

Symmetry, Direct Measurement, and Torgerson's Conjecture

Louis Narens *

Department of Cognitive Sciences
University of California, Irvine
email: lnarens@uci.edu

December 29, 2004

Abstract

Two forms of direct measurement are considered in the article: A strong form in which ratio productions named by number words are interpreted veridically as the numerical ratios they name; and a weak form in which the ratio productions named by number words may have interpretations as ratios that are different from numerical ratios they name. Both forms assume that the responses to instructions to produce ratios are represented numerically by ratios, and thus the word "ratio" (and supposedly the participants concept associated with it) is being "directly" represented. The strong form additionally "directly represents" the number mentioned in the instruction by itself. The article provides an axiomatic theory for the numerical representations produced by both forms. This theory eliminates the need for assuming anything is being "directly represented," allowing for a purely behavioral approach to ratio production data. It isolates two critical axioms for empirical testing. An measurement-theoretic explanation is provided for the puzzling empirical phenomena that subjects do not distinguish between ratios and differences in a variety of direct measurement tasks.

*The author thanks R. Duncan Luce and Ragnar Steingrimsson for helpful comments on earlier versions of the article.

1 Introduction

Falmagne's *Elements of Psychophysical Theory* appeared in 1985. While covering much of the theory about the measurement of sensation, it only touched briefly on the topic of direct measurement of sensation—a topic with a long and controversial history and enormous methodological, theoretical, and experimental literatures. I would guess that part of the reason for direct measurement's light treatment in *Elements* was that in 1985 there was very little mathematical theory concerning it that matched high level of scientific and mathematical rigor characteristic of Falmagne's *Elements*.

In 1985, many of us interested in axiomatic approaches to measurement believed that direct measurement methods were founded on unsound methodologies and theories. Researchers using direct measurement techniques did not appear aware of—or at least did not acknowledge—the strong structural assumptions implicit in their measurement techniques. Some of these assumptions had empirical import about the objects being measured that could, in principle, conflict with the numerical assignments produced by direct measurement.

Within the last ten years, Narens (1994, 1996) and Luce (2002, 2004) developed axiomatic theories that make explicit some of the structural assumptions inherent in representing direct measurement data, and have worked with colleagues to test the most important of these. This article explores two classical issues in psychophysics—the presumed ratio scalability of subjective intensity and Torgerson's Conjecture—in light of the above axiomatic theories and some of the empirical results they generated.

2 Representational Theory of Measurement

The principal method of measurement used throughout the article is the *Representational Theory*. To state the Representational Theory, some preliminary concepts and notation are needed:

Definition 1 Throughout the article \mathbb{R} stands for the real numbers, \mathbb{I} for the integers, \mathbb{R}^+ for the positive real numbers, \mathbb{I}^+ for the positive integers, and $*$ for the operation of function composition.

A *measurement scale* \mathcal{S} is a nonempty set of functions from a nonempty set of objects X into \mathbb{R} . Throughout the article it is assumed that X is a set of physical objects that is totally ordered by \preceq and that each measuring function φ in \mathcal{S} is a strictly \preceq -increasing function from X onto either \mathbb{R} or \mathbb{R}^+ ; that is, φ is an isomorphism of $\langle X, \preceq \rangle$ onto either $\langle \mathbb{R}, \leq \rangle$ or $\langle \mathbb{R}^+, \leq \rangle$. Measuring functions in \mathcal{S} are also sometimes called *scales*. \mathcal{S} is said to be a:

- *ratio scale* if and only if for all ψ in \mathcal{S} ,

$$\mathcal{S} = \{r\psi \mid r \in \mathbb{R}^+\},$$

- *interval scale* if and only if for all ψ in \mathcal{S} ,

$$\mathcal{S} = \{r\psi + s \mid r \in \mathbb{R}^+ \text{ and } s \in \mathbb{R}\},$$

- *log-interval scale*¹ if and only if for all ψ in \mathcal{S} ,

$$\mathcal{S} = \{s\psi^r \mid r \in \mathbb{R}^+ \text{ and } s \in \mathbb{R}^+\},$$

- *translation scale*² if and only if for all ψ in \mathcal{S} ,

$$\mathcal{S} = \{\psi + r \mid r \in \mathbb{R}\}.$$

The *Representational Theory of Measurement* measures objects through a scale \mathcal{S} of isomorphisms.³ Structure is added to X in terms of qualitative relations and functions, R_1, \dots, R_n . In practice R_1, \dots, R_n are usually observable. This gives rise to a *qualitative structure* $\mathfrak{X} = \langle X, \preceq, R_1, \dots, R_n \rangle$. Measurement consists of finding numerical structures $\mathfrak{N} = \langle N, \leq, S_1, \dots, S_n \rangle$, where $N = \mathbb{R}^+$ or $N = \mathbb{R}$, and using the measurement scale,

$$\mathcal{S} = \{\psi \mid \psi \text{ is an isomorphism of } \mathfrak{X} \text{ onto } \mathfrak{N}\}.$$

In principle any numerical structure isomorphic to X will suffice for producing a representational measurement scale; however, in practice numerical structures with well-understood relations are employed.

The relations \preceq, R_1, \dots, R_n are called the *primitives* of \mathfrak{X} . In some cases qualitative structures have infinitely many primitives. Isomorphisms of \mathfrak{X} onto itself are called *symmetries* (or *automorphisms*). Symmetries play a major role throughout the article.

Let $\psi \in \mathcal{S}$. Each relation or function R on \mathfrak{X} is mapped by ψ onto a relation or function on N denoted by $\psi(R)$. For example, the binary relation R on X is mapped onto the relation $\psi(R) = S$, where S is the binary relation on N such that for all r and s in N , $S(r, s)$ holds if and only if for some x and y in X , $R(x, y)$ holds and $\psi(x) = r$ and $\psi(y) = s$. \square

¹A log-interval scale can be viewed as a variant of an interval scale by a change of variable, i.e., $t \rightarrow e^t$.

²A translation scale can be viewed as a variant of a ratio scale by a change of variable, i.e., $t \rightarrow \log t$.

³Sometimes the representational theory uses a weaker form of structural preserving functions called *homomorphisms* as measuring functions. For the purposes of this article, only isomorphisms are needed.

3 Plateau's Theory

The Belgian physicist M. J. Plateau was the first to provide a psychophysical theory of subjective intensity based on direct measurement (Plateau, 1872). He provided eight artists with two disks—one painted black and the other white—and instructed them to paint a gray disk midway between them. He reported that the resulting eight gray disks were almost identical, despite the fact that they were painted under different conditions of illumination. Plateau assumed each artist mixed his gray paint to obtain a gray such that the ratio of the subjective intensity of white to gray equaled the subjective intensity of gray to black. Assuming a similar result for any pair of gray disks, then yields,

$$\frac{\psi(d)}{\psi(m)} = \frac{\psi(m)}{\psi(e)},$$

where d and e are the disks provided for midway judgment, m is the midway disk painted by the artist, and ψ is a function that measures subjective grayness. Because by physics, the *ratios* of physical light of gray disks (or of white and gray disks, black and gray disks, etc.) do not vary with illumination (and thus, for example, the ratios of physical light from the gray and white disks were the same in each artist's studio), Plateau concluded that his experiment established the following law for his stimuli:

Preserved Equal Ratios: For all gray disks (including black and white) d and e ,

$$\frac{\varphi(d)}{\varphi(m)} = \frac{\varphi(m)}{\varphi(e)} \text{ iff } \frac{\psi(d)}{\psi(m)} = \frac{\psi(m)}{\psi(e)},$$

where m is the gray disk produced midway between d and e , and φ is a function that measures the physical intensity of grays.

Plateau then showed the Preserved Equal Ratios implied that

$$\psi = r\varphi^s$$

for some positive r and s , that is, the Preserved Equal Ratios implied subjective intensity is a power function of physical intensity. The following is an argument showing this under the necessary assumptions that φ and ψ are onto \mathbb{R}^+ and are strictly monotonically related, that is, for all x and y in X ,

$$\varphi(x) < \varphi(y) \text{ iff } \psi(x) < \psi(y).$$

The argument, which uses a well-known functional equation, goes as follows:

Suppose the Preservation of Equal Ratios. Let x, y, u , and v be arbitrary elements of X such that

$$\frac{\varphi(x)}{\varphi(y)} = \frac{\varphi(u)}{\varphi(v)} \text{ iff } \frac{\psi(x)}{\psi(y)} = \frac{\psi(u)}{\psi(v)}. \quad (1)$$

Defined Ψ on \mathbb{R}^+ as follows: For all z in X ,

$$\Psi(\varphi(z)) = \psi(z).$$

It then follows from the assumptions about φ and ψ that Ψ is a strictly increasing function from \mathbb{R}^+ onto \mathbb{R}^+ . Let β and r be arbitrary positive reals. Because φ is onto \mathbb{R}^+ , let a, b, c , and d be elements of X such that

$$r = \varphi(a), \beta = \varphi(b), 1 = \varphi(c), \text{ and } d = \varphi(\beta r).$$

Because

$$\frac{r}{1} = \frac{\beta r}{\beta},$$

it follows that

$$\frac{\varphi(a)}{\varphi(c)} = \frac{\varphi(d)}{\varphi(b)}.$$

Thus by Equation 1,

$$\frac{\psi(a)}{\psi(c)} = \frac{\psi(d)}{\psi(b)},$$

and therefore by the definition Ψ ,

$$\frac{\Psi(\varphi(a))}{\Psi(\varphi(c))} = \frac{\Psi(\varphi(d))}{\Psi(\varphi(b))},$$

that is,

$$\frac{\Psi(r)}{\Psi(1)} = \frac{\Psi(\beta r)}{\Psi(\beta)}. \quad (2)$$

Letting

$$K(\beta) = \frac{\Psi(\beta)}{\Psi(1)},$$

Equation 2 becomes

$$K(\beta)\Psi(r) = \Psi(\beta r). \quad (3)$$

Equation 3 is a well-known, elementary functional equations whose solutions Ψ have the form $\Psi(r) = \lambda r^\gamma$, where λ and γ are fixed positive reals and r is an arbitrary positive real. Thus, by the definition of Ψ , for all z in X ,

$$\psi(z) = \lambda \varphi(z)^\gamma$$

for some positive λ and γ . \square

Plateau's argument for a power law has two major holes: Both of these involve the use and interpretation of "midway:" First, he did not check whether the artists' midway grays had the mathematical structural properties that specify a midway operation, the most important of which for empirical purposes being bisymmetry and commutativity:

A binary operation \ominus on X is said to satisfy *bisymmetry* if and only if for all x, y, u , and v in X ,

$$\ominus[\ominus(x, y), \ominus(u, v)] = \ominus[\ominus(x, u), \ominus(y, v)],$$

and \ominus is said to satisfy *commutativity* if and only if for all x and y in X ,

$$\ominus(x, y) = \ominus(y, x).$$

The second problem with Plateau's approach is that, even if his empirical midway operation satisfied all the mathematical, structural properties of a "midway" operation, (see Definition 2 below), then there is still no objective evidence to argue that the participant is using the midway operation to produce subjectively equal ratios, as oppose, for example, subjectively equal differences. Plateau apparently realized this, for Falmagne (1985) notes,

In a footnote in his paper, we read "Fechner's formula leads to this consequence that, when the overall illumination increases, the differences in sensation remain constant; it seemed to me more rational, in order to explain the invariance of the general effect of the picture, to postulate *a priori* the constancy of the ratios and not the differences of the sensations." (Plateau, 1872, pp. 382–383.)
[Translation by Flamgne, 1985; the emphasis is Falmagne's.]

4 Bisection Operations

Operations that have the formal properties of the *geometric mean* \sqrt{xy} on \mathbb{R}^+ (or equivalently the formal properties of the *arithmetic mean* $\frac{1}{2}(x + y)$ on \mathbb{R}) are called *bisection operations*. The following is an axiomatization of them.

Definition 2 $\langle X, \preceq, \ominus \rangle$ is said to be a *bisection structure with bisection operation* \ominus if and only if \ominus is a binary operation on X and the following five conditions hold for all u, x, y , and z in X :

Continuum: $\langle X, \preceq \rangle$ and $\langle \mathbb{R}^+, \leq \rangle$ are isomorphic.

Solvability: If $u \ominus z \preceq y \preceq x \ominus z$ then there exists v in X such that $v \ominus z = y$.

Monotonicity: $x \preceq y$ if and only if $x \ominus z \preceq y \ominus z$.

Bisymmetry: $(u \ominus x) \ominus (y \ominus z) = (u \ominus y) \ominus (x \ominus z)$.

Commutativity: $x \ominus y = y \ominus x$. \square

Theorem 1 Suppose $\langle X, \preceq, \ominus \rangle$ is a bisection structure and \mathcal{S} is the set of isomorphisms of $\langle X, \preceq, \ominus \rangle$ onto $\langle \mathbb{R}^+, \leq, B \rangle$, where B is the geometric mean. Then \mathcal{S} is a log-interval scale.

Proof. Follows from Theorem 10 of Section 6.9 of Krantz, Luce, Suppes, and Tversky (1971). \square

Given the ambiguity about whether a given bisection operation should be given a geometric or arithmetic representation, a natural line of inquiry is to present subjects with two bisection tasks that yield functions M_R and M_D on a set of stimuli such that for all x, y, z, u, v , and w in X , $M_R(x, z) = y$ if and only if the ratio of subjective intensities of x to y is the same as the ratio of subjective intensities of y to z , and $M_D(u, w) = v$ if and only if the difference of subjective intensities of u to v is the same as the difference of subjective intensities of v to w . Surprisingly, several empirical studies show $M_R = M_D$. Pfanzagl (1968) comments:

Other inquiries have shown that the values of the arithmetic scale are linearly related to the logarithms of the geometric scale (Torgerson, [1961]; Ekman, 1962; Ekman and Künnapas, 1962[a,b]). The natural explanation of this phenomenon is that in these cases the subjects are unable to distinguish between arithmetic and geometric bisection: Regardless whether the subjects are asked to bisect a given interval from a to b [$a \phi b$] such that the ratio $a : a \phi b$ equals the ratio $a \phi b : b$ or such that the interval from a to $a|b$ [the midpoint of the interval from a to b] equal the interval from $a|b$ to b , they always perform the same operation. This is also suggested by experiments of Garner (1954). If this were true, [by a previous theorem] a logarithmic relationship would exist between the arithmetic and geometric scales. Intuitively this is obvious: If both operations are in fact identical and the operation is one time mapped into the arithmetic mean and the other time into the geometric mean, the values of the first scale are related to the logarithms of the values of the second scale. (pg. 127)

The experiments of Garner (1954) are discussed below. The following theoretical assumptions are inspired by similar, but different, assumptions Garner made:

1. M_1 and M_2 are bisection operations.
2. There is a true subjective intensity scale \mathcal{S} which measures $\langle X, \preceq, M_1 \rangle$ and $\langle X, \preceq, M_2 \rangle$.
3. \mathcal{S} is a log-interval scale.

It follows from Assumptions 1 and 2 and Theorem 1 that M_1 and M_2 are isomorphic. Assumption 3 is used in to show $M_1 = M_2$:

Theorem 2 *Assume Statements 1 to 3 just above. Then $M_1 = M_2$.*

Proof. Let ψ be in \mathcal{S} . Because \mathcal{S} is a log-interval scale, it follows by Theorem 3.12 of Luce and Narens (1985) that r and s in the real interval $(0, 1)$ can be found such that for all x and y in X ,

$$\psi(M_1(x, y)) = \begin{cases} \psi(x)^r \psi(y)^{1-r} & \text{if } \psi(x) \leq \psi(y) \\ \psi(x)^s \psi(y)^{1-s} & \text{if } \psi(x) \geq \psi(y). \end{cases}$$

Because M_1 is commutative, $r = s$. Let \ominus' be the operation defined on \mathbb{R}^+ by $u \ominus' v = u^r v^{1-r}$. Then ψ is an isomorphism of $\langle X, \preceq, M_1 \rangle$ onto $\langle \mathbb{R}^+, \leq, \ominus' \rangle$. By Theorem 1, $\langle X, \preceq, M_1 \rangle$ is isomorphic to $\langle \mathbb{R}^+, \leq, \ominus \rangle$, where \ominus is the geometric mean operation. Thus $\langle \mathbb{R}^+, \leq, \ominus' \rangle$ and $\langle \mathbb{R}^+, \leq, \ominus \rangle$ are isomorphic, and it is not difficult to show that this can only happen when $r = \frac{1}{2}$. Thus ψ^{-1} is an isomorphism of $\langle \mathbb{R}^+, \leq, \ominus \rangle$ onto $\langle X, \preceq, M_1 \rangle$. A similar argument shows that ψ^{-1} is an isomorphism of $\langle \mathbb{R}^+, \leq, \ominus \rangle$ onto $\langle X, \preceq, M_S \rangle$. Thus $M_1 = M_2$. \square

Theorem 2 provides a theory for the experimental results mentioned in the above quote of Pfanzagl's: Subjects are measuring the stimuli on (or with respect to) a log-linear scale (or an equivalent scale, for example an interval scale). However, most psychophysical researchers believe or assume—some explicitly and others implicitly—subjects are using a ratio scale.

5 Torgerson's Conjecture

Torgerson (1961) discusses several experiments involving direct judgments of subjective intensity and concludes,

These results are all consistent with the notion that the subject perceives only a single quantitative relation between stimuli. When this relation is interpreted as either a psychological distance or a psychological ratio, it can be shown that the subjective magnitudes obey the properties of the corresponding commutative group—the addition group for the distance interpretation and the multiplication group the ratio interpretation. (pg. 205)

I will call this conclusion *Torgerson's Conjecture*, because it is my and others' view that it does not follow from the experimental data Torgerson presented. In particular, the data were not rich enough to show that "the subjective magnitudes obey the properties of the corresponding commutative group—the addition group for the distance interpretation and the multiplication group the ratio interpretation." Theoretical and experimental results presented below suggest Torgerson's Conjecture is correct for the the experimental situations cited in his article.

It is important to distinguish Torgerson's Conjecture from the proposition that "the judgments of equal ratios corresponded to judgments of equal differences" (used sometimes by Torgerson, 1961) and as well as the proposition "the ordering of ratios is the same as the ordering of differences" (used Birbaum and Mellers, 1978, and other researchers). These latter two propositions are about the equality of two 4-ary relations on stimuli, whereas Torgerson's Conjecture is about the equality of two 2-ary relations on stimuli.

6 Garner's Experiments

Torgerson used experimental results from Garner (1954) to support his conclusion. In his experiments, Garner used fractionation and equisection methods to obtain his stimuli. *Fractionation* methods produce constant ratio sequences and *equisection* methods produce constant difference sequences. Fractionation data are obtained by supplying an initial stimulus t_1 and instructing observers to provide a stimulus t_2 that is a particular fraction f (which could > 1 or < 1 or an integer, for example, one-half or one-quarter) that is in subjective intensity f of t_1 , and then provide stimulus t_3 that is f in terms of subjective intensity of t_2 , and so on. There are variants of this method, for example, providing the observer with two stimuli that provide a fixed subjective intensity ratio (in place of the fraction f) and then asking the observer to adjust another stimuli to a standard so that it and the standard produce the same subjective ratio. For equisection judgments, observers are asked to provide a sequence of stimuli such that the intervals between adjacent stimuli are equal in subjective intensity. Garner had subjects produce both fractionation and equisections in a loudness experiment. The relevant result for this discussion is summarize by Torgerson as follows:

Several years ago, Garner (1954) tried to get subjects first to set a variable stimulus between two standard stimuli so that the successive differences were equal, and second, to set the variable so that the successive ratios were equal. That is, first, so that $V - S_1 = S_2 - V$, and second, so that $V/S_1 = S_2/V$. For most of his

subjects—thirteen out of eighteen—the value set for the variable was the same in the two conditions: Equal subjective intervals were also equal subjective ratios. (*Torgerson, 1961, pg. 204*)

I consider the following assumptions and definitions to be a fair interpretation and idealization of the above result of Garner (1954) and assumptions Garner made:

1. *Physical Assumption* $\langle X, \preceq \rangle$ is a continuum.
2. *(Empirical Assumption)* To each pair of stimuli a and b and for each stimuli c the subject can adjust a stimuli d to satisfy the command, “Find a stimulus d so that the ratio of the loudness of a to b is the same as the ratio of the loudness of c to d .” These adjustments produce a strictly increasing function R_{ab} on $\langle X, \preceq \rangle$ defined by,

$$R_{ab}(c) = d.$$

3. *(Definition)* M_R is defined on X as follows: For all x, y, z in X ,

$$M_R(x, z) = y \text{ iff } R_{xy}(y) = z.$$

4. *(Empirical Assumption)* To each pair of stimuli (x, z) the subject can adjust a stimulus y to satisfy the command, “Find a stimulus y so that the difference of the loudness of x to y is the same as the difference of the loudness of y to z .” When such a y is produced, we write $M_D(x, z) = y$.
5. *(Theoretical Assumption)* Let $\mathfrak{X} = \langle X, \preceq, R_{ab} \rangle_{a,b \in X}$. Then there exist a function ψ and a numerical structure, $\mathfrak{N} = \langle \mathbb{R}^+, \leq, T_{ab} \rangle_{a,b \in X}$, such that (i) ψ is an isomorphism of \mathfrak{X} onto \mathfrak{N} , and (ii) for each a and b in X there exists $c_{ab} \in \mathbb{R}^+$ such that for each s in \mathbb{R}^+ ,

$$T_{ab}(s) = c_{ab}s.$$

Statement 5 corresponds a direct measurement assumption made by Garner that the T_{ab} can be interpreted as ratios. Statement 5 is needed to insure the T_{ab} are multiplications by positive reals. An immediate consequence of Statement 5 is that $s\psi$ is an isomorphism of \mathfrak{X} onto \mathfrak{N} for each s in \mathbb{R}^+ , and it is not difficult to show that the set of isomorphisms of \mathfrak{X} onto \mathfrak{N} is a ratio scale. Under the assumption of Statements 1 to 4, an empirical assumption that is logically equivalent to Statement 5 can be given. The equivalent states that the functions in $\mathcal{R} = \{R_{ab} \mid a \in X \text{ and } b \in X\}$ commute and that \mathcal{R} is closed under function composition. Thus, because Statements 1 to 4 do not

use direct measurement, Statements 1 to 5 can be reformulated to yield and equivalent system of assumptions that do not use direct measurement.⁴

From Statements 1 to 5 and the additional assumption that $\langle M_D, \preceq \rangle$ is a bisection structure, it does not follow that $M_R = M_D$, even with the additional assumption that the set of isomorphisms of

$$\langle X, \preceq, R_{ab}, M_R, M_D \rangle_{a,b \in X} \text{ onto } \langle \mathbb{R}^+, \leq, T_{ab}, M_R^*, M_D^* \rangle_{a,b \in X}$$

is a ratio scale. This follows, because there are such situations where M_R^* is the geometric mean function and M_D^* is arithmetic mean function and therefore $M_R^* \neq M_D^*$. Thus in the context of Statements 1 to 5, the observance of the equality $M_R = M_D$ is an empirical fact, not an analytic conclusion. However, if functions representing subjective differences between stimuli are introduced, then it becomes an analytic conclusion:

6. (*Empirical Assumption*) To each pair of stimuli a and b and for each stimuli c the subject can adjust a stimuli d to satisfy the command, “Find a stimulus d so that the difference of the loudness of a to b is the same as the difference of the loudness of c to d .” These adjustments produce a strictly increasing function D_{ab} on $\langle X, \preceq \rangle$ defined by $D_{ab}(c) = d$.
7. (*Empirical Assumption*) For all x, y, z in X ,

$$M_D(x, z) = y \text{ iff } D_{xy}(y) = z.$$

8. (*Empirical Assumption*) For all a, b, x, y , and z in X ,

$$R_{ab}(M_D(x, z)) = M_D(R_{ab}(x), R_{ab}(z)).$$

9. (*Theoretical Assumption*) There exists a numerical structure

$$\langle \mathbb{R}^+, \leq, R_{ab}^*, D_{ab}^* \rangle_{a,b \in X}$$

such that the set of isomorphisms from

$$\langle X, \preceq, R_{ab}, D_{ab} \rangle_{a,b \in X} \text{ onto } \langle \mathbb{R}^+, \leq, R_{ab}^*, D_{ab}^* \rangle_{a,b \in X}$$

is a ratio scale.

Theorem 3 *Assume Statements 1 to 4 and 6 to 9. Then the following three statements are true:*

1. $M_R = M_D$.

⁴The proof of this is similar to the proof of Theorem 6 below.

2. Torgerson's Conjecture: (i) For all a and b in X , there exist c and d in X such that $D_{ab} = R_{cd}$; and (ii) for all a and b in X , there exist c and d in X such that $R_{ab} = D_{cd}$.
3. For each a and b in X there exist s in and t in \mathbb{R}^+ such that R_{ab}^* is the function that is multiplication by s and D_{ab}^* is the function that is multiplication by t .

Proof. Statement 1 follows from Statement 2. Statements 2 and 3 are immediate consequences of the following theorem. \square

Theorem 4 Suppose $\langle X, \preceq, V_j \rangle_{j \in J}$ is such that $\langle X, \preceq \rangle$ is a continuum, F is a strictly \preceq -increasing function from X onto X , V_j is a set, function, or relation on X for each j in J , and \mathcal{S} is a ratio scale of isomorphisms of

$$\langle X, \preceq, F, V_j \rangle_{j \in J} \text{ onto } \langle \mathbb{R}^+, \leq, F^*, V_j^* \rangle_{j \in J}.$$

Then there exists r in \mathbb{R}^+ such that for all t in \mathbb{R}^+ ,

$$F^*(t) = rt.$$

Proof. Because \mathcal{S} is a ratio scale, for all s and t in \mathbb{R}^+ ,

$$F^*(st) = sF^*(t),$$

and it is well-known that the only solution this functional equation is $F^*(t) = rt$ for some r in \mathbb{R}^+ . \square

The following theorem is an immediate consequence of Theorem 4.

Theorem 5 Suppose $\mathfrak{X} = \langle X, \preceq, F, G, V_j \rangle_{j \in J}$ is a relational structure and $\langle X, \preceq \rangle$ is a continuum, and F and G are strictly \preceq -increasing functions from X onto X . Suppose $\mathfrak{N} = \langle \mathbb{R}^+, \leq, F^*, G^*, V_j^* \rangle_{j \in J}$ is a numerical structure and \mathcal{S} is a ratio scale of isomorphisms of \mathfrak{X} onto \mathfrak{N} . Then F and G commute, that is $F * G = G * F$, where $*$ is the operation of function composition.

Proof. By Theorem 4, F^* and G^* are multiplications by positive reals and thus F^* and G^* commute. By isomorphism, F and G commute. \square

The following theorem shows that under the assumption of Statements 1 to 4 and 6 to 8, the conclusion of Theorem 3 follows with theoretical assumption Statement 9 replaced by an empirical condition (Statement (ii) of Theorem 6 below).

Theorem 6 Assume Statements 1 to 4 and 6 to 8. Let $\mathcal{R} = \{R_{ab} \mid a \in X \text{ and } b \in X\}$. Then the following two statements are logically equivalent:

- (i) *Statement 9.*
- (ii) *The elements of \mathcal{R} commute, \mathcal{R} is closed under function composition, and for all a and b in X , R_{ab} is a symmetry of \mathfrak{X} (Definition 1).*

Proof. We first note that it follows from Statement 2 that \mathcal{R} is *homogeneous*, that is, for each x and y in X , there exists R in \mathcal{R} such that $R(x) = y$.

Assume (i). By Statement 9, let $\mathfrak{N} = \langle \mathbb{R}^+, \leq, R_{ab}^*, D_{ab}^* \rangle_{a,b \in X}$ be a numerical structure and ψ be an isomorphism of \mathfrak{X} onto \mathfrak{N} . Let $\mathcal{R}^* = \{R_{ab}^* \mid a \in X \text{ and } b \in X\}$. Because \mathcal{R} is homogeneous, it follows by isomorphism that \mathcal{R}^* is homogeneous. Because by Statement 1 $\langle X, \preceq \rangle$ is a continuum, it then follows that \mathcal{R}^* is the set of all multiplications by positive reals. Because \mathcal{R}^* is the set of multiplications by all positive reals, it follows that the elements of \mathcal{R}^* commute and are closed under function composition. By isomorphism, the elements of \mathcal{R} commute and are closed under function composition. It is easy to verify that for each α and β in \mathcal{S} , that $\alpha * \beta^{-1}$ is an isomorphism of \mathfrak{N} onto itself and all isomorphisms of \mathfrak{N} onto itself can be obtained in this way. Because \mathcal{S} is a ratio scale, it is easy to show that

$$\mathcal{A} = \{\alpha * \beta^{-1} \mid \alpha \in \mathcal{S} \text{ and } \beta \in \mathcal{S}\}$$

is the set of multiplication by positive reals. Because \mathcal{R}^* is also the set of all multiplications by positive reals, $\mathcal{A} = \mathcal{R}^*$. By isomorphism, \mathcal{R} is the set of symmetries of \mathcal{X} .

Assume (ii). Then by (ii) elements of \mathcal{R} commute, \mathcal{R} is closed under the operation of function composition, $*$, and elements of \mathcal{R} are symmetries of X . It has already been shown that \mathcal{R} is homogeneous. By statement 2, $R_{ab} * R_{ba}$ is the identity function on X for a and b in X . It then follows by Theorem 4.2 of Narens (1981) (or Lemma 4.4 and Theorem 4.3 of Narens, 1985) that there exists a ratio scale of isomorphisms of X onto some numerical structure \mathfrak{N} having the form $\mathfrak{N} = \langle \mathbb{R}^+, \leq, R_{ab}^*, D_{ab}^* \rangle_{a,b \in X}$. \square

7 Magnitude Estimation

S. S. Stevens (1948, 1950) introduced direct measurement methods generically called “magnitude estimation.” *Ratio Magnitude Estimation* methods collect data in the form (x, \mathbf{p}, y) , where x and y are intensity stimuli and \mathbf{p} is a word that describes the real number p . In practice, \mathbf{p} is an integer or a fraction, but theoretically it could name any real number. Instructions like “Adjust the stimulus y so that y is p times x ” are typically used to gather the data (x, \mathbf{p}, y) . In direct measurement methodologies, the data are measured through procedures that take seriously the semantics of the instructions

or at least a good part of the semantics; for example Stevens measures using functions ψ so that such that for all x and y in X ,

$$\frac{\psi(y)}{\psi(x)} = p \text{ iff } (x, p, y),$$

and Garner measures using functions ψ so that there exists a positive real c such that for all x and y in X ,

$$\frac{\psi(y)}{\psi(x)} = c \text{ iff } (x, p, y).$$

Narens (1996) presented a theory for measuring magnitude estimation data using the representational theory of measurement. This axiomatic theory is based on the primitive $f_p(x) = y$, where p is fixed and x and y vary.

7.1 Narens (1996) Theory

Definition 3 $\mathfrak{S} = \langle X, \succeq, f_1, \dots, f_n, \dots \rangle_{n \in \mathbb{I}^+}$ is said to be a *Stevens' magnitude structure* if and only if $\langle X, \preceq \rangle$ is a continuum, for each n in \mathbb{I}^+ , f_n is a \preceq -strictly increasing function from X onto X , and the following five conditions hold, where p is an arbitrary positive integer and t, x, y , and z are arbitrary elements of X :

- (1) $x \preceq f_p(x)$.
- (2) f_1 is the identity function on X .
- (3) For all t in X , $f_p(t) \prec f_q(t)$ iff $p < q$.
- (4) For all x and t in X , if $t \prec x$, then there exist m in \mathbb{I}^+ and z in X such that

$$t \prec z \prec x \text{ and } f_{m+1}(t) = f_m(z).$$

- (5) (*Multiplicative Property*) If $r = q \cdot p$, then $f_r = f_q * f_p$.

Conditions (1), (2), (3), and (5) are straightforward and are very reasonable idealizations for Stevens' methods of ratio estimations. Condition (4) is a condition like an Archimedean axiom (of measurement theory).

Theorem 7 Let $\mathfrak{S} = \langle X, \succeq, f_1, \dots, f_n, \dots \rangle_{n \in \mathbb{I}^+}$. Then the following two statements are equivalent:

1. \mathfrak{S} is a Stevens Magnitude Structure.
2. There exists an isomorphism ψ from \mathfrak{S} onto $\mathfrak{N} = \langle \mathbb{R}^+, \leq, T_n \rangle_{n \in \mathbb{I}^+}$, where for each n in \mathbb{I}^+ , T_n is the function on \mathbb{R}^+ such that for all x in \mathbb{R}^+ , $T_n(x) = nx$.

Proof. Statement 1 implies Statement 2 by Theorem 12 of Narens (1996). Assume Statement 2. Let ψ be an isomorphism from \mathfrak{S} onto \mathfrak{N} . Then Conditions 1 to 4 of Definition 3 immediately follow by isomorphism. To show Condition (5), let p and q be arbitrary elements of \mathbb{I}^+ . Then $T_p * T_q = T_{pq}$, and thus by the isomorphism ψ^{-1} , $f_p * f_q = f_{pq}$. \square

Theorem 8 *Let \mathfrak{S} and \mathfrak{N} be as in Theorem 7. Then the set of isomorphisms of \mathfrak{X} onto \mathfrak{N} is a ratio scale.*

Proof. Theorem 12 of Narens (1996). \square

Narens (1996) conjectured that the multiplicative principle was too strong of a restriction to hold empirically, and suggested that it be replaced by a weaker principle it implies. The weaker principle is called the “Commutative Property.”

Definition 4 $\mathfrak{M} = \langle X, \succeq, f_1, \dots, f_n, \dots \rangle_{n \in \mathbb{I}^+}$ is said to be a *magnitude structure* if and only if $\langle X, \preceq \rangle$ is a continuum, for each n in \mathbb{I}^+ , f_n is a strictly increasing \preceq -function from X onto X , Conditions (1) to (4) of Definition 3 hold, and the following condition holds:

$$(5') \text{ (Commutative Property)} \quad f_p * f_q = f_q * f_p.$$

Theorem 9 *The following two statements are logically equivalent.*

1. \mathfrak{M} is a magnitude structure.
2. There exists a numerical structure $\mathfrak{N} = \langle \mathbb{R}^+, \leq, M_1, \dots, M_n, \dots \rangle_{n \in \mathbb{I}^+}$ such that (i) M_1 is multiplication by 1 and for each $n \neq 1$ in \mathbb{I}^+ , M_n is the function of multiplication by some real number c_n , and (ii) the set of isomorphisms from \mathfrak{M} onto \mathfrak{N} is a ratio scale.

Proof. Theorems 4 and 5 of Narens (1996). \square

It is easy to reformulate Definitions 3 and 4 so that they apply to fraction names of the form $\frac{1}{n}$, $n \in \mathbb{I}^+$, and show the obvious reformulations of Theorems 7, 8, and 9 for such fractions. The following is another reformulation for a situation involving a name t for each real number t .

Definition 5 $\mathfrak{X} = \langle X, \preceq, f_t \rangle_{t \in \mathbb{R}^+}$ is said to be a *full magnitude structure* if and only if $\langle X, \preceq \rangle$ is a continuum, for each t in \mathbb{R}^+ , f_t is a \preceq -strictly increasing function from X onto X , and the following three conditions hold, where s and t are arbitrary elements of \mathbb{R}^+ and x, y, z are arbitrary elements of X :

- (1) Either $x \prec f_t(x)$ for all x in X , or $x = f_t(x)$ for all x in X , or $f_t(x) \prec x$ for all x in X .

- (2) For all u in X , $f_s(u) \prec f_t(u)$ iff $s < t$.
- (3) (*Commutative Property*) $f_s * f_t = f_t * f_s$.
- (4) (*Homogeneity*) For all u and v in X , there exists r in \mathbb{R}^+ such that $f_r(u) = v$. \square

A numerical representing structure for a full magnitude structure \mathfrak{X} with f_t interpreted via isomorphism as multiplication by a positive real is called a “full multiplicative representing structure:”

Definition 6 $\mathfrak{N} = \langle \mathbb{R}^+, \leq, M_t \rangle_{t \in \mathbb{R}^+}$ is said to be a *a full multiplicative representing structure* if and only if for each t in \mathbb{R}^+ , M_t is the function of multiplication by some real number c_t . \square

Theorem 10 Suppose $\mathfrak{X} = \langle X, \succeq, f_t \rangle_{t \in \mathbb{R}^+}$ is a full magnitude structure. Then there exists a full multiplicative representing structure $\mathfrak{N} = \langle \mathbb{R}^+, \leq, M_t \rangle_{t \in \mathbb{R}^+}$ such that the set of isomorphisms of \mathfrak{X} onto \mathfrak{N} is a ratio scale.

Proof. One shows that $\{f_t | t \in \mathbb{R}^+\}$ is a commutative homogeneous group of symmetries of \mathfrak{X} and then use Lemma 4.4 and Theorem 4.3 of Narens (1985).⁵ \square

⁵Let $\mathcal{G} = \{f_t | t \in \mathbb{R}^+\}$. Then $\langle \mathcal{G}, *\rangle$ is a group: Let f and g be arbitrary elements of \mathcal{G} and x be an element of X . Let y be an arbitrary element of X and by homogeneity, let θ in \mathcal{G} be such that $\theta(x) = y$.

By homogeneity, let h in \mathcal{G} be such that $h(x) = (f * g)(x)$. To show \mathcal{G} is closed under function composition, it needs only to be shown that $h(y) = (f * g)(y)$. By the Commutative Property (Definition 5),

$$\begin{aligned} (f * g)(y) &= (f * g)(\theta(x)) = (f * g * \theta)(x) \\ &= (f * \theta * g)(x) = (\theta * f * g)(x) = \theta[f * g](x) \\ &= \theta[h(x)] = (\theta * h)(x) = (h * \theta)(x) \\ &= h[\theta(x)] = h(y). \end{aligned}$$

By homogeneity, let k in \mathcal{G} be such that $k(x) = f^{-1}(x)$. To show f^{-1} is in \mathcal{G} , it needs only to be shown that $k(y) = f^{-1}(y)$. By the Commutative Property (Definition 5),

$$\begin{aligned} (f^{-1})(y) &= (f^{-1})(\theta(x)) = (f^{-1} * \theta)(x) \\ &= \theta[f^{-1}(x)] = \theta[k(x)] = (\theta * k)(x) \\ &= (k * \theta)(x) = k[\theta(x)] = k(y). \end{aligned}$$

Because f and f^{-1} is in \mathcal{G} and \mathcal{G} is closed under function composition, the identity function on X is in \mathcal{G} .

It follows from the definition of “full magnitude structure” (Definition 5) that each element of \mathcal{G} is a symmetry of \mathfrak{X} . Theorem 10 then follows from Lemma 4.4 and Theorem 4.3 of Narens (1985).

Theorem 11 Suppose $\mathfrak{X} = \langle X, \succeq, f_t \rangle_{t \in \mathbb{R}^+}$ is a full magnitude structure. Then the following two statements are equivalent:

1. (Multiplicative Property) For all r, s , and t in \mathbb{R}^+ , $r \cdot s = t$ if and only if $f_r * f_s = f_t$.
2. There exists a full multiplicative representing structure $\mathfrak{N} = \langle \mathbb{R}^+, \leq, M_t \rangle_{t \in \mathbb{R}^+}$ such that for each $t \in \mathbb{R}^+$, M_t is multiplication by t , and the set of isomorphisms of \mathfrak{X} onto \mathfrak{N} is a ratio scale.

Proof. Easy consequence of Theorem 10. \square

Note that the Multiplicative Property of Theorem 11 implies commutativity as given in Definition 5.

8 Empirical Tests

8.1 Multiplicative and Commutative Properties

Ellermeier and Faulhammer (2000) tested the Multiplicative and Commutative Properties in a standard loudness magnitude production paradigm and found the Multiplicative Property to fail and the Commutative Property to hold. For example, they found that when subjects were asked to produce a stimulus x that was three times as loud as the stimulus t , and then later were asked to produce a stimulus y that was twice as loud as x , and at an unrelated time to produce a stimulus z that was 6 times as loud as t , that $z \neq y$. Results of Ellermeier and Faulhammer indicate that y was generally much larger than z in the sense that the y was about the production of twelve times t , thus producing a failure of the Multiplicative Condition. However, the productions of (1) a stimulus x that was three times as loud a stimulus t followed later by the production of a stimulus y that was twice as loud as x , and (2) the production of c that was twice as loud as t followed at a later time by the production of d that was three times as loud as c , generally yielded y and d of approximately the same size, confirming the holding of the Commutative Property. A parallel study in vision by Peißner (1999) yielded similar results. Steingrimsson and Luce (2005b) also observed the holding of the Commutativity property in a loudness experiment for instructions producing stimuli that were two times and three times other stimuli. Zimmer (2004) in a loudness production experiment also found Multiplicative Property to fail and the Commutative Property to hold. Her tests were for producing stimuli that were one-half as loud, one-third as loud, and one-sixth as loud as other stimuli.

8.2 Commutativity of Ratios with Differences

Most psychophysicists assume that psychological intensities like loudness are measurable by a ratio scale. Theorem 5 shows that when strictly increasing functions f and g on a continuum of stimuli are measured by a common ratio scale of isomorphisms, they both are represented as multiplications. Because multiplications commute, it follows by isomorphism that the functions f and g commute. This suggests that a ratio magnitude function R produced through ratio production instructions should commute with a difference magnitude function D produced through difference production instructions. The following experiment of Ellermeier, Narens, and Dielmann (2003) tested this conclusion.

The subjects were asked to make ratio and difference productions. The ratio production instructions were the same as in Ellermeier and Faulhammar (2000): Subjects were presented a tone and asked to make the second tone p times as loud as the first one (denoted by $R_p(t) = x$, where t is the first tone and x is the result of the adjustment) for $p = 2, 3$. Ellermeier, et al. (2003) page 72, write the following about their difference instruction: "... we cannot simply say for difference productions 'adjust the second tone so that the loudness difference is p' without providing a unit. Therefore, the following *difference matching* instruction $D_{a,b}$, $a \prec b$, was implemented where $D_{a,b}(x) = y$ holds if and only if the subject adjusts a stimulus y such that 'the difference in loudness between y and x is the same as the difference between b and a' ."

Two different choices of a, b were used in the experiment: 50, 58 dB SPL and 50, 62 dB SPL. Starting from the base level of 65 dB SPL this yields four tests of the commutativity of R_p and $D_{a,b}$ illustrated in Figure 1.

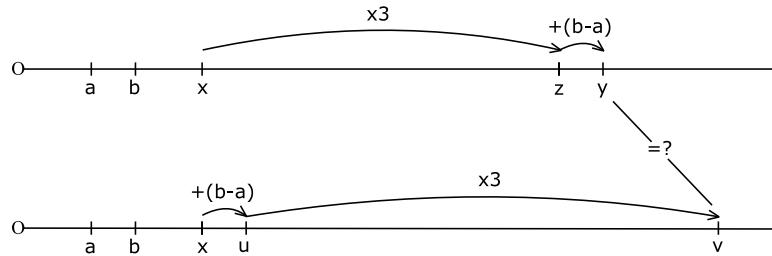


Figure 1. Test of the commutativity of R_3 and $D_{a,b}$. If commutativity holds, the two orders of chaining these operations should not make a difference, and multiple productions of y and v should be statistically indistinguishable.

Four the the six subjects satisfied commutativity between their ratio and difference productions.⁶

9 Theoretical Implications

In psychophysics, it is often assumed that judgments of intensity within a modality are made on the same scale. (Some researchers assume a common scale across modalities.) For example, Garner states:

1. It is assumed that there is a true loudness scale, which functions for all kinds of loudness judgments.
2. It is assumed that all fractionation judgments (for a given verbalized fraction) are made with the same ratio or fraction, regardless of loudness level. (*Garner, 1954, pg. 17*)

In the representational theory such “same scale” or “common scale” assumptions correspond to assuming that the judgments are primitive relations of measurement structure, for example, “all fractionation judgments (for a given verbalized fraction) are made with the same ratio or fraction” corresponds to a primitive R (“the same ratio”) that is a function from stimuli onto stimuli.

The assumption made by many experimenters that ratio productions can be represented by numerical ratios, implies, given a rich enough set of ratio productions, that the ratio productions can be measured on a ratio scale.

The assumption that ratio productions can be measured on a ratio scale, implies by Theorem 4 that such productions, when taken as primitives of the qualitative measurement structure and measured on a ratio scale, are necessarily represented by numerical ratios. However Theorem 4 shows more: Any strictly \preceq -increasing function from the set of stimuli onto itself that is a primitive of the qualitative measurement structure is represented by a numerical ratio. In particular, difference productions will be represented as ratios. Thus Torgerson’s conjecture is a consequence of ratio and difference productions being primitives of the same qualitative structure that is measurable by a ratio scale (Theorem 5).

With regards to measurement on a ratio scale, there are three theories to be tested:

- (1) Ratio productions are measurable on a ratio scale.
- (2) Difference productions are measurable on a ratio scale (or equivalently, difference productions are measurable on a translation scale (Definition 1)).

⁶That two subjects were not in line with this outcome is consistent with the observation that some subjects sometimes distinguish perceptual ratios and differences (e.g. Schneider, 1980; Birnbaum, 1982).

(3) Ratio and difference productions are measurable on a common ratio scale.

The above studies of Ellermeier and Faulhammer (2000), Peissner (1999), Steingrimsson and Luce (2005b), and Zimmer (2004) provides empirical confirmation for (1), and the study of Ellermeier, Narens, and Dielmann (2003) provides empirical confirmation for (3) and therefore also (1) and (2). There have been many other empirical studies showing that subjects' behavior for ratio judgments is the same for difference judgments; but these studies do not provide credible evidence that either kind of judgment can be represented as a ratios or differences.

Because of the empirical failure of the Multiplicative Condition (Ellermeier and Faulhammer 2000, Peissner 1999, and Zimmer 2004), direct measurement methodologies that rely on interpreting number names in instructions as numbers for measuring stimuli should be looked upon with incredulity, especially since the empirical studies were performed in domains that were regarded by the leading proponents of direct measurement techniques as being best suited for those techniques. However, the data produced by direct measurement instructions can still be an important tool in rigorous psychophysical research if analyzed by other methodologies—the above studies and theory involving the Commutative Property being a testament to this point. In my opinion, given the failing of the Multiplicative Property and the holding of the Commutative Property, the mapping of number names onto the ratios they name via a ratio scale is probably of little value for basic psychophysics, although it still may be of importance in parts of applied psychophysics where communication about subjective intensity is important. However, preserved relationships among functions obtained through the use of number names in instructions can be very valuable for basic psychophysics.

10 Combined Monaural and Binaural Judgments

In the past, direct measurement studies of loudness have been conducted using monaural and binaural presentations. These studies suffer, like almost all of the direct measurement studies in the literature, from false or non-grounded assumptions about how the number words used in the instructions to participants are to be interpreted in the numerical representation of the data. In recent theoretical and empirical articles, Luce and Steingrimsson studied loudness productions in which sounds were presented to the left ear only, the right ear only, and simultaneously to both ears, including productions that were responses to direct measurement instructions (Luce 2002, 2004; Steingrimsson

and Luce, 2005a, 2005b). They took care to explicitly state how number words in the instructions to participants are related to the theoretical models they were investigating, and tested empirically their assumptions about the relationship. This section investigates the role of the Commutativity Property for loudness in situations where combined monaural and binaural judgments are made.

Throughout the section, let $\langle X, \preceq \rangle$ be a continuum of sound intensities, and let the notation (x, y) stand for the simultaneous presentation of the sound x to the left ear and the sound y to the right ear. $(x, 0)$ will stand for presenting the sound x in the left ear and no sound in the right ear, and $(0, y)$ will stand for presenting the sound y in the right ear and no sound in the left ear.⁷

Ratio magnitude production instructions l, p and r, p and p can be given for, respectively, the sounds presented only in left ear, sounds presented only in the right ear, and pairs of physically equivalent sounds (z, z) presented simultaneously to both ears. This gives rise to the three qualitative structures,

$$\mathfrak{X}_l = \langle X, \preceq, f_{l,p} \rangle_{p \in \mathbb{R}^+}, \quad \mathfrak{X}_r = \langle X, \preceq, f_{r,p} \rangle_{p \in \mathbb{R}^+}, \quad \text{and} \quad \mathfrak{X} = \langle X, \preceq, f_p \rangle_{p \in \mathbb{R}^+},$$

for measuring the loudness of sounds, respectively, for the left ear only, the right ear only, and both ears simultaneously. For \mathfrak{X}_l , “ $f_{l,p}(x) = z$ ” is to be interpreted as “ z is the adjusted stimulus to the instruction, ‘Find the stimulus that is p times x ’,” and a similar interpretation holds for \mathfrak{X}_r and $f_{r,p}$; and for \mathfrak{X} , “ $f_p(x) = z$ ” is to be interpreted as “ (z, z) is the adjusted stimulus to the instruction, ‘Find the stimulus z that is p times (x, x) ’.”

Assume \mathfrak{X}_l , \mathfrak{X}_r , and \mathfrak{X} are measurable by ratio scales of isomorphisms. (The testing of the commutativity of the primitive ratio production functions of each structure provides via Theorem 4 a method for testing this assumption.)

It follows from the above ratio scale assumption and Theorem 5 that for \mathfrak{X}_l and \mathfrak{X}_r to be measurable on the same ratio scale of isomorphisms, the primitive ratio production functions of \mathfrak{X}_l must commute with the primitive ratio production functions on \mathfrak{X}_r . Similar results hold for the pairs $\mathfrak{X}_l, \mathfrak{X}$ and $\mathfrak{X}_r, \mathfrak{X}$. These commutativity conditions are natural and simple tests to conduct. To my knowledge no experiments of this type has yet been performed. This is unfortunate because such experiments bear directly on the following assumptions of Garner, which many psychophysicists appear to me to use as an intuitive basis—along with ratio scalability—for guiding their loudness research:

1. It is assumed that there is a true loudness scale, which functions for all kinds of loudness judgments.

⁷Some experiments involving ratio judgments p , where p is a name of a fraction < 1 and x is not too far from perceptual threshold, 0 is taken to be the element of X that is at threshold. However, throughout this section it will be assumed that 0 represents no sound.

2. It is assumed that all fractionation judgments (for a given verbalized fraction) are made with the same ratio or fraction, regardless of loudness level. (*Garner, 1954, pg. 17*)

If the above suggested commutativity tests were to fail, I believe that the concept of a “loudness scale” for “all kinds of loudness judgments” would be in great jeopardy. Commutativity between the ratio production primitives of \mathfrak{X}_l and \mathfrak{X} produce the following interesting result:

Theorem 12 *Suppose \sim is a new primitive binary relation on X , where $x \sim z$ stands for, “The stimulus x when presented to the left ear is judged by the participant to be the same loudness as (z, z) , that is, the same loudness as the stimulus z presented simultaneously to both ears.” Suppose \sim is monotonic in the following sense: For all x, y, u , and v in X , if $x \sim (y, y)$ and $u \sim (v, v)$, then*

- (i) $x \preceq u$ if and only if $y \preceq v$, and
- (ii) the function defined by $h(x) = y$ if and only if $x \sim (y, y)$ is a strictly increasing function from X onto X .

Let \mathfrak{C} be the structure,

$$\mathfrak{C} = \langle X, \preceq, \sim, f_{l,p} \rangle_{p \in \mathbb{R}^+}.$$

Suppose \mathcal{S} is a ratio scale of isomorphisms of \mathfrak{C} onto

$$\mathfrak{N} = \langle \mathbb{R}^+, \leq, \sim^*, g_{l,p} \rangle_{p \in \mathbb{R}^+}$$

and \mathfrak{X}_l is homogeneous, that is, for each x and y in X there exists t in \mathbb{R}^+ such that $f_{l,t}(x) = y$. Then there exists a ratio production primitive $f_{l,q}$ of \mathfrak{X}_l such that for all x in X ,

$$f_{l,q}(x) \sim (x, x).$$

Proof. Let $\psi \in \mathcal{S}$. Let α be an arbitrary symmetry of \mathfrak{C} and x and y be arbitrary elements of X . Then

$$h(x) = y \text{ iff } x \sim (y, y) \text{ iff } \alpha(x) \sim (\alpha(y), \alpha(y)) \text{ iff } h(\alpha(x)) = \alpha(y),$$

that is, h is invariant under the symmetries of \mathfrak{X} . Let $\psi(h) = h^*$. Then by isomorphism, h^* is a strictly increasing function from \mathbb{R}^+ onto \mathbb{R}^+ that is invariant under the symmetries of \mathfrak{N} , that is, h^* is invariant under multiplications by positive reals. Thus for each positive real number s , $h^*(sw) = sh^*(w)$ for all real w . The well-known solution to this functional equation is $h^* =$ multiplication by a positive real number. Thus h^* is a symmetry of \mathfrak{N} , and

therefore by isomorphism, h is a symmetry of \mathfrak{X} . Let u and v be elements of X such that $h(u) = v$. By the assumed homogeneity of \mathfrak{X}_l , let $f_{l,q}$ be a primitive production function of \mathfrak{X}_l such that

$$f_{l,q}(u) = v.$$

By Theorem 4, ψ maps $f_{l,q}$ onto a function $g_{l,q}$ that is multiplication by a positive real. Because by isomorphism,

$$h^*(\psi(u)) = \psi(v) \text{ and } g_{l,q}(\psi(u)) = \psi(v),$$

it follows that $h^* = g_{l,q}$, because two multiplications that have the same value on a positive real are the same multiplication. By isomorphism, $h = f_{l,q}$, and thus, by the definition of h , $f_{l,q}(x) \sim (x, x)$ for each x in X . \square

11 Discussion

It is important to distinguish *direct measurement data* from *direct measurement*. Direct measurement data can be used in various ways, for example, as input into a direct measurement procedure, or as in this article, a means for testing qualitative axioms of a representational measurement theory for magnitude production.

Two forms of direct measurement are considered in the article: A strong form in which ratio productions named by number words are interpreted veridically as the numerical ratios they name; and a weak form in which the ratio productions named by number words may have interpretations as ratios that are different from numerical ratios they name. Both forms assume that the responses to instructions to produce ratios are represented numerically by ratios, and thus the word “ratio” (and supposedly the participants concept associated with it) is being “directly” represented. The strong form additionally “directly represents” the number mentioned in the instruction by itself. The previous sections show that both forms are testable, with the Multiplicative Condition being the appropriated test for the strong form, and the Commutative Condition being the appropriate test for the weak form. Subsection 8.1 cites experimental literature that is confirmatory for the Commutative Condition and disconfirmatory for the Multiplicative Condition.

Section 7 provides an axiomatic theory for the numerical representations produced by both forms of direct measurement. The axiomatic theory eliminates the need for assuming anything is being “directly represented,” and thus allows for a purely behavioral approach to ratio production data. It isolates the Multiplicative and Commutative Conditions as the critical axioms for empirical testing.

There is a sizable literature showing that subjects do not distinguish between ratios and differences in a variety of direct measurement tasks. However, except for one, they do not provide rigorous experimental tests for the part of Torgerson's Conjecture which require ratio and difference judgments to be simultaneously represented as ratios. The exception is the study of Ellermeier, Narens, and Dielmann (2003), which shows commutativity between ratio and difference production functions. By Theorem 5, such commutativity is a necessary condition for production functions to be simultaneously represented as ratios.

A consequence of Theorem 5 is that for a ratio production function R and a difference production function D to be represented by the same ratio scale, then R and D must commute. Section 10 uses this result to formulate testable relationships between judgments of loudness for sounds presented to a single ear, say the left ear, and judgments of loudness for sounds simultaneously presented to both ears: For such judgments to be measured by a common ratio scale it is necessary for the left ear loudness production functions commute with the both ears loudness production functions. Also under the assumption of a common scale, a remarkably simple relationship was shown between loudness of sounds z presented to left ear only and the simultaneous presentation z to the left and right ears: There is a ratio production function f for sounds presented to left ear such that for each sound z that is presented simultaneously in each ear, the sound $f(z)$ presented to left ear only is judged to be the same loudness as the simultaneous presentation.

The above results demonstrate the power of the assumption that subjective intensity is ratio scalable, particularly when subjective intensities under different instructions or conditions are represented on a common ratio scale. For ratio production paradigms, the ratio scalability of the behavior results from the commutativity of the ratio production functions. As discussed later, I do not view this commutativity corresponding to part of the participant's judging of ratios; Instead I view it as a result of a construction due to the experimenter or as an axiom about observable behavior formulated by a theoretician. Similarly, I view the ratio production functions as being a product of the experiment, and without additional strong assumptions about the participant's processing of intensity and judgmental strategy, I do not see them having phenomenological correlates in the participant. I consider the following to be plausible assumptions about the participant's perception and representation of intensity and judgmental strategy:

It is assumed that the perceptions of stimuli is totally ordered by \preceq' in terms of their subjective intensity and for all stimuli x and y , $x \preceq y$ if and only if $\pi(x) \preceq' \pi(y)$, where $\pi(z)$ stands for the perception of the stimulus z . Pairs of stimuli share an experienceable phenomenological relation, or an EPR

for short, that the participant is able to assign a lexical item, for example, a number word, or a phenomenological item, for example, the perceptions of two fixed sounds that has the phenomenological relation under consideration. Such pairs are called *named* pairs, with the lexical item or phonemonological item being the *name* of their EPR. The EPRs are assumed to be \preceq' -strictly increasing functions from the set of perceptions of stimuli on itself, and it is assumed that for each pair of perceptions of stimuli, there is an EPR satisfying the pair.

Note that the above assumes that an ERPs occurs in consciousness in conjunction with an associated pair of stimuli. No assumption is made as to whether ERPs can occur in other ways.

Call a group of functions from a nonempty set onto itself a *similarity group* if and only if it is isomorphic to the group of multiplications by positive reals. Let X' be the set of perceptions of X . It follows from the above plausible assumptions and results of Narens (1994) that for each EPR, P , there is a unique similarity group \mathcal{G}_P such that each element of \mathcal{G}_P is a symmetry of $\langle X', \preceq', P \rangle$ ⁸. By methods similar to ones presented in earlier sections of the article, the following two statements can be shown:

1. For EPRs P and Q to have a common similarity group of symmetries, it is necessary and sufficient for the elements of \mathcal{G}_P to commute with the elements of \mathcal{G}_Q .
2. Let \mathcal{P} be the set of EPRs. Then the symmetries in \mathcal{G}_P and \mathcal{G}_Q commute for all P and Q in \mathcal{P} if and only if the symmetries of

$$\mathfrak{X}' = \langle X', \preceq', T \rangle_{T \in \mathcal{P}}$$

form a similarity group.

Suppose each EPR arises from the comparison of percept pairs used in making a magnitude production judgment. Let $\langle \mathcal{G}, * \rangle$ be the symmetry group of \mathfrak{X}' . It then follows by the plausible assumptions above, the assumption that the behavioral production functions commute, and theorems of previous sections that

- (1) $\langle \mathcal{G}, * \rangle$ is a similarity group, and
- (2) $\mathcal{G} = \mathcal{P}$.

Although elements of \mathcal{P} are part of the participant's phenomenological experience, the set \mathcal{P} is not. The status of $*$ (applied to elements of \mathcal{G}) in relation to the participant's phenomenological experience is unclear at this time, although

⁸It follows from the same results of Narens (1994) that there are additional symmetries of $\langle X', \preceq', P \rangle$ that are not in \mathcal{G}_P .

it appears to me likely that the above plausible assumptions together with the right empirical tests would likely contradict the proposition, “* is part of the participants phenomenological experience.”

The above considerations lead naturally to the following question: “Is there some principled evolutionary reason for $\langle \mathcal{G}, * \rangle$ being a similarity group?” A satisfactory answer to this appears to me to be difficult. $\langle \mathcal{G}, * \rangle$ being a similarity group has desirable properties for the efficient processing of information (Narens, 1997a), and it matches up with physical similarity groups that occur in our natural perceptual environment (Narens, 1997b). However, these features alone do not get to the heart of the matter. What is needed is (i), a mathematical formulation of the evolutionary problem that is to be solved by having $\langle \mathcal{G}, * \rangle$ be a similarity group; (ii), the imbedding of the evolutionary problem of (i) in an evolutionary game; and (iii), showing that $\langle \mathcal{G}, * \rangle$ being a similarity group is one of a few, or better, *the* equilibrium solution to the evolutionary game under a wide variety of evolutionary dynamics. (iii) appears to me to be the most difficult step, because natural formulations of (ii) require infinite dimensional spaces, and thus the well-developed mathematical methods for dynamics in finite dimensional spaces or locally compact situations do not apply.

REFERENCES

- Birnbaum, M. H. (1982). Controversies in psychological measurement. In B. Wegener (Ed.), *Social attitudes and psychological measurement*, 401–485. Hillsdale, N.J.: Erlbaum.
- Birnbaum, M. H. and Mellers, B. A. (1978). Measurement and the mental map. *Perception & Psychophysics*, 23, 403–408.
- Ekman, G. (1962). Measurement of moral judgments: A comparison of scaling methods. *Perceptual and Motor Skills*, 15, 3–9.
- Ekman, G., and Künnapas, T. (1962a). Scales of aesthetic value. *Perceptual and Motor Skills*, 14, 19–26.
- Ekman, G., and Künnapas, T. (1962b). Measurement of aesthetic value by “direct” and “indirect” methods. *Scan. J. Psychol.*, 3, 33–39.
- Ellermeier, W. and Faulhammer, G. (2000). Empirical evaluation of axioms fundamental to Stevenss ratio-scaling approach: I. Loudness production. *Perception & Psychophysics*, 62(8), 1505–1511.
- Ellermeier, W., Narens, L., and Dielmann, B. (2003). Perceptual ratios, differences, and the underlying scale. In B. Berglund and E. Borg (Eds.), *Fechner Day 2003. Proceedings of the 19th annual meeting of the International Society for Psychophysics*, 71–76. Stockholm, Sweden: International Society for Psychophysics.
- Falmagne, J.-C. (1985). *Elements of Psychophysical Theory*. New York: Oxford University Press.
- Garner, W. R. (1954). A technique and a scale for loudness measurement. *J. Acoust. Soc. Am.*, 26, 73–88.
- Krantz, D. H., R. D. Luce, P. Suppes, & A. Tversky (1971). *Foundations of Measurement*, Vol. I. New York: Academic Press.
- Levelt, W., J., M., Riemersma, J., B., and Bunt, A., A. (1972). Binaural additivity of loudness. *British Journal of Mathematical and Statistical Psychology*, 25, 51–68.
- Luce, R. D. (2002). A psychophysical theory of intensity proportions, joint presentations, and matches. *Psychological Review*, 109, 520–532.
- Luce, R., D. (2004). Symmetric and asymmetric matching of joint presentations. *Psychological Review*, 111, 446–454.
- Luce, R. D. and Narens, L. (1985). Classification of concatenation structures by scale type. *Journal of Mathematical Psychology*, 29, 1–72.

- Narens, L. (1985). *Abstract Measurement Theory*. Cambridge, Mass.: The MIT Press.
- Narens, L. (1994). The measurement theory of dense threshold structures. *Journal of Mathematical Psychology*, 38, 301–321.
- Narens, L. (1996). A theory of magnitude estimation. *Journal of Mathematical Psychology*, 40, 109–129.
- Narens, L. (1997a). On subjective intensity and its measurement. In A. A. J. Marley (Ed.), *Choice, Decision, and Measurement: Essays in Honor of R. Duncan Luce*, 189–206. Mahwah: Lawrence Erlbaum and Associates.
- Narens, L. (1997b). Surmising Cognitive Universals for Extraterrestrial Intelligences. In C. B. Cosmovici, S. Bowyer, and D. Werthimer (Eds.). *Astronomical and Biochemical Origins and the Search for Life in the Universe*. Bologna, Italy: Editrice Compositori.
- Peißner, M. (1999). *Experimente zur direkten Skalierbarkeit von gesehenen Helligkeiten*. Unpublished master's thesis, Universität Regensburg.
- Pfanzagl, J. (1968). *Theory of Measurement*. New York: John Wiley & Sons.
- Plateau, M., J. (1872). Sur la mesure des sensations physiques, et sur la loi qui lie l'intensité de ces sensations à l'intensité de la cause excitante. *Bulletin de l'Académie Royale des Sciences, es Lettres et des Beaux Arts de Belgique*, 33, 376–388.
- Schneider, B. A. (1980). A technique for the nonmetric analysis of paired comparisons of psychological intervals. *Psychometrika*, 45, 357–372.
- Steingrimsson, R., and Luce, R. D. (2005a). Empirical evaluation of a model of global psychophysical judgments: I. Behavioral properties of summations and productions. *Journal of Mathematical Psychology*, under revision.
- Steingrimsson, R., and Luce, R. D. (2005b). Empirical evaluation of a model of global psychophysical judgments: II. Behavioral properties linking summations and productions. *Journal of Mathematical Psychology*, under revision.
- Torgerson, W. S. (1961). Distances and ratios in psychological scaling. *Acta Psychologica*, 19, 201–205.
- Zimmer, K. (2004). Examining the validity of numerical ratios in loudness fractionation. *Perception & Psychophysics*, in press.