathbf{a}, \mathbf{b}) = G_1(\mathbf{b}, \mathbf{a})$. Such geometric constructs are called *oriented distances*. We call $G_1(\mathbf{a}, \mathbf{b})$ the *oriented Fechnerian distance of the first kind* from $\mathbf{a}$ to $\mathbf{b}$.

By repeating the whole construction with $\phi^{(2)}(\mathbf{x}(t), \mathbf{x}(t + dt))$ in place of $\phi^{(1)}(\mathbf{x}(t), \mathbf{x}(t + dt))$ we get the psychometric lengths $L^{(2)}[\mathbf{a} \to \mathbf{x} \to \mathbf{b}]$ *of the second kind* (for arcs $\mathbf{x}(t)$ connecting $\mathbf{a}$ to $\mathbf{b}$), and, as their infima, the oriented Fechnerian distances $G_2(\mathbf{a}, \mathbf{b})$ of the second kind (from $\mathbf{a}$ to $\mathbf{b}$).

2.2. Multidimensional Fechnerian Scaling and Multidimensional Scaling

The following observation provides additional justification for computing the oriented Fechnerian distances in the way just outlined.

A metric d (symmetrical or oriented) on some set $\mathbf{S}$ is called *intrinsic* if $d(\mathbf{a}, \mathbf{b})$ for any $\mathbf{a}, \mathbf{b} \in \mathbf{S}$ equals the infimum of the lengths of all "allowable" arcs connecting $\mathbf{a}$ and $\mathbf{b}$ (i.e., arcs with some specified properties). The oriented Fechnerian distances $G_1(\mathbf{a}, \mathbf{b})$ and $G_2(\mathbf{a}, \mathbf{b})$ are intrinsic in this sense, provided the allowable arcs are defined as smooth arcs. In reference to the classical MDS, all Minkowski metrics are (symmetrical) intrinsic metrics, in the same sense.

---

transformation" Φ which makes the limit ratios

$$\lim_{s \to 0+} \frac{\Phi\left[\phi^{(\iota)}(\mathbf{x}(t), \mathbf{x}(t+s))\right]}{s} \qquad (\iota = 1, 2)$$

nonvanishing, finite, and continuous in $(\mathbf{x}(t), \dot{\mathbf{x}}(t))$, for all arcs. (Actually, this is the "First Main Theorem of Fechnerian Scaling," a consequence of some simpler assumptions.) As it turns out (Dzhafarov, 2002d), together with Nonconstant Self-Dissimilarity, this implies that $\Phi(h)/h \to k > 0$ as $h \to 0+$. That is, Φ is a scaling transformation in the small and can therefore be omitted from formulations, on putting $k = 1$ with no loss of generality. The uniqueness of extending $\Phi(h) = h$ to arbitrary values of $h \in [0, 1]$ is analyzed in Dzhafarov and Colonius (2005b). In this chapter, $\Phi(h) = h$ is assumed tacitly.

Assume now that the discrimination probabilities $\psi(\mathbf{x}, \mathbf{y})$ on $\mathbf{E}$ (with the same meaning as in the previous subsection) can be obtained from some symmetrical intrinsic distance d on $\mathbf{E}$ by means of (4), with β being a continuous increasing function. It is sufficient to assume that (4) holds for small values of d only. Then, as proved in Dzhafarov (2002b),

$$d(\mathbf{a}, \mathbf{b}) = G_1(\mathbf{a}, \mathbf{b}) = G_2(\mathbf{a}, \mathbf{b})$$

for all $\mathbf{a}, \mathbf{b} \in \mathbf{E}$. In other words, $\psi(\mathbf{x}, \mathbf{y})$ cannot monotonically and continuously depend on any (symmetrical) intrinsic metric other than the Fechnerian one. The latter in this case is symmetrical, and its two kinds G_1 and G_2 coincide.[6]

The classical MDS, including its modification proposed in Shepard and Carroll (1966), falls within this category of models. In the context of continuous stimulus spaces, therefore, Fechnerian Scaling and MDS are not simply compatible, the former is in fact a necessary consequence of the latter (under the assumption of intrinsicality, and without confining the class of metrics d to Minkowski ones). Fechnerian computations, however, are applicable in a much broader class of cases, including those where the probability-distance hypothesis is false (as we know it generally to be).

It should be noted for completeness that some nonclassical versions of MDS are based on Tversky's (1977) or Krumhansl's (1978) schemes rather than on the probability-distance hypothesis, and they have the potential of handling nonconstant self-dissimilarity or asymmetry (e.g., DeSarbo, Johnson, Manrai, Manrai, & Edwards, 1992; Weeks & Bentler, 1982). We do not review these approaches here. Certain versions of MDS can be viewed as intermediate between the classical MDS and Fechnerian Scaling. Shepard and Carroll (1966) discussed MDS methods where only sufficiently small distances are monotonically related to pairwise dissimilarities. More recently, this idea was implemented in two algorithms where large distances are obtained by cumulating small distances within stimulus sets viewed as manifolds embedded in Euclidean spaces (Roweis & Saul, 2000; Tenenbaum,

---

[6] This account is somewhat simplistic: Because the probability-distance hypothesis implies Constant Self-Dissimilarity, the theorem proved in Dzhafarov (2002b) is compatible with Fechnerian distances computed with Φ other than identity function (see Footnote 5). We could avoid mentioning this by positing in the formulation of the probability-distance hypothesis that $\beta(h)$ in (4) has a nonzero finite derivative at $h = 0+$. With this assumption, psychometric increments, hence also Fechnerian distances, are unique up to multiplication by a positive constant. Equation $d \equiv G_1 \equiv G_2$, therefore, could more generally be written as $d \equiv kG_1 \equiv kG_2$ $(k > 0)$. Throughout this chapter, we ignore the trivial distinction between different multiples of Fechnerian metrics. (It should also be noted that in Dzhafarov, 2002b, intrinsic metrics are called internal, and a single distance G is used in place of G_1 and G_2.)

de Silva, & Langford, 2000). When applied to discrimination probabilities, these modifications of MDS cannot handle nonconstant self-dissimilarity, but the idea of cumulating small differences can be viewed as the essence of Fechnerian Scaling.

2.3. Overall Fechnerian Distances in Continuous Spaces

The asymmetry of the oriented Fechnerian distances creates a difficulty in interpretation. It is easy to understand that in general, $\psi(\mathbf{x},\mathbf{y}) \neq \psi(\mathbf{y},\mathbf{x})$: stimulus $\mathbf{x}$ in the two cases belongs to two different observation areas and can therefore be perceived differently (the same being true for y). In $G_1(\mathbf{a},\mathbf{b})$, however, $\mathbf{a}$ and $\mathbf{b}$ belong to the same (first) observation area, and the noncoincidence of $G_1(\mathbf{a},\mathbf{b})$ and $G_1(\mathbf{b},\mathbf{a})$ prevents one from interpreting either of them as a reasonable measure of perceptual dissimilarity between $\mathbf{a}$ and $\mathbf{b}$ (in the first observation area, "from the point of view" of a given perceiver). The same consideration applies, of course, to G_2. In MDFS, this difficulty is resolved by taking as a measure of perceptual dissimilarity the *overall Fechnerian distances* $G_1(\mathbf{a},\mathbf{b})+G_1(\mathbf{b},\mathbf{a})$ and $G_2(\mathbf{a},\mathbf{b})+G_2(\mathbf{b},\mathbf{a})$. What justifies this particular choice of symmetrization is the remarkable fact that

$$G_1(\mathbf{a},\mathbf{b}) + G_1(\mathbf{b},\mathbf{a}) = G_2(\mathbf{a},\mathbf{b}) + G_2(\mathbf{b},\mathbf{a}) = G(\mathbf{a},\mathbf{b}), \tag{7}$$

where the overall Fechnerian distance $G(\mathbf{a},\mathbf{b})$ (we need not now specify of which kind) can be easily checked to satisfy all properties of a metric (Dzhafarov, 2002d; Dzhafarov & Colonius, 2005a).

On a moment's reflection, (7) makes perfect sense. We wish to obtain a measure of perceptual dissimilarity between $\mathbf{a}$ and $\mathbf{b}$, and we use the procedure of pairwise presentations with same-different judgments to achieve this goal. The meaning of (7) is that in speaking of perceptual dissimilarities among stimuli, one can abstract away from this particular empirical procedure. Caution should be exercised, however: the observation-area-invariance of the overall Fechnerian distance is predicated on the *canonical form* of Regular Minimality. In a more general case, as explained in Section 3.6, $G_1(\mathbf{a},\mathbf{b}) + G_1(\mathbf{b},\mathbf{a})$ equals $G_2(\mathbf{a}',\mathbf{b}') + G_2(\mathbf{b}',\mathbf{a}')$ if $\mathbf{a}$ and $\mathbf{a}'$ (as well as $\mathbf{b}$ and $\mathbf{b}'$) are PSEs, not necessarily physically identical.

Equation (7) is an immediate consequence of the following proposition (Dzhafarov, 2002d; Dzhafarov & Colonius, 2005a): for any smooth arcs $\mathbf{a}\to\mathbf{x}\to\mathbf{b}$ and $\mathbf{b}\to\mathbf{y}\to\mathbf{a}$,

$$\begin{aligned} &L^{(1)}[\mathbf{a}\to\mathbf{x}\to\mathbf{b}] + L^{(1)}[\mathbf{b}\to\mathbf{y}\to\mathbf{a}] \\ &= L^{(2)}[\mathbf{a}\to\mathbf{y}\to\mathbf{b}] + L^{(1)}[\mathbf{b}\to\mathbf{x}\to\mathbf{a}]. \end{aligned} \tag{8}$$

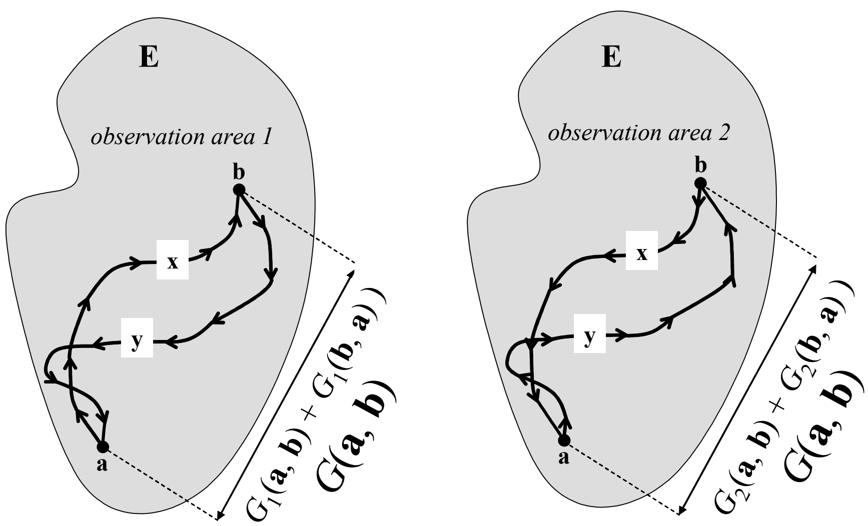


Fig. 3: Illustration for the Second Main Theorem: the psychometric length of the first kind of a closed loop from **a** to **b** and back equals the psychometric length of the second kind for the same loop traversed in the opposite direction. This leads to the equality of the overall Fechnerian distances in the two observation areas.

Put differently, the psychometric length of the first kind for any *closed loop* containing **a** and **b** equals the psychometric length of the second kind for the same closed loop but traversed in the opposite direction.

Together (8) and its corollary (7) constitute what we call the Second Main Theorem of Fechnerian Scaling (see Fig. 3). This theorem plays a critical role in extending the continuous theory to discrete and other, more complex object spaces (Dzhafarov & Colonius, 2005b).

3. FECHNERIAN SCALING OF DISCRETE OBJECT SETS (FSDOS)

The mathematical simplicity of this special case of Fechnerian Scaling allows us to present it in a greater detail than we did MDFS.

3.1. Discrete Object Spaces

Recall that a space of stimuli (or objects) is a set $\mathbf{S}$ of all objects of a particular kind endowed with a discrimination probability function $\psi(\mathbf{x}, \mathbf{y})$. For any $\mathbf{x}, \mathbf{y} \in \mathbf{S}$, we define *psychometric increments* of the first and second kind (or, in the first and second observation areas) as, respectively,

$$\begin{aligned} \phi^{(1)}(\mathbf{x}, \mathbf{y}) &= \psi(\mathbf{x}, \mathbf{y}) - \psi(\mathbf{x}, \mathbf{x}), \\ \phi^{(2)}(\mathbf{x}, \mathbf{y}) &= \psi(\mathbf{y}, \mathbf{x}) - \psi(\mathbf{x}, \mathbf{x}). \end{aligned} \tag{9}$$

Psychometric increments of both kinds are positive due to (a canonical form of) Regular Minimality, (3). A space $\mathbf{S}$ is called *discrete* if, for any $\mathbf{x} \in \mathbf{S}$,

$$\inf_{\mathbf{y}} \left[\phi^{(1)}(\mathbf{x}, \mathbf{y})\right] > 0, \qquad \inf_{\mathbf{y}} \left[\phi^{(2)}(\mathbf{x}, \mathbf{y})\right] > 0.$$

In other words, the psychometric increments of either kind from $\mathbf{x}$ to other stimuli cannot fall below some positive quantity. Intuitively, other stimuli cannot "get arbitrarily close" to $\mathbf{x}$. Clearly, stimuli in a discrete space cannot be connected by arcs (continuous images of intervals of reals).

3.2. Main Idea

To understand how Fechnerian computations can be made in discrete spaces, let us return for a moment to continuous spaces $\mathbf{E}$ discussed in the previous section.

Consider a smooth arc $\mathbf{x}(t)$,

$$\mathbf{x}\colon [\alpha, \beta] \to \mathbf{E}, \mathbf{x}(\alpha) = \mathbf{a}, \mathbf{x}(\beta) = \mathbf{b},$$

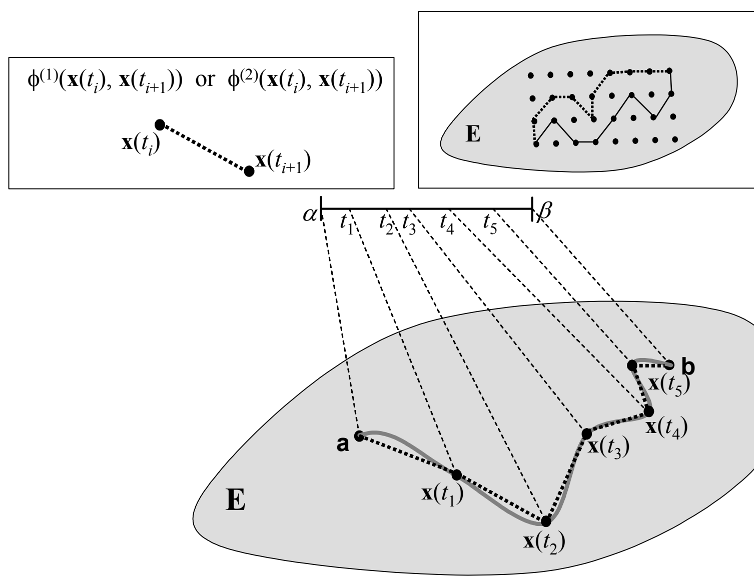


Fig. 4: The psychometric length of the first (second) kind of an arc can be approximated by the sum of psychometric increments of the first (second) kind chained along the arc. The right insert shows that if **E** is represented by a dense grid of points, the Fechnerian computations involve taking all possible chains leading from one point to another through successions of immediately neighboring points.

as shown in Fig. 4. We know that its psychometric length $L^{(\iota)}[\mathbf{a} \to \mathbf{x} \to \mathbf{b}]$ of the ιth kind ($\iota = 1, 2$) is obtained by cumulating psychometric differentials (6) of the same kind along this arc. It is also possible, however, to approximate $L^{(\iota)}[\mathbf{a} \to \mathbf{x} \to \mathbf{b}]$ by subdividing $[\alpha, \beta]$ into

$$\alpha = t_0, t_1, ..., t_k, t_{k+1} = \beta$$

and computing the sum of the chained psychometric increments

$$L^{(1)}[\mathbf{x}(t_0), \mathbf{x}(t_1), ..., \mathbf{x}(t_{k+1})] = \sum_{i=0}^{k} \phi^{(\iota)}(\mathbf{x}(t_i), \mathbf{x}(t_{i+1})) . \qquad (10)$$

As shown in Dzhafarov and Colonius (2005a), by progressively refining the partitioning, $\max_i \{t_{i+1} - t_i\} \to 0$, this sum can be made as close as one wishes to the value of $L^{(\iota)}[\mathbf{a} \to \mathbf{x} \to \mathbf{b}]$.

In practical computations, $\mathbf{E}$ (which, we recall, is an open connected region of n-dimensional vectors of reals) can be represented by a sufficiently dense discrete grid of points. In view of the result just mentioned, the oriented Fechnerian distance $G_\iota(\mathbf{a}, \mathbf{b})$ ($\iota = 1, 2$) between any $\mathbf{a}$ and $\mathbf{b}$ in this case can be approximated by (a) considering all possible chains of successive neighboring points leading from $\mathbf{a}$ to $\mathbf{b}$, (b) computing sums (10) for each of these chains, and (c) taking the smallest value.

This almost immediately leads to the algorithm for Fechnerian computations in discrete spaces. The main difference is that in discrete spaces, we have no physical ordering of stimuli to rely on, and the notion of "neighboring points" is not defined. In a sense, every point in a discrete space can be viewed as a "potential neighbor" of any other point. Consequently, in place of "all possible chains of successive neighboring points leading from $\mathbf{a}$ to $\mathbf{b}$," one has to consider simply *all possible chains of points leading from* $\mathbf{a}$ *to* $\mathbf{b}$ (see Fig. 5).

3.3. Illustration

Returning to our toy example (matrix TOY_0, reproduced here for the reader's convenience together with L_0 and G_0), let us compute the Fechnerian distance between, say, objects **D** and **B**.

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

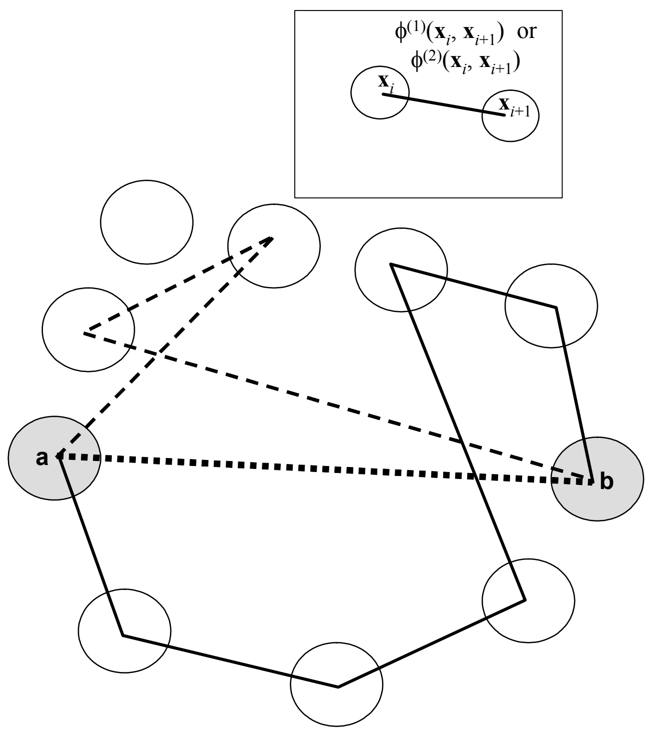


Fig. 5: In a discrete space (10 elements whereof are shown in an arbitrary spatial arrangement), Fechnerian computations are performed by taking sums of psychometric increments (of the first or second kind, as shown in the inset) for all possible chains leading from one point to another.

L_0	A	B	C	D
A	A	ACBA	ACA	ADA
B	BACB	B	BCB	BDCB
C	CAC	CBC	C	CDC
D	DAD	DCBD	DCD	D

G_0	A	B	C	D
A	0	1.3	1	1
B	1.3	0	0.9	1.1
C	1	0.9	0	0.7
D	1	1.1	0.7	0

The whole stimulus space here consists of four stimuli, $\{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}\}$, and we have five different chains in this space which are comprised of distinct (nonrecurring) objects and lead from $\mathbf{D}$ to $\mathbf{B}$:

$$\mathbf{DB}, \mathbf{DAB}, \mathbf{DCB}, \mathbf{DACB}, \mathbf{DCAB}.$$

We begin by computing their psychometric lengths of the first kind, $L^{(1)}[\mathbf{DB}]$, $L^{(1)}[\mathbf{DAB}]$, and so forth. By analogy with (10), $L^{(1)}[\mathbf{DCAB}]$, for example, is computed as

$$\begin{aligned} L^{(1)}[\mathbf{DCAB}] &= \phi^{(1)}(\mathbf{D}, \mathbf{C}) + \phi^{(1)}(\mathbf{C}, \mathbf{A}) + \phi^{(1)}(\mathbf{A}, \mathbf{B}) \\ &= [\psi(\mathbf{D}, \mathbf{C}) - \psi(\mathbf{D}, \mathbf{D})] + [\psi(\mathbf{C}, \mathbf{A}) - \psi(\mathbf{C}, \mathbf{C})] \\ &\quad + [\psi(\mathbf{A}, \mathbf{B}) - \psi(\mathbf{A}, \mathbf{A})] \\ &= [0.7 - 0.5] + [1.0 - 0.5] + [0.8 - 0.1] = 1.4. \end{aligned}$$

We have used here the definition of $\phi^{(1)}(\mathbf{x}, \mathbf{y})$ given in (9). Repeating this procedure for all our five chains, we will find out that the smallest value is provided by

$$\begin{aligned} L^{(1)}[\mathbf{DCB}] &= \phi^{(1)}(\mathbf{D}, \mathbf{C}) + \phi^{(1)}(\mathbf{C}, \mathbf{B}) \\ &= [\psi(\mathbf{D}, \mathbf{C}) - \psi(\mathbf{D}, \mathbf{D})] + [\psi(\mathbf{C}, \mathbf{B}) - \psi(\mathbf{C}, \mathbf{C})] \\ &= [0.7 - 0.5] + [0.6 - 0.5] = 0.3. \end{aligned}$$

Note that this value is smaller than the length of the one-link chain ("direct connection") $\mathbf{DB}$:

$$L^{(1)}[\mathbf{DB}] = \phi^{(1)}(\mathbf{D}, \mathbf{B}) = \psi(\mathbf{D}, \mathbf{B}) - \psi(\mathbf{D}, \mathbf{D}) = 1.0 - 0.5 = 0.5.$$

The chain $\mathbf{DCB}$ can be called a *geodesic chain* connecting $\mathbf{D}$ to $\mathbf{B}$. (Generally, there can be more than one geodesic chain, of the same length, for a given pair of stimuli, but in our toy example, all geodesics are unique.) Its length is taken to be the oriented Fechnerian distance of the first kind from $\mathbf{D}$ to $\mathbf{B}$,

$$G_1(\mathbf{D}, \mathbf{B}) = 0.3.$$

Consider now the same five chains but viewed in the opposite direction, that is, all chains in $\{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}\}$ leading from $\mathbf{B}$ to $\mathbf{D}$, and compute for these chains the psychometric lengths of the first kind: $L^{(1)}[\mathbf{BD}]$, $L^{(1)}[\mathbf{BAD}]$, and so forth. Having done this, we find out that this time, the shortest chain is the one-link chain $\mathbf{BD}$, with the length

$$L^{(1)}[\mathbf{BD}] = \phi^{(1)}(\mathbf{B}, \mathbf{D}) = \psi(\mathbf{B}, \mathbf{D}) - \psi(\mathbf{B}, \mathbf{B}) = 0.9 - 0.1 = 0.8.$$

The geodesic chain from $\mathbf{B}$ to $\mathbf{D}$ therefore is $\mathbf{BD}$, and the oriented Fechnerian distance of the first kind from $\mathbf{B}$ to $\mathbf{D}$ is

$$G_1(\mathbf{B}, \mathbf{D}) = 0.8.$$

Using the same logic as for continuous stimulus spaces, we now compute the (symmetrical) overall Fechnerian distance between $\mathbf{D}$ and $\mathbf{B}$ by adding the two oriented distances "to and fro,"

$$G(\mathbf{D}, \mathbf{B}) = G(\mathbf{B}, \mathbf{D}) = G_1(\mathbf{D}, \mathbf{B}) + G_1(\mathbf{B}, \mathbf{D}) = 0.3 + 0.8 = 1.1.$$

This is the value we find in cells $(\mathbf{D}, \mathbf{B})$ and $(\mathbf{B}, \mathbf{D})$ of matrix G_0. The concatenation of the two geodesic chains, $\mathbf{DCB}$ and $\mathbf{BD}$, forms the *geodesic loop* between $\mathbf{D}$ and $\mathbf{B}$, which we find in cells $(\mathbf{D}, \mathbf{B})$ and $(\mathbf{B}, \mathbf{D})$ of matrix L_0. This loop, of course, can be written in three different ways depending on which of its three distinct elements we choose to begin and end with. The convention adopted in matrix L_0 is to begin and end with the row object: $\mathbf{DCBD}$ in cell $(\mathbf{D}, \mathbf{B})$ and $\mathbf{BDCB}$ in cell $(\mathbf{B}, \mathbf{D})$.

Note that the overall Fechnerian distance $G(\mathbf{D}, \mathbf{B})$ and the corresponding geodesic loop could also be found by computing psychometric lengths for all 25 possible closed loops containing objects $\mathbf{D}$ and $\mathbf{B}$ in space $\{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}\}$ and finding the smallest. This, however, would be a more wasteful procedure.

The reason we do not need to add the qualification "of the first kind" to the designations of the overall Fechnerian distance $G(\mathbf{D}, \mathbf{B})$ and the geodesic loop $\mathbf{DCBD}$ is that precisely the same value of $G(\mathbf{D}, \mathbf{B})$ and the same geodesic loop (only traversed in the opposite direction) are obtained if the computations are performed with psychometric increments of the second kind.

For chain $\mathbf{DCAB}$, for example, the psychometric length of the second kind, using the definition of $\phi^{(2)}$ in (9), is computed as

$$\begin{aligned} L^{(2)}[\mathbf{DCAB}] &= \phi^{(2)}(\mathbf{D}, \mathbf{C}) + \phi^{(2)}(\mathbf{C}, \mathbf{A}) + \phi^{(2)}(\mathbf{A}, \mathbf{B}) \\ &= [\psi(\mathbf{C}, \mathbf{D}) - \psi(\mathbf{D}, \mathbf{D})] + [\psi(\mathbf{A}, \mathbf{C}) - \psi(\mathbf{C}, \mathbf{C})] \\ &\quad + [\psi(\mathbf{B}, \mathbf{A}) - \psi(\mathbf{A}, \mathbf{A})] \\ &= [1.0 - 0.5] + [0.6 - 0.5] + [0.8 - 0.1] = 1.3. \end{aligned}$$

Repeating this computation for all our five chains leading from $\mathbf{D}$ to $\mathbf{B}$, the shortest chain is found to be $\mathbf{DB}$, with the length

$$L^{(2)}[\mathbf{DB}] = \phi^{(2)}(\mathbf{D},\mathbf{B}) = \psi(\mathbf{B},\mathbf{D}) - \psi(\mathbf{D},\mathbf{D}) = 0.9 - 0.5 = 0.4,$$

taken to be the value of $G_2(\mathbf{D},\mathbf{B})$, the oriented Fechnerian distance form $\mathbf{D}$ to $\mathbf{B}$ of the second kind. For the same five chains but viewed as leading from $\mathbf{B}$ to $\mathbf{D}$, the shortest chain is $\mathbf{BCD}$, with the length

$$\begin{aligned} L^{(2)}[\mathbf{BCD}] &= \phi^{(2)}(\mathbf{B},\mathbf{C}) + \phi^{(2)}(\mathbf{C},\mathbf{D}) \\ &= [\psi(\mathbf{C},\mathbf{B}) - \psi(\mathbf{B},\mathbf{B})] + [\psi(\mathbf{D},\mathbf{C}) - \psi(\mathbf{C},\mathbf{C})] \\ &= [0.6 - 0.1] + [0.7 - 0.5] = 0.7 \end{aligned}$$

taken to be the value of $G_2(\mathbf{B},\mathbf{D})$, the oriented Fechnerian distance form $\mathbf{B}$ to $\mathbf{D}$ of the second kind. Their sum is

$$G(\mathbf{D},\mathbf{B}) = G(\mathbf{B},\mathbf{D}) = G_2(\mathbf{D},\mathbf{B}) + G_2(\mathbf{B},\mathbf{D}) = 0.4 + 0.7 = 1.1,$$

precisely the same value for the overall Fechnerian distance as before (although the oriented distances are different). The geodesic loop obtained by concatenating the geodesic chains $\mathbf{DB}$ and $\mathbf{BCD}$ is also the same as we find in matrix L_0 in cells $(\mathbf{D},\mathbf{B})$ and $(\mathbf{B},\mathbf{D})$, but read from right to left: $\mathbf{DBCD}$ in cell $(\mathbf{D},\mathbf{B})$ and $\mathbf{BCDB}$ in cell $(\mathbf{B},\mathbf{D})$.

The complete formulation of the convention adopted in L_0 therefore is as follows: the geodesic loop in cell $(\mathbf{x},\mathbf{y})$ begins and ends with $\mathbf{x}$ and is read from left to right for the computations of the first kind, and from right to left for the computations of the second kind (yielding one and the same result, the overall Fechnerian distance between $\mathbf{x}$ and $\mathbf{y}$).

3.4. Procedure of Fechnerian Scaling of Discrete Object Sets

It is clear that any finite set $S = \{\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_N\}$ endowed with probabilities $p_{ij} = \psi(\mathbf{s}_i, \mathbf{s}_j)$ forms a discrete space in the sense of our formal definition. As this case is of the greatest interest in empirical applications, in the following we confine our discussion to finite object spaces. All our statements, however, unless specifically qualified, apply to discrete object spaces of arbitrary cardinality.

The procedure shown later is described as if one knew the probabilities p_{ij} on the population level. If sample sizes do not warrant this approximation, the procedure should ideally be repeated with a large number of matrices p_{ij} that are statistically retainable given a matrix of frequency estimates $\hat{p}_{ij}$. We return to this issue in the concluding section.

The computation of Fechnerian distances G_{ij} among $\{\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_N\}$ proceeds in several steps. The first step in the computation is to check for Regular Minimality: for any i and all $j \neq i$,

$$p_{ii} < \min\{p_{ij}, p_{ji}\}.$$

If Regular Minimality is violated (on the population level), FSDOS will not work. Put differently, given a matrix of frequency estimates $\hat{\psi}(\mathbf{s}_i, \mathbf{s}_j)$, one should use statistically retainable matrices of probabilities p_{ij} that do satisfy Regular Minimality; and if no such matrices can be found, FSDOS is not applicable. The theory of Fechnerian Scaling treats Regular Minimality as the defining property of discrimination. If it is not satisfied, something can be wrong in the procedure: for collective perceivers, for example, substantially different groups of people could be responding to different pairs of stimuli (violating thereby the requirement of having a "single perceiver"), or the semantic meaning of the responses "same" and "different" could vary from one pair of stimuli to another. (Alternatively, of course, the theory of Fechnerian Scaling may be wrong itself, which would be a preferable conclusion if regular Minimality was found to be violated systematically, or at least not very rarely.)

Having Regular Minimality verified, we compute psychometric increments of the first and second kind,

$$\phi^{(1)}(\mathbf{s}_i, \mathbf{s}_j) = p_{ij} - p_{ii},$$
$$\phi^{(2)}(\mathbf{s}_i, \mathbf{s}_j) = p_{ji} - p_{ii},$$

which are positive for all $j \neq i$.

Consider now a chain of stimuli $\mathbf{s}_i = \mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k = \mathbf{s}_j$ leading from $\mathbf{s}_i$ to $\mathbf{s}_j$, with $k \geq 2$. The psychometric length of the first kind for this chain, $L^{(1)}[\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k]$, is defined as the sum of the psychometric increments $\phi^{(1)}(\mathbf{x}_m, \mathbf{x}_{m+1})$ taken along this chain,

$$L^{(1)}[\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k] = \sum_{m=1}^{k} \phi^{(1)}(\mathbf{x}_m, \mathbf{x}_{m+1}).$$

The set of different psychometric lengths across all possible chains of distinct elements connecting $\mathbf{s}_i$ to $\mathbf{s}_j$ being finite, it contains a minimum value $L^{(1)}_{\min}(\mathbf{s}_i, \mathbf{s}_j)$. (The consideration can always be confined to chains $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ of distinct elements, because if $\mathbf{x}_l = \mathbf{x}_m$ $(l < m)$, the length $L^{(1)}$ cannot increase if the subchain $(\mathbf{x}_{l+1}, ..., \mathbf{x}_m)$ is removed.) This value is called the oriented Fechnerian distance of the first kind from object $\mathbf{s}_i$ to object $\mathbf{s}_j$:

$$G_1(\mathbf{s}_i, \mathbf{s}_j) = L^{(1)}_{\min}(\mathbf{s}_i, \mathbf{s}_j).$$

It is easy to prove that the oriented Fechnerian distance satisfies all properties of a metric, except for symmetry: (A) $G_1(\mathbf{s}_i, \mathbf{s}_j) \geq 0$; (B) $G_1(\mathbf{s}_i, \mathbf{s}_j) = 0$ if and only if $i = j$; (C) $G_1(\mathbf{s}_i, \mathbf{s}_j) \leq G_1(\mathbf{s}_i, \mathbf{s}_m) + G_1(\mathbf{s}_m, \mathbf{s}_j)$; but in general, $G_1(\mathbf{s}_i, \mathbf{s}_j) \neq G_1(\mathbf{s}_j, \mathbf{s}_i)$.[7] In according with the general logic of Fechnerian Scaling, $G_1(\mathbf{s}_i, \mathbf{s}_j)$ is interpreted as the oriented Fechnerian distance from $\mathbf{s}_i$ to $\mathbf{s}_j$ in the first observation area.

Any chain from $\mathbf{s}_i$ to $\mathbf{s}_j$ whose elements are distinct and whose length equals $G_1(\mathbf{s}_i, \mathbf{s}_j)$ is a geodesic chain from $\mathbf{s}_i$ to $\mathbf{s}_j$. There may be more than one geodesic chain for given $\mathbf{s}_i, \mathbf{s}_j$. (Note that in the case of infinite discrete sets mentioned in footnote 7 geodesic chains need not exist.)

The oriented Fechnerian distances $G_2(\mathbf{s}_i, \mathbf{s}_j)$ of the second kind (in the second observation area) and the corresponding geodesic chains are computed analogously, using the chained sums of psychometric increments $\phi^{(2)}$ instead of $\phi^{(1)}$.

As argued earlier (Section 2.3), the order of two stimuli in a given observation area has no operational meaning, and we add the two oriented distances, "to and fro," to obtain the (symmetrical) overall Fechnerian distances

$$
\begin{aligned}
G_{ij} &= G_1(\mathbf{s}_i, \mathbf{s}_j) + G_1(\mathbf{s}_j, \mathbf{s}_i) = G_{ji}, \\
G_{ij} &= G_2(\mathbf{s}_i, \mathbf{s}_j) + G_2(\mathbf{s}_j, \mathbf{s}_i) = G_{ji}.
\end{aligned}
$$

G_{ij} clearly satisfies all the properties of a metric.

The validation for this procedure (and for writing G_{ij} without indicating observation area) is provided by the fact that

$$G_1(\mathbf{s}_i, \mathbf{s}_j) + G_1(\mathbf{s}_j, \mathbf{s}_i) = G_2(\mathbf{s}_i, \mathbf{s}_j) + G_2(\mathbf{s}_j, \mathbf{s}_i), \tag{11}$$

that is, the distance G_{ij} between the ith and the jth objects does not depend on the observation area in which these objects are taken. This fact is a consequence of the following statement, which is of interest on its own sake: for any two chains $\mathbf{s}_i = \mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k = \mathbf{s}_j$ and $\mathbf{s}_i = \mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_l = \mathbf{s}_j$

[7] Properties (A) and (B) trivially follow from the fact that for $i \neq j$, $G_1(\mathbf{s}_i, \mathbf{s}_j)$ is the smallest of several positive quantities, $L^{(1)}[\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k]$. Property (C) follows from the observation that the chains leading from $\mathbf{s}_i$ to $\mathbf{s}_j$ through a fixed $\mathbf{s}_k$ form a proper subset of all chains leading from $\mathbf{s}_i$ to $\mathbf{s}_j$. For an infinite discrete $\mathbf{S}$, $L^{(1)}_{\min}(\mathbf{a}, \mathbf{b})$ $(\mathbf{a}, \mathbf{b} \in \mathbf{S})$ need not exist and should be replaced with $L^{(1)}_{\inf}(\mathbf{a}, \mathbf{b})$, the infimum of $L^{(1)}[\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k]$ for all finite chains of distinct elements with $\mathbf{a} = \mathbf{x}_1$ and $\mathbf{x}_k = \mathbf{b}$ $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k \in \mathbf{S})$. The argument for properties (A) and (B) then should be modified: for $\mathbf{a} \neq \mathbf{b}$, $G_1(\mathbf{a}, \mathbf{b}) > 0$ because $L^{(1)}_{\inf}(\mathbf{a}, \mathbf{b}) \geq \inf_{\mathbf{x}} \left[\phi^{(1)}(\mathbf{a}, \mathbf{x})\right]$, and by definition of discrete object spaces, $\inf_{\mathbf{x}} \left[\phi^{(1)}(\mathbf{a}, \mathbf{x})\right] > 0$.

(connecting $\mathbf{s}_i$ to $\mathbf{s}_j$),

$$\begin{aligned} &L^{(1)}\left[\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k\right] + L^{(1)}\left[\mathbf{y}_l, \mathbf{y}_{l-1}, ..., \mathbf{y}_1\right] \\ &= L^{(2)}\left[\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_l\right] + L^{(2)}\left[\mathbf{x}_k, \mathbf{x}_{k-1}, ..., \mathbf{x}_1\right]. \end{aligned} \tag{12}$$

As the proof of this statement is elementary, it may be useful to present it here. Denoting $p'_{ij} = \psi\left(\mathbf{x}_i, \mathbf{x}_j\right)$ and $p''_{ij} = \psi\left(\mathbf{y}_i, \mathbf{y}_j\right)$,

$$\begin{aligned} &L^{(1)}\left[\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k\right] + L^{(1)}\left[\mathbf{y}_l, \mathbf{y}_{l-1}, ..., \mathbf{y}_1\right] \\ &= \sum_{m=1}^{k-1}\left(p'_{m,m+1} - p'_{ii}\right) + \sum_{m=1}^{l-1}\left(p''_{m+1,m} - p''_{m+1,m+1}\right), \\ &L^{(2)}\left[\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_l\right] + L^{(2)}\left[\mathbf{x}_k, \mathbf{x}_{k-1}, ..., \mathbf{x}_1\right] \\ &= \sum_{m=1}^{l-1}\left(p''_{m+1,m} - p''_{m,m}\right) + \sum_{m=1}^{k-1}\left(p'_{m,m+1} - p'_{m+1,m+1}\right). \end{aligned}$$

Subtracting the second equation from the first,

$$\begin{aligned} &\left(L^{(1)}\left[\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k\right] - L^{(2)}\left[\mathbf{x}_k, \mathbf{x}_{k-1}, ..., \mathbf{x}_1\right]\right) \\ &+ \left(L^{(1)}\left[\mathbf{y}_l, \mathbf{y}_{l-1}, ..., \mathbf{y}_1\right] - L^{(2)}\left[\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_l\right]\right) \\ &= \left(\sum_{m=1}^{k-1}\left(p'_{m,m+1} - p'_{ii}\right) - \sum_{m=1}^{k-1}\left(p'_{m,m+1} - p'_{m+1,m+1}\right)\right) \\ &+ \left(\sum_{m=1}^{l-1}\left(p''_{m+1,m} - p''_{m+1,m+1}\right) - \sum_{m=1}^{l-1}\left(p''_{m+1,m} - p''_{m,m}\right)\right) \\ &= \left(p'_{kk} - p'_{11}\right) + \left(p''_{11} - p''_{kk}\right). \end{aligned}$$

But $p'_{11} = p''_{11} = p_{ii}$ and $p'_{kk} = p''_{kk} = p_{jj}$, where, we recall, $p_{ij} = \psi\left(\mathbf{s}_i, \mathbf{s}_j\right)$. The difference therefore is zero, and (12) is proved. Equation (11) follows as a corollary, on observing

$$\begin{aligned} G_1\left(\mathbf{s}_i, \mathbf{s}_j\right) + G_1\left(\mathbf{s}_j, \mathbf{s}_i\right) =& \inf L^{(1)}\left[\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k\right] + \inf L^{(1)}\left[\mathbf{y}_l, \mathbf{y}_{l-1}, ..., \mathbf{y}_1\right] \\ =& \inf\left\{L^{(1)}\left[\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k\right] + L^{(1)}\left[\mathbf{y}_l, \mathbf{y}_{l-1}, ..., \mathbf{y}_1\right]\right\} \\ =& \inf\left\{L^{(2)}\left[\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_l\right] + L^{(2)}\left[\mathbf{x}_k, \mathbf{x}_{k-1}, ..., \mathbf{x}_1\right]\right\} \\ =& \inf L^{(2)}\left[\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_l\right] + \inf L^{(2)}\left[\mathbf{x}_k, \mathbf{x}_{k-1}, ..., \mathbf{x}_1\right] \\ =& G_2\left(\mathbf{s}_j, \mathbf{s}_i\right) + G_2\left(\mathbf{s}_i, \mathbf{s}_j\right). \end{aligned}$$

Together (11) and (12) provide a simple version of the Second Main Theorem of Fechnerian Scaling, mentioned earlier, when discussing MDFS.

An equivalent way of defining the overall Fechnerian distances G_{ij} is to consider all *closed loops* $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n, \mathbf{x}_1$ $(n \geq 2)$ containing two given stimuli $\mathbf{s}_i, \mathbf{s}_j$: G_{ij} is the shortest of the psychometric lengths computed for all such loops. Note that the psychometric length of a loop depends on the direction in which it is traversed: generally,

$$L^{(1)}(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n, \mathbf{x}_1) \neq L^{(1)}(\mathbf{x}_1, \mathbf{x}_n, ..., \mathbf{x}_2, \mathbf{x}_1),$$
$$L^{(2)}(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n, \mathbf{x}_1) \neq L^{(2)}(\mathbf{x}_1, \mathbf{x}_n, ..., \mathbf{x}_2, \mathbf{x}_1).$$

The result just demonstrated tells us, however, that

$$L^{(1)}(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n, \mathbf{x}_1) = L^{(2)}(\mathbf{x}_1, \mathbf{x}_n, ..., \mathbf{x}_2, \mathbf{x}_1),$$

that is, any closed loop in the first observation area has the same length as the same closed loop traversed in the opposite direction in the second observation area. In particular, if $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n, \mathbf{x}_1$ is a geodesic (i.e., shortest) loop containing the objects $\mathbf{s}_i, \mathbf{s}_j$ in the first observation area (obviously, the concatenation of the geodesic chains connecting $\mathbf{s}_i$ to $\mathbf{s}_j$ and $\mathbf{s}_j$ to $\mathbf{s}_i$), then the same loop is a geodesic loop in the second observation area, if traversed in the opposite direction, $\mathbf{x}_1, \mathbf{x}_n, ..., \mathbf{x}_2, \mathbf{x}_1$.

The computational procedure of FSDOS is summarized in the form of a detailed algorithm presented in the Appendix at the end of this chapter.

3.5. Two Examples

We used the procedure just described to compute Fechnerian distances and geodesic loops among 36 Morse codes with pairwise discrimination probabilities reported in Rothkopf (1957), and among 32 Morse-code-like signals data with discrimination probabilities reported in Wish (1967). For typographic reasons only, small subsets of these stimulus sets are shown in matrices *Ro* and Wi in Section 1.3, chosen because they form "self-contained" *subspaces*: any two elements of such a subset can be connected by a geodesic loop lying entirely within the subset. The Fechnerian distances and geodesic loops are presented here for these subsets only: for matrix *Ro*, they are

G_{Ro}	B	C	0	1	2	3	4	5	6	7	8	9
B	0	95	151	142	118	95	97	16	57	77	140	157
0	151	133	0	48	105	160	150	147	127	99	61	73
1	142	114	48	0	57	132	164	147	125	128	106	121
2	118	116	105	57	0	100	123	105	129	142	158	161
3	95	143	160	132	100	0	68	95	127	145	165	169
4	97	152	150	164	123	68	0	106	138	160	171	174
5	16	109	147	147	105	95	106	0	41	61	124	143
6	57	122	127	125	129	127	138	41	0	44	92	118
7	77	107	99	128	142	145	160	61	44	0	63	83
8	140	136	61	106	158	165	171	124	92	63	0	26
9	157	156	73	121	161	169	174	143	118	83	26	0

L_{Ro}	B	0	1	2	3	4	5	6	7	8	9
B	B	B0B	B1B	BX25B	B35B	B4B	B5B	B565B	B5675B	B567875B	B975B
0	0B0	0	010	01210	030	040	050	0670	070	080	090
1	1B1	101	1	121	131	141	151	161	171	1081	10901
2	25BX2	21012	212	2	232	242	252	2562	272	21082	292
3	35B3	303	313	323	3	343	35B3	363	3673	383	393
4	4B4	404	414	424	434	4	45B4	4564	474	484	494
5	5B5	505	515	525	5B35	5B45	5	565	5675	567875	5975
6	65B56	6706	616	6256	636	6456	656	6	676	6786	678986
7	75B567	707	717	727	7367	747	7567	767	7	787	7897
8	875B5678	808	8108	82108	838	848	875678	8678	878	8	898
9	975B9	909	90109	929	939	949	9759	986789	9789	989	9

and for matrix Wi they are[8]

[8] In the complete 32×32 matrix reported in Wish (1967); but outside the 10×10 submatrix Wi, there are two violations of Regular Minimality, both due to a single value, $\hat{p}_{TV} = 0.03$: this value is the same as $\hat{p}_{VV}$ and smaller than $\hat{p}_{TT} = 0.06$ (using the labeling of stimuli described in Section 1.3); see also Footnote 4. As Wish's data are used here for illustration purposes only, we simply replaced $\hat{p}_{TV} = 0.03$ with $p_{TV} = 0.07$, putting $p_{ij} = \hat{p}_{ij}$ for the rest of the data. Chi-square deviation of thus defined matrix of p_{ij} from the matrix of $\hat{p}_{ij}$ is negligibly small.

G_{Wi}	S	U	W	X	0	1	2	3	4	5
S	0	32	72	89	57	119	112	128	119	138
U	32	0	76	79	89	107	80	116	107	128
W	72	76	0	30	119	55	122	67	58	79
X	89	79	30	0	123	67	94	39	45	49
0	57	89	119	123	0	113	71	143	95	132
1	119	107	55	67	113	0	109	31	72	108
2	112	80	122	94	71	109	0	116	92	74
3	128	116	67	39	143	31	116	0	84	77
4	119	107	58	45	95	72	92	84	0	68
5	138	128	79	49	132	108	74	77	68	0

L_{Wi}	S	U	W	X	0	1	2	3	4	5
S	S	SUS	SWS	SUXS	S0S	SU1WS	SU2US	SUX3XS	SUX4WS	SUX5XS
U	USU	U	UWU	UXWU	US0SU	U1WU	U2U	UX31WU	UX4WU	UX5XWU
W	WSW	WUW	W	WXW	WS0W	W1W	W2XW	WX31W	WX4W	WX5XW
X	XSUX	XWUX	XWX	X	X0X	X31WX	X2X	X3X	X4X	X5X
0	0S0	0SUS0	0WS0	0X0	0	010	020	0130	040	0250
1	1WSU1	1WU1	1W1	1WX31	101	1	121	131	141	135X31
2	2USU2	2U2	2XW2	2X2	202	212	2	232	242	252
3	3XSUX3	31WUX3	31WX3	3X3	3013	313	323	3	3X4X3	35X3
4	4WSUX4	4WUX4	4WX4	4X4	404	414	424	4X3X4	4	454
5	5XSUX5	5XWUX5	5XWX5	5X5	5025	5X3135	525	5X35	545	5

Recall our convention on presenting geodesic loops. Thus, in matrix L_{Ro}, the geodesic chain from letter B to digit 8 in the first observation area is $B \rightarrow 5 \rightarrow 6 \rightarrow 7 \rightarrow 8$ and that from 8 to B is $8 \rightarrow 7 \rightarrow 5 \rightarrow B$. In the second observation area, the geodesic chains should be read from right to left: $8 \leftarrow 7 \leftarrow 5 \leftarrow B$ from B to 8, and $B \leftarrow 5 \leftarrow 6 \leftarrow 7 \leftarrow 8$ from 8 to B. The oriented Fechnerian distances (lengths of the geodesic chains) are

A more comprehensive procedure should have involved a repeated generation of statistically retainable p_{ij} matrices subject to Regular Minimality, as discussed in the concluding section.

$G_1(B,8) = .70$, $G_1(8,B) = .70$, $G_2(B,8) = .77$, and $G_2(8,B) = .63$. The lengths of the closed loops in both observation areas add up to the same value, $G(8,B) = 1.40$, as they should.

Note that Fechnerian distances G_{ij} are not monotonically related to discrimination probabilities p_{ij}: there is no functional relation between the two because the computation of G_{ij} for any *given* (i,j) involves p_{ij} values for *all* (i,j). And, the oriented Fechnerian distances $G_1(s_i,s_j)$ and $G_2(s_i,s_j)$ are not monotonically related to psychometric increments $p_{ij} - p_{ii}$ and $p_{ji} - p_{ii}$, due to the existence of longer-than-one-link geodesic chains. There is, however, a strong positive correlation between p_{ij} and G_{ij}: 0.94 for Rothkopf's data and 0.89 for Wish's data (the Pearson correlation for the entire matrices, 36×36 and 32×32). This indicates that the probability-distance hypothesis, even if known to be false mathematically, may still be acceptable as a crude approximation. We may expect consequently that MDS-distances could provide crude approximations to the Fechnerian distances. That the adjective "crude" cannot be dispensed with is indicated by the relatively low values of Kendall's correlation between p_{ij} and G_{ij}: 0.76 for Rothkopf's data and 0.68 for Wish's data.

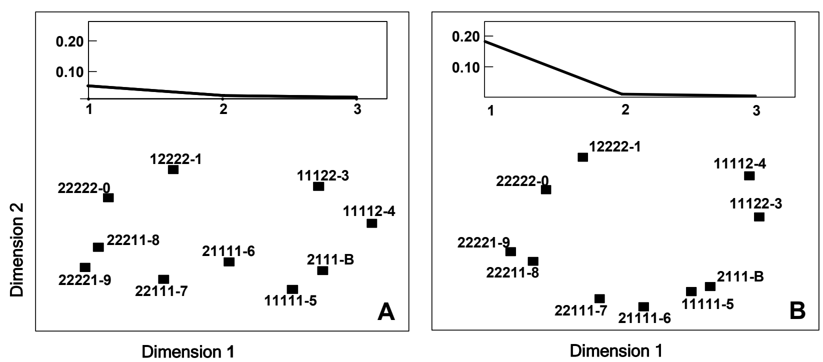


Fig. 6: Two-dimensional Euclidean representations for discrimination probabilities (nonmetric MDS, Panel A) and for Fechnerian distances in matrix G_{Ro} (metric MDS, Panel B). The MDS program used is PROXSCAL 1.0 in SPSS 11.5, minimizing raw stress. Sequence of "1"s and "2"s preceding a dash is the Morse code for the symbol following the dash. Insets are scree plots (normalized raw stress versus number of dimensions).

MDS can be used in conjunction with FSDOS, as a follow-up analysis once Fechnerian distances have been computed. A nonmetric version of MDS can be applied to Fechnerian distances (as opposed to discrimination

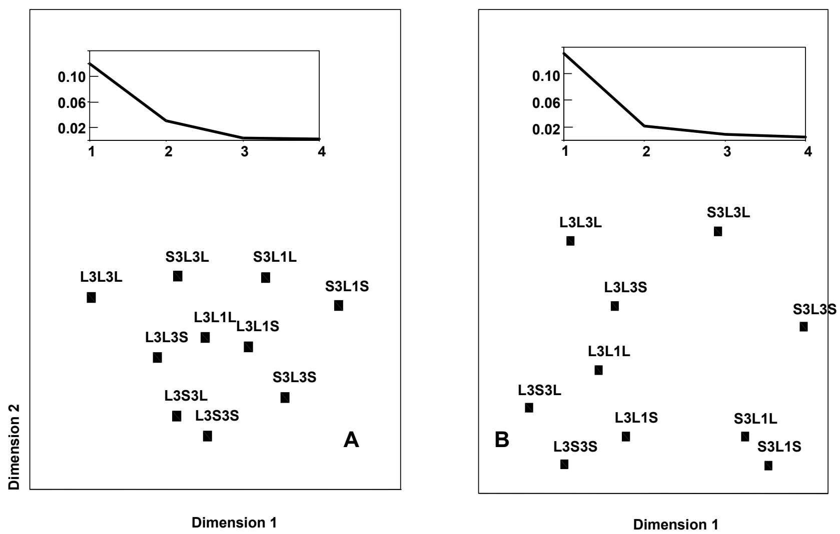


Fig. 7: Same as Fig. 6, but for discrimination probabilities (nonmetric MDS, Panel A) and for Fechnerian distances in matrix G_{Wi} (metric MDS, Panel B). L stands for long tone, S for short tone, whereas digits 1 an 3 show the lengths of the two pauses.

probabilities directly) simply to provide a rough graphical representation for matrices like Ro and Wi. More interestingly, a *metric* version of MDS can be applied to Fechnerian distances to test the hypothesis that Fechnerian distances, not restricted a priori to any particular class (except for being intrinsic), de facto belong to a class of Euclidean metrics (or, more generally, Minkowski ones), at least approximately; the degree of approximation for any given dimensionality is measured by the achieved stress value. Geometrically, metric MDS on Fechnerian distances is an attempt to *isometrically embed* the discrete object space into a low-dimensional Euclidean (or Minkowskian) space. Isometric embedment (or *immersion*) means mapping without distorting pairwise distances. Figures 6 and 7 provide a comparison of the metric MDS on Fechnerian distances (matrices Ro, Wi) with nonmetric MDS performed on discrimination probabilities directly (matrices G_{Ro}, G_{Wi}). Using the value of normalized raw stress as our criterion, the two-dimensional solution is almost equally good in both analyses. Therefore, to the extent that we consider the traditional MDS solution acceptable, we can view the Fechnerian distances in these two cases as being approximately Euclidean. The configurations of points obtained by performing the metric MDS on Fechnerian distances and nonmetric MDS on discrimination probabilities are more similar in Fig. 6 than in Fig. 7, indicating that MDS-distances provide a better approximation to Fechnerian distances in the former case. This may reflect the fact that the correlation between the probabilities and Fechnerian distances for Rothkopf's data is higher than for Wish's data (0.94 vs. 0.89). A detailed comparison of the configurations provided by the two analyses, as well as such related issues as interpretation of axes, are, however, beyond the scope of this chapter.

3.6. General Form of Regular Minimality

In continuous stimulus spaces, it often happens that Regular Minimality does not hold in a canonical form: for a fixed value of $\mathbf{x}$, $\psi(\mathbf{x}, \mathbf{y})$ achieves its minimum not at $\mathbf{y} = \mathbf{x}$ but at some other value of $\mathbf{y}$. It has been noticed since Fechner (1860), for example, that when one and the same stimulus is presented twice in a succession, the second presentation often seems larger (bigger, brighter, etc.) than the first: this is the classical phenomenon of "time error." It follows that in a successive pair of unidimensional stimuli, (x, y), the two elements maximally resemble each other when y is physically smaller than x. Other examples were discussed in Chapter 1. Although it is possible that in discrete stimulus spaces Regular Minimality always holds in a canonical form, it need not be so a priori.

Returning once again to our toy example, assume that matrix TOY_0 was the result of a canonical relabeling of matrix 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

with the correspondence table

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

where $\mathcal{O}_1$ and $\mathcal{O}_2$, as usual, denote the two observation areas (row stimuli and column stimuli). Having performed the Fechnerian analysis on TOY_0 and having computed the matrices L_0 and G_0, it makes sense now to return to the original labeling (using the table of correspondences above) and present the Fechnerian distances and geodesic loops separately for the first and the second observation areas:

L_{11}	a	b	c	d
a	a	acba	aca	ada
b	bacb	b	bcb	bdcb
c	cac	cbc	c	cdc
d	dad	dcbd	dcd	d

G_{11}	a	b	c	d
a	0	1.3	1	1
b	1.3	0	0.9	1.1
c	1	0.9	0	0.7
d	1	1.1	0.7	0

L_{12}	c	d	b	a
c	c	cbdc	cbc	cac
d	dcbd	d	dbd	dabd
b	bcb	bdb	b	bab
a	aca	abda	aba	a

G_{12}	c	d	b	a
c	0	1.3	1	1
d	1.3	0	0.9	1.1
b	1	0.9	0	0.7
a	1	1.1	0.7	0

Denoting, as indicated in Section 1.2, the overall Fechnerian distances in the first and second observation areas by $G^{(1)}(\mathbf{a},\mathbf{b})$ and $G^{(2)}(\mathbf{a},\mathbf{b})$, respectively, not to be confused with the oriented Fechnerian distances $G_1(\mathbf{a},\mathbf{b})$ and $G_2(\mathbf{a},\mathbf{b})$,

$$G^{(1)}(\mathbf{a},\mathbf{b}) = G_1(\mathbf{a},\mathbf{b}) + G_1(\mathbf{b},\mathbf{a}) = G^{(1)}(\mathbf{b},\mathbf{a}),$$
$$G^{(2)}(\mathbf{a},\mathbf{b}) = G_2(\mathbf{a},\mathbf{b}) + G_2(\mathbf{b},\mathbf{a}) = G^{(2)}(\mathbf{b},\mathbf{a}).$$

We see, for instance, that $G^{(1)}(\mathbf{a},\mathbf{b})$ is 1.3, whereas $G^{(2)}(\mathbf{a},\mathbf{b})$ is 0.7, reflecting the fact that $\mathbf{a},\mathbf{b}$ are perceived differently in the two observation areas. On the other hand, $G^{(2)}(\mathbf{c},\mathbf{d})$ is 1.3., the same as $G^{(1)}(\mathbf{a},\mathbf{b})$. This reflects the fact that $\mathbf{c}$ and $\mathbf{d}$ in $\mathcal{O}_2$ are the PSEs for, respectively, $\mathbf{a}$ and $\mathbf{b}$ in $\mathcal{O}_1$. Moreover, the geodesic loop containing $\mathbf{c},\mathbf{d}$ (in $\mathcal{O}_2$) is obtained from the geodesic loop containing $\mathbf{a},\mathbf{b}$ (in $\mathcal{O}_1$) by replacing every element of the latter loop by its PSE.

4. CONCLUDING REMARKS ON FECHNERIAN SCALING OF DISCRETE OBJECT SETS

We confine these concluding remarks to FSDOS only because this is the case of Fechnerian Scaling we presented in a relatively comprehensive way. With some technical caveats and modifications, the discussion to follow also applies to MDFS and the more general theory of continuous and "discrete-continuous" stimulus spaces presented in Dzhafarov and Colonius (2005a, 2005b).

4.1. Statistical Issues

In some applications, the number of replications from which frequency estimates of $p_{ij} = \psi(\mathbf{s}_i, \mathbf{s}_j)$ are obtained can be made sufficiently large to ignore statistical issues and treat FSDOS as being performed on essentially a population level. To a large extent, this is how the theory of FSDOS is presented in this chapter. The questions of finding the joint sampling distribution for Fechnerian distances $\hat{G}_{ij}$ $(i, j = 1, 2, ..., N)$ or joint confidence intervals for G_{ij} are beyond the scope of this chapter. We can, however, outline a general approach.

The estimators $\hat{P}_{ij}$ of the probabilities p_{ij} are obtained as

$$\hat{P}_{ij} = \frac{1}{R_{ij}} \sum_{k=1}^{R_{ij}} X_{ijk},$$

where $\{X_{ij1}, ..., X_{ijR_{ij}}\}$ are random variables representing binary responses ($1 = different$, $0 = same$). The index k may represent chronological trial numbers for $(\mathbf{s}_i, \mathbf{s}_j)$, different examples of this pair, different respondents, or some combination thereof. Random variables X_{ijk} and $X_{i'j'k'}$ can be treated as stochastically independent, provided $(i, j, k) \neq (i', j', k')$. Strictly speaking, X_{ijk} and $X_{i'j'k'}$ are *unrelated* random variables, they do not have a joint distribution (i.e., there is no pairing scheme for potential realizations of these two variables). Unrelated random variables, however (with no pairing scheme), can always be treated as independent (all-to-all pairing).[9]

Assuming that $\Pr[X_{ijk} = 1]$ does not vary too much as a function of k (i.e., ignoring such factors as fatigue, learning, and individual differences), $\hat{P}_{ij}$ may be viewed as independent normally distributed variables with means p_{ij} and variances $p_{ij}(1 - p_{ij})/R_{ij}$, from which it would follow that the joint distribution of the psychometric lengths of all chains with distinct elements is asymptotically multivariate normal, with both the means and covariances being known functions of true probabilities p_{ij}. The problem then is reduced to finding the (asymptotic) joint sampling distribution of the minima of psychometric lengths with common terminal points. Realistically, the problem is more likely to be dealt with by means of Monte Carlo simulations.

Monte Carlo is also likely to be used for constructing joint confidence intervals for G_{ij}, given a matrix of $\hat{p}_{ij}$. The procedure consists of repeatedly replacing the latter with matrices of p_{ij} that are subject to Regular Minimality and deviate from $\hat{p}_{ij}$ less than some critical value (e.g., by the conventional chi-square criterion), and computing Fechnerian distances from each of these matrices.

4.2. Choice of Object Set

In some cases, as with Rothkopf's (1957) Morse codes, the set S of stimuli used in an experiment or computation may contain all objects of a given kind. If such a set is too large or infinite, however, one can only use a subset S' of the entire S. This gives rise to a problem: for any two

[9] In psychometric applications, it is customary to treat random variables obtained from one and the same group of observers responding to different treatments as being paired by the observer,that is, having a joint distribution and being potentially interdependent. This is not a mathematical necessity, however, but merely an indication of what one is interested in. Let $R_{ij} = R_{i'j'} = R$, and let K be the random variable attaining values $(1, ..., R)$ with (say) equal probabilities. The question of traditional interest then can be formulated as that of finding $\Pr[X_{ijK} = 1 \text{ and } X_{i'j'K} = 1]$ (the probability that responses randomly chosen from the two cells are 1 given that they are by one and the same observer), which need not decompose as $\Pr[X_{ijK} = 1]\Pr[X_{i'j'K} = 1]$ although X_{ijk} and $X_{i'j'k}$ are independent for every k. In this context, however, the relevant question is different: what is $\Pr[X_{ijK} = 1 \text{ and } X_{i'j'K'} = 1]$ (the probability that responses randomly chosen from the two cells are 1)? Here, K and K' are independent random variables attaining values $(1, ..., R_{ij})$ and $(1, ..., R_{i'j'})$, respectively: in this case, R_{ij} and $R_{i'j'}$ need not be the same, and all computations are invariant with respect to all possible permutations of the third index in all sets $\{X_{ij1}, ..., X_{ijR_{ij}}\}$.

stimuli $\mathbf{a},\mathbf{b}\in S'$, the Fechnerian distance $G(\mathbf{a},\mathbf{b})$ will generally depend on what other stimuli are included in S'. Thus, adding a new object s_{N+1} to a subset $\{\mathbf{s}_1,\mathbf{s}_2,...,\mathbf{s}_N\}$ may change the pairwise discrimination probabilities $\psi(\mathbf{s}_i,\mathbf{s}_j)$ within the old subset $(i,j=1,2,...,N)$. This generally happens in a psychophysical experiment, when pairs of stimuli are presented repeatedly to a single observer. In a group experiment with each pair presented just once, or for the "paper-and-pencil" perceivers (as in our example with statistical models), adding $\mathbf{s}_{N+1}$ may not change $\psi(\mathbf{s}_i,\mathbf{s}_j)$ within $\{\mathbf{s}_1,\mathbf{s}_2,...,\mathbf{s}_N\}$, but it will still add new chains with which to connect any given stimuli $\mathbf{s}_i,\mathbf{s}_j$ $(i,j=1,2,...,N)$; as a result, the minimum psychometric lengths $L^{(\iota)}_{\min}(\mathbf{s}_i,\mathbf{s}_j)$ and $L^{(\iota)}_{\min}(\mathbf{s}_j,\mathbf{s}_i)$ $(\iota=1,2)$ will generally decrease.[10]

A formal approach to this issue is to simply state that the Fechnerian distance between two given stimuli is a relative concept: $G(\mathbf{a},\mathbf{b})$ shows how far apart the two stimuli are "from the point of view" of a given perceiver and *with respect to a given object set*. This approach may be sufficient in a variety of applications, especially in psychophysical experiments with repeated presentations of pairs to a single observer: one might hypothesize that the observer in such a situation gets adapted to the immediate context of the stimuli in play, effectively confining to it the subjective "universe of possibilities." A discussion of this "adaptation to subspace" hypothesis can be found in Dzhafarov and Colonius (2005a).

Like in many other situations involving sampling, however (including, for example, sampling of respondents in a group experiment), one may only be interested in a particular subset $\mathbf{S}'$ of stimuli to the extent that it is representative of the entire set $\mathbf{S}$ of stimuli of a particular kind. In this case, one faces two distinctly different questions. The first question is empirical: is $\mathbf{S}'$ large enough (well chosen enough) for its further enlargements not to lead to noticeable changes in discrimination probabilities within $\mathbf{S}'$? This question is not FSDOS-specific, any other analysis of discrimination probabilities (e.g., MDS) will have to address it, too. The second question is computational, and it is FSDOS-specific: provided the first question is answered in the affirmative, is $\mathbf{S}'$ large (well chosen) enough for its further enlargements not to lead to noticeable changes in Fechnerian distances within $\mathbf{S}'$? A detailed discussion being outside the scope of this chapter, we can only mention what seems to be an obvious approach: the affirmative answer to the second question can be given if one can show, by means of an

[10] This decrease must not be interpreted as a decrease in subjective dissimilarity. Fechnerian distances are determined up to multiplication by an arbitrary positive constant, which means that only *relative* Fechnerian distances $G(\mathbf{a},\mathbf{b})/G(\mathbf{c},\mathbf{d})$ are meaningfully interpretable. Adding a new object to a subset may very well increase $G(\mathbf{a},\mathbf{b})$ with respect to some or even all other distances.

appropriate version of subsampling, that the exclusion of a few stimuli from $\mathbf{S}'$ does not lead to changes in Fechnerian distances within the remaining subset.

4.3. Other Empirical Procedures

The procedure of pairwise presentations with same–different judgments is the focal empirical paradigm for FSDOS. With some caution, however, FSDOS can also be applied to other empirical paradigms, such as the *identification* paradigm: all stimuli $\{\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_N\}$ are associated with rigidly fixed, normative reactions $\{R_1, R_2, ..., R_N\}$ (e.g., fixed names, if the perceiving system is a person or group of people), and the stimuli are presented one at a time. Such an experiment results in (estimates of) the stimulus-response confusion probabilities $\eta\,(R_j|\mathbf{s}_i)$ with which reaction R_j (normatively reserved for $\mathbf{s}_j$) is given to object $\mathbf{s}_i$. FSDOS here can be applied under the additional assumption that $\eta\,(R_j|\mathbf{s}_i)$ can be interpreted as $1-\psi\,(\mathbf{s}_i, \mathbf{s}_j)$. The Regular Minimality property here means that each object $\mathbf{s}_i$ has a single modal reaction R_j (in the canonical form, R_i), and then any other object evokes R_j less frequently than does $\mathbf{s}_i$. Thus understood, Regular Minimality is satisfied, for example, in the data reported in Shepard (1957, 1958). We reproduce here one of the matrices from this work (matrix Sh, rows are stimuli, columns normative responses, entries conditional probabilities of responses given stimuli), together with the matrix of Fechnerian distances (G_{Sh}). Geodesic loops are not shown because the space $\{\mathbf{A}, \mathbf{B}, ..., \mathbf{I}\}$ here turns out to be a "Fechnerian simplex": a geodesic chain from $\mathbf{a}$ to $\mathbf{b}$ in this space is always the one-link chain $\mathbf{a} \rightarrow \mathbf{b}$.[11]

Sh	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**
A	0.678	0.148	0.054	0.03	0.025	0.02	0.016	0.011	0.016
B	0.167	0.544	0.066	0.077	0.053	0.015	0.045	0.018	0.015
C	0.06	0.07	0.615	0.015	0.107	0.067	0.022	0.03	0.014
D	0.015	0.104	0.016	0.542	0.057	0.005	0.163	0.032	0.065
E	0.037	0.068	0.12	0.057	0.46	0.075	0.057	0.099	0.03
F	0.027	0.029	0.053	0.015	0.036	0.715	0.015	0.095	0.014
G	0.011	0.033	0.015	0.145	0.049	0.016	0.533	0.052	0.145
H	0.016	0.027	0.031	0.046	0.069	0.096	0.053	0.628	0.034
I	0.005	0.016	0.011	0.068	0.02	0.021	0.061	0.018	0.78

[11] For the identification paradigm the construction of sampling distributions and confidence intervals mentioned in Section 4.1 should be modified, as the probability estimators within rows are no longer stochastically independent: $\sum_{j=1}^{N} \eta\,(R_j|\mathbf{s}_i) = 1$.

G_{Sh}	A	B	C	D	E	F	G	H	I
A	0	0.907	1.179	1.175	1.076	1.346	1.184	1.279	1.437
B	0.907	0	1.023	0.905	0.883	1.215	0.999	1.127	1.293
C	1.179	1.023	0	1.126	0.848	1.21	1.111	1.182	1.37
D	1.175	0.905	1.126	0	0.888	1.237	0.767	1.092	1.189
E	1.076	0.883	0.848	0.888	0	1.064	0.887	0.92	1.19
F	1.346	1.215	1.21	1.237	1.064	0	1.217	1.152	1.46
G	1.184	0.999	1.111	0.767	0.887	1.217	0	1.056	1.107
H	1.279	1.127	1.182	1.092	0.92	1.152	1.056	0	1.356
I	1.437	1.293	1.37	1.189	1.19	1.46	1.107	1.356	0

In a variant of the identification procedure, the reactions may be *preference ranks* for stimuli $\{\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_N\}$, R_1 designating, say, the most preferred object, R_N the least preferred. Suppose that Regular Minimality holds in the following sense: each object has a modal (most frequent) rank, each rank has a modal object, and R_j is the modal rank for $\mathbf{s}_i$ if and only if $\mathbf{s}_i$ is the modal object for R_j. Then the frequency rank R_j that is assigned to stimulus $\mathbf{s}_i$ can be taken as an estimate of $1 - \psi(\mathbf{s}_i, \mathbf{s}_j)$, and the data be subjected to FSDOS. The fact that these and similar procedures are used in a variety of areas (psychophysics, neurophysiology, consumer research, educational testing, political science), combined with the great simplicity of the algorithm for FSDOS, makes one hope that its potential application sphere may be very large.

4.4. Transformation of Discrimination Probabilities

This is probably the most difficult of the open problems remaining in Fechnerian Scaling. If $\psi(\mathbf{x}, \mathbf{y})$ satisfies Regular Minimality, then so does

$$\phi(\mathbf{x}, \mathbf{y}) = \varphi[\psi(\mathbf{x}, \mathbf{y})],$$

for any strictly increasing transformation φ. Regular Minimality is the only prerequisite for FSDOS, and the latter makes no critical use of the fact that the values of $\psi(\mathbf{x}, \mathbf{y})$ are probabilities, or even that they are confined to the interval $[0, 1]$. The question arises, therefore: Is there a principled way of choosing the "right" transformation $\varphi[\psi(\mathbf{x}, \mathbf{y})]$ of $\psi(\mathbf{x}, \mathbf{y})$? In particular, is it justifiable to use the "raw" discrimination probabilities?

One possible approach to this issue is to relate it to another issue: to that of the possibility of experimental manipulations or spontaneous changes of context that change discrimination probabilities but leave intact subjective dissimilarities among the stimuli. In other words, we may relate the issue of possible transformations of discrimination probabilities to that of *response bias*.

Suppose that according to some theory of response bias, discrimination probability functions can be presented as $\psi_{\mathcal{B}}(\mathbf{x},\mathbf{y})$, where $\mathcal{B}$ is value of response bias, varying within some abstract set (of reals, real-valued vectors, functions, etc.). Intuitively, this means that although $\psi_{\mathcal{B}_1}(\mathbf{x},\mathbf{y})$ and $\psi_{\mathcal{B}_2}(\mathbf{x},\mathbf{y})$ for two distinct response bias values may be different, the difference is not in "true" subjective dissimilarities but merely in the "overall readiness" of the perceiver to respond "different" rather than "same." If Fechnerian distances are to be interpreted as "true" subjective dissimilarities, one should expect then that Fechnerian metrics corresponding to $\psi_{\mathcal{B}_1}(\mathbf{x},\mathbf{y})$ and $\psi_{\mathcal{B}_2}(\mathbf{x},\mathbf{y})$ are identical (up to multiplication by positive constants). This may or may not be true for Fechnerian metrics computed directly from $\psi_{\mathcal{B}}(\mathbf{x},\mathbf{y})$, and if it is not, it may be true for Fechnerian metrics computed from some transformation $\varphi[\psi_{\mathcal{B}}(\mathbf{x},\mathbf{y})]$ thereof. The solution for the problem of what transformations of discrimination probabilities one should make use of can now be formulated as follows: choose $\phi_{\mathcal{B}}(\mathbf{x},\mathbf{y}) = \varphi[\psi_{\mathcal{B}}(\mathbf{x},\mathbf{y})]$ so that $G(\mathbf{a},\mathbf{b})$ computed from $\phi_{\mathcal{B}}(\mathbf{x},\mathbf{y})$ is invariant (up to positive scaling) with respect to $\mathcal{B}$.

The approach proposed is, of course, open-ended, as the solution now depends on one's theory of response bias, independent of Fechnerian Scaling. Thus, if one adopts Luce's (1963) or Blackwell's (1953) linear model of bias, φ is essentially the identity function and one should deal with "raw" discrimination probabilities. If one adopts the conventional d' measure of sensitivity, φ can be chosen as the inverse of the standard normal integral,

$$\psi(\mathbf{x},\mathbf{y}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\varphi[\psi(\mathbf{x},\mathbf{y})]} e^{-z^2/2} \mathrm{d}z.$$

We do not know which model of response bias should be preferred.

Another approach to the problem of choosing the "right" transformation φ, which we mention without elaborating, is through adopting a model for computing discrimination probabilities from Fechnerian distances (and, possibly, other functions of stimuli). Thus, in Chapter 1, we discussed a "quadrilateral dissimilarity" model and its mathematically equivalent "uncertainty blobs" counterpart. According to this model, if we assume the canonical form of Regular Minimality, $\psi(\mathbf{x},\mathbf{y})$ (hence also $\varphi[\psi(\mathbf{x},\mathbf{y})]$) is a strictly increasing transformation of

$$S(\mathbf{x},\mathbf{y}) = R_1(\mathbf{x}) + 2D(\mathbf{x},\mathbf{y}) + R_2(\mathbf{y}),$$

where $D(\mathbf{x},\mathbf{y})$ is some intrinsic metric and R_1, R_2 some positive functions subject to certain constrains. It is easy to show that $D(\mathbf{x},\mathbf{y})$ will generally be different from the Fechnerian metric $G(\mathbf{x}, \mathbf{y})$ computed from thus generated $\psi(\mathbf{x}, \mathbf{y})$. The two intrinsic metrics may coincide, however, if $G(\mathbf{x}, \mathbf{y})$ is computed from $\varphi[\psi(\mathbf{x}, \mathbf{y})]$ rather than $\psi(\mathbf{x}, \mathbf{y})$. This suggests the following solution for the problem of what transformations of discrimination probabilities one should make use of: choose $\phi(\mathbf{x}, \mathbf{y}) = \varphi[\psi(\mathbf{x}, \mathbf{y})]$ so that $G(\mathbf{a}, \mathbf{b})$ computed from $\phi(\mathbf{x}, \mathbf{y})$ coincide with $D(\mathbf{x}, \mathbf{y})$ in the "quadrilateral dissimilarity" model.

APPENDIX: ALGORITHM OF FECHNERIAN SCALING OF DISCRETE OBJECT SETS

Given: a set of objects $\{\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_N\}$ and $N \times N$ matrix of discrimination probabilities $\psi(\mathbf{s}_i, \mathbf{s}_j)$ (referred to later as the original matrix).

1. Check the matrix for Regular Minimality: for $i = 1, ..., N$, the ith row should contain a single minimum $\psi(\mathbf{s}_i, \mathbf{s}_j)$ in cell (i, j), and this value should also be a single minimum in the jth column.
   - The row object $\mathbf{s}_i$ and the column object $\mathbf{s}_j$ forming such a cell, are points of subjective equality (PSE) for each other.
2. Form the table of mutual PSEs (row object vs. column object):
$$(\mathbf{s}_1, \mathbf{s}_{j_1}), (\mathbf{s}_2, \mathbf{s}_{j_2}), ..., (\mathbf{s}_N, \mathbf{s}_{j_N}).$$
   - $(j_1, j_2, ..., j_N)$ is a complete permutation of $(1, 2, ..., N)$.
3. Relabel the objects by assigning the same but otherwise arbitrary labels to mutual PSEs:
$$(\mathbf{s}_1, \mathbf{s}_{j_1}) \to (\mathbf{a}_1, \mathbf{a}_1), (\mathbf{s}_2, \mathbf{s}_{j_2}) \to (\mathbf{a}_2, \mathbf{a}_2), ..., (\mathbf{s}_N, \mathbf{s}_{j_N}) \to (\mathbf{a}_N, \mathbf{a}_N).$$
4. Form the matrix $\{\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_N\} \times \{\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_N\}$, with PSEs comprising the main diagonal.
   - Denote $\psi(\mathbf{a}_i, \mathbf{a}_j) = p_{ij}$ $(i, j = 1, ..., N)$.
   - Regular minimality now is satisfied in the canonical form: $p_{ii} < \min\{p_{ij}, p_{ji}\}$ for all $j \neq i$.
5. Compute the matrix of psychometric increments of the first kind,
$$\phi^{(1)}(\mathbf{a}_i, \mathbf{a}_j) = p_{ij} - p_{ii}.$$

6. For every ordered pair $(\mathbf{a}_i, \mathbf{a}_j)$, compute the smallest value of

$$L^{(1)}(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k) = \sum_{m=1}^{k-1} \phi^{(1)}(\mathbf{x}_m, \mathbf{x}_{m+1})$$

across all possible chains $\mathbf{a}_i = \mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k = \mathbf{a}_j$ $(k = 1, ..., N)$ whose elements are distinct.
   - This minimum value, $L^{(1)}_{\min}(\mathbf{a}_i, \mathbf{a}_j)$, is the oriented Fechnerian distance $G_1(\mathbf{a}_i, \mathbf{a}_j)$, of the first kind.
   - Any chain at which this minimum is achieved is a Fechnerian geodesic chain from $\mathbf{a}_i$ to $\mathbf{a}_j$.
   - [Simple heuristics can significantly reduce the combinatorial search for $G_1(\mathbf{a}_i, \mathbf{a}_j)$.]
7. From the $N \times N$ matrix of $G_1(\mathbf{a}_i, \mathbf{a}_j)$, compute the overall Fechnerian distances

$$G_{ij} = G_1(\mathbf{a}_i, \mathbf{a}_j) + G_1(\mathbf{a}_j, \mathbf{a}_i) = G_{ji}.$$

   - The concatenation of a geodesic chain from $\mathbf{a}_i$ to $\mathbf{a}_j$ with that from $\mathbf{a}_j$ to $\mathbf{a}_i$ forms a geodesic loop between $\mathbf{a}_i$ and $\mathbf{a}_j$ whose length $L^{(1)}$ equals G_{ij}.

8. (Alternatively or additionally, for verification purposes.) Perform Steps 5, 6, 7 with $\phi^{(2)}(\mathbf{a}_i, \mathbf{a}_j) = p_{ji} - p_{ii}$ replacing $\phi^{(1)}(\mathbf{a}_i, \mathbf{a}_j)$ to obtain oriented Fechnerian distances $G_2(\mathbf{a}_i, \mathbf{a}_j)$, of the second kind, overall Fechnerian distances $G_{ij} = G_2(\mathbf{a}_i, \mathbf{a}_j) + G_2(\mathbf{a}_j, \mathbf{a}_i) = G_{ji}$, and the corresponding geodesic chains and loops between $\mathbf{a}_i$ and $\mathbf{a}_j$.
   - Overall Fechnerian distances should be the same,

$$G_2(\mathbf{a}_i, \mathbf{a}_j) + G_2(\mathbf{a}_j, \mathbf{a}_i) = G_1(\mathbf{a}_i, \mathbf{a}_j) + G_1(\mathbf{a}_j, \mathbf{a}_i).$$

   - Geodesic chains and loops are the same, but read in the opposite direction.
9. In the matrix of overall Fechnerian distances, relabel the objects back,

$$\{\mathbf{a}_1 \to \mathbf{s}_1, \mathbf{a}_2 \to \mathbf{s}_2, ..., \mathbf{a}_N \to \mathbf{s}_N\}$$

and

$$\{\mathbf{a}_1 \to \mathbf{s}_{j_1}, \mathbf{a}_2 \to \mathbf{s}_{j_2}, ..., \mathbf{a}_N \to \mathbf{s}_{j_N}\},$$

to obtain, separately, the matrix of Fechnerian distances $G^{(1)}_{ij}$ for the row objects of the original matrix and the matrix of Fechnerian distances $G^{(2)}_{ij}$ for the column objects of the original matrix.
   - $G^{(1)}_{ij} = G^{(2)}_{i'j'}$ if and only if $(\mathbf{s}_i, \mathbf{s}_{i'})$ and $(\mathbf{s}_j, \mathbf{s}_{j'})$ are pairs of mutual PSEs.

10. In the matrix of geodesic loops, relabel all the objects back, as in the previous step, to obtain the geodesic loops between the row objects of the original matrix, and separately, the geodesic loops between the column objects of the original matrix.
    - A loop $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n, \mathbf{x}_1$ is a geodesic loop between the row objects $\mathbf{s}_i$ and $\mathbf{s}_j$ if and only if the corresponding loop of PSEs $\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_n, \mathbf{y}_1$ traversed in the opposite direction (i.e., $\mathbf{y}_1, \mathbf{y}_n, ..., \mathbf{y}_2, \mathbf{y}_1$) is a geodesic loop between the column objects $\mathbf{s}_{i'}$ and $\mathbf{s}_{j'}$ that are PSEs for $\mathbf{s}_i$ and $\mathbf{s}_j$, respectively.

Remark 1. No relabeling is needed if Regular Minimality in the original matrix holds in the canonical form to begin with. The matrices of Fechnerian distances and geodesic loops for the row and column objects then coincide (except that the geodesic loops for the column objects should be read in the opposite direction).

Remark 2. The original matrix of probabilities $\psi(\mathbf{s}_i, \mathbf{s}_j)$ can be any matrix that satisfies Regular Minimality and whose values are statistically compatible with the empirical estimates $\hat{\psi}(\mathbf{s}_i, \mathbf{s}_j)$. The algorithm does not work if no such matrix can be found. With large sample sizes, $\psi(\mathbf{s}_i, \mathbf{s}_j)$ can be simply identified with $\hat{\psi}(\mathbf{s}_i, \mathbf{s}_j)$, with smaller sample sizes, one may need to try a large set of matrices, $\psi(\mathbf{s}_i, \mathbf{s}_j)$ statistically compatible with given $\hat{\psi}(\mathbf{s}_i, \mathbf{s}_j)$, and to replicate the algorithm with each of these to eventually obtain joint confidence intervals for Fechnerian distances.

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

References

Blackwell, H. R. (1953). Psychophysical thresholds: experimental studies of methods of measurement. *Engineering Research Bulletin*, No. 36. Ann Arbor: University of Michgan Press.

Borg, I., & Groenen, P. (1997). *Modern multidimensional scaling.* New York: Springer-Verlag.

Corter, J. E. (1996). *Tree models of similarity and association.* Beverly Hills, CA: Sage.

DeSarbo, W. S., Johnson, M. D., Manrai, A. K., Manrai, L. A., & Edwards, E. A. (1992) TSCALE: A new multidimensional scaling procedure based on Tversky's contrast model. *Psychometrika*, 57, 43–70.

Dzhafarov, E. N. (2002a). Multidimensional Fechnerian scaling: Regular variation version. *Journal of Mathematical Psychology,* 46, 226–244.

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

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

Dzhafarov, E. N. (2002d). 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., & Colonius, H. (1999). Fechnerian metrics in unidimensional and multidimensional stimulus spaces. *Psychonomic Bulletin and Review,* 6, 239–268.

Dzhafarov, E. N., & Colonius, H. (2001). Multidimensional Fechnerian scaling: Basics. *Journal of Mathematical Psychology,* 45, 670–719.

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.

Everitt, B. S., & Rabe-Hesketh, S. (1997). *The analysis of proximity data.* New York: Wiley.

Fechner, G. T. (1860). *Elemente der Psychophysik* [Elements of psychophysics]. Leipzig, Germany: Breitkopf & Härtel.

Hartigan, J. A. (1975). *Clustering algorithms.* New York:Wiley.

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.

Kruskal, J. B., & Wish, M. (1978). *Multidimensional scaling.* Beverly Hills, CA: Sage.

Luce, R. D. (1963). A threshold theory for simple detection experiments. *Psychological Review*, 70, 61–79.

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

Roweis, S. T., & Saul, L. K. (2000). Nonlinear dimensionality reduction by locally linear embedding. *Science,* 290, 2323-2326.

Sankoff, D. & Kruskal, J. (1999). *Time warps, string edits, and macromolecules.* Stanford, CA: CSLI Publications.

Semple, C., & Steele, M. (2003). *Phylogenetics.* Oxford, England: Oxford University Press.

Shepard, R. N. (1957). Stimulus and response generalization: A stochastic model relating generalization to distance in psychological space. *Psychometrika*, 22, 325–345.

Shepard, R. N. (1958). Stimulus and response generalization: Tests of a model relating generalization to distance in psychological space. *Journal of Experimental Psychology*, 55, 509–523.

Shepard, R. N., and Carroll, J. D. (1966). Parametric representation of nonlinear data structures. In P. R. Krishnaiah (Ed.), *Multivariate analysis* (pp. 561–592). New York, NY: Academic Press.

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

Tenenbaum, J. B., de Silva, V., & Langford, J. C. (2000). A global geometric framework for nonlinear dimensionality reduction. *Science,* 290, 2319–2323.

Tversky, A. (1977). Features of similarity. *Psychological Review*, 84, 327–352.

Weeks, D. G., & Bentler, P. M. (1982). Restricted multidimensional scaling models for asymmetric proximities. *Psychometrika*, 47, 201–208.

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