4
Referential Duality and Representational Duality in the Scaling of Multidimensional and Infinite-Dimensional Stimulus Space

Jun Zhang

University of Michigan

1. INTRODUCTION

Traditional theories of geometric representations for stimulus spaces (see, e.g., Shepard, 1962a, 1962b) rely on the notion of a "distance" in some multidimensional vector space $\mathbb{R}^n$ to describe the subjective proximity between various stimuli whose features are represented by the axes of the space. Such a distance is often viewed as induced by a "norm" of the vector space, defined as a real-valued function $\mathbb{R}^n \to \mathbb{R}$ and denoted $||\cdot||$, that satisfies the following conditions for all $x, y \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$: (i) $||x|| \geq 0$ with equality holding if and only if $x = 0$; (ii) $||\alpha x|| = |\alpha| \cdot ||x||$; (iii) $||x+y|| \leq ||x|| + ||y||$. The distance measure or "metric" induced by such a norm is defined as

$$\Delta(x, x') = ||x - x'|| \ .$$

Such a metric is continuous with respect to (x, x') and it satisfies the axioms of (i) non-negativity: $\Delta(x, x') \geq 0$, with 0 attained if and only if $x = x'$; (ii) symmetry: $\Delta(x, x') = \Delta(x', x)$; and (iii) triangle inequality: $\Delta(x, x') + \Delta(x', x'') \geq \Delta(x, x'')$, for any triplet x, x', x''. The norm $||\cdot||$ may also be used to define an inner product $\langle \cdot, \cdot \rangle$:

$$\langle x, x' \rangle = \frac{1}{2}(||x + x'||^2 - ||x||^2 - ||x'||^2) \ .$$

The inner product operation on the vector space allows one to define the angle between two vectors (and hence orthogonality), which allows one to project one vector onto another (and hence onto a subspace).

However, such a norm-based approach in defining a metric (distance) is fundamentally dissatisfying because it cannot capture the asymmetry intrinsic to comparative judgments that give rise to our sense of similarity/dissimilarity of stimuli. This asymmetry arises from the different status of a fixed reference stimulus (a *referent*, for short) and a variable comparison stimulus (a *probe*, for short), for example, between a stimulus in perception and one in memory during categorization, between the current state and a goal state during planning, between the status quo payoff and the uncertainty about possible gains or losses during decision making, and so forth. A well-cited example of such asymmetry is Tversky's (1977) demonstration that Red China was judged to be more similar to North Korea than was North Korea to Red China, making the norm-based metric Δ highly questionable. As an alternative to the norm-based approach mentioned earlier, Dzhafarov and Colonius (1999, 2001) introduced instead (a generalized version of) Finslerian metric functions defined on the tangent space of a stimulus manifold. Asymmetric or "oriented" Fechnerian distances are constructed from these metric functions which are derived from the (not necessarily symmetric) discrimination probabilities. However, in recent developments (Dzhafarov & Colonius, 2005a, 2005b), the oriented Fechnerian distances are not interpreted as subjective dissimilarities and serve only as an intermediate step in computing a symmetrized "overall Fechnerian distance," which is taken to be a measure of subjective similarity (for explanations, see Chapter 2 of this volume). Therefore, the output of Dzhafarov and Colonius's framework is still a symmetric measure of subjective similarity, though its input (discrimination probability) is generally asymmetric.

In this chapter, we discuss mathematical notions that are specifically aimed at expressing the asymmetric status of a referent and a probe in a direct and natural way. One such notion is "duality," intuitively, the quality or character of being dichotomous or twofold. A related notion is "conjugacy," referring to objects that are, in some sense, inversely or oppositely related to each other. The referent and probe in a comparative judgment are dual to each other in an obvious way: dialectically speaking, neither of them can exist without the other coexisting. The referent and probe are also conjugate to each other, because the two stimuli can switch their roles if we change our frame of reference (all these notions are given precise mathematical meanings later).

The experimental paradigm to which our analysis applies can be described as follows. Two types of stimuli, one assigned the role of a comparison stimulus (probe) and the other the role of a reference stimulus (referent), are presented to the participants, who are to make some judgment about their similarity. As an example, suppose a participant is to make a same-different judgment on the "value" of two gambles, one involving a guaranteed payoff of x utility units and the other involving a probabilistic payoff in which the participant will receive either y units or 0 with fixed probability known to the participant. In the literature on gambles, x is called the "certainty equivalent" (CE) value of a probabilistic 0/y outcome (with given probabilities).

In the experimental paradigm on which we are focusing, the value of the referent is fixed whereas that of the probe is varied within a block of trials; the variation across trials for the latter can be either random or in an ascending/descending (including "staircase") order. In our example with gambles, for a given probability of receiving y units in the second gamble, the experimenter can either hold that y value fixed and have the value of certainty equivalent x change over a series of trials (which we call the *forward procedure*), or conversely, hold x fixed and have the value of y change (which we call the *reverse procedure*). Because these two procedures are conjugate to each other, the forward/reverse terminology can itself be reversed, in which case the assignment of reference and comparison status to x and y will be exchanged as well. To fix the notation and terminology, and for this purpose only, we pick one stimulus as a referent and call this the forward procedure, and we refer to the reverse procedure as the *conjugate* one.

Two stimuli, one assigned as a referent and the other as a probe, generally would invoke substantively different mental representations; the two mutually conjugate procedures which assign the referent and probe roles to the two stimuli differently, generally would invoke distinct psychological processes. Thus, in our example with gambles, it is natural to assume that the comparative process where a fixed value of x is used as a reference for the evaluation of a series of probabilistic gambles with variable payoffs y is different from the process where one gamble with probabilistic payoff y is used as a fixed reference for the evaluation of the varied values of x — the asymmetry in this scenario may reveal some fundamental difference in the participant's mental representations of risky and risk-free outcomes as well as in the underlying psychological processes dependent on their actual assigned role as a referent or a probe.

The goal of this chapter is to investigate the duality that arises from the distinct roles played by a referent and a probe in a comparative judgment, and to formulate some basic measures related to the asymmetry (dual symmetry, to be precise) in comparing a pair of stimuli. In Section 2, we investigate the principle of "regular cross-minimality" along with the property of "nonconstant self-similarity," the notions analogous to "regular minimality" and "nonconstant self-similarity" proposed by Dzhafarov (2002) in a somewhat different context (see Chapters 1 and 2 of this volume, where the latter is called "nonconstant self-dissimilarity"). A particular representation for psychometric functions is proposed, capturing the dual nature of the status of a referent and of a probe, in both forward and reverse procedures, through the use of a conjugate pair of strictly convex functions. The resulting "psychometric differentials" (a terminology borrowed from Dzhafarov and Colonius's Fechnerian scaling theory) are *bidualistic*, namely, they exhibit both the duality of assigning the referent-probe status to stimuli (*referential duality*) and the duality of selecting one of the two mutually conjugate representations (*representational duality*); we consider this to be a distinct improvement over the symmetric distance induced by a norm. In Section 3 we construct a family of dually symmetric psychometric differentials characterized as "divergence functions" indexed by real numbers; this is done both in the multidimensional setting and in the infinite-dimensional setting. Section 4 deals with the differential geometric structure induced by these divergence functions. It is demonstrated, in particular, that the referential duality and the representational duality are reflected in the family of affine connections defined on a stimulus manifold, together with one and the same Riemannian metric. The chapter closes with a brief discussion of the implications of this formulation of duality and conjugacy that combines the mathematical tools of convex analysis, function space, and differential geometry. The materials presented in this chapter, including detailed proofs for most propositions and corollaries stated herein, have previously been published elsewhere (Zhang, 2004a, 2004b).

2. DUAL SCALING BETWEEN REFERENCE AND COMPARISON STIMULUS SPACES

We consider a comparison task in which two types of stimuli are being compared with each other, one serving as a referent and the other as a probe. Let $\Psi_y(x)$ denote, in (an arbitrarily defined) *forward procedure*, a quantity monotonically related to the discrimination probability with which x, as a comparison stimulus, is judged to be dissimilar in magnitude to y, a reference stimulus. By abuse of language, we refer to $\Psi_y(x)$ as a psychometric function, although it need not be a probability function. Similarly, $\Phi_x(y)$ denotes, in the *reverse procedure*, a quantity monotonically related to the discrimination probability describing the value of y, now a comparison stimulus, is judged to be dissimilar in magnitude to the value of x, now a reference stimulus. Here and later, $x = [x^1, \cdots, x^n]$ and $y = [y_1, \cdots, y_n]$ are assumed to be (contravariant and covariant forms of) vectors comprising

certain subsets $\mathcal{X} \subseteq \mathbb{R}^n$ and $\mathcal{Y} \subseteq (\mathbb{R}^n)^*$ of some multidimensional vector space $\mathbb{R}^n$ and its dual $(\mathbb{R}^n)^*$, respectively. In this context, the dualistic assignment of the reference and comparison stimulus status to x and to y (and hence the dualism of two psychometric procedures, forward and reverse) is referred to as the *referential duality*.

2.1. Regular cross-minimality and positive diffeomorphism in stimulus mappings

By analogy with Dzhafarov's (2002) regular minimality principle, proposed in a different context, we require that $\Psi_y(x)$ and $\Phi_x(y)$ satisfy the principle of *regular cross-minimality*. The essence of this principle is as follows — *if*, corresponding to a particular value of the reference stimulus $\hat{y}$, there exists a unique value of the comparison stimulus $x = \hat{x}$ such that

$$\hat{x} = \text{argmin}_x \Psi_{\hat{y}}(x) ,$$

then when the entire procedure is reversed, that is, when $\hat{x}$ is being held fixed and y varies, the psychometric function $\Phi_{\hat{x}}(y)$ thus obtained would have its unique minimum value at $y = \hat{y}$:

$$\hat{y} = \text{argmin}_y \Phi_{\hat{x}}(y) .$$

In other words, when the reference stimulus ($\hat{y}$ for the forward procedure, $\hat{x}$ for the reverse procedure) is fixed and the comparison stimulus is varied (x in $\Psi_{\hat{y}}(x)$, y in $\Phi_{\hat{x}}(y)$), the corresponding psychometric functions indexpsychometric!function achieve their global minima at values $x = \hat{x}, y = \hat{y}$ such that

$$\Psi_{\hat{y}}(x) \geq \min_x \Psi_{\hat{y}}(x) = \Psi_{\hat{y}}(\hat{x}) ,$$
$$\Phi_{\hat{x}}(y) \geq \min_y \Phi_{\hat{x}}(y) = \Phi_{\hat{x}}(\hat{y}) .$$

A precise statement of the principle of regular cross-minimality is with reference to the existence of a pair of mutually inverse functions (see Dzhafarov & Colonius, 2005a, and Chapter 1 in this volume):

AXIOM 1 (REGULAR CROSS-MINIMALITY). *There exist functions $\psi : \mathcal{X} \to \mathcal{Y}$ and $\phi : \mathcal{Y} \to \mathcal{X}$ ($\mathcal{X} \subseteq \mathbb{R}^n$, $\mathcal{Y} \subseteq (\mathbb{R}^n)^*$) such that*

(i) $\Psi_{\hat{y}}(x) > \Psi_{\hat{y}}(\phi(\hat{y})),\ \forall x \neq \phi(\hat{y})$;
(ii) $\Phi_{\hat{x}}(y) > \Phi_{\hat{x}}(\psi(\hat{x})),\ \forall y \neq \psi(\hat{x})$;
(iii) $\phi = \psi^{-1}$.

It follows that ϕ and ψ must be bijective (one-to-one and onto). Below, we further assume they are sufficiently smooth and are "curl-less," therefore allowing a pair of convex functions ("potentials") to induce them.

AXIOM 2 (POSITIVE DIFFEOMORPHISM). *The mappings ψ and ϕ have symmetric and positive-definite Jacobians, with*

$$\frac{\partial \psi_i}{\partial x^j} = \frac{\partial \psi_j}{\partial x^i}\,, \qquad \frac{\partial \phi^i}{\partial y_j} = \frac{\partial \phi^j}{\partial y_i}\,, \tag{1}$$

where the subscript (or superscripts) i, j attached to the vector-valued map ψ (or ϕ) denote its i-th and j-th components.

An immediate consequence of Axiom 1 (which says that ψ, ϕ are mutually inverse functions) and Axiom 2 (which says that ψ, ϕ have symmetric, positive-definite Jacobians) is as follows:

COROLLARY 1. *There exists a pair of strictly convex functions $\Psi : \mathcal{X} \to \mathbb{R}$ and $\Phi : \mathcal{Y} \to \mathbb{R}$ such that*

(i) *they are conjugate to each other*[1]

$$\Phi^* = \Psi \longleftrightarrow \Phi = \Psi^*, \tag{2}$$

where $*$ denotes convex conjugation operation (to be explained later);

(ii) *they induce ψ, ϕ via*

$$\psi = \nabla\Psi \longleftrightarrow \phi = \nabla\Phi\,, \tag{3}$$

where $(\nabla\Psi)(x) = [\partial\Psi/\partial x^1, \cdots, \partial\Psi/\partial x^n]$, $(\nabla\Phi)(y) = [\partial\Phi/\partial y_1, \cdots, \partial\Phi/\partial y_n]$ denote the gradient of the functions Ψ and Φ, respectively.

Proof. Symmetry of the derivatives of ψ, ϕ, (1), allows us to write them in the form of (3) using some functions Ψ, Φ. Positive-definiteness further implies that Ψ, Φ are strictly convex. That $\phi = \psi^{-1}$ (from Axiom 1), or equivalently $(\nabla\Phi) = (\nabla\Psi)^{-1}$, implies (2) (apart from a constant). ◇

Note that in the unidimensional case, ψ, ϕ are simply mappings of reals to reals and hence (1) is naturally satisfied for strictly increasing (that is, order-preserving) functions ψ and ϕ.

When the mappings $\psi = \phi^{-1} \leftrightarrow \phi = \psi^{-1}$ between the two stimulus spaces $\mathcal{X}$ and $\mathcal{Y}$ are associated with a pair of conjugate convex functions $\Psi = \Phi^*$ and $\Phi = \Psi^*$, we say that the comparison stimulus and the reference stimulus are *conjugate-scaled*. The Jacobians of the mappings ψ and ϕ between two conjugate-scaled stimulus spaces are simply

$$\frac{\partial^2\Psi}{\partial x^i \partial x^j} \quad \text{and} \quad \frac{\partial^2\Phi}{\partial y_i \partial y_j}\,,$$

[1] Here and throughout this chapter, the $\leftrightarrow$ sign (or $\longleftrightarrow$ if in a displayed equation) is to be read as "and equivalently," so that "equality A $\leftrightarrow$ equality B" means that "equality A holds and equivalently equality B holds as well."

which can be shown to be matrix inverses of each other. Two conjugate-scaled representations are dual to each other; we call this *representational duality.*

Now we explain the meaning of the conjugation operation in convex analysis (see, e.g., Roberts & Varberg, 1973). A function $\Psi : \mathcal{X} \subseteq \mathbb{R}^n \to \mathbb{R}$ is called strictly convex if for $x \neq x'$ and any $\lambda \in (0, 1)$,

$$(1-\lambda)\Psi(x) + \lambda\Psi(x') > \Psi((1-\lambda)x + \lambda x')\,. \tag{4}$$

An equality sign replaces the inequality sign shown earlier if and only if $x = x'$ when $\lambda \in (0, 1)$, or if $\lambda \in \{0, 1\}$. For a strictly convex function Ψ, its conjugate $\Psi^* : \mathcal{Y} \subseteq (\mathbb{R}^n)^* \to \mathbb{R}$ is defined as

$$\Psi^*(y) \equiv \langle(\nabla\Psi)^{-1}(y), y\rangle - \Psi((\nabla\Psi)^{-1}(y))\,, \tag{5}$$

where $\langle x, y\rangle$ is the (Euclidean) inner product of two vectors $x \in \mathcal{X}$, $y \in \mathcal{Y}$ defined as

$$\langle x, y\rangle = \sum_{i=1}^{n} x^i y_i\,,$$

which is a bilinear form mapping $\mathcal{X} \times \mathcal{Y} \to \mathbb{R}$. It can be shown that Ψ^* is also a strictly convex function, with

$$\nabla\Psi^* = (\nabla\Psi)^{-1}$$

and

$$(\Psi^*)^* = \Psi\,.$$

The convex conjugation operation is associated with a pair of dual vector spaces $\mathcal{X}$ and $\mathcal{Y}$. Substituting $y = (\nabla\Psi)(x) \leftrightarrow x = (\nabla\Psi^*)(y)$ into (5) yields the relationship

$$\Psi(x) + \Psi^*(\nabla\Psi(x)) - \langle x, \nabla\Psi(x)\rangle = 0 \tag{6}$$

between a convex function $\Psi(\cdot)$ and its convex conjugate $\Psi^*(\cdot)$, called the *Fenchel* duality in convex analysis (see Rockafellar, 1970).

2.2. Psychometric differential and reference-representation biduality

Axiom 1 allows us to introduce a non-negative quantity, called the "psychometric differential." For each of the two psychometric functions, let $\mathcal{A}_\Psi(\cdot,\cdot) : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}_+$ and $\mathcal{A}_\Phi(\cdot,\cdot) : \mathcal{Y} \times \mathcal{X} \to \mathbb{R}_+$ (where $\mathbb{R}_+ = \mathbb{R}^+ \cup \{0\}$) denote:

$$\mathcal{A}_\Psi(x, \hat{y}) = \Psi_{\hat{y}}(x) - \Psi_{\hat{y}}(\phi(\hat{y}))\,, \tag{7}$$

$$\mathcal{A}_\Phi(y, \hat{x}) = \Phi_{\hat{x}}(y) - \Phi_{\hat{x}}(\psi(\hat{x}))\,. \tag{8}$$

Psychometric differential is a more interesting quantity to study than the psychometric functions themselves. This is because the two psychometric functions $\Psi_{\hat{y}}(x)$ and $\Phi_{\hat{x}}(y)$, for the forward and reverse procedures respectively, being only monotonically related to the discrimination probabilities, can always contain an additive function of the reference stimulus value (denoted with the "hat"), say, $P(\hat{y})$ in the former case and $Q(\hat{x})$ in the latter case, so that their self-similarity, that is, $\Psi_{\hat{y}}(\phi(\hat{y}))$ and $\Phi_{\hat{x}}(\psi(\hat{x}))$, need not be constant but may be a function of the reference stimulus value. Stated in another way, the property of "nonconstant self-similarity" does not impose any additional constraints on the possible forms of the two psychometric functions; in this respect, the situation is very different from that in Dzhafarov and Colonius's theory, where the combination of regular minimality and nonconstant self-(dis)similarity is shown to greatly restrict the possible forms of the (single) discrimination probability function. However, not to be too unconstrained, we impose a further restriction on the psychometric differential (and indirectly on the psychometric functions).

AXIOM 3 (REFERENCE-REPRESENTATION BIDUALITY). *The two psychometric differentials as defined in (7) and (8) satisfy*

$$\mathcal{A}_{\Psi}(x,y) = \mathcal{A}_{\Phi}(y,x)\ .$$

Axiom 3 postulates that the referential duality and representational duality themselves are "dual," that is, when one switches the referent-probe role assignment between two stimuli as well as the conjugate-scaled representations of these stimuli, the psychometric differential remains unchanged. In other words, the asymmetry embodied in the referent-probe status of the two stimuli is linked to the asymmetry in the scaling of these stimuli by a pair of conjugate convex functions. Axiom 3 is essential to our theory; it restricts the possible forms of psychometric differentials (and hence psychometric functions).

PROPOSITION 2. *The following form of the psychometric differentials* [2],

$$\begin{aligned} \mathcal{A}_{\Psi}(x,y) = \Psi(x) - \langle x,y\rangle + \Psi^*(y) &= \mathcal{A}_{\Psi^*}(y,x)\ , \\ \mathcal{A}_{\Phi}(y,x) = \Phi(y) - \langle x,y\rangle + \Phi^*(x) &= \mathcal{A}_{\Phi^*}(x,y)\ , \end{aligned} \tag{9}$$

where Ψ and Φ are conjugate convex functions, satisfies Axioms 1, 2 and 3.

Proof. Clearly

$$\mathcal{A}_{\Phi}(y,x) = \mathcal{A}_{\Psi^*}(y,x) = \Psi^*(y) - \langle x,y\rangle + (\Psi^*)^*(x) = \mathcal{A}_{\Psi}(x,y)\ ,$$

[2] Note that the subscripts Ψ and Φ when A was first introduced in (7) and (8) refer to the psychometric functions $\Psi_y(x)$ and $\Phi_x(y)$. In the statement of this proposition, $\Psi(\cdot)$ and $\Phi(\cdot)$ are single-variable convex functions, not to be confused with the two-variable psychometric functions. As a result of Proposition 2, we can subsequently treat the Ψ, Φ in the subscripts of A as $\Psi(\cdot), \Phi(\cdot)$.

because $(\Psi^*)^* = \Psi$. Axiom 3 is therefore satisfied. Regular cross-minimality (Axiom 1) is satisfied because

$$\hat{x} = \text{argmin}_x \Psi_{\hat{y}}(x) = \text{argmin}_x \mathcal{A}_{\Psi}(x,\hat{y}) = (\nabla\Psi)^{-1}(\hat{y}) = (\nabla\Phi)(\hat{y})$$

and

$$\hat{y} = \text{argmin}_y \Phi_{\hat{x}}(y) = \text{argmin}_y \mathcal{A}_{\Phi}(y,\hat{x}) = (\nabla\Phi)^{-1}(\hat{x}) = (\nabla\Psi)(\hat{x})$$

mutually imply each other. Because of the strict convexity of Φ and Ψ, the Jacobians of the mappings $\phi = \nabla\Phi, \psi = \nabla\Psi$ are symmetric and positive-definite (Axiom 2). ◇

Throughout the rest of the chapter, we assume that psychometric differentials $\mathcal{A}_{\Psi}(x,y)$ are representable in form (9). As such $\mathcal{A}_{\Psi}(x,y) = \mathcal{A}_{\Phi}(y,x)$ measures the difference between x, y assigned to a reference stimulus and a comparison in the forward procedure and scaled by Ψ and vice versa in the reverse procedure and scaled by Φ (the term *differential* is really a misnomer in our usage because its value need not be infinitesimally small). Because the mapping between the two spaces is homeomorphic, we can express the psychometric differential $\mathcal{A}$ in an alternative way, using functions of which both arguments are defined either in $\mathcal{X}$ alone or in $\mathcal{Y}$ alone:

$$\begin{aligned} \mathcal{D}_{\Psi}(x,\hat{x}) &= \mathcal{A}_{\Psi}(x,(\nabla\Psi)(\hat{x}))\ , \\ \mathcal{D}_{\Psi^*}(y,\hat{y}) &= \mathcal{A}_{\Psi^*}(y,(\nabla\Psi^*)(\hat{y}))\ . \end{aligned}$$

This is an analogue of the "canonical transformation" used in Dzhafarov and Colonius's theory (see Chapter 1). Writing them out explicitly,

$$\mathcal{D}_{\Psi}(x,\hat{x}) = \Psi(x) - \Psi(\hat{x}) - \langle x-\hat{x}, (\nabla\Psi)(\hat{x})\rangle\ , \tag{10}$$

$$\mathcal{D}_{\Psi^*}(y,\hat{y}) = \Psi^*(y) - \Psi^*(\hat{y}) - \langle(\nabla\Psi^*)(\hat{y}), y-\hat{y}\rangle\ . \tag{11}$$

$\mathcal{D}_{\Psi}$ (or $\mathcal{D}_{\Psi^*}$), which is the psychometric differential in an alternative form, is a measure of dissimilarity between a probe x (respectively, y) and a referent represented by $\hat{x}$ (respectively, $\hat{y}$). Loosely speaking, we say that D_{Ψ} (or D_{Ψ^*}) provides a "scaling" of stimuli in $\mathcal{X}$ (or $\mathcal{Y}$).

COROLLARY 3. *The two psychometric differentials $\mathcal{D}_{\Psi}$ and $\mathcal{D}_{\Psi^*}$ satisfy the reference-representation biduality*

$$\begin{aligned} \mathcal{D}_{\Psi}(x,\hat{x}) &= \mathcal{D}_{\Psi^*}((\nabla\Psi)(\hat{x}),(\nabla\Psi)(x))\ , \\ \mathcal{D}_{\Psi^*}(y,\hat{y}) &= \mathcal{D}_{\Psi}((\nabla\Psi^*)(\hat{y}),(\nabla\Psi^*)(y))\ . \end{aligned}$$

Proof. By straightforward application of (6) to the definition of $\mathcal{D}_{\Psi}$ (or $\mathcal{D}_{\Psi^*}$). ◇

Expressions (10) and (11) are formally identical because $(\Psi^*)^* = \Psi$. Hence, either Ψ or Ψ^* can be viewed as the "original" convex function with the other being derived by means of conjugation. Similarly, either $\mathbb{R}^n$ or $(\mathbb{R}^n)^*$ can be viewed as the "original" vector space with the other being its dual space, because $((\mathbb{R}^n)^*)^* = \mathbb{R}^n$. The function subscript in $\mathcal{D}$ specifies the stimulus space ($\mathcal{X}$ or $\mathcal{Y}$) whereas the two function arguments of $\mathcal{D}(\cdot,\cdot)$ are always occupied by (comparison stimulus, reference stimulus), in that order.

2.3. Properties of psychometric differentials

PROPOSITION 4. *The psychometric differential $\mathcal{D}_\Psi(x,x')$ satisfies the following properties:*

(i) *Non-negativity: For all $x, x' \in \mathcal{X}$,*

$$\mathcal{D}_\Psi(x,x') \geq 0$$

with equality holding if and only if $x = x'$.

(ii) *Conjugacy: For all $x, x' \in \mathcal{X}$,*

$$\mathcal{D}_\Psi(x,x') = \mathcal{D}_{\Psi^*}((\nabla\Psi)(x'), (\nabla\Psi)(x)) .$$

(iii) *Triangle (or generalized cosine) relation: For any three points $x, x', x'' \in \mathcal{X}$,*

$$\mathcal{D}_\Psi(x,x') + \mathcal{D}_\Psi(x',x'') - \mathcal{D}_\Psi(x,x'') = \langle x - x', (\nabla\Psi)(x'') - \nabla\Psi(x') \rangle .$$

(iv) *Quadrilateral relation: For any four points $x, x', x'', x''' \in \mathcal{X}$,*

$$\begin{aligned}&\mathcal{D}_\Psi(x,x') + \mathcal{D}_\Psi(x''',x'') - \mathcal{D}_\Psi(x,x'') - \mathcal{D}_\Psi(x''',x') \\ &= \langle x - x''', (\nabla\Psi)(x'') - (\nabla\Psi)(x') \rangle .\end{aligned}$$

As a special case, when $x''' = x'$ so that $\mathcal{D}_\Psi(x''',x') = 0$, the aforementioned equality reduces to the triangle relation (iii).

(v) *Dual representability: For any two points $x, x' \in \mathcal{X}$,*

$$\mathcal{D}_\Psi(x,x') = \mathcal{A}_\Psi(x, (\nabla\Psi)(x')) = \mathcal{A}_{\Psi^*}((\nabla\Psi^*)^{-1}(x'), x) .$$

This is another statement of the conjugacy relation (ii).

Proof. Parts (ii) and (v) are simply Corollary 3. The proof for parts (iii) and (iv) is through direct substitution. Part (i) is a well-known property of a strictly convex function Ψ. ◇

Note that Proposition 4 can be reformulated and proved for psychometric differentials presented in the $\mathcal{A}$-form, (9). These properties of a psychometric differential make it very different from a norm-induced metric Δ traditionally used to model dissimilarity between two stimuli (see the Introduction). The non-negativity property for Δ is retained: $\mathcal{D}_\Psi(x,x') \geq 0$ with 0 attained if and only if $x = x'$. However, the symmetry property for Δ is replaced by the bidualistic relation $\mathcal{D}_\Psi(x,x') = \mathcal{D}_{\Psi^*}((\nabla\Psi)(x'), (\nabla\Psi)(x))$ with Ψ^* satisfying $(\Psi^*)^* = \Psi$. In lieu of the triangle inequality for Δ, we have the triangle (generalized cosine) relation for $\mathcal{D}_\Psi$; in this sense $\mathcal{D}_\Psi$ can be viewed as generalizing the notion of a squared distance.

2.4. Extending to the infinite-dimensional case with conjugate scaling

Let us consider now how to extend the psychometric differentials (10) and (11) that are defined for stimuli in multidimensional vector spaces to stimuli in infinite-dimensional function spaces. A stimulus sometimes may be represented as a function, that is, a point in an infinite-dimensional space of functions. An example is the representation of a human face by means of a function relating grey-level or elevation above a plane of pixels to the two-dimensional coordinates of these pixels (Townsend, Solomon, & Smith, 2001). There, all grey-level or elevation image functions satisfying certain regularity conditions form a function space, and the set X of pixels on which the image functions are defined form a support of the function space on X, which is always measurable. Here we denote functions on X by $p(\zeta), q(\zeta)$ where $p, q : X \to \mathbb{R}$. In the infinite-dimensional space of face-representing functions, p and q are just two different faces defined on the pixel grid X.

To construct psychometric differentials on infinite-dimensional function spaces, we first look at a special case in the multidimensional setting when the stimulus dimensions are "noninteracting," in the following sense:

$$\Psi(x) = \sum_{i=1}^{n} f(x^i) .$$

Here f is a smooth, strictly convex function $\mathbb{R} \to \mathbb{R}$. In this case,

$$\nabla\Psi = [f'(x^1), \cdots, f'(x^n)] ,$$

where f' is the ordinary derivative. The psychometric differential then becomes

$$\mathcal{D}_\Psi(x,\hat{x}) = \sum_{i=1}^{n} \mathcal{D}_f(\hat{x}^i, x^i) ,$$

where

$$\mathcal{D}_f(x^i,\hat{x}^i) = f(x^i) - f(\hat{x}^i) - (x^i - \hat{x}^i) f'(\hat{x}^i)$$

is defined for each individual dimension $i = 1, ..., n$.

Recall that the convex conjugate $f^* : \mathbb{R} \to \mathbb{R}$ of f is defined as

$$f^*(t) = t\,(f')^{-1}(t) - f((f')^{-1}(t)) ,$$

with $(f^*)^* = f$ and $(f^*)' = (f')^{-1}$. So $\mathcal{D}_f$ possesses all of the properties stated in Proposition 4. In particular,

$$\mathcal{D}_f(x^i, \hat{x}^i) = \mathcal{D}_{f^*}(f'(\hat{x}^i), f'(x^i)) .$$

This excursion to the psychometric differential in the noninteracting multidimensional case suggests a way of constructing the psychometric differential in the infinite-dimensional case, that is, by replacing the summation across dimensions with integration over the support X,

$$\mathcal{D}_f(p,q) = \int_X \{f(p(\zeta)) - f(q(\zeta)) - (p(\zeta) - q(\zeta))f'(q(\zeta))\}\, d\mu , \qquad (12)$$

where $d\mu \equiv \mu(d\zeta)$ is some measure imposed on X. (Here and later, when dealing with an integral $\int_X(\cdot)d\mu$, we assume that it is finite.) Just like its multidimensional counterpart, $\mathcal{D}_f$ satisfies the bidualistic relation

$$\mathcal{D}_f(p,q) = \mathcal{D}_{f^*}(f'(q), f'(p)) \longleftrightarrow \mathcal{D}_{f^*}(p,q) = \mathcal{D}_f((f')^{-1}(q), (f')^{-1}(p)) . \qquad (13)$$

In its original ($\mathcal{A}$) form (see Section 2.2), the psychometric differential for the infinite-dimensional function space is

$$\mathcal{A}_f(p,q) = \mathcal{D}_f(p, (f')^{-1}(q)) = \mathcal{D}_{f^*}(q, f'(p)) ,$$

or written explicitly,

$$\mathcal{A}_f(p,q) = \int_X \{f(p(\zeta)) + f^*(q(\zeta)) - p(\zeta)\, q(\zeta)\}\, d\mu . \qquad (14)$$

It satisfies

$$\mathcal{A}_f(p,q) = \mathcal{A}_{f^*}(q,p) .$$

In the infinite-dimensional case, we have the additional freedom of "scaling" the p, q functions. To be concrete, we need to introduce the notion of *conjugate scaling* of functions $p(\zeta), q(\zeta)$. For a strictly increasing function $\rho : \mathbb{R} \to \mathbb{R}$, we call $\rho(\alpha)$ the *ρ-scaled representation* (of a real number α here). Clearly, a ρ-scaled representation is order invariant. For a smooth, strictly convex function f (with its conjugate f^*), we call the τ-scaled representation (of α) *conjugate* to its ρ-scaled representation with respect to f if

$$\tau(\alpha) = f'(\rho(\alpha)) = ((f^*)')^{-1}(\rho(\alpha)) \longleftrightarrow \rho(\alpha) = (f')^{-1}(\tau(\alpha)) = (f^*)'(\tau(\alpha)) . \qquad (15)$$

In this case, we also say that (ρ, τ) form an *ordered pair* of conjugate scales (with respect to f). Note that any two strictly increasing functions ρ, τ form an ordered pair of conjugate scales for some f. This is because the composite functions $\tau(\rho^{-1}(\cdot))$ and $\rho(\tau^{-1}(\cdot))$, which are mutually inverse, are always strictly increasing, so we may construct a pair of strictly convex and mutually conjugate functions (for some constants c and c^*)

$$f(t) = \int_c^t \tau(\rho^{-1}(s))\, ds$$

and

$$f^*(t) = \int_{c^*}^t \rho(\tau^{-1}(s))\, ds ,$$

to be associated with the (ρ, τ) scale by satisfying (15).

For a function $p(\zeta)$, we can construct a ρ-scaled representation $\rho(p(\zeta))$ and a τ-scaled representation $\tau(p(\zeta))$, denoted for brevity $\rho_p = \rho(p(\zeta))$ and $\tau_p = \tau(p(\zeta))$, respectively; they are both defined on the same support X as is $p(\zeta)$. Similar notations apply to ρ_q, τ_q with respect to the function $q(\zeta)$. With the notion of conjugate scaling, the reference-representation biduality of psychometric differential acquires the form (compare this with (13))

$$\mathcal{D}_f(\rho_p, \rho_q) = \mathcal{D}_{f^*}(\tau_q, \tau_p) .$$

2.5. The psychometric differential as a divergence function

Mathematically, the notion of a psychometric differential on a multidimensional vector space coincides with the so-called "divergence function," also known under various other names, such as "objective function," "loss function," and "contrast function," encountered in contexts entirely different from ours: in the fields of convex optimization, machine learning, and information geometry. In the form $\mathcal{D}_\Psi(x, x')$, the psychometric differential is known as the "Bregman divergence" (Bregman, 1967), an essential quantity in the area of convex optimization (Bauschke, 2003; Bauschke, Borwein, & Combettes, 2003; Bauschke & Combettes, 2003a, 2003b). This form of divergence is also referred to as the "geometric divergence" (Kurose, 1994), due to its significance in the hypersurface realization in affine differential geometric study of statistical manifolds. In its $\mathcal{A}$-form, the psychometric differential coincides with the "canonical divergence" first encountered in the analysis of the exponential family of probability distributions using information-theoretic methods (Amari, 1982, 1985). Henceforth, we will refer to $\mathcal{A}$ as such.[3]

[3] In information geometry, $\mathcal{A}$ is called "canonical" because its form is uniquely given in a dually flat space using a pair of biorthogonal coordinates (see Amari &

In the infinite-dimensional case, the psychometric differential in the form (12) is essentially the "U-divergence" recently proposed in the machine learning context (Murata, Takenouchi, Kanamori, & Eguchi, 2004). If we put $f(t) = t \log t - t$ $(t > 0)$, which is strictly convex, then $\mathcal{D}_f(p, q)$ acquires the form of the familiar Kullback-Leibler divergence between two probability densities p and q:

$$K(p,q) = \int_X \left\{ q - p - p \log \frac{q}{p} \right\} d\mu = K^*(q,p) . \tag{16}$$

As another example, consider the so-called "α-embedding" (here parameterized with $\lambda = (1+\alpha)/2$),

$$l^{(\lambda)} = \frac{1}{1-\lambda} p^{1-\lambda} ,$$

for $\lambda \in (0,1)$. In this case, one can put

$$f(t) = \frac{1}{\lambda}((1-\lambda)t)^{\frac{1}{1-\lambda}} , \quad f^*(t) = \frac{1}{1-\lambda}(\lambda t)^{\frac{1}{\lambda}} ,$$

so that $\rho(p) = l^{(\lambda)}(p)$ and $\tau(p) = l^{(1-\lambda)}(p)$ form an ordered pair of conjugate scales with respect to f. Under α-embedding, the canonical divergence (14) becomes

$$A^{(\lambda)}(p,q) = \frac{1}{\lambda(1-\lambda)} \int_X \{(1-\lambda)p + \lambda q - p^{1-\lambda} q^{\lambda}\} \, d\mu . \tag{17}$$

This is an important form of divergence, called "α-divergence" ($\alpha = 2\lambda - 1$). It is known that the α-divergence reduces to the Kullback-Leibler divergence, (16), when $\lambda \in \{0,1\}$ as a limiting case.

3. SCALING STIMULUS SPACE BY A FAMILY OF DIVERGENCE FUNCTIONS

The asymmetric divergence functions $\mathcal{D}_\Psi$ (for multidimensional spaces) and $\mathcal{D}_f$ (for infinite-dimensional spaces) investigated in the previous section are induced by smooth and strictly convex (but otherwise arbitrary) functions $\Psi : \mathbb{R}^n \to \mathbb{R}$ in the former case and $f : \mathbb{R} \to \mathbb{R}$ in the latter case. In this

Nagaoka, 2000, p. 61). On the other hand, $\mathcal{D}$ is the analogue of the "canonically transformed" psychometric function in the theory of Dzhafarov and Colonius (see Chapter 1 of this volume). One should not confuse these two usages of "canonical."

section, we show that any such Ψ (or f) induces a family of divergence functions that include $\mathcal{D}_\Psi$ (respectively, $\mathcal{D}_f$) as a special case. Although stimuli as vectors in a multidimensional space and stimuli as functions in an infinite-dimensional space are different, when there exists a parametric representation of a function, the divergence functional on the infinite-dimensional function space becomes, through a pullback to the parameter space, a divergence function on the multidimensional vector space.

3.1. Divergence on multidimensional vector space

Consider the vector space $\mathbb{R}^n$ where each point represents a stimulus. Recall that a function Ψ defined on a nonempty, convex set $\mathcal{X} \subseteq \mathbb{R}^n$ is called "strictly convex" if the inequality (4) is satisfied for any two distinct points $x, x' \in \mathcal{X}$ and any real number $\lambda \in (0,1)$. Also recall that the inequality sign is replaced by equality when (i) $x = x'$, for any $\lambda \in \mathbb{R}$; or (ii) $\lambda \in \{0,1\}$ for all $x, x' \in \mathcal{X}$.

PROPOSITION 5. *For any smooth, strictly convex function Ψ and any real number λ, the expression*

$$\mathcal{D}_\Psi^{(\lambda)}(x,x') = \frac{1}{\lambda(1-\lambda)}((1-\lambda)\Psi(x) + \lambda\Psi(x') - \Psi((1-\lambda)x + \lambda x'))$$

defines a parametric family (indexed by $\lambda \in \mathbb{R}$) of divergence functions.

Proof. See Proposition 1 of Zhang (2004a). ◇

Note the asymmetry of each divergence function, $\mathcal{D}_\Psi^{(\lambda)}(x,x') \neq \mathcal{D}_\Psi^{(\lambda)}(x',x)$. At the same time,

$$\mathcal{D}_\Psi^{(\lambda)}(x,x') = \mathcal{D}_\Psi^{(1-\lambda)}(x',x) , \tag{18}$$

indicating that the referential duality (in assigning to x or x' the referent or probe status) is reflected in the $\lambda \leftrightarrow 1-\lambda$ duality.

Two important special cases are as follows:

$$\lim_{\lambda \to 1} \mathcal{D}_\Psi^{(\lambda)}(x,x') = \mathcal{D}_\Psi(x,x') ,$$
$$\lim_{\lambda \to 0} \mathcal{D}_\Psi^{(\lambda)}(x,x') = \mathcal{D}_\Psi(x',x) .$$

Therefore

$$\mathcal{D}_\Psi^{(1)}(x,x') = \mathcal{D}_{\Psi^*}^{(1)}(\nabla\Psi(x'), \nabla\Psi(x)) = \mathcal{D}_{\Psi^*}^{(0)}(\nabla\Psi(x), \nabla\Psi(x')) = \mathcal{D}_\Psi^{(0)}(x',x) . \tag{19}$$

Here the (convex) conjugate scaled mappings $y = (\nabla\Psi)(x) \leftrightarrow x = (\nabla\Psi^*)(y)$ reflect the representational duality, in the choice of representing the stimulus as a vector in the original vector space $\mathcal{X}$, versus in the dual vector space $\mathcal{Y}$ (the gradient space). The aforementioned equation (19) states very concisely that when $\lambda \in \{0,1\}$, the referential duality and the representational duality are themselves dualistic — in other words, the canonical divergence functions exhibit the reference-representation biduality.

Note that $\mathcal{D}_{\Psi}^{(\lambda)}$, as a parametric family of divergence functions that reduces to $\mathcal{D}_{\Psi}$ as its special case, is not the only family capable of doing so. For instance, we may introduce

$$\tilde{D}_{\Psi}^{(\lambda)}(x,x') = (1-\lambda)\,\mathcal{D}_{\Psi}^{(0)}(x,x') + \lambda\,\mathcal{D}_{\Psi}^{(1)}(x,x') \;=\; (1-\lambda)\,\mathcal{D}_{\Psi}(x',x) + \lambda\,\mathcal{D}_{\Psi}(x,x') \,.$$

It turns out that these two families of divergence functions $\mathcal{D}_{\Psi}^{(\lambda)}$ and $\tilde{D}_{\Psi}^{(\lambda)}$ agree with each other up to the third order in their Taylor expansions in terms of x and y. However, $\mathcal{D}_{\Psi}^{(\lambda)} \neq \tilde{D}_{\Psi}^{(\lambda)}$ unless $\lambda \in \{0,1\}$; the reason lies in the fact that their Taylor expansions differ in the fourth and higher order terms. In particular, the self-dual elements ($\lambda = 1/2$) of those two families differ:

$$\mathcal{D}_{\Psi}^{(1/2)}(x,x') = 2\left(\Psi(x) + \Psi(x') - 2\Psi\left(\frac{x+x'}{2}\right)\right) \,,$$

$$\tilde{D}_{\Psi}^{(1/2)}(x,x') = \frac{1}{2}\,\langle x - x', (\nabla\Psi)(x) - (\nabla\Psi)(x')\rangle \,.$$

In Section 4 it will be shown that divergence functions induce a Riemannian metric and a pair of conjugate connections. The Riemannian structure induced by $\mathcal{D}_{\Psi}^{(\lambda)}$ and $\tilde{D}_{\Psi}^{(\lambda)}$ turns out to be identical.

3.2. Divergence on infinite-dimensional function space

By analogy with (4), in function spaces, we introduce a strictly convex function $f : \mathbb{R} \to \mathbb{R}$, which satisfies

$$(1-\lambda) f(\alpha) + \lambda f(\beta) > f((1-\lambda)\,\alpha + \lambda\,\beta)$$

for all $\alpha \neq \beta$. This convex function f allows us to introduce a family of divergence functionals on a function space whose elements are all functions $X \to \mathbb{R}$.

PROPOSITION 6. *Let $p(\zeta), q(\zeta)$ denote two functions on X, $f : \mathbb{R} \to \mathbb{R}$ a smooth, strictly convex function, and $\lambda \in \mathbb{R}$. Then the following expression gives a family of divergence functionals under ρ-scaling*

$$\mathcal{D}_f^{(\lambda)}(\rho_p, \rho_q) = \frac{1}{\lambda(1-\lambda)} \int_X \{(1-\lambda)\, f(\rho_p) + \lambda\, f(\rho_q) - f((1-\lambda)\,\rho_p + \lambda\rho_q)\}\, d\mu \tag{20}$$

where $\rho : \mathbb{R} \to \mathbb{R}$ is a strictly increasing function, $d\mu(\zeta) \equiv \mu(d\zeta)$ is a measure on X.

Proof. This is analogous to Proposition 5, after integration with respect to the support X. ◇

As special cases,

$$\lim_{\lambda\to 1} \mathcal{D}_f^{(\lambda)}(\rho_p, \rho_q) = \mathcal{D}_f(\rho_p, \rho_q) \;=\; \mathcal{D}_{f^*}(\tau_q, \tau_p) \,,$$

$$\lim_{\lambda\to 0} \mathcal{D}_f^{(\lambda)}(\rho_p, \rho_q) = \mathcal{D}_f(\rho_q, \rho_p) \;=\; \mathcal{D}_{f^*}(\tau_p, \tau_q) \,,$$

where $\mathcal{D}_f$ is given in (12). The reference-representation biduality here can be presented as

$$\begin{aligned}\mathcal{D}_f^{(1)}(\rho_p, \rho_q) \;&=\; \mathcal{D}_f^{(0)}(\rho_q, \rho_p) \;=\; \mathcal{D}_{f^*}^{(1)}(\tau_q, \tau_p) \\ &=\; \mathcal{D}_{f^*}^{(0)}(\tau_p, \tau_q) = \mathcal{A}_f(\rho_p, \tau_q) = \mathcal{A}_{f^*}(\tau_q, \rho_p) \,.\end{aligned}$$

An example of $\mathcal{D}_f^{(\lambda)}$ is the α-divergence (17), $\lambda = (1+\alpha)/2$. Putting $\rho(t) = \log t$, $f(t) = e^t$, and hence $\tau(t) = t$, it is easily seen that the functional (20) reduces to (17).

3.3. Connection between the multidimensional and infinite-dimensional cases

When the functions $p(\cdot), q(\cdot)$ in Proposition 6 belong to a parametric family $h(\cdot|\theta)$ with $\theta = [\theta^1, \cdots, \theta^n]$, so that $p(\zeta) = h(\zeta|\theta_p)$, $q(\zeta) = h(\zeta|\theta_q)$, the divergence *functional* taking in functions p, q as its arguments can be viewed as a divergence *function* of their parametric representation θ_p, θ_q. In other words, through parameterizing the functions representing stimuli, we arrive at a divergence function over a multidimensional vector space. We now investigate conditions under which this divergence function have the same form as that stated in Proposition 5.

PROPOSITION 7. *Let $f : \mathbb{R} \to \mathbb{R}$ be strictly convex, and (ρ, τ) be an ordered pair of conjugate scales with respect to f. Suppose the ρ-scaled representation $\rho(h(\zeta)) \equiv \rho_h(\zeta)$ of the stimulus function $h(\zeta)$, $\zeta \in X$, can be represented as*

$$\rho_h = \langle \theta, \lambda(\zeta) \rangle \,, \tag{21}$$

where $\lambda(\zeta) = [\lambda_1(\zeta), \cdots, \lambda_n(\zeta)]$, with its components representing n linearly independent basis functions, and $\theta = [\theta^1, \cdots, \theta^n]$ is a vector whose components are real numbers. Then

(i) *the function*

$$\Psi(\theta) = \int_X f(\rho_h)\, d\mu$$

is strictly convex;

(ii) *denote* $\eta = [\eta_1, \cdots, \eta_n]$ *as the projection of the* τ*-scaled representation of* $h(\zeta)$, $\tau(h(\zeta)) \equiv \tau_h(\zeta)$, *on* $\lambda(\zeta)$,

$$\eta = \int_X \tau_h \, \lambda(\zeta) \, d\mu \ , \tag{22}$$

and denote

$$\tilde{\Psi}(\theta) = \int_X f^*(\tau_h) \, d\mu \ ,$$

then the function $\tilde{\Psi}((\nabla \Psi)^{-1}(\cdot)) \equiv \Psi^*(\cdot)$ *is the convex conjugate of* $\Psi(\cdot)$;

(iii) *the* θ *and* η *parameters are related to each other via*

$$(\nabla \Psi)(\theta) = \eta \ , \qquad (\nabla \Psi^*)(\eta) = \theta \ ;$$

(iv) *the divergence functionals* $\mathcal{D}_f^{(\lambda)}$ *becomes the divergence functions* $\mathcal{D}_\Psi^{(\lambda)}$,

$$\mathcal{D}_f^{(\lambda)}(\rho_p, \rho_q) = \mathcal{D}_\Psi^{(\lambda)}(\theta_p, \theta_q) \ ;$$

(v) *the canonical divergence functional* $\mathcal{A}_f$ *becomes the canonical divergence function* $\mathcal{A}_\Psi$,

$$\mathcal{A}_f(\rho_p, \tau_q) = \Psi(\theta_p) + \Psi^*(\eta_q) - \langle \theta_p, \eta_q \rangle \ = \ \mathcal{A}_\Psi(\theta_p, \eta_q) \ .$$

Proof. Parts (iv) and (v) are natural consequences of part (i), by substituting the expression of $\Phi(\theta)$ for the corresponding term in the definition of $\mathcal{D}$ and $\mathcal{A}$. Parts (i) to (iii) were proved in Proposition 9 of Zhang (2004a).◇

The parameter θ in $h(\cdot|\theta)$ can be viewed as the "natural parameter" (borrowing the terminology from statistics) of parameterized functions representing stimuli. In information geometry, it is well known that an exponential family of density functions can also be parameterized by means of the "expectation parameter," which is dual to the natural parameter; this is our parameter η here. We have thus generalized the duality between the natural parameter and the expectation parameter from the exponential family to stimuli under arbitrary ρ- and τ-embeddings. Proposition 7 specifies the sufficient condition, (21), under which we can use one of the vectors, θ (natural parameter) or η (expectation parameter), to represent an individual stimulus function $h(\zeta)$, and under which the pullback of $\mathcal{D}_f^{(\lambda)}(\rho(\cdot), \rho(\cdot))$ and $\mathcal{D}_{f^*}^{(\lambda)}(\tau(\cdot), \tau(\cdot))$ in the multidimensional parameter space gives rise to the form of divergence functions presented by Proposition 5.

4. BIDUALISTIC RIEMANNIAN STRUCTURE OF STIMULUS MANIFOLDS

A divergence function, while measuring distance of two points in the large, induces a dually symmetric Riemannian structure in the small with a metric g and a pair of conjugate connections Γ, Γ^*. They are given by what we refer to here as the *Eguchi relations* (Eguchi, 1983, 1992):

$$g_{ij}(x) = - \left. \frac{\partial^2 \mathcal{D}_\Psi^{(\lambda)}(x', x'')}{\partial x'^i \partial x''^j} \right|_{x'=x''=x} , \tag{23}$$

$$\Gamma_{ij,k}^{(\lambda)}(x) = - \left. \frac{\partial^3 \mathcal{D}_\Psi^{(\lambda)}(x', x'')}{\partial x'^i \partial x'^j \partial x''^k} \right|_{x'=x''=x} , \tag{24}$$

$$\Gamma_{ij,k}^{*(\lambda)}(x) = - \left. \frac{\partial^3 \mathcal{D}_\Psi^{(\lambda)}(x', x'')}{\partial x''^i \partial x''^j \partial x'^k} \right|_{x'=x''=x} . \tag{25}$$

Later, we explicitly give the Riemannian metric and the pair of conjugate connections induced by the divergence functions $\mathcal{D}_\Psi^{(\lambda)}$ and the divergence functionals $\mathcal{D}_f^{(\lambda)}$.

4.1. Riemannian structure on multidimensional vector space

In the proposition to follow, $\Psi_{ij}(x)$, $\Psi_{ijk}(x)$ denote, respectively, second and third partial derivatives of $\Psi(x)$,

$$\Psi_{ij}(x) = \frac{\partial^2 \Psi(x)}{\partial x^i \partial x^j} \ , \qquad \Psi_{ijk}(x) = \frac{\partial^3 \Psi(x)}{\partial x^i \partial x^j \partial x^k} \ ,$$

and $\Psi^{ij}(x)$ is the matrix inverse of $\Psi_{ij}(x)$.

PROPOSITION 8. *The divergence functions* $\mathcal{D}_\Psi^{(\lambda)}(x, x')$ *induce on the stimulus manifold a metric* g *and a pair of conjugate connections* $\Gamma^{(\lambda)}, \Gamma^{*(\lambda)}$ *with*

(i) *the metric tensor given by*

$$g_{ij}(x) = \Psi_{ij}(x) \ ;$$

(ii) *the conjugate connections given by*

$$\Gamma_{ij,k}^{(\lambda)}(x) = (1 - \lambda) \, \Psi_{ijk}(x) \ , \qquad \Gamma_{ij,k}^{*(\lambda)}(x) = \lambda \, \Psi_{ijk}(x) \ ;$$

(iii) *the Riemann-Christoffel curvature for the connection* $\Gamma^{(\lambda)}_{ij,k}$ *given by*

$$R^{(\lambda)}_{ij\mu\nu}(x) = \lambda(1-\lambda)\sum_{l,k}(\Psi_{il\nu}(x)\Psi_{jk\mu}(x) - \Psi_{il\mu}(x)\Psi_{jk\nu}(x))\Psi^{lk}(x) \ .$$

Proof. The proof for parts (i) and (ii) and for part (iii) follows, respectively, those in Proposition 2 and in Proposition 3 of Zhang (2004a). ◇

According to Proposition 8, the metric tensor g_{ij}, which is symmetric and positive semidefinite due to the strict convexity of Ψ, is independent of λ, whereas the affine connections are λ-dependent, satisfying the dualistic relation

$$\Gamma^{*(\lambda)}_{ij,k}(x) = \Gamma^{(1-\lambda)}_{ij,k}(x) \ . \qquad (26)$$

When $\lambda = 1/2$, the self-conjugate connection $\Gamma^{(1/2)} = \Gamma^{*(1/2)} \equiv \Gamma^{LC}$ is the Levi-Civita connection, related to the metric tensor by

$$\Gamma^{(1/2)}_{ij,k}(x) = \Gamma^{LC}_{ij,k}(x) \equiv \frac{1}{2}\left(\frac{\partial g_{ik}(x)}{\partial x^j} + \frac{\partial g_{kj}(x)}{\partial x^i} - \frac{\partial g_{ij}(x)}{\partial x^k}\right) \ .$$

For any λ, the mutually conjugate connections $\Gamma^{(\lambda)}_{ij,k}$ and $\Gamma^{(1-\lambda)}_{ij,k}$ satisfy the relation

$$\frac{1}{2}\left(\Gamma^{(\lambda)}_{ij,k}(x) + \Gamma^{(1-\lambda)}_{ij,k}(x)\right) = \Gamma^{LC}_{ij,k}(x) \ .$$

When $\lambda \in \{0,1\}$, all components of the Riemann-Christoffel curvature tensor vanish, in which case $\Gamma^{(0)}_{ij,k}(x) = 0 \leftrightarrow \Gamma^{*(1)}_{ij,k}(x) = 0$, or $\Gamma^{(1)}_{ij,k}(x) = 0 \leftrightarrow \Gamma^{*(0)}_{ij,k}(x) = 0$. The manifold in these cases is said to be "dually flat" (Amari, 1985; Amari & Nagaoka, 2000) and the divergence functions defined on it is the unique, canonical divergence studied in Sections 2.2 and 2.3.

Note that the referential duality exhibited by the divergence functions $\mathcal{D}^{(\lambda)}_{\Psi}$, (18), is reflected in the conjugacy of the affine connections $\Gamma \leftrightarrow \Gamma^*$, (26). On the other hand, the one-to-one mapping $y = (\nabla\Psi)(x) \leftrightarrow x = (\nabla\Psi^*)(y)$ between the space $\mathcal{X}$ and its dual $\mathcal{Y}$ indicates that we may view $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ as two coordinate representations for one and the same underlying manifold. Later, we investigate this representational duality. We will relate the Riemannian structures (metric, dual connnections, Rieman-Christoffel curvature) expressed in x and in y.

PROPOSITION 9. *Denote the Riemannian metric, the connection, and the Riemann-Christoffel curvature tensor induced by* $\mathcal{D}^{(\lambda)}_{\Psi^*}(y,y')$ *as, respectively,* $\tilde{g}^{mn}(y)$, $\tilde{\Gamma}^{(\lambda)mn,l}(y)$, *and* $\tilde{R}^{(\lambda)klmn}(y)$, *whereas the analogous quantities without the tilde sign are induced by* $\mathcal{D}^{(\lambda)}_{\Psi}(x,x')$, *as in Proposition 8. Then, as long as*

$$y = (\nabla\Psi)(x) \longleftrightarrow x = (\nabla\Psi^*)(y) \ ,$$

(i) *the metric tensors are related as*

$$\sum_l g_{il}(x)\tilde{g}^{ln}(y) = \delta^n_i \ ;$$

(ii) *the affine connections are related as*

$$\tilde{\Gamma}^{(\lambda)mn,l}(y) = -\sum_{i,j,k}\tilde{g}^{im}(y)\tilde{g}^{jn}(y)\tilde{g}^{kl}(y)\Gamma^{(\lambda)}_{ij,k}(x) \ ;$$

(iii) *the Riemann-Christoffel curvatures are related as*

$$\tilde{R}^{(\lambda)klmn}(y) = \sum_{i,j,\mu,\nu}\tilde{g}^{ik}(y)\tilde{g}^{jl}(y)\tilde{g}^{\mu m}(y)\tilde{g}^{\nu n}(y)R^{(\lambda)}_{ij\mu\nu}(x) \ .$$

Proof. See Proposition 5 of Zhang (2004a). ◇

4.2. Riemannian structure on infinite-dimensional function space

Similar to the multidimensional case, we can compute the Riemannian geometry of the stimulus manifold for the infinite-dimensional (function space) case. For ease of comparison, we assume here that the stimulus functions $h(\zeta)$ have a parametric representation $h(\zeta|\theta)$, so strictly speaking, "infinite-dimensional space" is a misnomer.

PROPOSITION 10. *The family of divergence functions* $\mathcal{D}^{(\lambda)}_f(\rho(h(\cdot|\theta_p)), \rho(h(\cdot|\theta_q)))$ *induce a Riemannian structure on the stimulus manifold for each* $\lambda \in \mathbb{R}$, *with*

(i) *the metric tensor given as*

$$g_{ij}(\theta) = \int_X \left\{ f''(\rho(h(\zeta|\theta)))\frac{\partial\rho(h(\zeta|\theta))}{\partial\theta^i}\frac{\partial\rho(h(\zeta|\theta))}{\partial\theta^j}\right\} d\mu \ ;$$

(ii) *the conjugate connections given as*

$$\begin{aligned}\Gamma^{(\lambda)}_{ij,k}(\theta) &= \int_X \{(1-\lambda)\, f'''(\rho(h(\zeta|\theta)))\, A_{ijk}(\zeta|\theta) \\ &\quad + f''(\rho(h(\zeta|\theta)))\, B_{ijk}(\zeta|\theta)\}\, d\mu \ , \\ \Gamma^{*(\lambda)}_{ij,k}(\theta) &= \int_X \{\lambda\, f'''(\rho(h(\zeta|\theta)))\, A_{ijk}(\zeta|\theta) \\ &\quad + f''(\rho(h(\zeta|\theta)))\, B_{ijk}(\zeta|\theta)\}\, d\mu \ .\end{aligned}$$

Here A_{ijk}, B_{ijk} denote

$$A_{ijk}(\zeta|\theta) = \frac{\partial\rho(h(\zeta|\theta))}{\partial\theta^i}\frac{\partial\rho(h(\zeta|\theta))}{\partial\theta^j}\frac{\partial\rho(h(\zeta|\theta))}{\partial\theta^k} ,$$
$$B_{ijk}(\zeta|\theta) = \frac{\partial^2\rho(h(\zeta|\theta))}{\partial\theta^i\partial\theta^j}\frac{\partial\rho(h(\zeta|\theta))}{\partial\theta^k} .$$

Proof. The proof is by straightforward application of (23) to (25). See Proposition 7 of Zhang (2004a) for details. ⋄

Note that the strict convexity of f implies $f'' > 0$, and thereby guarantees the positive semidefiniteness of g_{ij}. Clearly, the conjugate connections satisfy $\Gamma^{*(\lambda)}_{ijk}(\theta) = \Gamma^{(1-\lambda)}_{ijk}(\theta)$, and hence reflect referential duality.

As a special case, if we set $f(t) = e^t$ and $\rho(t) = \log t$, then Proposition 10 gives the Fisher information metric

$$\int_X \left\{ h(\zeta|\theta)\,\frac{\partial\log(h(\zeta|\theta))}{\partial\theta^i}\,\frac{\partial\log(h(\zeta|\theta))}{\partial\theta^j} \right\} d\mu ,$$

and the α-connections associated with the α-divergence mentioned earlier (with $\alpha = 2\lambda - 1$):

$$\int_X \left\{ h(\zeta|\theta)\left((1-\lambda)\,\frac{\partial\log h(\zeta|\theta)}{\partial\theta^i}\,\frac{\partial\log h(\zeta|\theta)}{\partial\theta^j} + \frac{\partial^2\log h(\zeta|\theta)}{\partial\theta^i\partial\theta^j} \right) \times\frac{\partial\log h(\zeta|\theta)}{\partial\theta^k} \right\} d\mu .$$

Therefore, the Riemannian structure derived here generalizes the core concepts of classic parametric information geometry as summarized in Amari (1985) and Amari and Nagaoka (2000).

For the next proposition, recall the notion of conjugate scaling of functions we introduced in Section 2.4.

PROPOSITION 11. *Under conjugate ρ- and τ-scaling (with respect to some strictly convex function f),*

(i) *the metric tensor is given by*

$$g_{ij}(\theta) = \int_X \left\{ \frac{\partial\rho(h(\zeta|\theta))}{\partial\theta^i}\frac{\partial\tau(h(\zeta|\theta))}{\partial\theta^j} \right\} d\mu$$
$$= \int_X \left\{ \frac{\partial\tau(h(\zeta|\theta))}{\partial\theta^i}\frac{\partial\rho(h(\zeta|\theta))}{\partial\theta^j} \right\} d\mu ;$$

(ii) *the conjugate connections are given by*

$$\Gamma^{(\lambda)}_{ij,k}(\theta) = \int_X \left\{ (1-\lambda)\frac{\partial^2\tau(h(\zeta|\theta))}{\partial\theta^i\partial\theta^j}\frac{\partial\rho(h(\zeta|\theta))}{\partial\theta^k} +\lambda\frac{\partial^2\rho(h(\zeta|\theta))}{\partial\theta^i\partial\theta^j}\frac{\partial\tau(h(\zeta|\theta))}{\partial\theta^k} \right\} d\mu ,$$
$$\Gamma^{*(\lambda)}_{ij,k}(\theta) = \int_X \left\{ \lambda\frac{\partial^2\tau(h(\zeta|\theta))}{\partial\theta^i\partial\theta^j}\frac{\partial\rho(h(\zeta|\theta))}{\partial\theta^k} +(1-\lambda)\frac{\partial^2\rho(h(\zeta|\theta))}{\partial\theta^i\partial\theta^j}\frac{\partial\tau(h(\zeta|\theta))}{\partial\theta^k} \right\} d\mu .$$

Proof. See Proposition 8 of Zhang (2004a). ⋄

Proposition 11 casts the metric and conjugate connections in dualistic forms with respect to any pair of conjugate scales (ρ, τ). This leads to the following corollary.

COROLLARY 12. *The metric tensor $\tilde{g}_{ij}$ and the dual affine connection $\tilde{\Gamma}^{(\lambda)}_{ij,k}, \tilde{\Gamma}^{*(\lambda)}_{ij,k}$ induced on the stimulus manifold by the divergence functions $\mathcal{D}^{(\lambda)}_{f^*}(\tau(h(\cdot|\theta_p)), \tau(h(\cdot|\theta_q)))$ are related to, respectively, g_{ij}, $\Gamma^{(\lambda)}_{ij,k}$, and $\Gamma^{*(\lambda)}_{ij,k}$ induced by $\mathcal{D}^{(\lambda)}_{f}(\rho(h(\cdot|\theta_p)), \rho(h(\cdot|\theta_q)))$ as*

$$\tilde{g}_{ij}(\theta) = g_{ij}(\theta) ,$$

with

$$\tilde{\Gamma}^{(\lambda)}_{ij,k}(\theta) = \Gamma^{*(\lambda)}_{ij,k}(\theta) , \qquad \tilde{\Gamma}^{*(\lambda)}_{ij,k}(\theta) = \Gamma^{(\lambda)}_{ij,k}(\theta) .$$

Proof. See Corollary 3 of Zhang (2004a). ⋄

Combining Proposition 11 with Corollary 12, we get

$$\Gamma^{*(\lambda)}_{ij,k}(\theta) = \tilde{\Gamma}^{(\lambda)}_{ij,k}(\theta) .$$

This is the *reference-representation biduality* for the Riemannian structure of an infinite-dimensional stimulus space (after parameterization).

4.3. Connection between the multidimensional and infinite-dimensional cases

We have shown in Proposition 7 that when (21) holds, the divergence functionals $\mathcal{D}^{(\lambda)}_f$ become the divergence functions $\mathcal{D}^{(\lambda)}_\Psi$ on the finite-dimensional parameter space. This correspondence also holds for the Riemannian geometries they induce.

PROPOSITION 13. *Under representation (21), the metric tensor and the conjugate connections on a stimulus manifold as given by Proposition 10,* *acquire the form given in Proposition 8:*

$$g_{ij}(\theta) = \frac{\partial^2 \Psi(\theta)}{\partial\theta^i \partial\theta^j}\,, \quad \Gamma^{(\lambda)}_{ij,k}(\theta) = (1-\lambda)\frac{\partial^3 \Psi(\theta)}{\partial\theta^i \partial\theta^j \partial\theta^k}\,, \quad \Gamma^{*(\lambda)}_{ij,k}(\theta) = \lambda\frac{\partial^3 \Psi(\theta)}{\partial\theta^i \partial\theta^j \partial\theta^k}\,.$$

Proof. See Proposition 9 of Zhang (2004a). ◇

According to Proposition 7, θ and η are mutually orthogonal coordinates. They are related to the covariant and contravariant representations of the metric tensor (see Proposition 9):

$$\frac{\partial\eta_i}{\partial\theta^j} = g_{ij}(\theta)\,, \quad \frac{\partial\theta^i}{\partial\eta_j} = \tilde{g}^{ij}(\eta)\,.$$

5. SUMMARY AND DISCUSSION

Understanding the intrinsic asymmetry during referent-probe comparisons was the motivation for this exposition. It was our goal to find a proper mathematical formalism to express the asymmetric difference between a reference stimulus (referent) and a comparison stimulus (probe) in such a way that the assignment of the referent-probe status itself can be "arbitrarily made." This is to say, subject to a change of the representation ("scaling") of the two stimuli, the roles of the reference and the comparison stimuli in expressing the asymmetric difference are "exchangeable" within the formalism. This was the kind of duality we were looking for, namely, to account for the referent-probe asymmetry by scaling the stimulus representations.

To this end, we have made use of some tools in convex analysis and differential geometry. We characterized the asymmetric distance between a reference stimulus and a comparison stimulus by proposing a dually symmetric psychometric differential function measuring the directed difference between them. The principle of regular cross-minimality (Axiom 1) allowed us to establish a one-to-one mapping between the two stimulus spaces. Further requiring the mapping to be symmetric with positive-definite Jacobian (Axiom 2) led to (convex) conjugate scaled representations of these mappings. Making full use of the machineries of convex analysis, we constructed a family of psychometric differentials between any two points (one as referent and one as probe) based on the fundamental inequality of a convex function. The family of psychometric differentials is indexed by the parameter $\lambda \in \mathbb{R}$, with $\lambda \in \{0,1\}$ cases specializing to the canonical divergence which satisfy the reference-representation biduality (Axiom 3). So what we have accomplished here is to show the convex conjugation (under Legendre transformation) to be the precise mathematical expression of the representational duality and the convex mixture coefficient as expressing

the referential duality. We also showed that this kind of biduality — referential duality and representational duality — is fundamentally coupled in defining the Riemannian geometry (metric, connection, curvature, etc.) of the stimulus manifold. The referential duality is revealed as the conjugacy of the connection pair, whereas the representational duality is revealed as the choice of the contravariant and covariant form of a vector to represent the stimulus (for the multidimensional case), or as the choice from a pair of monotone transformations ρ, τ to "scale" the stimulus function (for the infinite-dimensional case).

Throughout our investigation of dual scaling of the stimulus space, in terms of either the psychometric differential (divergence function) in the large or the resultant Riemannian geometry in the small, our treatment has been both in the multidimensional vector space setting and in the infinite-dimensional function space setting. Through an appropriate affine submanifold embedding, namely (21), the infinite-dimensional forms of the divergence measure and of the geometry reduce to the corresponding multidimensional forms. Submanifold embedding not only provides a unified view of duality independent of whether stimuli are defined in the multidimensional space or in the infinite-dimensional space, but also is a window for more intuitive understanding of the kind of dual Riemannian geometry (involving conjugate connections) studied here. It is known in affine differential geometry (Nomizu & Sasaki, 1994; Simon, Schwenk-Schellschmidt, & Viesel, 1991) that conjugate connections arise from characterizing the different ways that hypersurfaces can be embedded into a higher dimensional space. Future research will elaborate how these geometric intuitions may be applied to the bona fide infinite-dimensional function space as well (not merely the parameterized version as done here), and explain how referential duality and representational duality become dualistic themselves.

References

Amari, S. (1982). Differential geometry of curved exponential families — Curvatures and information loss. *Annals of Statistics*, 10, 357-385.

Amari, S. (1985). *Differential geometric methods in statistics* (Lecture Notes in Statistics), 28, New York: Springer-Verlag.

Amari, S., & Nagaoka, H. (2000). *Method of information geometry* (AMS monograph). New York: Oxford University Press.

Bauschke, H. H. (2003). Duality for Bregman projections onto translated cones and affine subspaces. *Journal of Approximation Theory*, 121, 1-12.

Bauschke, H. H., Borwein, J. M., & Combettes, P. L. (2003). Bregman monotone optimization algorithms. *SIAM Journal on Control and Optimization*, 42, 596-636.

Bauschke, H. H. & Combettes, P. L. (2003a). Construction of best Bregman approximations in reflexive Banach spaces. *Proceedings of the American Mathematical Society*, 131, 3757-3766.
Bauschke, H. H. & Combettes, P. L. (2003b). Iterating Bregman retractions. *SIAM Journal on Optimization*, 13, 1159-1173.
Bregman, L. M. (1967). The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. *USSR Computational Mathematics and Physics*, 7, 200-217.
Dzhafarov, E. N. (2002). Multidimensional Fechnerian scaling, pairwise comparisons, regular minimality, and nonconstant self-similarity. *Journal of Mathematical Psychology*, 46, 583-608.
Dzhafarov, E. N., & Colonius, H. (1999). Fechnerian metrics in unidimensional and multidimensional stimulus spaces. *Psychological 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.
Eguchi, S. (1983). Second order efficiency of minimum contrast estimators in a curved exponential family. *Annals of Statistics*, 11, 793-803.
Eguchi, S. (1992). Geometry of minimum contrast. *Hiroshima Mathematical Journal*, 22, 631-647.
Kurose, T. (1994). On the divergences of 1-conformally flat statistical manifolds. *Tôhoko Mathematical Journal*, 46, 427-433.
Murata, N., Takenouchi, T., Kanamori, T., & Eguchi, S. (2004). Information geometry of U-boost and Bregman divergence. *Neural Computation*, 16, 1437-1481.
Nomizu, K., & Sasaki, T. (1994). *Affine differential geometry*. Cambridge, UK: Cambridge University Press.
Roberts, A. W., & Varberg, D. E. (1973). *Convex functions*. New York: Academic.
Rockafellar, R. T. (1970). *Convex analysis*. Princeton, NJ: Princeton University Press.
Shepard, R. N. (1962a). The analysis of proximities, multidimensional scaling with an unknown distance function. I. *Psychometrika*, 27, 125-140.
Shepard, R. N. (1962b). The analysis of proximities, multidimensional scaling with an unknown distance function. II. *Psychometrika*, 27, 219-246.
Simon, U., Schwenk-Schellschmidt, A., & Viesel, H. (1991). *Introduction to the affine differential geometry of hypersurfaces*. Tokyo: Science University of Tokyo.
Townsend, J. T., Solomon, B., & Smith, J. S. (2001). The perfect Gestalt, Infinite dimensional Riemannian face spaces and other aspects of face perception. In M. J. Wenger, & J. T. Townsend (Eds), *Computational, geometric, and process perspectives of facial cognition, contexts and challenges* (pp 39-82). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Tversky, A. (1977). Features of similarity. *Psychological Review*, 84, 327-352.
Zhang, J. (2004a). Divergence function, duality, and convex analysis. *Neural Computation*, 16, 159-195.
Zhang, J. (2004b). Dual scaling between comparison and reference stimuli in multidimensional psychological space. *Journal of Mathematical Psychology*, 48, 409-424.

SCIENTIFIC PSYCHOLOGY SERIES

Stephen W. Link and James T. Townsend, Series Editors

MONOGRAPHS

Louis Narens • *Theories of Meaningfulness*

R. Duncan Luce • *Utility of Gains and Losses: Measurement-Theoretical and Experimental Approaches*

William R. Uttal • *The War Between Mentalism and Behaviorism: On the Accessibility of Mental Processes*

William R. Uttal • *Toward a New Behaviorism: The Case Against Perceptual Reductionism*

Gordon M. Redding and Benjamin Wallace • *Adaptive Spatial Alignment*

John C. Baird • *Sensation and Judgment: Complementarity Theory of Psychophysics*

John A. Swets • *Signal Detection Theory and ROC Analysis in Psychology and Diagnostics: Collected Papers*

William R. Uttal • *The Swimmer: An Integrated Computational Model of a Perceptual–Motor System*

Stephen W. Link • *The Wave Theory of Difference and Similarity*

EDITED VOLUMES

Christian Kaernbach, Erich Schröger, and Hermann Müller • *Psychophysics Beyond Sensation: Laws and Invariants of Human Cognition*

Michael Wenger and James Townsend • *Computational, Geometric, and Process Perspectives on Facial Cognition: Contests and Challenges*

Jonathan Grainger and Arthur M. Jacobs • *Localist Connectionist Approaches to Human Cognition*

Cornilia E. Dowling, Fred S. Roberts, and Peter Theuns • *Recent Progress in mathematical Psychology*

F. Gregory Ashby • *Multidimensional Models of Perception and Cognition*

Hans-Georg Geissler, Stephen W. Link, and James T. Townsend • *Cognition, Information Processing, and Psychophysics: Basic Issues*

TEXTBOOKS

Norman H. Anderson • *Empirical Direction in Design and Analysis*

MEASUREMENT AND REPRESENTATION OF SENSATIONS

Edited by

Hans Colonius
Universität Oldenburg

Ehtibar N. Dzhafarov
Purdue University

LEA

2006

LAWRENCE ERLBAUM ASSOCIATES, PUBLISHERS

Mahwah, New Jersey London