Available online at www.sciencedirect.com

SCIENCE DIRECT

Journal of Mathematical Psychology 47 (2003) 205–219

Journal of Mathematical Psychology

http://www.elsevier.com/locate/jmp

Thurstonian-type representations for "same-different" discriminations: Probabilistic decisions and interdependent images

Ehtibar N. Dzhafarov*

Department of Psychological Sciences, Purdue University, 703 Third Street, West Lafayette, IN 47907-2004, USA

Received 13 February 2002; revised 30 October 2002

Abstract

A general Thurstonian-type representation (with stochastically interdependent images and probabilistic decisions) for a "same-different" discrimination probability function $\psi(\mathbf{x}, \mathbf{y})$ is a model in which the two stimuli $\mathbf{x}, \mathbf{y}$ are mapped into two generally interdependent random images $P(\mathbf{x})$ and $Q(\mathbf{y})$ taking on their values in some "perceptual" space; and the realizations of these two random images in a given trial determine the probability with which $\mathbf{x}$ and $\mathbf{y}$ in this trial are judged to be different. While stochastically interdependent, $P(\mathbf{x})$ and $Q(\mathbf{y})$ are selectively attributed to (influenced by), respectively, $\mathbf{x}$ and $\mathbf{y}$, which is understood as the possibility of conditioning $P(\mathbf{x})$ and $Q(\mathbf{y})$ on some random variable R that renders them stochastically independent, with their conditional distributions selectively depending on, respectively, $\mathbf{x}$ and $\mathbf{y}$. A general Thurstonian-type representation is considered "well-behaved" if the conditional probability with which $P(\mathbf{x})$ and $Q(\mathbf{y})$, given a value of the conditioning random variable R, fall within two given subsets of the perceptual space, possess appropriately defined bounded directional derivatives with respect to $\mathbf{x}$ and $\mathbf{y}$. It is shown that no such well-behaved Thurstonian-type representation can account for $\psi(\mathbf{x}, \mathbf{y})$ possessing two basic properties: regular minimality and nonconstant self-similarity. At the same time, an alternative to Thurstonian-type modeling (a model employing "uncertainty blobs" in stimulus spaces instead of random variables in perceptual spaces) is readily available that predicts these two properties "automatically".

1. Introduction

This paper deals with "same-different" discrimination probabilities

$$\psi(\mathbf{x}, \mathbf{y}) = \Pr[\mathbf{y} \text{ is discriminated from } \mathbf{x}]$$

and their Thurstonian-type representations (models), in which stimuli $\mathbf{x}, \mathbf{y}$ are mapped into random images $P(\mathbf{x}), Q(\mathbf{y})$ taking on their values in a hypothetical "perceptual" space $\mathfrak{R}$. In a companion paper (Dzhafarov, 2003) it is shown that if a discrimination probability function $\psi(\mathbf{x}, \mathbf{y})$ possesses two basic properties, regular minimality and nonconstant self-similarity, then it cannot be accounted for by a well-behaved Thurstonian-type model with deterministic decisions and stochastically independent images. The practical significance of this result lies in the fact that the "well-behavedness" is a weak constraint unlikely to be violated in any conceivable Thurstonian-type representation constructed to fit empirical data. In the present paper the notion of well-behavedness and the main conclusion arrived at in Dzhafarov (2003) are extended to Thurstonian-type models with (generally) *probabilistic decisions* and (generally) *stochastically interdependent images*.

For the convenience of reference, and to some extent imitating Thurstone's famous "cases" (Thurstone, 1927a,b), I introduce four "varieties" of Thurstonian-type representations for discrimination probabilities, as shown in the table below:

Thurstonian-type representations	Independent Images	Interdependent Images
Deterministic decisions	Deterministic-Independent	Deterministic-Interdependent
Probabilistic decisions	Probabilistic-Independent	Probabilistic-Interdependent

*Fax: 765-496-1264.
E-mail address: ehtibar@purdue.edu (E.N. Dzhafarov).

0022-2496/03/$ - see front matter
doi:10.1016/S0022-2496(02)00014-7

The four varieties are not mutually exclusive, rather we have

$$\text{Deterministic-Independent} \quad \subset \begin{cases} \text{Deterministic-Interdependent} \\ \quad \text{Probabilistic-Independent} \end{cases} \subset \text{Probabilistic-Interdependent},$$

with $\subset$ informally standing for "is a special case of".

In the deterministic-decision varieties (Deterministic-Independent and Deterministic-Interdependent) of Thurstonian-type models, the response "$\mathbf{x}$ and $\mathbf{y}$ are different" is chosen in a given trial if and only if the realizations p, q of $P(\mathbf{x})$ and $Q(\mathbf{y})$ in this trial fall within a certain area $\mathfrak{S} \subseteq \mathfrak{R} \times \mathfrak{R}$, called the "decision area". This means that the (p, q)-pairs within the area $\mathfrak{S}$ evoke the response "different" with probability 1, whereas the (p, q)-pairs outside this area evoke this response with probability 0. In the probabilistic-decision varieties (Probabilistic-Independent and Probabilistic-Interdependent) this decision scheme is generalized: every (p, q)-pair is associated with a certain probability $\sigma(p, q)$ of evoking the response "different".

In the independent-images varieties (Deterministic-Independent and Probabilistic-Independent) of Thurstonian-type models, the images $P(\mathbf{x})$ and $Q(\mathbf{y})$ of stimuli $\mathbf{x}, \mathbf{y}$ are stochastically independent and *selectively influenced* by (or *selectively attributed* to) their respective stimuli: that is, the distribution of $P(\mathbf{x})$ does not depend on $\mathbf{y}$, while the distribution of $Q(\mathbf{y})$ does not depend on $\mathbf{x}$. In the interdependent-images varieties (Deterministic-Interdependent and Probabilistic-Interdependent), $P(\mathbf{x})$ and $Q(\mathbf{y})$ are generally stochastically interdependent, but the selective attribution of these random variables to, respectively, $\mathbf{x}$ and $\mathbf{y}$ should still be preserved.

As pointed out in Dzhafarov (2003), this selective attribution is taken as an inherent feature of any Thurstonian-type model. Without it one would have no justification for writing $P = P(\mathbf{x})$ and calling it an image of $\mathbf{x}$ (rather than a response to the pair $\mathbf{x}, \mathbf{y}$). This position leads to a non-trivial conceptual problem: *how should one understand the selectiveness in the influence of* $\mathbf{x}$ *and* $\mathbf{y}$ *upon, respectively,* $P(\mathbf{x})$ *and* $Q(\mathbf{y})$ *when the latter are not stochastically independent*? This problem is considered in Dzhafarov (1999, 2001), but the approach adopted in the present paper is based on the general solution proposed in Dzhafarov (in press). The essence of this solution, when applied to $P(\mathbf{x})$ and $Q(\mathbf{y})$, is that one can find a random variable R (whose distribution does not depend on $\mathbf{x}$ or $\mathbf{y}$) such that the two random variables $P(\mathbf{x})$ and $Q(\mathbf{y})$ are conditionally independent given any value of R, and the conditional distribution of $P(\mathbf{x})$ does not depend on $\mathbf{y}$, while the conditional distribution of $Q(\mathbf{y})$ does not depend on $\mathbf{x}$.

Remark 1.1. As mentioned in Dzhafarov (2003), A.A.J. Marley (pers. comm., 2002) pointed out that the term "Thurstonian-type representations" may have too narrow connotations to be appropriate for constructs as general as those considered in the present work. His tentative suggestion was to replace it with the term "*random-image representations*".

2. Plan of the paper and notation conventions

The development presented in this paper can be summarized as follows:

1. Section 3 provides a recapitulation of the basic notions introduced in Dzhafarov (2002c, 2003): the properties of regular minimality and nonconstant self-similarity, the notion of a patch of a discrimination probability function, the near-smoothness property, and the well-behavedness of a Thurstonian-type representation (of the Deterministic-Independent variety). The section also presents the two main results obtained in Dzhafarov (2003): that a patch possessing the regular minimality and nonconstant self-similarity properties cannot be near-smooth, while the well-behavedness of a Thurstonian-type representation for a patch implies its near-smoothness.
2. In Section 4 the notion of well-behavedness (defined in the same way as for the Deterministic-Independent variety) is applied to Thurstonian-type models of the Probabilistic-Independent variety, and it is shown (Theorem 5.1) that these models imply the near-smoothness property, because of which they cannot account for "same-different" discrimination probability functions subject to the regular minimality and nonconstant self-similarity constraints.
3. Section 6 describes the general approach to the problem of selective influence under stochastic interdependence adopted in this paper.
4. In Section 7 this approach is applied to Thurstonian-type representations of the most general (Probabilistic-Interdependent) variety, allowing one to naturally extend the notion of well-behavedness to such models. It is shown then (Theorem 8.1) that even in this, most general version, a well-behaved Thurstonian-type model predicts the near-smoothness property, because of which it cannot account for a discrimination probability function possessing the properties of regular minimality and nonconstant self-similarity.

5. In Section 9 it is shown that the definition of well-behavedness for Thurstonian-type representations of the Probabilistic-Interdependent variety can be significantly relaxed without affecting any of the results obtained in this paper.
6. In the Conclusion, I briefly discuss the implications of the inadequacy of the well-behaved Thurstonian-type models for discrimination probabilities and outline a certain alternative to such models ("uncertainty blobs" of stimuli).
7. The development is aided by an appendix containing two lemmas labeled A.1 and A.2.

The notation conventions are the same as in the companion paper.

Boldface lowercase letters $(\mathbf{x}, \mathbf{y}, \mathbf{u}, \ldots)$ denote real-valued vectors; their components, if shown, are superscripted, e.g., $\mathbf{x} = (x^1, \ldots, x^n)$, $\mathbf{u} = (u^1, \ldots, u^n)$.

Uppercase Gothic letters $(\mathfrak{R}, \mathfrak{C}, \mathfrak{U}, \mathfrak{B}$, etc.) denote sets, lowercase Gothic letters $(\mathfrak{p}, \mathfrak{q}, \mathfrak{r}, \ldots)$ denote subsets of the "perceptual space" $\mathfrak{R}$ or the "conditioning space" $\mathfrak{C}$.

Uppercase Greek letters Σ and Ω denote sets of subsets.

Lowercase and uppercase italics designate real-valued quantities, except for letters P, Q, p, q that are reserved to denote random images (P, Q) and their values (p, q), and letters R, r reserved for, respectively, the conditioning random variable and its values.

3. Basic definitions and facts

Although this paper extends the results obtained in Dzhafarov (2003), the development to follow is self-contained, with all requisite definitions and facts stated explicitly if not in detail.

3.1. Two basic properties of discrimination probabilities

Stimuli $\mathbf{x}$ and $\mathbf{y}$ in $\psi(\mathbf{x}, \mathbf{y})$ belong to an *open connected subset* $\mathfrak{M}$ of $\mathrm{Re}^n (n \geqslant 1)$. The pair $(\mathbf{x}, \mathbf{y})$ is ordered, due to the fact that $\mathbf{x}$ and $\mathbf{y}$ belong to two distinct *observation areas* (spatial and/or temporal intervals).

It is assumed that for a certain homeomorphic mapping $\mathbf{h} : \mathfrak{M} \to \mathfrak{M}$ (one-to-one, onto, continuous together with its inverse),

$$\arg\min_{\mathbf{y}} \psi(\mathbf{x}, \mathbf{y}) = \mathbf{h}(\mathbf{x}), \arg\min_{\mathbf{x}} \psi(\mathbf{x}, \mathbf{y}) = \mathbf{h}^{-1}(\mathbf{y}). \quad (1)$$

This is called the *regular minimality* property of $\psi(\mathbf{x}, \mathbf{y})$. One also says in this case that $\psi(\mathbf{x}, \mathbf{y})$ possesses *regular minima*. This property is a weakened version of the so-called First Assumption of multidimensional Fechnerian scaling (see Dzhafarov, 2002a, c; Dzhafarov & Colonius, 2001). Stimuli $\mathbf{x}, \mathbf{h}(\mathbf{x})$ are *points of subjective equality* with respect to each other. If $\mathbf{h}$ is the identity $\mathfrak{M} \to \mathfrak{M}$, the regular minimality property acquires its simplest form

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

If the minimum value $\psi(\mathbf{x}, \mathbf{h}(\mathbf{x}))$ of $\psi(\mathbf{x}, \mathbf{y})$ is generally different for different $\mathbf{x}$, then $\psi(\mathbf{x}, \mathbf{y})$ is said to have the property of *nonconstant self-similarity*.

The two properties of $\psi(\mathbf{x}, \mathbf{y})$ are corroborated by available empirical data (Dzhafarov, 2002c; Indow, 1998; Indow, Robertson, von Grunau, & Fielder, 1992; Krumhansl, 1978; Rothkopf, 1957; Tversky, 1977; Zimmer & Colonius, 2000), and I consider them fundamental for "same-different" discriminations. For a detailed analysis of these properties the reader should consult Dzhafarov (2002c).

3.2. Patches of discrimination probabilities

A stimulus–direction pair $(\mathbf{s}, \mathbf{u})$, $\mathbf{s} \in \mathfrak{M} \subseteq \mathrm{Re}^n$, $\mathbf{0} \neq \mathbf{u} \in \mathrm{Re}^n$, is called a *line element*. Having chosen $(\mathbf{s}, \mathbf{u})$ and a sufficiently small $a > 0$, all $(\mathbf{x}, \mathbf{y}) = (\mathbf{s} + \mathbf{u}x, \mathbf{h}(\mathbf{s} + \mathbf{u}y))$ with $(x, y) \in [-a, a]^2$ belong to $\mathfrak{M} \times \mathfrak{M}$, and the restriction of $\psi(\mathbf{x}, \mathbf{y})$ to this square area of stimulus pairs, viewed as a function of (x, y), is called a *patch* of $\psi(\mathbf{x}, \mathbf{y})$ at the line element $(\mathbf{s}, \mathbf{u})$. This patch is denoted as $\psi_{(\mathbf{s},\mathbf{u})}(x, y)$, or, when $(\mathbf{s}, \mathbf{u})$ is fixed or arbitrary, simply as $\psi(x, y)$. (The precise value of a is never important: it can always be taken as small as one wishes.) With this local parametrization, the segment of *corresponding* stimuli $\mathbf{x}$ and $\mathbf{y} = \mathbf{h}(\mathbf{x})$ is encoded by $-a \leqslant x = y \leqslant a$, and the "central" pair $(\mathbf{s}, \mathbf{h}(\mathbf{s}))$ is represented by $x = y = 0$.

The set of all patches $\psi(x, y)$ of $\psi(\mathbf{x}, \mathbf{y})$ may not cover the entire function $\psi(\mathbf{x}, \mathbf{y})$, but it will cover $\psi(\mathbf{x}, \mathbf{y})$ in a sufficiently small vicinity of its minima $\psi(\mathbf{x}, \mathbf{h}(\mathbf{x}))$, which is all that the subsequent development requires. By abuse of language, x and y are conveniently referred to as *stimuli* (rather than parametric representations of stimuli $\mathbf{s} + \mathbf{u}x$ and $\mathbf{h}(\mathbf{s} + \mathbf{u}y)$). The following two statements are almost obvious (see Dzhafarov, 2003).

If $\psi(\mathbf{x}, \mathbf{y})$ possesses regular minima, then so does any of its patches $\psi(x, y)$, but in the simplest form

$$\arg\min_{y} \psi(x, y) = x, \arg\min_{x} \psi(x, y) = y,$$

or, equivalently,

$$\psi(x, x) < \begin{cases} \psi(x, y), \\ \psi(y, x), \end{cases} \quad x \in [-a, a], \quad y \in [-a, a]. \quad (2)$$

If $\psi(\mathbf{x}, \mathbf{y})$ possesses the nonconstant self-similarity property, then, at least at some line elements $(\mathbf{s}, \mathbf{u})$,

$$\psi(x, x) \not\equiv \mathrm{const}, \quad x \in [-a, a] \quad (3)$$

for all $a > 0$. Such patches are called *typical*, and the line elements $(\mathbf{s}, \mathbf{u})$ at which the patches are typical are called *typical line elements*.

3.3. Near-smoothness

A patch $\psi(x, y)$ is called *near-smooth* if it is both right- and left-differentiable in both x and y, with the unilateral derivatives $\frac{\partial}{\partial x\pm}\psi(x, y)$ being continuous in y and $\frac{\partial}{\partial y\pm}\psi(x, y)$ being continuous in x. All smooth (continuously differentiable) functions are near-smooth. Simple examples of nonsmooth but near-smooth functions are $|x| + |y|$, $|xy|$, $1 - \exp[-(|x| - |y|)^2]$, etc. $(-a \leqslant x, y \leqslant a)$. A simple example of a unilaterally differentiable but not near-smooth function is $|x - y|$.

The following theorem is proved in Dzhafarov (2003).

Theorem 3.1. *Let a discrimination probability function* $\psi(\mathbf{x}, \mathbf{y})$ *be subject to both the regular minimality and nonconstant self-similarity constraints. Then a typical patch* $\psi(x, y)$ *of this function cannot be near-smooth.*

3.4. Thurstonian-type representations of the Deterministic-Independent variety

A Thurstonian-type representation of the Deterministic-Independent variety for a patch $\psi(x, y)$ is defined by the construct

$$\{\mathfrak{R}, A_x, B_y, \mathfrak{S}\}, \tag{4}$$

where $\mathfrak{R}$ is a perceptual space (an arbitrary set); $A_x(\mathfrak{p}), \mathfrak{p} \in \Sigma_A$, is a probability measure defined on a sigma-algebra Σ_A of subsets of $\mathfrak{R}$ and associated with the *random image* $P(x)$ of x,

$$A_x(\mathfrak{p}) = \Pr[P(x) \in \mathfrak{p}],$$

the probabilistic measure $B_y(\mathfrak{q}), \mathfrak{q} \in \Sigma_B$ is defined analogously, with

$$B_y(\mathfrak{q}) = \Pr[Q(y) \in \mathfrak{q}],$$

and $\mathfrak{S} \subseteq \mathfrak{R} \times \mathfrak{R}$ (called the *decision area*) is defined by

$$\psi(x, y) = \Pr[(P(x), Q(y)) \in \mathfrak{S}] = \int_{(p,q) \in \mathfrak{S}} dA_x(p)\, dB_y(q). \tag{5}$$

3.5. Well-behavedness

Definition 3.1. The probability measure $A_x(\mathfrak{p})$ is called *well-behaved* (in the absolute, or narrow sense) if $\frac{\partial}{\partial x+}A_x(\mathfrak{p})$ and $\frac{\partial}{\partial x-}A_x(\mathfrak{p})$ exist and are bounded for all $(\mathfrak{p}, x) \in \Sigma_A \times [-a, a]$, that is,

$$\left|\frac{\partial}{\partial x\pm}A_x(\mathfrak{p})\right| < \text{const}.$$

The well-behavedness of $B_y(\mathfrak{q})$ is defined analogously, and a Thurstonian-type representation of the Deterministic-Independent variety for a patch $\psi(x, y)$, $(x, y) \in [-a, a]^2$, is called *well-behaved* (in the absolute, or narrow sense) if both $A_x(\mathfrak{p})$ and $B_y(\mathfrak{q})$ in it are well-behaved.

As argued in Dzhafarov (2003), the well-behavedness is a weak regularity feature. Most of the existing models for "same-different" discriminations use multivariate or univariate normal distributions for A_x, B_y (Dai, Versfeld, & Green, 1996; Ennis, 1992; Ennis, Palen, & Mullen, 1988; Luce & Galanter, 1963; Sorkin, 1962; Suppes & Zinnes, 1963; Thomas, 1996, 1999; Zinnes & MacKay, 1983), and any such a model is well-behaved provided the relationship between stimuli and the means and covariances of the normal distributions is assumed to be sufficiently smooth.

3.6. Main result for the Deterministic-Independent variety

The following theorem is proved in Dzhafarov (2003).

Theorem 3.2. *A patch* $\psi(x, y)$ *that has a well-behaved Thurstonian-type representation of Deterministic-Independent variety is near-smooth.*

Relating this result to Theorem 3.1, one arrives at the

Main Conclusion (For the Deterministic-Independent variety). A typical patch $\psi(x, y)$ (satisfying the regular minimality and nonconstant self-similarity conditions (2) and (3)) does not have a well-behaved Thurstonian-type representation of the Deterministic-Independent variety. As a result, no discrimination probability function $\psi(\mathbf{x}, \mathbf{y})$ with regular minima and nonconstant self-similarity allows for a Thurstonian-type representation of the Deterministic-Independent variety that is well-behaved at any of the typical line elements $(\mathbf{s}, \mathbf{u})$.

Remark 3.6.1. This conclusion and Theorem 3.2 from which it follows are valid under a more general definition of well-behavedness than Definition 3.1 (the "well-behavedness in the relative, or broad sense"). For the general, Probabilistic-Interdependent variety of Thurstonian-type representations this generalized notion is defined in Section 9. Due to its greater simplicity and intuitiveness, however, all the results in this paper are first established for the absolute, or narrow meaning of well-behavedness.

4. Thurstonian-type representations of Probabilistic-Independent variety

A Thurstonian-type representation of the Probabilistic-Independent variety for a patch $\psi(x, y)$, $(x, y) \in [-a, a]^2$, of a discrimination probability function

$\psi(\mathbf{x}, \mathbf{y})$ is defined by the construct

$$\{\mathfrak{R}, A_x, B_y, \sigma\}, \tag{6}$$

where the perceptual space $\mathfrak{R}$ and the probability measures $A_x(\mathfrak{p})(\mathfrak{p} \in \Sigma_A)$ and $B_y(\mathfrak{q})(\mathfrak{q} \in \Sigma_B)$ are understood in the same way as in (4), but the decision area $\mathfrak{S}$ in (4) and (5) is replaced by a *decision probability function*

$$\sigma : \mathfrak{R} \times \mathfrak{R} \to [0, 1],$$

such that for any pair of values $P(x) = p, Q(y) = q$, the probability with which (p, q) evokes the response "different" is $\sigma(p, q)$. No restrictions are imposed on $\sigma(p, q)$, except for its measurability (with respect to the product measure $AB_{xy} = A_x \times B_y$, defined on the sigma-algebra Σ_{AB} generated by $\Sigma_A \times \Sigma_B$). We have, therefore,

$$\psi(x, y) = \sigma(P(x), Q(y)) = \int_{(p,q) \in \mathfrak{R}^2} \sigma(p, q)\, dA_x(p)\, dB_y(q). \tag{7}$$

Remark 4.1. The Deterministic-Independent variety of Thurstonian-type representations is obtained as a special case, by putting

$$\sigma(p, q) = \chi_{\mathfrak{S}}(p, q) = \begin{cases} 1 & \text{if } (p, q) \in \mathfrak{S}, \\ 0 & \text{if } (p, q) \notin \mathfrak{S}. \end{cases}$$

By Fubini's theorem (see, e.g., Hewitt & Stromberg, 1965, pp. 384–385),

$$\psi(x, y) = \int_{p \in \mathfrak{R}} B^*(p, y) dA_x(p) = \int_{q \in \mathfrak{R}} A^*(q, x)\, dB_y(q), \tag{8}$$

where

$$B^*(p, y) = \int_{q \in \mathfrak{R}} \sigma(p, q)\, dB_y(q),$$

$$A^*(q, x) = \int_{p \in \mathfrak{R}} \sigma(p, q)\, dA_x(p),$$

and $p \to B^*(p, y)$ and $q \to A^*(q, x)$ are, respectively, A-measurable and B-measurable functions.

The well-behavedness (in the absolute, or narrow sense) for the Probabilistic-Independent variety is defined in precisely the same way as for the Deterministic-Independent Variety (Definition 3.1): the unilateral derivatives $\frac{\partial}{\partial x\pm} A_x(\mathfrak{p})$ and $\frac{\partial}{\partial y\pm} B_y(\mathfrak{q})$ exist and are bounded by a constant c,

$$\left|\frac{\partial}{\partial x\pm} A_x(\mathfrak{p})\right| \leqslant c, \left|\frac{\partial}{\partial y\pm} B_y(\mathfrak{q})\right| \leqslant c, \tag{9}$$

for all $\mathfrak{p} \in \Sigma_A$, $\mathfrak{q} \in \Sigma_B$, and $(x, y) \in [-a, a]^2$.

Remark 4.2. As in Dzhafarov (2003), the reader who wishes to overlook measure-theoretic technicalities may think of Thurstonian-type representations in terms of the following simple example, ignoring all references to sigma-algebras and measurability. Consider $\mathfrak{R}$ which is a finite set of states $\{1, \ldots, k\}$. The distributions of $P(x)$ and $Q(y)$ in this case are defined by

$$\begin{cases} dA_x(i) = \alpha(i, x) = \Pr[P(x) = i], \\ dB_y(i) = \beta(i, y) = \Pr[Q(y) = i], \end{cases} \quad i = 1, \ldots, k,$$

so that $A_x(\mathfrak{p}) = \sum_{i \in \mathfrak{p}} \alpha(i, x)$, $B_y(\mathfrak{q}) = \sum_{i \in \mathfrak{q}} \beta(i, y)$, where $\mathfrak{p}$ and $\mathfrak{q}$ belong to $\Sigma_A = \Sigma_B$, which is simply the set of all 2^k subsets of $\{1, \ldots, k\}$. The decision probability function σ may be any function mapping the pairs $(i,j)(i,j = 1, \ldots, k)$ into the interval $[0, 1]$. Relation (7) acquires the form

$$\psi(x, y) = \sum_{i=1}^{k} \sum_{j=1}^{k} \sigma(i,j)\alpha(i, x)\beta(j, y).$$

The well-behavedness here means that $\frac{\partial}{\partial x\pm}\alpha(i, x)$ and $\frac{\partial}{\partial y\pm}\beta(i, y)$ exist and are bounded for all $i = 1, \ldots, k$ and all $(x, y) \in [-a, a]^2$.

Thurstonian-type representations of the Probabilistic-Independent variety can also be defined for the entire discrimination probability function $\psi(\mathbf{x}, \mathbf{y})$ rather than for its patches. This is done by replacing $A_x(\mathfrak{p})$ ($\mathfrak{p} \in \Sigma_A$, $x \in [-a, a]$) and $B_y(\mathfrak{q})$ ($\mathfrak{q} \in \Sigma_B$, $y \in [-a, a]$) with $A_{\mathbf{x}}(\mathfrak{p})$ ($\mathfrak{p} \in \Sigma_A$, $\mathbf{x} \in \mathfrak{M}$) and $B_{\mathbf{y}}(\mathfrak{q})(\mathfrak{q} \in \Sigma_B$, $\mathbf{y} \in \mathfrak{M})$. Relationship (7) then transforms into

$$\psi(\mathbf{x}, \mathbf{y}) = \sigma(P(\mathbf{x}), Q(\mathbf{y})) = \int_{(p,q) \in \mathfrak{R}^2} \sigma(p, q) dB_{\mathbf{y}}(q) dA_{\mathbf{x}}(p).$$

Clearly, this relationship cannot hold true if (7) fails to hold for even a single patch $\psi(x, y)$. This explains the significance of the result presented in the next section, that no typical patch $\psi(x, y)$ can have a well-behaved Thurstonian-type representation of the Probabilistic-Independent variety.

5. Near-Smoothness Theorem for Probabilistic-Independent variety

Theorem 5.1. *A patch $\psi(x, y)$ that has a well-behaved Thurstonian-type representation of the Probabilistic-Independent variety is near-smooth.*

Proof. We prove that the derivatives $\frac{\partial}{\partial y\pm}\psi(x, y)$ exist and are continuous in x. The proof that $\frac{\partial}{\partial x\pm}\psi(x, y)$ exist and are continuous in y is obtained by symmetrical argument.

Part 1: Existence: In accordance with (8),

$$\psi(x, y) = \int_{p \in \mathfrak{R}} B^*(p, y)\, dA_x(p),$$

where

$$B^*(p,y)=\int_{q\in\mathfrak{R}}\sigma(p,q)\,dB_y(q).$$

We first prove that $y\to B^*(p_0,y)$ possesses unilateral derivatives bounded on $[-a,a]$, for any fixed $p=p_0$.

Choose an arbitrary $0<\varepsilon<1$, and construct a sequence of functions $\{\varphi_i(q)\}_{i=1}^{\infty}$ as follows (by induction).

Induction base: Since $q\to\sigma(p_0,q)$ is B-measurable and bounded, $0\leqslant\sigma(p_0,q)\leqslant 1$, one can find (see, e.g., Hewitt & Stromberg, 1965, pp. 172–173) a B-measurable (so-called "simple") function

$$\varphi_1(q)=\sum_{j=1}^{n_1}b_{1j}\chi_{\mathfrak{q}_{1j}}(q),$$

with $\chi_{\mathfrak{q}_{1j}}$ being characteristic functions of pairwise disjoint B-measurable sets $\mathfrak{q}_{1j}$ ($j=1,\dots,n_1$), such that, for all $q\in\mathfrak{R}$,

$$0\leqslant\sigma(p_0,q)-\varphi_1(q)\leqslant\varepsilon.$$

Clearly, $0\leqslant\varphi_1(q)\leqslant 1$, or, equivalently, $\{0\leqslant b_{1j}\leqslant 1\}_{j=1}^{n_1}$.

Induction step: Assuming $\{\varphi_i(p)\}_{i=1}^{k-1}$ have been constructed ($k>1$), and that

$$0\leqslant\sigma(p_0,q)-\sum_{i=1}^{k-1}\varphi_i(q)\leqslant\varepsilon^{k-1},$$

one can use the same argument as in the induction base above to find a B-measurable simple function

$$\varphi_k(q)=\sum_{j=1}^{n_k}b_{kj}\chi_{\mathfrak{q}_{kj}}(q),$$

with $\chi_{\mathfrak{q}_{kj}}$ being characteristic functions of pairwise disjoint B-measurable sets $\mathfrak{q}_{kj}$ ($j=1,\dots,n_k$), such that, for all $q\in\mathfrak{R}$,

$$0\leqslant\left[\sigma(p_0,q)-\sum_{i=1}^{k-1}\varphi_i(q)\right]-\varphi_k(q)\leqslant\varepsilon^{k}.$$

Clearly, $0\leqslant\varphi_k(q)\leqslant\varepsilon^{k-1}$, or, equivalently, $\{0\leqslant b_{kj}\leqslant\varepsilon^{k-1}\}_{j=1}^{n_k}$.

This completes the inductive definition of $\{\varphi_i(q)\}_{i=1}^{\infty}$. By construction,

$$\sum_{i=1}^{\infty}\varphi_i(q)=\sigma(p_0,q).$$

Form now the sequence of functions $\{f_i(y)\}_{i=1}^{\infty}$ defined by

$$f_i(y)=\int_{q\in\mathfrak{R}}\varphi_i(q)\,dB_y(q).$$

Using the construction logic of Lebesgue integrals and the measurability of $q\to\varphi_i(q)$,

$$B^*(p_0,y)=\int_{q\in\mathfrak{R}}\sigma(p_0,q)\,dB_y(q)$$

$$=\int_{q\in\mathfrak{R}}\sum_{i=1}^{\infty}\varphi_i(q)\,dB_y(q)=\sum_{i=1}^{\infty}f_i(y).$$

Now,

$$f_i(y)=\int_{q\in\mathfrak{R}}\varphi_i(q)\,dB_y(q)=\sum_{j=1}^{n_i}\int_{q\in\mathfrak{R}}b_{ij}\chi_{\mathfrak{q}_{ij}}(q)dB_y(q)$$

$$=\sum_{j=1}^{n_i}b_{ij}B_y(\mathfrak{q}_{ij}),$$

whence

$$\frac{\partial}{\partial y\pm}f_i(y)=\sum_{j=1}^{n_i}b_{ij}\frac{\partial}{\partial y\pm}B_y(\mathfrak{q}_{ij}).$$

Let the indexation by j be so arranged that $\frac{\partial}{\partial y\pm}B_y(\mathfrak{q}_{ij})\geqslant 0$ for $j=1,\dots,k_i$, and $\frac{\partial}{\partial y\pm}B_y(\mathfrak{q}_{ij})<0$ for the remainder. Rewriting the last equation as

$$\frac{\partial}{\partial y\pm}f_i(y)=\sum_{j=1}^{k_i}b_{ij}\frac{\partial}{\partial y\pm}B_y(\mathfrak{q}_{ij})+\sum_{j=k_i+1}^{n_i}b_{ij}\frac{\partial}{\partial y\pm}B_y(\mathfrak{q}_{ij}),$$

and recalling that $0\leqslant b_{ij}\leqslant\varepsilon^{i-1}$ for all $j=1,\dots,n_i$, we have

$$\left|\frac{\partial}{\partial y\pm}f_i(y)\right|\leqslant\max\left\{\sum_{j=1}^{k_i}\varepsilon^{i-1}\frac{\partial}{\partial y\pm}B_y(\mathfrak{q}_{ij}),\right.$$
$$\left.-\sum_{j=k_i+1}^{n_i}\varepsilon^{i-1}\frac{\partial}{\partial y\pm}B_y(\mathfrak{q}_{ij})\right\}$$

$$=\varepsilon^{i-1}\max\left\{\frac{\partial}{\partial y\pm}B_y\left(\bigcup_{j=1}^{k_i}\mathfrak{q}_{ij}\right),-\frac{\partial}{\partial y\pm}B_y\left(\bigcup_{j=k_i+1}^{n_i}\mathfrak{q}_{ij}\right)\right\}.$$

Since $B_y(\mathfrak{q})$ is well-behaved,

$$\left|\frac{\partial}{\partial y\pm}B_y(\mathfrak{q})\right|\leqslant c,\quad \mathfrak{q}\in\Sigma_B,$$

and applying this to the previous inequality we get

$$\left|\frac{\partial}{\partial y\pm}f_i(y)\right|\leqslant c\varepsilon^{i-1}.$$

Since the sequence $\{c\varepsilon^{i-1}\}_{i=1}^{\infty}$ is summable, we invoke Lemma A.1 (with M being the counting measure on $r=1,2,\dots$) to obtain

$$\frac{\partial}{\partial y\pm}B^*(p_0,y)=\frac{\partial}{\partial y\pm}\sum_{i=1}^{\infty}f_i(y)=\sum_{i=1}^{\infty}\frac{\partial}{\partial y\pm}f_i(y)$$

with

$$\left|\frac{\partial}{\partial y\pm}B^*(p_0,y)\right|\leqslant c\sum_{i=1}^{\infty}\varepsilon^{i-1}=\frac{c}{1-\varepsilon}.$$

Since ε is an arbitrary number between in $(0,1)$, the last inequality should hold for any such number, whence it follows that

$$\left|\frac{\partial}{\partial y\pm}B^*(p_0,y)\right|\leqslant c.$$

Repeating this derivation for all $p \in \Re$, we will have established that $\frac{\partial}{\partial y\pm} B^*(p, y)$ exist and are bounded by c on $\Re \times [-a, a]$.

Part 2: Continuity: This part is essentially identical to the continuity part of the proof of Theorem 8.1 in Dzhafarov (2003). Returning to the opening equation of the present proof, we can invoke Lemma A.1 once again (this time with M being A_x) to obtain

$$\frac{\partial}{\partial y\pm} \psi(x, y) = \int_{p \in \Re} \frac{\partial}{\partial y\pm} B^*(p, y) dA_x(p).$$

Being the limits of A-measurable functions

$$\frac{B^*(p, y \pm \delta) - B^*(p, y)}{\delta}, \quad \delta \to 0+,$$

$\frac{\partial}{\partial y\pm} B^*(p, y)$ are A-measurable. Fix y, and rewrite, for simplicity, $\frac{\partial}{\partial y\pm} B^*(p, y)$ as $b(p)$. Since $b(p)$ is A-measurable and bounded, there is (see, e.g., Hewitt & Stromberg, 1965, pp. 172–173) a sequence of A-measurable simple functions

$$\eta_i(p) = \sum_{j=1}^{n_i} a_{ij} \chi_{\mathfrak{p}_{ij}}(p),$$

with $\chi_{\mathfrak{p}_{ij}}$ being characteristic functions of pairwise disjoint A-measurable sets $\mathfrak{p}_{ij}$, such that $\varphi_i(p)$ converges to $b(p)$ uniformly. This means that, for some function $n(\varepsilon)$,

$$b(p) - \varepsilon \leqslant \eta_i(p) \leqslant b(p) + \varepsilon$$

for all $i > n(\varepsilon)$. But then, for any $x \in [-a, a]$,

$$\int_{p \in \Re} [b(p) - \varepsilon] dA_x(p) \leqslant \int_{p \in \Re} \eta_i(p) dA_x(p)$$

$$\leqslant \int_{p \in \Re} [b(p) + \varepsilon] dA_x(p),$$

which is equivalent to

$$\int_{p \in \Re} b(p) dA_x(p) - \varepsilon \leqslant \int_{p \in \Re} \eta_i(p) dA_x(p)$$

$$\leqslant \int_{p \in \Re} b(p) dA_x(p) + \varepsilon.$$

By the construction logic of Lebesgue integrals,

$$\int_{p \in \Re} b(p) dA_x(p) = \lim_{i \to \infty} \int_{p \in \Re} \eta_i(p) dA_x(p),$$

and the inequalities above indicate that this convergence is uniform on $x \in [-a, a]$. Now,

$$\int_{p \in \Re} \eta_i(p) dA_x(p) = \sum_{j=1}^{n_i} \int_{p \in \Re} a_{ij} \chi_{\mathfrak{p}_{ij}}(p) dA_x(p)$$

$$= \sum_{j=1}^{n_i} a_{ij} A_x(\mathfrak{p}_{ij}),$$

which is continuous in x, because so is $A_x(\mathfrak{p})$ for any A-measurable $\mathfrak{p}$. The limit of uniformly converging continuous functions being continuous, we have proved that $\frac{\partial}{\partial y\pm} \psi(x, y)$ are continuous in x. $\square$

The following immediate consequence of this theorem (extracted from the existence part of its proof) is made use of later, in the context of Thurstonian-type representations of the Probabilistic-Interdependent variety.

Corollary 5.1. $\left|\frac{\partial}{\partial x\pm} \psi(x, y)\right| \leqslant c, \left|\frac{\partial}{\partial y\pm} \psi(x, y)\right| \leqslant c.$

Proof. From

$$\frac{\partial}{\partial y\pm} \psi(x, y) = \int_{p \in \Re} \frac{\partial}{\partial y\pm} B^*(p, y) dA_x(p)$$

and

$$\left|\frac{\partial}{\partial y\pm} B^*(p_0, y)\right| \leqslant c,$$

it follows that

$$\left|\frac{\partial}{\partial y\pm} \psi(x, y)\right| \leqslant \int_{p \in \Re} c dA_x(p) = c.$$

This completes the proof. $\square$

Relating Theorem 5.1 to Theorem 3.1, we arrive at the following generalization of the main conclusion formulated at the end of Section 3.

Main Conclusion (*For the Probabilistic-Independent variety*). A typical patch $\psi(x, y)$ (satisfying the regular minimality and nonconstant self-similarity conditions (2) and (3)) does not have a well-behaved Thurstonian-type representation of the Probabilistic-Independent variety. As a result, no discrimination probability function $\psi(\mathbf{x}, \mathbf{y})$ with regular minima and nonconstant self-similarity allows for a Thurstonian-type representation of the Probabilistic-Independent variety that is well-behaved at any of the typical line elements $(\mathbf{s}, \mathbf{u})$.

6. Notion of selective influence

We turn now to the analysis of the possibility that the random images $P(x)$ and $Q(y)$ of the stimuli x and y are not stochastically independent. As pointed out in the Introduction, it is considered an inherent property of any Thurstonian-type model that $P(x)$ and $Q(y)$ can be selectively attributed to (or, are selectively influenced by) the respective stimuli x and y. This selectiveness provides a justification for writing $P = P(x)$ and calling it an image of x. The models in which (x, y) as a pair is mapped into a single image $R(x, y)$, even if this image itself can be presented as a pair, $R(x, y) = (P(x, y), Q(x, y))$, are not included in the class of Thurstonian-type models. In particular, I do not include in this class the models like that of Takane and Sergent

(1983), in which every pair of stimuli being presented evokes a single random variable interpretable as the "subjective difference" between the two stimuli.

The question, posed in the Introduction, of how one should understand the selectiveness in the influence of x and y upon, respectively, random variables $P(x)$ and $Q(y)$ when the latter are stochastically interdependent, has been primarily discussed in the literature in the context of information processing architectures and response time decompositions (Dzhafarov, 1992, 1997; Dzhafarov & Rouder, 1996; Dzhafarov & Schweickert, 1995; Townsend, 1984; Townsend & Schweickert, 1989; Townsend & Thomas, 1994). $P(x)$ and $Q(y)$ in this context would typically represent durations of two processes believed to be selectively influenced by two distinct "factors", x and y.

Historically, the first general solution for the problem of selective influence under stochastic interdependence is proposed in Townsend (1984). Its mathematical theory is given in Dzhafarov (1999). This solution, however, is not suitable in the present context, as it is generally incompatible with the following *marginal selectivity* condition (Townsend & Schweickert, 1989) in the dependence of (P, Q) on (x, y): the marginal distribution of P depends on x but not on y, while the marginal distribution of Q depends on y but not on x. As argued in Dzhafarov (2001, in press), this condition should be viewed as necessary but not sufficient for (P, Q) being selectively influenced by (x, y), respectively.

The understanding of Thurstonian-type representations with interdependent images adopted in this paper is based on the theory of selective influence proposed in Dzhafarov (in press). According to this theory, random variables (P, Q) are selectively influenced by stimuli (x, y), respectively (equivalently, P and Q are random images *of* x and y, respectively) if they can be presented as

$$P = f(R, R_P, x), \quad Q = g(R, R_Q, y), \tag{10}$$

where (R, R_P, R_Q) are mutually independent random variables whose distributions do not depend on x or y, and f, g are arbitrary measurable functions. This representation provides an "explanation" for why the P and Q are stochastically interdependent (both f and g depend on a common variable R), and why they are nevertheless selectively attributed to x and y (f does not depend on y, and g does not depend on x).

All applicable restrictions on possible distributions of R, R_P, R_Q are immaterial for the present discussion. Let the random variable R (called the "common source of randomness" for P and Q) take on its values in some space $\mathfrak{C}$ and be associated with (represented by) a probability measure C defined on some sigma-algebra Σ_C of subsets of $\mathfrak{C}$,

$$C(\mathfrak{r}) = \Pr[R \in \mathfrak{r}], \quad \mathfrak{r} \in \Sigma_C.$$

Remark 6.1. It should be noted at this point that the term "random image" (or, more general, "random variable") in this paper, as well as in Dzhafarov (2003), is not used in the standard "Kolmogorovian" sense, as a real-valued measurable function on a sample space. Rather a random variable is taken to be the *identity function* (trivially measurable) from a sample space onto itself. Thus understood it is a logically redundant notion, used only because it is more intuitive to speak of a random variable taking on some value or falling within some subset than to speak of this value or subset as "occurring". In particular, the random variable R just introduced is the identity function from some set $\mathfrak{C}$ onto $\mathfrak{C}$, associated with a probability measure C defined on a sigma-algebra Σ_C. The values of R, therefore, are simply elements of $\mathfrak{C}$, and the term "distribution of R" is understood as synonymous with the measure function $C(\mathfrak{r})$, $\mathfrak{r} \in \Sigma_C$. The random images $P(x)$ and $Q(y)$ are understood analogously, the arguments x and y indicating that the distributions (= measures) $A_x(\mathfrak{p})$ and $B_y(\mathfrak{q})$ depend on, respectively, x and y.

Representation (10) is not explicitly used in the subsequent development. Rather the development is based on the following obvious consequence of (10): conditional upon any value r of the "common source of randomness" R, the random variables P, Q are mutually independent, with their distributions depending on x and y, respectively. As shown in Dzhafarov (in press), this consequence is in fact equivalent to the representability of P, Q by (10). In other words, one can selectively attribute P, Q to, respectively, x, y and write

$$P = P(x), \quad Q = Q(y)$$

if and only if one can find a random variable R conditioned upon whose values P, Q are mutually independent, with their distributions depending on x and y, respectively. In this case one can define *conditional probability measures* $A_{x,r}(\mathfrak{p})$, with $\mathfrak{p} \in \Sigma_A$, and $B_{y,r}(\mathfrak{q})$, with $\mathfrak{q} \in \Sigma_B$, such that

$$\Pr[P \in \mathfrak{p}, Q \in \mathfrak{q}] = \int_{r \in \mathfrak{C}} A_{x,r}(\mathfrak{p}) B_{y,r}(\mathfrak{q}) dC(r), \tag{11}$$

and (as a consequence)

$$A_x(\mathfrak{p}) = \Pr[P \in \mathfrak{p}] = \int_{r \in \mathfrak{C}} A_{x,r}(\mathfrak{p}) dC(r),$$

$$B_y(\mathfrak{q}) = \Pr[Q \in \mathfrak{q}] = \int_{r \in \mathfrak{C}} B_{y,r}(\mathfrak{q}) dC(r).$$

These three equations lay the foundation for the construction below.

Remark 6.2. As shown in Dzhafarov (in press), the notion of selective influence (attribution) is restrictive: not every joint distribution $\Pr[P \in \mathfrak{p}, Q \in \mathfrak{q}]$ with marginal selectivity allows for representation (11).

7. Thurstonian-type representations of Probabilistic-Interdependent variety

A Thurstonian-type representation of the Probabilistic-Interdependent variety (or *general* Thurstonian-type representation) for a patch $\psi(x, y)$, $(x, y) \in [-a, a]^2$, of a discrimination probability function $\psi(\mathbf{x}, \mathbf{y})$ is defined by the construct

$$\{\mathfrak{R}, \mathfrak{C}, A_{x,r}, B_{y,r}, C, \sigma\}, \tag{12}$$

with the following meaning of the terms:

(i) The perceptual space $\mathfrak{R}$ and the decision probability function $\sigma(p, q)$ are understood in the same way as for the Probabilistic-Independent variety (Section 4).

(ii) C is a probability measure associated with a *conditioning* random variable R taking on its values in some space $\mathfrak{C}$,

$$C(\mathfrak{r}) = \Pr[R \in \mathfrak{r}], \quad \mathfrak{r} \in \Sigma_C,$$

where Σ_C is some sigma-algebra of subsets of $\mathfrak{C}$.

(iii) $A_{x,r}(\mathfrak{p}), B_{y,r}(\mathfrak{q})$ are probability measures conditional upon values r of R and defined on respective sigma-algebras Σ_A, Σ_B (independent of r and of x, y),

$$A_{x,r}(\mathfrak{p}) = \Pr[P(x) \in \mathfrak{p} | R = r], \quad \mathfrak{p} \in \Sigma_A, \quad r \in \mathfrak{C}, \quad x \in [-a, a],$$

$$B_{y,r}(\mathfrak{q}) = \Pr[Q(y) \in \mathfrak{q} | R = r], \quad \mathfrak{q} \in \Sigma_B, \quad r \in \mathfrak{C}, \quad y \in [-a, a].$$

Remark 7.1. The conditional measures $A_{x,r}(\mathfrak{p}), B_{y,r}(\mathfrak{q})$ may be allowed to be undefined on, respectively, subsets $\mathfrak{c}_A(x)$ and $\mathfrak{c}_B(y)$ of $\mathfrak{C}$ such that the C-measure of $\bigcup_{x \in [-a,a]} \mathfrak{c}_A(x) \cup \bigcup_{y \in [-a,a]} \mathfrak{c}_B(y)$ is zero. This exceptional set, however, need not be mentioned because it is always possible to additionally define the conditional measures on this set in an arbitrary fashion, making them, in particular, to comply with the definition of well-behavedness given below (e.g., by making them independent of x and y on this set).

(iv) the relationship between $\psi(x, y)$ and $\{\mathfrak{R}, \mathfrak{C}, \mathfrak{A}_{x,r}, B_{y,r}, C, \sigma\}$ is given by

$$\psi(x, y) = \int_{r \in \mathfrak{C}} \left[\int_{(p,q) \in \mathfrak{R}^2} \sigma(p, q)\, dA_{x,r}(p) dB_{y,r}(q) \right] dC(r). \tag{13}$$

Remark 7.2. The Probabilistic-Independent variety of Thurstonian-type representations is obtained as a special case, by putting

$$A_{\mathbf{x},r} \equiv A_{\mathbf{x}}, \quad B_{\mathbf{y},r} \equiv B_{\mathbf{y}},$$

in which case, irrespective of C, (13) reduces to (7).

Remark 7.3. The Deterministic-Interdependent variety of Thurstonian-type representations is obtained as a special case, by putting

$$\sigma(p, q) = \chi_{\mathfrak{S}}(p, q) = \begin{cases} 1 & \text{if } (p, q) \in \mathfrak{S}, \\ 0 & \text{if } (p, q) \notin \mathfrak{S}. \end{cases}$$

This case is not considered separately, because its treatment is not any simpler than that of the general Probabilistic-Interdependent case.

Unlike for the Probabilistic-Independent variety, Definition 3.1 of well-behavedness can no longer be adopted verbatim, but its modification suggests itself trivially.

Definition 7.1. The conditional probability measure $A_{x,r}(\mathfrak{p})$ is called *well-behaved* (in the absolute, or narrow sense) if $\frac{\partial}{\partial x+} A_{x,r}(\mathfrak{p})$ and $\frac{\partial}{\partial x-} A_{x,r}(\mathfrak{p})$ exist and are bounded by a constant c for all $(\mathfrak{p}, x, r) \in \Sigma_A \times [-a, a] \times \mathfrak{C}$, that is,

$$\left| \frac{\partial}{\partial x \pm} A_{x,r}(\mathfrak{p}) \right| < c.$$

The well-behavedness of $B_{y,r}(\mathfrak{q})$ is defined analogously, and a Thurstonian-type representation of the Probabilistic-Interdependent variety for a patch $\psi(x, y)$, $(x, y) \in [-a, a]^2$, is called *well-behaved* (in the absolute, or narrow sense) if both $A_{x,r}(\mathfrak{p})$ and $B_{y,r}(\mathfrak{q})$ in it are well-behaved.

In the same way as for the Probabilistic-Independent or Deterministic-Independent varieties, Thurstonian-type representations of the Probabilistic-Interdependent variety can also be defined for the entire discrimination probability function $\psi(\mathbf{x}, \mathbf{y})$ rather than for its patches, by replacing $A_{x,r}(\mathfrak{p})$ and $B_{y,r}(\mathfrak{q})$ with, respectively, $A_{\mathbf{x},r}(\mathfrak{p})$ ($\mathfrak{p} \in \Sigma_A$, $r \in \mathfrak{C}$, $\mathbf{x} \in \mathfrak{M}$) and $B_{\mathbf{y},r}(\mathfrak{q})$ ($\mathfrak{q} \in \Sigma_B, r \in \mathfrak{C}$, $\mathbf{y} \in \mathfrak{M}$). Relationship (13) then transforms into

$$\psi(\mathbf{x}, \mathbf{y}) = \int_{r \in \mathfrak{C}} \int_{(p,q) \in \mathfrak{R}^2} \sigma(p, q)\, dA_{\mathbf{x},r}(p)\, dB_{\mathbf{y},r}(q)\, dC(r).$$

Again, this relationship cannot hold true if (13) fails to hold for some of the patches $\psi(x, y)$, and this justifies our focusing on (13): it is shown in the next section that this representation cannot hold for any typical patch $\psi(x,y)$.

Remark 7.4. For the discrete finite example $\mathfrak{R} = \{1,\ldots,k\}$ considered in Remark 4.2, we have

$$\begin{cases} dA_{x,r}(p) = \alpha(i|r,x) = \Pr[P(x) = i|R = r], \\ dB_{y,r}(q) = \beta(i|r,y) = \Pr[Q(y) = i|R = r], \end{cases} \quad i = 1,\ldots,k,$$

relation (13) acquires the form

$$\psi(x,y) = \int_{r\in\mathfrak{C}} \sum_{i=1}^{k} \sum_{j=1}^{k} \sigma(i,j)\alpha(i|r,x)\beta(j|r,y)dC(r),$$

and the well-behavedness means that $\frac{\partial}{\partial x\pm}\alpha(i|r,x)$ and $\frac{\partial}{\partial y\pm}\beta(i|r,y)$ exist and are bounded for all $i = 1,\ldots,k$, all $r\in\mathfrak{C}$, and all $(x,y)\in[-a,a]^2$. Note that even in this simple case, with $\mathfrak{R} = \{1,\ldots,k\}$, the random variable R may take on its values in an arbitrary space $\mathfrak{C}$.

8. General Near-Smoothness theorem

The proof of the theorem below is remarkably short and simple. This is explained by the fact that all of its complexity has been absorbed by the notion of selective influence and Theorem 5.1 with its corollary for the Probabilistic-Independent variety.

Theorem 8.1. *A patch $\psi(x,y)$ that has a well-behaved general Thurstonian-type representation is near-smooth.*

Proof. Denote

$$\psi_r(x,y) = \int_{(p,q)\in\mathfrak{R}^2} \sigma(p,q)dA_{x,r}(p)dB_{y,r}(q).$$

By Theorem 5.1, $\frac{\partial}{\partial y\pm}\psi_r(x,y)$ exist, are continuous in x, and (by Corollary 5.1) dominated by a constant c. Applying Lemmas A.1 and A.2 to

$$\psi(x,y) = \int_{r\in\mathfrak{C}} \psi_r(x,y)dC(r),$$

we get

$$\frac{\partial}{\partial y\pm}\psi(x,y) = \int_{r\in\mathfrak{C}} \frac{\partial}{\partial y\pm}\psi_r(x,y)dM(r)$$

and $\frac{\partial}{\partial y\pm}\psi(x,y)$ are continuous in x. That $\frac{\partial}{\partial x\pm}\psi(x,y)$ exist and are continuous in y is proved analogously. □

Relating this result to Theorem 3.1, we arrive at our final conclusion.

Main Conclusion (**general**). A typical patch $\psi(x,y)$ (satisfying the regular minimality and nonconstant self-similarity conditions (2) and (3)) does not have a well-behaved Thurstonian-type representation (of any variety). As a result, no discrimination probability function $\psi(\mathbf{x},\mathbf{y})$ with regular minima and nonconstant self-similarity allows for a Thurstonian-type representation (of any variety) that is well-behaved at any of the typical line elements $(\mathbf{s},\mathbf{u})$.

9. "Relativization" of well-behavedness

The logic by which the definition of well-behavedness can be relaxed without affecting the validity of Theorems 5.1 and 8.1 is the same as for the Deterministic-Independent variety of Thurstonian-type representations (see Dzhafarov, 2003). Due to this fact the generalized definitions below are presented without much elaboration. Also, they are only presented for the most general, Probabilistic-Interdependent variety of Thurstonian-type representations. Their specialization to the Probabilistic-Independent variety is obtained by "crossing out" all references to the conditioning measure C and its space $\mathfrak{C}$.

Recall that the proof of Theorem 8.1 is based on the proof of Theorem 5.1. On inspecting these proofs one observes that the requirement that $\frac{\partial}{\partial x\pm}A_{x,r}(\mathfrak{p})$ and $\frac{\partial}{\partial y\pm}B_{y,r}(\mathfrak{q})$ exist and be bounded on, respectively, $(\mathfrak{p},x,r)\in\Sigma_A\times[-a,a]\times\mathfrak{C}$ and $(\mathfrak{q},y,r)\in\Sigma_B\times[-a,a]\times\mathfrak{C}$, is only needed to the extent it implies a significantly weaker requirement, the existence and boundedness of

$$D_A^{\pm}(q,x,r) = \frac{\partial}{\partial x\pm}\int_{p\in\mathfrak{R}} \sigma(p,q)\,dA_{x,r}(p), \quad (q,x,r)\in\mathfrak{R}\times[-a,a]\times\mathfrak{C},$$

$$D_B^{\pm}(p,y,r) = \frac{\partial}{\partial y\pm}\int_{q\in\mathfrak{R}} \sigma(p,q)\,dB_{y,r}(q), \quad (p,y,r)\in\mathfrak{R}\times[-a,a]\times\mathfrak{C}.$$

This requirement is weaker because these integrals contain a specific function $\sigma(p,q)$ rather than the characteristic functions $\chi_{\mathfrak{p}}(p)$ and $\chi_{\mathfrak{q}}(q)$ for all possible $\mathfrak{p}\in\Sigma_A$ and $\mathfrak{q}\in\Sigma_B$. This observation leads to the following relaxation of Definition 7.1.

Definition 9.1. The conditional probability measure $A_{x,r}(\mathfrak{p})$ is *well-behaved with respect to* $\sigma(p,q)$ (a decision probability function) if the left- and right-hand derivatives

$$D_A^{\pm}(q,x,r) = \frac{\partial}{\partial x\pm}\int_{p\in\mathfrak{R}} \sigma(p,q)\,dA_{x,r}(p)$$

exist and are bounded on $(q,x,r)\in\mathfrak{R}\times[-a,a]\times\mathfrak{C}$.

The well-behavedness of $B_{y,r}(\mathfrak{q})$ with respect to $\sigma(p,q)$ is defined analogously.

A Thurstonian-type representation of the Probabilistic-Interdependent variety is called *well-behaved with respect to* $\sigma(p,q)$ if both $A_{x,r}(\mathfrak{p})$ and $B_{y,r}(\mathfrak{q})$ in it are well-behaved with respect to $\sigma(p,q)$.

The most general definition of well-behavedness obtainable by analyzing the proofs is also the least intuitive. It is stated here for completeness only and without any details of the argument leading to it. The interested reader can reconstruct this argument by appropriately adapting the analogous argument presented in Dzhafarov (2003) when constructing the notion of well-behavedness in the relative sense for Thurstonian-type representations of the Deterministic-Independent variety.

Definition 9.2. Given a Thurstonian-type representation $\{\mathfrak{R}, \mathfrak{C}, A_{x,r}, B_{y,r}, C, \sigma\}$ for a patch $\psi(x, y)$, $(x, y) \in [-a, a]^2$, the conditional probability measure $A_{x,r}(\mathfrak{p})$ is *well-behaved with respect to* $(\sigma, B_{y,r}, C)$ if

(i) for all $(q, x, r) \in \mathfrak{R} \times [-a, a] \times \mathfrak{C}$, there exist

$$D_A^{\pm}(q, x, r) = \frac{\partial}{\partial x\pm} \int_{p \in \mathfrak{R}} \sigma(p, q)\, dA_{x,r}(p);$$

(ii) for some function $c(x, r) \geqslant 0$ and for all $q \in \mathfrak{R}$,

$$|D_A^{\pm}(q, x, r)| \leqslant c(x, r);$$

(iii) for some function $g(q, r)$, and for all $(q, x, r) \in \mathfrak{R} \times [-a, a] \times \mathfrak{C}$,

$$|D_A^{\pm}(q, x, r)| \leqslant g(q, r),$$
$$\int_{r \in \mathfrak{C}} \int_{q \in \mathfrak{R}} g(q, r) dB_{y,r}(q) < \infty.$$

The well-behavedness of $B_{y,r}(\mathfrak{q})$ with respect to $(\sigma, A_{x,r}, C)$ is defined in a symmetrical fashion.

A Thurstonian-type representation $\{\mathfrak{R}, \mathfrak{C}, A_{x,r}, B_{y,r}, C, \sigma\}$ for a patch $\psi(x, y)$, $(x, y) \in [-a, a]^2$, is *well-behaved in the relative* (*or broad*) *sense* if $A_{x,r}(\mathfrak{p})$ is well-behaved with respect to $(\sigma, B_{y,r}, C)$ and $B_{y,r}(\mathfrak{q})$ is well-behaved with respect to $(\sigma, A_{x,r}, C)$.

Definitions intermediate between the two given in this section are possible but need not be discussed. The point being made is that even though the definition of well-behavedness in the absolute sense seems innocuous as it is (see Dzhafarov, 2003, for a comprehensive justification), the near-smoothness theorems and the ensuing inadequacy of well-behaved Thurstonian-type representations remain valid under even weaker assumptions. Moreover, since the definition of well-behavedness in the relative sense has been derived by inspecting specific proofs, its assumptions are still merely sufficient for the validity of our main conclusion, rather than both sufficient and necessary.

10. Conclusion

10.1. Summary and general remarks

This is what we know about discrimination probabilities and their Thurstonian-type representations.

1. As shown in Dzhafarov (2003), if the probability measures associated with random images of stimuli are not constrained in any way (in particular, allowed to be singular), a Thurstonian-type representation can be found for any discrimination probability function $\psi(\mathbf{x}, \mathbf{y})$. Moreover, this representation can be found within the most restrictive, Deterministic-Independent variety of Thurstonian-type models. Thus the general idea of the Thurstonian-type representability for $\psi(\mathbf{x}, \mathbf{y})$ (even if one confines oneself to stochastically independent perceptual images and deterministic decisions) is not a falsifiable assumption, but rather a theoretical language providing an alternative description for $\psi(\mathbf{x}, \mathbf{y})$.
2. If, however, the probability measures in question are constrained to be "well-behaved" (Definitions 3.1, 7.1, 9.1, 9.2), then even the most general, Probabilistic-Interdependent variety of Thurstonian-type models cannot represent a discrimination probability function $\psi(\mathbf{x}, \mathbf{y})$ subject to regular minimality and nonconstant self-similarity (Sections 3.1 and 3.2).
3. These two properties are considered fundamental for discrimination probability functions (Dzhafarov, 2002c, 2003), while the well-behavedness is a relatively weak constraint that is unlikely to be violated in realistically conceivable Thurstonian-type models designed to fit empirical data. The Thurstonian-type modeling, therefore, is a theoretical language that is not particularly well suited for dealing with "same-different" discriminations.

As argued in Dzhafarov (1993), different theoretical languages, although no empirical evidence can reject one of them in favor of another, are not necessarily equally "good": one of them may, for example, be conceptually more economic or transparent than another, or the formulation of a problem in one of them may be more suggestive of its possible solutions than its formulation in another. It seems natural to expect from a good theoretical language that if the empirical entities it is designed to describe have some robust inbuilt properties, the language should be able to depict these property easily, without much ingenuity involved and convoluted constructions used. As an example from a different area, any receiver operating characteristic (ROC) curve for detection probabilities (probability of hits versus probability of false alarms) can be generated by a standard signal detectability scheme, with two distributions on a real axis and a variable cut-off point. This scheme, therefore, is not a falsifiable model but a theoretical language providing an alternative description for an ROC curve. The list of robust inbuilt features of an ROC curve includes its being increasing, lying above the identity line, and having a nonincreasing slope. The signal detectability models would hardly enjoy their present wide acceptance if it was not easy to choose pairs of distribution functions that would generate ROC curves with these properties (quite aside from the question of how well these theoretical curves fit empirical data). By contrast, Thurstonian-type modeling of discrimination probabilities definitely fails this "easiness of capturing basic properties" test. With all due caveats, therefore, I suggest that the inadequacy of well-behaved Thurstonian-type models for discrimination probabilities should lead one to critically weigh up the soundness of the entire enterprise of Thurstonian-type modeling, with stimuli mapped into random variables.

10.2. "Uncertainty Blobs": an alternative to Thurstonian-type representations

In Dzhafarov (2003) I mention two alternatives to Thurstonian-type modeling. One consists in abandoning the idea of stochastic images representing individual stimuli, and in assuming instead that every pair of stimuli $(\mathbf{x}, \mathbf{y})$ presented for a comparison is mapped into a single random variable (or process) $R(\mathbf{x}, \mathbf{y})$ interpretable as a measure of subjective difference between the two stimuli (as it is done in the model by Takane & Sergent, 1983). The other alternative consists in abandoning random representations altogether, and considering instead a *deterministic dissimilarity function*, imposed directly on a stimulus space, such that the dissimilarity of $\mathbf{y}$ from $\mathbf{x}$ is mapped into $\psi(\mathbf{x}, \mathbf{y})$ by a fixed monotonic transformation. I conclude this paper by constructing a model illustrating the latter approach. This model simultaneously predicts the regular minimality property in its simplest form,

$$\arg\min_{\mathbf{y}} \psi(\mathbf{x}, \mathbf{y}) = \mathbf{x}, \arg\min_{\mathbf{x}} \psi(\mathbf{x}, \mathbf{y}) = \mathbf{y}, \tag{14}$$

and the nonconstant self-similarity property,

$$\psi(\mathbf{x}, \mathbf{x}) \not\equiv \text{const}.$$

A general form of regular minimality is obtained from this model by a trivial modification, mentioned later.

According to the so-called "*probability-distance hypothesis*", whose analysis is given in Dzhafarov (2002b), $\psi(\mathbf{x}, \mathbf{y})$ is an increasing function f of $D(\mathbf{x}, \mathbf{y})$, a certain ("subjective") metric imposed on the stimulus space $\mathfrak{M} \subseteq \mathrm{Re}^n$. To avoid technicalities and simplify the discussion, let $\mathfrak{M}$ form a *G-space* (space with geodesics) with respect to the metric D (Aleksandrov & Berestovskii, 1995; Busemann, 1955). This means that

(i) the notion of a length $L[\mathbf{z}(t)]$ is defined for any piecewise continuously differentiable path $\mathbf{z}(t) (0 \leqslant t \leqslant 1)$ in $\mathfrak{M}$ that connects any $\mathbf{x} = \mathbf{z}(0)$ with any $\mathbf{y} = \mathbf{z}(1)$ (for details see Dzhafarov & Colonius, 1999, 2001);
(ii) for any distinct points $\mathbf{x}, \mathbf{y}$ there is a path $\overline{\mathbf{xy}}$ (called a *geodesic*) whose length among all paths connecting $\mathbf{x}$ with $\mathbf{y}$ is minimal, and this length $L[\overline{\mathbf{xy}}]$ equals $D(\mathbf{x}, \mathbf{y})$;
(iii) for every point there is a geodesic path passing through it;
(iv) any geodesic $\overline{\mathbf{xy}}$ can be uniquely extended beyond these two points.

The probability-distance hypothesis,

$$\psi(\mathbf{x}, \mathbf{y}) = f[D(\mathbf{x}, \mathbf{y})], \tag{15}$$

does not fare any better than the well-behaved Thurstonian-type representations: although it predicts the regular minimality property (14), it also predicts

$$\psi(\mathbf{x}, \mathbf{x}) = f[D(\mathbf{x}, \mathbf{x})] = f(0) = \text{const}.$$

Consider, however, the following modification of this hypothesis. Its essence consists in replacing $D(\mathbf{x}, \mathbf{y})$ in (15) with appropriately defined dissimilarity between two "blobs", some neighborhoods of $\mathbf{x}$ and $\mathbf{y}$ whose size slowly changes with changing $\mathbf{x}$ and $\mathbf{y}$. Specifically, let each stimulus $\mathbf{x}$ be associated with an "*uncertainty distance*" $R_i(\mathbf{x})$, where $i = 1$ or 2 depending on whether $\mathbf{x}$ belongs to the first or second observation areas. Intuitively, stimulus $\mathbf{x}$ belonging to the first observation area is represented by (or mapped into) the "*uncertainty blob*" of stimuli

$$\mathfrak{B}_1(\mathbf{x}) = \{\mathbf{x}' \in \mathfrak{M}\colon D(\mathbf{x}, \mathbf{x}') \leqslant R_1(\mathbf{x})\},$$

which plays a role analogous to the random image of $\mathbf{x}$ in a Thurstonian-type model. The stimulus blob

$$\mathfrak{B}_2(\mathbf{x}) = \{\mathbf{x}' \in \mathfrak{M}\colon D(\mathbf{x}, \mathbf{x}') \leqslant R_2(\mathbf{x})\}$$

represents $\mathbf{x}$ belonging to the second observation area. The distance $D(\mathbf{x}, \mathbf{y})$ here is assumed to be symmetrical, $D(\mathbf{x}, \mathbf{y}) = D(\mathbf{y}, \mathbf{x})$ (this is not assumed in Dzhafarov (2002b) where D in (15) is treated as an *oriented* distance).

We assume that $R_1(\mathbf{x})$ and $R_2(\mathbf{x})$ change with $\mathbf{x}$ *relatively slowly*, in the following sense: for all $\mathbf{x}, \mathbf{y}$,

$$\left.\begin{cases} |R_1(\mathbf{x}) - R_1(\mathbf{y})| \\ |R_2(\mathbf{x}) - R_2(\mathbf{y})| \end{cases}\right. < D(\mathbf{x}, \mathbf{y}). \tag{16}$$

This is, essentially, a form of the *Lipschitz condition* for $R_1(\mathbf{x})$ and $R_2(\mathbf{x})$.

The *dissimilarity* $D^*(\mathfrak{X}, \mathfrak{Y})$ between two stimulus subsets $\mathfrak{X}, \mathfrak{Y}$ of $\mathfrak{M}$ is defined as

$$D^*(\mathfrak{X}, \mathfrak{Y}) = \sup_{\mathbf{a} \in \mathfrak{X}, \mathbf{b} \in \mathfrak{Y}} D(\mathbf{a}, \mathbf{b}).$$

Then (refer to Fig. 1) the dissimilarity between $\mathfrak{B}_1(\mathbf{x})$ and $\mathfrak{B}_2(\mathbf{y})$ can easily be shown to equal

$$D^*[\mathfrak{B}_1(\mathbf{x}), \mathfrak{B}_2(\mathbf{y})] = R_1(\mathbf{x}) + D(\mathbf{x}, \mathbf{y}) + R_2(\mathbf{y}). \tag{17}$$

Indeed, for any $\mathbf{a} \in \mathfrak{B}_1(\mathbf{x}), \mathbf{b} \in \mathfrak{B}_2(\mathbf{y})$,

$$D(\mathbf{a}, \mathbf{b}) \leqslant D(\mathbf{a}, \mathbf{x}) + D(\mathbf{x}, \mathbf{y}) + D(\mathbf{y}, \mathbf{b}),$$

and the maximum value for the latter expression is $R_1(\mathbf{x}) + D(\mathbf{x}, \mathbf{y}) + R_2(\mathbf{y})$. To show that a geodesic of this maximum length exists for some $\mathbf{a} \in \mathfrak{B}_1(\mathbf{x}), \mathbf{b} \in \mathfrak{B}_2(\mathbf{y})$, connect $\mathbf{x}$ and $\mathbf{y}$ by a geodesic and extend it beyond $\mathbf{x}$ and $\mathbf{y}$ until it reaches the boundaries of $\mathfrak{B}_1(\mathbf{x})$ (at a point $\mathbf{a}$) and $\mathfrak{B}_2(\mathbf{y})$ (at a point $\mathbf{b}$).

The modification of the probability-distance hypothesis being proposed is this: $\psi(\mathbf{x}, \mathbf{y})$ is an increasing function f of the dissimilarity between $\mathfrak{B}_1(\mathbf{x})$ and $\mathfrak{B}_2(\mathbf{y})$,

$$\begin{aligned}\psi(\mathbf{x}, \mathbf{y}) &= f[D^*[\mathfrak{B}_1(\mathbf{x}), \mathfrak{B}_2(\mathbf{y})]] \\ &= f[R_1(\mathbf{x}) + D(\mathbf{x}, \mathbf{y}) + R_2(\mathbf{y})].\end{aligned} \tag{18}$$

For example, using a version of the well-known Shepard's (1987) idea, one could put

$$\psi(\mathbf{x}, \mathbf{y}) = 1 - \exp\left[-R_1(\mathbf{x}) - D(\mathbf{x}, \mathbf{y}) - R_2(\mathbf{y})\right].$$

Whatever the choice of f, as an immediate consequence of (16) and (18) we have (refer to Fig. 2)

$$\psi(\mathbf{x}, \mathbf{x}) = f[R_1(\mathbf{x}) + R_2(\mathbf{x})] < \begin{cases} \psi(\mathbf{x}, \mathbf{y}) = f[R_1(\mathbf{x}) + D(\mathbf{x}, \mathbf{y}) + R_2(\mathbf{y})], \\ \psi(\mathbf{y}, \mathbf{x}) = f[R_2(\mathbf{x}) + D(\mathbf{x}, \mathbf{y}) + R_1(\mathbf{y})], \end{cases}$$

which is the property of regular minimality (14). At the same time,

$$\psi(\mathbf{x}, \mathbf{x}) = f[R_1(\mathbf{x}) + R_2(\mathbf{x})]$$

is a quantity that generally varies with $\mathbf{x}$ (nonconstant self-similarity).

To obtain the general form of regular minimality, (1), one simply has to substitute $\psi(\mathbf{x}, \mathbf{h}(\mathbf{y}))$ for $\psi(\mathbf{x}, \mathbf{y})$ in (18), with $\mathbf{h}$ being some homeomorphism $\mathfrak{M} \to \mathfrak{M}$.

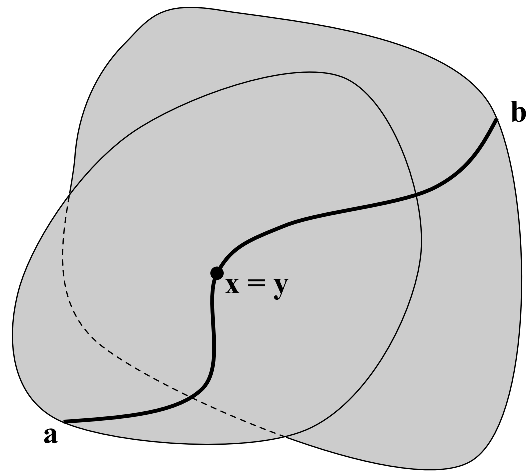

Fig. 2. The "uncertainty blobs" $\mathfrak{B}_1(\mathbf{x})$ and $\mathfrak{B}_2(\mathbf{y})$ when stimulus $\mathbf{x}$ (in the first observation area) coincides with stimulus $\mathbf{y}$ (in the second observation area). The dissimilarity between the two "blobs" is the length $R_1(\mathbf{x}) + R_2(\mathbf{y})$ of any geodesic passing through $\mathbf{x} = \mathbf{y}$ until it intersects with the boundaries of the two "blobs".

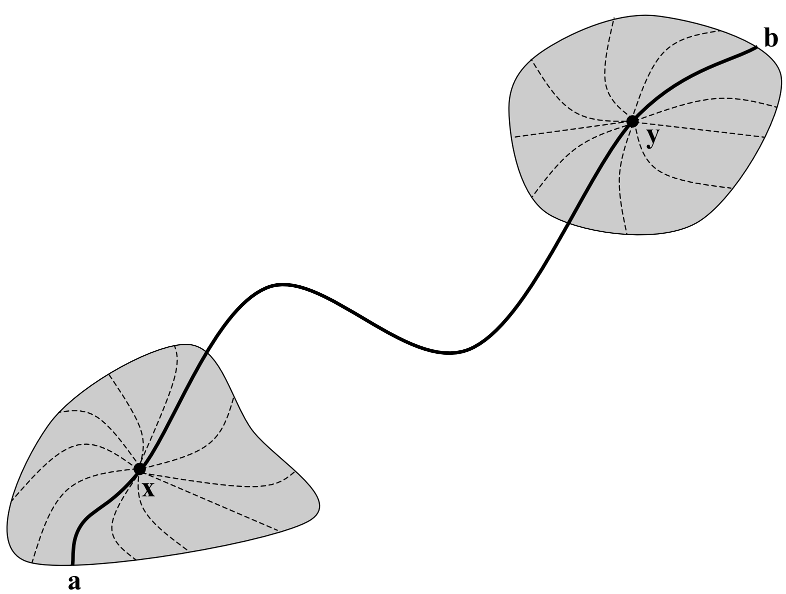

Fig. 1. The two shaded areas are "uncertainty blobs" $\mathfrak{B}_1(\mathbf{x})$ and $\mathfrak{B}_2(\mathbf{y})$ for, respectively, stimulus $\mathbf{x}$ (in the first observation area) and stimulus $\mathbf{y}$ (in the second observation area). The interrupted lines are the geodesics connecting the centers of the "blobs" with their boundary points: the length of any of these geodesic radii in $\mathfrak{B}_1(\mathbf{x})$ is $R_1(\mathbf{x})$, in $\mathfrak{B}_2(\mathbf{y})$ it is $R_2(\mathbf{y})$. The solid line $\overline{\mathbf{xy}}$ (connecting $\mathbf{x}$ with $\mathbf{y}$) is a geodesic of length $D(\mathbf{x}, \mathbf{y})$. The geodesic radii $\overline{\mathbf{xa}}$ and $\overline{\mathbf{yb}}$ are continuations of $\overline{\mathbf{xy}}$ (in a G-space every geodesic has a unique continuation at both ends). The length of the geodesic $\overline{\mathbf{axyb}}$ is $R_1(\mathbf{x}) + D(\mathbf{xy}) + R_2(\mathbf{y})$, and it is taken to measure the dissimilarity between $\mathfrak{B}_1(\mathbf{x})$ and $\mathfrak{B}_2(\mathbf{y})$.

No claim is being made here that the approach using the "uncertainty blobs" is empirically valid. My only point is to demonstrate that one can easily construct a class of models which employ no random images and no notion of a perceptual space, that are therefore conceptually more economic than Thurstonian-type representations, and that, in a sharp contrast with the latter, capture the fundamental properties of regular minimality and nonconstant self-similarity "automatically". In the least, the development of such models constitutes a viable alternative to searching for non-well-behaved Thurstonian-type representations capable of describing "same-different" discriminations.

Acknowledgments

This research has been supported by the NSF Grant SES-0001925. The author is deeply grateful to Hans Colonius, R. Duncan Luce, and A.A.J. Marley for discussions, criticism, and help.

Appendix A. Auxiliary facts

Lemma A.1. *Let*

$$F(y) = \int_{r\in\Re} f(r,y)dM(r), \quad y\in[-a,a],$$

where M is a sigma-finite measure. Let $y\to f(r,y)$ be both right-and left-differentiable on $[-a,a]$, with

$$\left|\frac{\partial}{\partial y\pm} f(r,y)\right| \leqslant g(r), \qquad \int_{r\in\Re} g(r)dM(r) = c < \infty.$$

Then $F(y)$ is both right- and left-differentiable (hence continuous) on $[-a,a]$, with

$$\frac{\partial}{\partial y\pm} F(y) = \int_{r\in\Re} \frac{d}{dx\pm} f(r,y)dM(r) \leqslant c.$$

Proof. See Lemmas A.3 and A.4 in Dzhafarov (2003). □

Lemma A.2. *Let*

$$F(x) = \int_{p\in\Re} f(p,x)\,dM(p), \quad x\in[-a,a],$$

where M is a sigma-finite measure. Let $x\to f(p,x)$ be continuous, and

$$|f(p,x)| \leqslant g(p), \qquad \int_{p\in\Re} g(p)dM(p) < \infty.$$

Then $F(x)$ is continuous.

Proof. An immediate consequence of the Lebesgue Dominated Convergence Theorem (e.g., Hewitt & Stromberg, 1965, pp. 172–173) applied to $f(p,x+\delta)\to f(p,x)$ as $\delta\to 0$. □

References

Aleksandrov, A. D., & Berestovskii, V. N. (1995). Finsler space, generalized. In M. Hazewinkel (Ed.), *Encyplopaedia of mathematics*, Vol. 4 (pp. 26–27). Dordrecht: Kluwer Academic Publishers.

Busemann, H. (1955). *The geometry of geodesics*. New York: Academic Press.

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

Dzhafarov, E. N. (1992). The structure of simple reaction time to step-function signals. *Journal of Mathematical Psychology*, *36*, 235–268.

Dzhafarov, E. N. (1993). Grice-representability of response time distribution families. *Psychometrika*, *58*, 281–314.

Dzhafarov, E. N. (1997). Process representations and decompositions of response times. In A. A. J. Marley (Ed.), *Choice, decision and measurement: Essays in honor of R. Duncan Luce* (pp. 255–278). Mahwah, NJ: Erlbaum.

Dzhafarov, E. N. (1999). Conditionally selective dependence of random variables on external factors. *Journal of Mathematical Psychology*, *43*, 123–157.

Dzhafarov, E. N. (2001). Unconditionally selective dependence of random variables on external factors. *Journal of Mathematical Psychology*, *45*, 421–451.

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: Pairwise comparisons, regular minimality, and nonconstant self-similarity. *Journal of Mathematical Psychology*, *46*, 583–608.

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

Dzhafarov, E. N. (in press). Selective influence through conditional independence, *Psychometrika*.

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., & Rouder, J. N. (1996). Empirical discriminability of two models for stochastic relationship between additive components of response time. *Journal of Mathematical Psychology*, *40*, 48–63.

Dzhafarov, E. N., & Schweickert, R. (1995). Decompositions of response times: An almost general theory. *Journal of Mathematical Psychology*, *39*, 285–314.

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

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

Hewitt, E., & Stromberg, K. (1965). *Real and abstract analysis*. New York: Springer.

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

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

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

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

Rothkopf, E. Z. (1957). A measure of stimulus similarity and errors in some paired-associate learning tasks. *Journal of Experimental Psychology*, *53*, 94–101.

Shepard, R. N. (1987). Towards a universal law of generalization for psychological sciences. *Science*, *237*, 1317–1323.

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

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

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

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

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

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

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

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

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

Townsend, J. T., & Thomas, R. D. (1994). Stochastic dependencies in parallel and serial models: Effects on systems factorial interactions. *Journal of Mathematical Psychology*, *38*, 1–34.

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

Zimmer, K., & Colonius, H. (2000). Testing a new theory of Fechnerian scaling: The case of auditory intensity discrimination. *Paper presented at the 140th Meeting of the Acoustical Society of America*.

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