Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Mathematical Psychology 50 (2006) 290–301
Journal of Mathematical Psychology
www.elsevier.com/locate/jmp

Symmetry, direct measurement, and Torgerson's conjecture

Louis Narens

Department of Cognitive Sciences, University of California, Irvine, USA

Received 29 December 2004; received in revised form 15 November 2005
Available online 21 February 2006

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 phenomenon that subjects do not distinguish between ratios and differences in a variety of direct measurement tasks.

Keywords: Measurement theory; Bisection; Symmetry; Direct measurement; Torgerson's conjecture; Psychophysics

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 the 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. Some of these structural assumptions were empirically investigated at the time (e.g., Birnbaum & Elmasian, 1977; Mellers, Davis, & Birnbaum, 1984), and modeled axiomatically (Miyamoto, 1983).

Within the last dozen 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 Togerson'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

E-mail address: lnarens@uci.edu.

0022-2496/$ - see front matter
doi:10.1016/j.jmp.2005.12.007

numbers, $\mathbb{I}^+$ for the positive integers, and $*$ for the operation of function composition.

A *measurement scale* $\mathscr{S}$ is a nonempty set of functions from a nonempty set of object X into $\mathbb{R}$. Throughout the article it is assumed that X is a set of physical objects that is totally ordered by $\preccurlyeq$ and that each measuring function φ in $\mathscr{S}$ is a strictly $\preccurlyeq$-increasing function from X onto either $\mathbb{R}$ or $\mathbb{R}^+$; that is, φ is an isomorphism of $\langle X, \preccurlyeq \rangle$ onto either $\langle \mathbb{R}, \leqslant \rangle$ or $\langle \mathbb{R}^+ \leqslant \rangle$. (In the literature, measuring functions in $\mathscr{S}$ are often called "scales.") $\mathscr{S}$ is said to be a:

- *ratio scale* if and only if for all ψ in $\mathscr{S}$,
  $\mathscr{S} = \{r\psi \mid r \in \mathbb{R}^+\}$,
- *interval scale* if and only if for all ψ in $\mathscr{S}$,
  $\mathscr{S} = \{r\psi + s \mid r \in \mathbb{R}^+ \text{ and } s \in \mathbb{R}\}$,
- *log-interval scale*[1] if and only if for all ψ in $\mathscr{S}$,
  $\mathscr{S} = \{s\psi^r \mid r \in \mathbb{R}^+ \text{ and } s \in \mathbb{R}^+\}$,
- *translation scale*[2] if and only if for all ψ in $\mathscr{S}$,
  $\mathscr{S} = \{\psi + r \mid r \in \mathbb{R}\}$.

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

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

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

The relations $\preccurlyeq, R_1, \ldots, R_n$ 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 $\psi \in \mathscr{S}$. 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 $\psi(y) = s$.

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

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

[3]Sometimes 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, even though 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. Generalization of this to all pairs 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 Midway Ratio Law: For all gray disks (including black and white) d and e,

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

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 Midway Ratio Law implied that

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

for some positive r and s, that is, the Preserved Midway Ratio Law 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,

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

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

Assume the Preserved Midway Ratio Law. Let x, y, u, and v be arbitrary elements of X. It will first be shown that,

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

Assume

$$\frac{\varphi(x)}{\varphi(y)} = \frac{\varphi(u)}{\varphi(v)}. \tag{2}$$

It will be shown that $\psi(x)/\psi(y) = \psi(u)/\psi(v)$. By Eq. (2) and because φ and ψ are onto the positive reals, let a and b in X be such that

$$\frac{\varphi(x)}{\varphi(a)} = \frac{\varphi(a)}{\varphi(v)} \quad \text{and} \quad \frac{\varphi(u)}{\varphi(b)} = \frac{\varphi(b)}{\varphi(y)}.$$

Then by the Preserved Midway Ratio Law,

$$\frac{\psi(x)}{\psi(a)} = \frac{\psi(a)}{\psi(v)} \quad \text{and} \quad \frac{\psi(u)}{\psi(b)} = \frac{\psi(b)}{\psi(y)}.$$

Thus

$$\varphi(x)\varphi(v) = \varphi(a)^2 \quad \text{and} \quad \varphi(u)\varphi(y) = \varphi(b)^2, \tag{3}$$

and

$$\psi(x)\psi(v) = \psi(a)^2 \quad \text{and} \quad \psi(u)\psi(y) = \psi(b)^2. \tag{4}$$

By Eq. (2), $\varphi(x)\varphi(v) = \varphi(u)\varphi(y)$. Thus it follows from Eq. (3) that $\varphi(a) = \varphi(b)$. Because φ is strictly increasing, it is a one-to-one function. Therefore $a = b$. Thus by Eq. (4),

$$\psi(x)\psi(v) = \psi(a)^2 = \psi(b)^2 = \psi(u)\psi(y),$$

and therefore, because $a = b$,

$$\frac{\psi(x)}{\psi(y)} = \frac{\psi(u)}{\psi(v)}.$$

The implication, $\psi(x)/\psi(y) = \psi(u)/\psi(v)$ implies $\varphi(x)/\varphi(y) = \varphi(u)/\varphi(v)$ follows by a similar argument. Thus Eq. (1) has been shown.

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

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

Because

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

it follows that

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

Thus by Eq. (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,

$$\frac{\Psi(r)}{\Psi(1)} = \frac{\Psi(\beta r)}{\Psi(\beta)}. \tag{5}$$

Letting

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

Eq. (5) becomes

$$K(\beta)\Psi(r) = \Psi(\beta r). \tag{6}$$

Eq. (6) 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.

Plateau's argument for a power law has two major holes. Both 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 $\ominus$, where $\ominus(x, y) = z$ if and only if z is "midway" between x and z. For empirical purposes, the most important properties are 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, 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 Falmagne, 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 axiomatizes them.

Definition 2. $\langle X, \preccurlyeq, \ominus \rangle$ 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:* $\langle X, \preccurlyeq \rangle$ and $\langle \mathbb{R}^+, \leqslant \rangle$ are isomorphic.
- *Solvability:* If $u \ominus z \preccurlyeq y \preccurlyeq x \ominus z$ then there exists v in X such that $v \ominus z = y$.
- *Monotonicity:* $x \preccurlyeq y$ if and only if $x \ominus z \preccurlyeq y \ominus z$.
- *Bisymmetry:* $(u \ominus x) \ominus (y \ominus z) = (u \ominus y) \ominus (x \ominus z)$.
- *Commutativity:* $x \ominus y = y \ominus x$.
- *Idempotence:* $x \ominus x = x$.

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

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

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 (Ekman, 1962; Ekman & Künnapas, 1962a, 1962b; Torgerson, 1961). 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. (p. 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 $\mathscr{S}$ which measures $\langle X, \preccurlyeq, M_1 \rangle$ and $\langle X, \preccurlyeq, M_2 \rangle$.
3. $\mathscr{S}$ 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 the following theorem to show that $M_1 = M_2$:

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

Proof. Let ψ be in $\mathscr{S}$. Because $\mathscr{S}$ 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,

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

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 $\langle X, \preccurlyeq, M_1 \rangle$ onto $\langle \mathbb{R}^+, \leqslant, \ominus' \rangle$. By Theorem 1, $\langle X, \preccurlyeq, M_1 \rangle$ is isomorphic to $\langle \mathbb{R}^+, \leqslant, \ominus \rangle$, where $\ominus$ is the geometric mean operation. Thus $\langle \mathbb{R}^+, \leqslant, \ominus' \rangle$ and $\langle \mathbb{R}^+, \leqslant, \ominus \rangle$ 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 $\langle \mathbb{R}^+, \leqslant, \ominus \rangle$ onto $\langle X, \preccurlyeq, M_1 \rangle$. A similar argument shows that ψ^{-1} is an isomorphism of $\langle \mathbb{R}^+, \leqslant, \ominus \rangle$ onto $\langle X, \preccurlyeq, M_2 \rangle$.Thus $M_1 = M_2$. □

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, many psychophysical researchers believe or assume—some explicitly and others implicitly—subjects are using a ratio scale.

5. Torgerson's conjecture

Torgerson (1961) examines several experiments involving direct judgments of subjective intensity. He 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. (p. 205)

It is important to distinguish Torgerson's conjecture, as formulated above, 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." These latter two propositions are about the equality or ordering of two 4-ary relations on stimuli. Torgerson's conjecture, however, is about the equality of pairs 2-ary relations on stimuli; that is, for all stimuli x and y and all relations α, if $x\alpha y$ stands for the judged ratio of y to x is p (where p is a particular number), then there exist a number q and a relation β such that,

$x\beta y$ iff the judged difference between x and y is q and

$\alpha = \beta$.

From an experimental point of view, judgments of the above 4-ary relations are of a different nature from the judgments of the above 2-ary relations; and from a mathematical point of view, the mathematical theory leading to the conclusion that "the judgments of equal ratios corresponded to judgments of equal differences" is very different for the two kinds of relations. Whether or not the above distinctions between 2-ary versus 4-ary relations are of practical import in experimental applications is not clear at this time.

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 stimulus 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 described 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, p. 204)

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

1. *Physical assumption* $\langle X, \preccurlyeq \rangle$ is a continuum.
2. (*Empirical assumption*) To each pair of stimuli a and b and for each stimulus c the subject can adjust a stimulus 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} from $\langle X, \preccurlyeq \rangle$ onto $\langle X, \preccurlyeq \rangle$ defined by

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

3. (*Empirical assumption*) (i) For all x and y in X, there exist a and b in X such that $R_{ab}(x) = y$; and (ii) there exists c and d in X such that for all z in X, $R_{cd}(z) = z$.
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 $\mathfrak{X} = \langle X, \preccurlyeq, R_{ab} \rangle_{a,b \in X}$. Then there exist a function ψ and a numerical structure, $\mathfrak{R} = \langle \mathbb{R}^+, \leqslant, T_{ab} \rangle_{a,b \in X}$, such that (i) ψ is an isomorphism of $\mathfrak{X}$ onto $\mathfrak{R}$, 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 $\mathfrak{R}$ for each s in $\mathbb{R}^+$, and it is not difficult to show that the set of isomorphisms of $\mathfrak{X}$ onto $\mathfrak{R}$ is a ratio scale. Under the assumption of Statements 1–4, an empirical assumption that is logically equivalent to Statement 5 can be given. The equivalent states that the functions in $\mathscr{R} = \{R_{ab} \mid a \in X \text{ and } b \in X\}$ commute and that $\mathscr{R}$ is closed under function composition. Thus, because Statements 1–4 do not use direct measurement, Statements 1–5 can be reformulated to yield an equivalent system of assumptions that do not use direct measurement.[4]

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

Definition: M_R is defined on X as follows: for all x, y, z in X,

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

From Statements 1–5 and the additional assumption that $\langle X, \preccurlyeq, M_D\rangle$ is a bisection structure, it does not follow that $M_R = M_D$, even with the additional assumption that for some $M_R^\star$ and $M_D^\star$ the set of isomorphisms of

$$\langle X, \preccurlyeq, R_{ab}, M_R, M_D\rangle_{a,b\in X} \quad \text{onto}$$
$$\langle \mathbb{R}^+, \leqslant, T_{ab}, M_R^\star, M_D^\star\rangle_{a,b\in X}$$

is a ratio scale. This is because there are such situations where $M_R^\star$ is the geometric mean function and $M_D^\star$ is arithmetic mean function and therefore $M_R^\star \neq M_D^\star$. Thus in the context of Statements 1–5, 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 stimulus c the subject can adjust a stimulus 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 $\langle X, \preccurlyeq\rangle$ defined by $D_{ab}(c) = d$.
7. (*Empirical assumption*) For all x, y, z in X,

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

8. (*Theoretical assumption*) There exists a numerical structure

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

   such that the set of isomorphisms from

   $$\langle X, \preccurlyeq, R_{ab}, D_{ab}\rangle_{a,b\in X} \quad \text{onto } \langle \mathbb{R}^+, \leqslant, R_{ab}^\star, D_{ab}^\star\rangle_{a,b\in X}$$

   is a ratio scale.

Theorem 3. *Assume Statements* 1–4 *and* 6–8. *Then the following three propositions are true*:

1. $M_R = M_D$.
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* $R_{ab}^\star$ *is the function that is multiplication by s and* $D_{ab}^\star$ *is the function that is multiplication by t.*

Proof. Proposition 1 follows from Proposition 2. Propositions 2 and 3 are immediate consequences of the following theorem. □

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

$$\langle X, \preccurlyeq, F, V_j\rangle_{j\in J} \quad \textit{onto } \langle \mathbb{R}^+, \leqslant, F^\star, V_j^\star\rangle_{j\in J}.$$

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

$$F^\star(t) = rt.$$

Proof. Because $\mathscr{S}$ is a ratio scale of isomorphisms, it follows that for all s and t in $\mathbb{R}^+$,

$$F^\star(st) = sF^\star(t),$$

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

The following theorem is an immediate consequence of Theorem 4.

Theorem 5. *Suppose* $\mathfrak{X} = \langle X, \preccurlyeq, F, G, V_j\rangle_{j\in J}$ *is a relational structure and* $\langle X, \preccurlyeq\rangle$ *is a continuum, and F and G are strictly* $\preccurlyeq$*-increasing functions from X onto X. Suppose* $\mathfrak{N} = \langle \mathbb{R}^+, \leqslant, F^\star, G^\star, V_j^\star\rangle_{j\in J}$ *is a numerical structure and* $\mathscr{S}$ *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, $F^\star$ and $G^\star$ are multiplications by positive reals and thus $F^\star$ and $G^\star$ commute. By isomorphism, F and G commute. □

The following theorem shows that under the assumption of Statements 1–4 and 6 and 7, the conclusion of Theorem 3 follows with the theoretical assumption Statement 8 replaced by an empirical condition (Statement (ii) of Theorem 6 below).

Theorem 6. *Assume Statements* 1–4 *and* 6 *and* 7. *Let* $\mathscr{R} = \{R_{ab} \mid a \in X \text{ and } b \in X\}$. *Then the following two statements are logically equivalent*:

(i) *Statement* 8.
(ii) *The elements of* $\mathscr{R}$ *commute,* $\mathscr{R}$ *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 3 that $\mathscr{R}$ is *homogeneous,* that is, for each x and y in X, there exists R in $\mathscr{R}$ such that $R(x) = y$.

Assume (i). By Statement 8, let

$$\mathfrak{X} = \langle X, \preccurlyeq, R_{ab}, D_{ab}\rangle_{a,b\in X}, \quad \mathfrak{N} = \langle \mathbb{R}^+, \leqslant, R_{ab}^\star, D_{ab}^\star\rangle_{a,b\in X},$$

$\mathscr{S}$ be the set of isomorphisms of $\mathfrak{X}$ onto $\mathfrak{N}$, and $\psi \in \mathscr{S}$. Let $\mathscr{R}^\star = \{R_{ab}^\star \mid a \in X \text{ and } b \in X\}$. By Theorem 4, each element of $\mathscr{R}^\star$ is a multiplication by a positive real. Because $\mathscr{R}$ is homogeneous, it follows by isomorphism that $\mathscr{R}^\star$ is homogeneous. Because by Statement 1 $\langle X, \preccurlyeq\rangle$ is a continuum, it then follows that $\mathscr{R}^\star$ is the set of all multiplications by positive reals. Because $\mathscr{R}^\star$ is the set of multiplications by all positive reals, it follows that the elements of $\mathscr{R}^\star$ commute and are closed under function composition. By isomorphism, the elements of $\mathscr{R}$ commute and are closed under function composition. It is easy to verify that for each α and β in $\mathscr{S}$, 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 $\mathscr{S}$ is a ratio scale, it is easy to show that

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

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

Assume (ii). Then the elements of $\mathscr{R}$ commute, $\mathscr{R}$ is closed under the operation of function composition, $*$, and elements of $\mathscr{R}$ are symmetries of $\mathfrak{X}$. It has already been shown that $\mathscr{R}$ is homogeneous. By part (ii) of Statement 2, the identity function on X is in $\mathscr{R}$. 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 $\langle X, \preccurlyeq, R_{ab}, D_{ab}\rangle_{a,b\in X}$ onto some numerical structure $\mathfrak{N}'$ of the form $\mathfrak{N}' = \langle \mathbb{R}^+, \leqslant, R'_{ab}, D'_{ab}\rangle_{a,b\in X}$. □

7. Magnitude estimation and production

Stevens (1948, 1950) introduced direct measurement methods of magnitude estimation and production. *Ratio magnitude estimation* and *production* methods generally collect data in the form $(x, \boldsymbol{p}, y)$, where x and y are intensity stimuli and $\boldsymbol{p}$ is a word that describes the real number p. In practice, $\boldsymbol{p}$ is an integer or a fraction, but theoretically it could name any real number. Instructions like "For the stimuli x and y, estimate the number $\boldsymbol{p}$ such that y is p times x" (*magnitude estimation*), or "Adjust the stimulus y so that y is p times x" (*magnitude production*) are typically used to gather the data $(x, \boldsymbol{p}, y)$. 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 for all x and y in X,

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

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

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

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

7.1. Narens (1996) theory

Definition 3. $\mathfrak{S} = \langle X, \preccurlyeq, f_{\boldsymbol{1}}, \ldots, f_{\boldsymbol{n}}, \ldots\rangle_{n\in\mathbb{I}^+}$ is said to be a *Stevens' magnitude structure* if and only if $\langle X, \preccurlyeq\rangle$ is a continuum, for each n in $\mathbb{I}^+$, $f_{\boldsymbol{n}}$ is a $\preccurlyeq$-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) $x \preccurlyeq f_{\boldsymbol{p}}(x)$.
(2) $f_{\boldsymbol{1}}$ is the identity function on X.
(3) For all t in X, $f_{\boldsymbol{p}}(t) \prec f_{\boldsymbol{q}}(t)$ iff $p < q$.
(4) For all x and t in X, if $t \prec x$, then
  (i) there exist n in $\mathbb{I}^+$ such that $x \prec f_{\boldsymbol{n}}(t)$, and
  (ii) there exist m in $\mathbb{I}^+$ and u in X such that

$$t \prec u \prec x \quad \text{and} \quad f_{\boldsymbol{m+1}}(t) = f_{\boldsymbol{m}}(u).$$

(5) (*Multiplicative Property*) If $r = q \cdot p$, then $f_{\boldsymbol{r}} = f_{\boldsymbol{q}} * f_{\boldsymbol{p}}$.

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* $\mathfrak{S} = \langle X, \preccurlyeq, f_{\boldsymbol{1}}, \ldots, f_{\boldsymbol{n}}, \ldots\rangle_{n\in\mathbb{I}^+}$. *Then the following two statements are equivalent*:

1. $\mathfrak{S}$ *is a Steven magnitude structure.*
2. *There exists an isomorphism* ψ *from* $\mathfrak{S}$ *onto* $\mathfrak{N} = \langle \mathbb{R}^+, \leqslant, T_n\rangle_{n\in\mathbb{I}^+}$, *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 1–4 of Definition 3 immediately follow by isomorphism. To show Condition (5), let p and q be arbitrary elements if $\mathbb{I}^+$. Then $T_{\boldsymbol{p}} * T_{\boldsymbol{q}} = T_{\boldsymbol{pq}}$, and thus by the isomorphism ψ^{-1}, $f_{\boldsymbol{p}} * f_{\boldsymbol{q}} = f_{\boldsymbol{pq}}$. □

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). □

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

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

(5′) (*Commutative Property*) $f_{\boldsymbol{p}} * f_{\boldsymbol{q}} = f_{\boldsymbol{q}} * f_{\boldsymbol{p}}$.

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

1. $\mathfrak{M}$ *is a magnitude structure.*

2. *There exists a numerical structure* $\mathfrak{N} = \langle \mathbb{R}^+, \leqslant, M_1, \ldots, M_n, \ldots \rangle_{n \in \mathbb{I}^+}$ *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). □

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

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

(1) Either $x \prec f_{\boldsymbol{t}}(x)$ for all x in X, or $x = f_{\boldsymbol{t}}(x)$ for all x in X, or $f_{\boldsymbol{t}}(x) \prec x$ for all x in X.
(2) For all u in $X, f_{\boldsymbol{s}}(u) \prec f_{\boldsymbol{t}}(u)$ iff $s < t$.
(3) (*Commutative Property*) $f_{\boldsymbol{s}} * f_{\boldsymbol{t}} = f_{\boldsymbol{t}} * f_{\boldsymbol{s}}$.
(4) (*Homogeneity*) For all u and v in X there exists r in $\mathbb{R}^+$ such that $f_{\boldsymbol{r}}(u) = v$.

A numerical representing structure for a full magnitude structure $\mathfrak{X}$ with $f_{\boldsymbol{t}}$ interpreted via isomorphism as multiplication by a positive real is called a "full multiplicative representing structure."

Definition 6. $\mathfrak{N} = \langle \mathbb{R}^+, \leqslant, M_t \rangle_{t \in \mathbb{R}^+}$ 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.

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

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

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 ask to produce a stimulus y that was twice as loud as x, and at an unrelated time to produce a stimulus z that was six 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 12 times t, thus producing a failure of the Multiplicative Condition. However, the productions of (i) 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 (ii) 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 (2005) 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

Many psychophysicists and psychophysical models 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

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

By homogeneity, let h in $\mathscr{G}$ be such that $h(x) = (f * g)(x)$. To show $\mathscr{G}$ 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 $\mathscr{G}$ be such that $k(x) = f^{-1}(x)$. To show f^{-1} is in $\mathscr{G}$, 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} are in $\mathscr{G}$ and $\mathscr{G}$ is closed under function composition, the identity function on X is in $\mathscr{G}$.

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

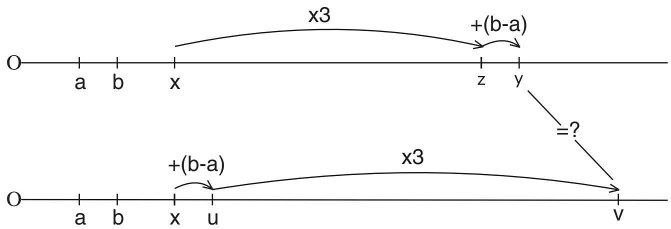

Fig. 1. Test of the commutativity of R_3 and $D_{a,b}$. If commutativity holds, the two orders of chaining these operations should not make a difference, and multiple productions of y and v should be statistically indistinguishable.

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 Faulhammer (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, p. 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 Fig. 1.

Four of 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, p. 17)

[6]That 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. Birnbaum, 1982; Schneider, 1980).

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 $\preccurlyeq$-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)).
(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 (2005) provides empirical confirmation for (1), and the study of Ellermeier et al. (2003) provides empirical confirmation for (3) and therefore also (1) and (2).[7]

Because of the empirical failure of the Multiplicative Condition (Ellermeier & Faulhammer, 2000; Peißner, 1999; Zimmer, 2005), direct measurement methodologies that rely on assigning a number to a stimulus because it corresponds to a number named in the instructions should be looked upon with incredulity, especially since the above 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

[7]These studies reach their conclusions through tests of critical qualitative axioms. Other researchers, for example, Birnbaum (1978, 1982), Hagerty and Birnbaum (1978), and Veit (1978) have used quantitative fits to the data to argue that the sole underlying operation of magnitude judgment is a subtractive operation.

theory involving the Commutative Property being a testament to this point. In my opinion, given the failing of the Multiplicative Property, the mapping of number names onto the ratios they name 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 an important consideration. However, the holding of the Commutative Property suggests that ratio scales obtained through the use of number names in instructions and proper measurement techniques may still be a valuable tool 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 nongrounded 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 & 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 $\langle X, \precsim \rangle$ 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.[8]

Ratio magnitude production instructions $l, \boldsymbol{p}$ and $r, \boldsymbol{p}$ and $\boldsymbol{p}$ 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,

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

for measuring the loudness of sounds, respectively, for the left ear only, the right ear only, and both ears simultaneously. For $\mathfrak{X}_l$, "$f_{l,\boldsymbol{p}}(x) = z$" 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 $f_{r,\boldsymbol{p}}$; and for $\mathfrak{X}$, "$f_{\boldsymbol{p}}(x) = z$" 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 have 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.
>
> 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, p. 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 11. *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) $x \precsim u$ if and only if $y \precsim v$, 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,

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

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

$$\mathfrak{N} = \langle \mathbb{R}^+, \leqslant, \sim^\star, g_{l,\boldsymbol{p}} \rangle_{p \in \mathbb{R}^+}$$

[8]Some experiments involving ratio judgments $\boldsymbol{p}$, where $\boldsymbol{p}$ 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.

and 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,*

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

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

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

that is, h is invariant under the symmetries of $\mathfrak{X}$. Let $\psi(h) = h^\star$. Then by isomorphism, $h^\star$ is a strictly increasing function from $\mathbb{R}^+$ onto $\mathbb{R}^+$ that is invariant under the symmetries of $\mathfrak{R}$, that is, $h^\star$ is invariant under multiplications by positive reals. Thus for each positive real number s, $h^\star(sw) = sh^\star(w)$ for all real w. The well-known solution to this functional equation is $h^\star$ = multiplication by a positive real number. Thus $h^\star$ is a symmetry of $\mathfrak{R}$, and therefore by isomorphism, h is a symmetry of $\mathfrak{X}$. By a hypothesis of the theorem, let $f_{l,q}$ be a primitive production function of $\mathfrak{X}_l$ such that

$$f_{l,q}(x) = y.$$

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

$$h^\star(\psi(x)) = \psi(y) \quad \text{and} \quad g_{l,q}(\psi(x)) = \psi(y),$$

it follows that $h^\star = g_{l,q}$, 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, $u \sim (f_{l,q}(u), f_{l,q}(u))$ for each u in X. □

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. Section 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, many of these rely on direct measurement methodologies and thus are flawed because they do not provide credible evidence that either kind of judgment can be represented as ratios or differences. Others (e.g., Birnbaum & Elmasian, 1977) provide credible evidence for the simultaneous representation of ratio and difference judgments by transforming the data to fit an appropriate model. The study of Ellermeier et al. (2003) is unique in that verifies empirically the holding of commutativity between qualitative ratio and difference production functions. By Theorem 5, this qualitative condition of commutativity is necessary 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. 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.

Acknowledgments

The author thanks R. Duncan Luce, Ragnar Steingrimsson, and an anonymous reviewer for helpful comments on earlier versions of the article.

References

Birnbaum, M.H. (1978). Differences and ratios in psychophysical measurement. In N.J. Castellan Jr., & F. Restle (Eds.), *Cognitive Theory* (*Vol. 3*). Hillsdale, NJ: Erlbaum.

Birnbaum, M. H. (1982). Controversies in psychological measurement. In B. Wegener (Ed.), *Social attitudes and psychological measurement* (pp. 401–485). Hillsdale, NJ: Erlbaum.

Birnbaum, H. H., & Elmasian, R. (1977). Loudness ratios and differences involve the same psychophysical operation. *Perception and Psychophysics*, *22*, 383–391.

Ekman, G. (1962). Measurement of moral judgments: A comparison of scaling methods. *Perceptual and Motor Skills*, *15*, 3–9.

Ekman, G., & Künnapas, T. (1962a). Scales of aesthetic value. *Perceptual and Motor Skills*, *14*, 19–26.

Ekman, G., & Künnapas, T. (1962b). Measurement of aesthetic value by "direct" and "indirect" methods. *Scandinavian Journal of Psychology*, *3*, 33–39.

Ellermeier, W., & Faulhammer, G. (2000). Empirical evaluation of axioms fundamental to Stevenss ratio-scaling approach: I. Loudness production. *Perception and Psychophysics*, *62*(8), 1505–1511.

Ellermeier, W., Narens, L., & Dielmann, B. (2003). Perceptual ratios, differences, and the underlying scale. In B. Berglund, & E. Borg (Eds.), *Fechner Day 2003. Proceedings of the 19th annual meeting of the international society for psychophysics* (pp. 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. *Journal of Acoustical Society America*, *26*, 73–88.

Hagerty, M., & Birnbaum, M. H. (1978). Nonmetric tests of ratio vs. subtractive theories of stimulus comparison. *Perception and Psychophysics*, *24*, 121–129.

Krantz, D. H., Luce, R.D., Suppes, P., & Tversky, A., (1971). *Foundations of measurement* (*Vol. I*). New York: Academic Press.

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., & Narens, L. (1985). Classification of concatenation structures by scale type. *Journal of Mathematical Psychology*, *29*, 1–72.

Mellers, B. A., Davis, D., & Birnbaum, M. H. (1984). Weight of evidence supports one operation for "ratios" and "differences" of heaviness. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 216–230.

Miyamoto, J. M. (1983). An axiomatization of the ratio/difference representation. *Journal of Mathematical Psychology*, *27*, 439–455.

Narens, L. (1985). *Abstract measurement theory*. Cambridge, MA: 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.

Peißner, M., 1999. *Experimente zur direkten Skalierbarkeit von gesehenen Helligkeiten.* Unpublished master's thesis, Universität Regensburg.

Pfanzagl, J. (1968). *Theory of measurement*. New York: Wiley.

Plateau, M. J. (1872). Sur la measure 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., & Luce, R. D. (2005a). Empirical evaluation of a model of global psychophysical judgments: I. Behavioral properties of summations and productions. *Journal of Mathematical Psychology*, submitted for publication.

Steingrimsson, R., & Luce, R. D. (2005b). Empirical evaluation of a model of global psychophysical judgments: II. Behavioral properties linking summations and productions. *Journal of Mathematical Psychology*, submitted for publication.

Torgerson, W. S. (1961). Distances and ratios in psychological scaling. *Acta Psychologica*, *19*, 201–205.

Veit, C. T. (1978). Ratio and subtractive processes in psychophysical judgment. *Journal of Experimental Psychology: General*, *107*, 81–107.

Zimmer, K. (2005). Examining the validity of numerical ratios in loudness fractionation. *Perception and Psychophysics*, *67*, 569–579.