3
Global Psychophysical Judgments of Intensity: Summary of a Theory and Experiments

R. Duncan Luce[1] and Ragnar Steingrimsson[2]

[1] University of California, Irvine
[2] New York University

This chapter has three thrusts: (a) It formulates in a common framework mathematical representations of two global sensory procedures: summation of intensity and the method of ratio production (Luce, 2002, 2004). Until recently, these two topics have not been treated together in the literature. (b) Although the psychophysical representations we arrive at include both free parameters and free functions, a message of this work, especially illustrated in Steingrimsson and Luce (2005a, 2005b),[3] is that one can evaluate the adequacy of the representations without ever estimating either the parameters or the functions, but rather by just evaluating parameter-free behavioral properties that give rise to the representations. (c) A closely related message is that, to the degree that the theory holds, no individual differences arise in the defining behavioral properties except, of course, for the fact that each person has his or her own sense of the relative intensities of two stimuli, that is, the subjective intensity ordering. At the same time, the potential exists for substantial individual differences in the representations in the following sense: there is a strictly increasing psychophysical function and a strictly increasing function that distorts numerical responses but that is not otherwise constrained without additional axioms. The work on the forms of these functions, although quite well developed, is not yet in final manuscript form. Nonetheless, we cover it in some detail in Sections 5 and 6. A number of interesting open problems are listed.

[3] In the remainder of the chapter, the four collaborative articles by Steingrimsson and Luce are identified as SL-I, SL-II, and so forth.

The chapter describes, without proof, the theory and discusses our joint experimental program to test that theory. As of November 2005, some of this work, SL-I and SL-II (see footnote 1), is published, whereas SL-III on the forms of the psychophysical function revision is in press, and SL-IV, on the forms of the weighting function, is nearing completion. Portions of all of these, including much of the experimental work, derive in part from Steingrimsson's (2002) University of California, Irvine dissertation. The network of results and the results of experimental testing reported are summarized in Fig. 4 later in Section 7.2.

We formulate the exposition in terms of loudness judgments about pure tones of the same frequency and phase. However, many other interpretations of the primitives are possible and each one has to be evaluated empirically in a separate experimental program of some magnitude. Some work on brightness summation across the two eyes is currently underway by the second author.

1. Primitives and Representations

1.1. Ordering of Joint Presentations

Let x denote the signal intensity less the threshold intensity of a pure tone presented to the left ear. We stress that we mean an intensity difference, not the more usual intensity ratio that leads to differences in dB. Let u denote an intensity less the threshold of a pure tone of the same frequency and phase presented to the right ear. Thus, 0 is the threshold intensity in each ear; intensities below threshold are set to 0. The notation (x, u) means the simultaneous presentation of x in the left ear and u in the right ear. This part of the model is, of course, an idealization—in reality, the threshold is a random variable which we idealize as a single number. In this connection and elsewhere, we rely on the position articulated shortly before his death by the youthful philosopher Frank Ramsey (1931/1964) in talking about decision making under uncertainty:

> Even in physics we cannot maintain that things that are equal to the same thing are equal to one another unless we take "equal" not as meaning "sensibly equal" but a fictitious or hypothetical relation. I do not want to discuss the metaphysics or epistemology of this process, but merely to remark that if it is allowable in physics it is allowable in psychology also. The logical simplicity characteristic of the relations dealt within a science is never attained by nature alone without any admixture of fiction. (p. 168/p. 70)

In the task we used, respondents were asked to judge if (x,u) is at least as loud as (y,v), which is denoted $(x,u) \succsim (y,v)$. The results we report show conditions such that this loudness ordering is reflected by a numerical mapping, called a psychophysical function, $\Psi : \mathbb{R}_+ \times \mathbb{R}_+ \xrightarrow{onto} \mathbb{R}_+$, where[4] $\mathbb{R}_+ := [0,\infty[$, that is strictly increasing in each variable and is order preserving, that is,

$$(x,u) \succsim (y,v) \Leftrightarrow \Psi(x,u) \geq \Psi(y,v), \tag{1}$$
$$\Psi(0,0) = 0. \tag{2}$$

And we assume that loudness and intensity agree to the extent that

$$(x,0) \succsim (y,0) \Leftrightarrow x \geq y,$$
$$(0,u) \succsim (0,v) \Leftrightarrow u \geq v.$$

Thus, $\Psi(x,0)$ and $\Psi(0,u)$ are each strictly increasing.

We assume that the respondent can always establish matches of three types to each stimulus:

$$(x,u) \sim (z_l,0),\ (x,u) \sim (0,z_r),\ (x,u) \sim (z_s,z_s), \tag{3}$$

where $\sim$ means equally loud. The left and right matches z_l and z_r are called asymmetric and z_s is called a symmetric match. Symmetric matches have the decided advantage of reducing the degree of localization change between (x,u) and the matching pair. The asymmetric matches encounter some difficulties, which we discuss in Section 5.1, and overcome in Section 5.2.

Note that each of the z's is a function of both x and u. To make that explicit and suggestive, we use an operator notation:

$$x \oplus_i u := z_i \quad (i = l,r,s). \tag{4}$$

It is not difficult to show that each of the $\oplus_i$ defined by (4) is, indeed, a binary operation that is defined for each pair (x,u) of intensities. The operator $\oplus_i$ may be referred to as a summation operator; however, one must not confuse $\oplus_i$ with $+$, that is, the addition of physical intensities. Some readers of our work have expressed discomfort over the fact that we can explore, for example, whether the operation is associative, that is,

$$x \oplus_i (y \oplus_i z) = (x \oplus_i y) \oplus_i z \quad (i = l,r,s), \tag{5}$$

despite the fact that the notation

$$(x,(y,z)) \sim ((x,y),z)$$

[4] The notation $A := B$ means that A is defined to be B.

is, itself, meaningless. Such a defined operator is, however, a familiar and commonly used trick in the theory of measurement to map something with two or more dimensions into a structure on a single dimension. See, for example, the treatment of conjoint measurement in Section 6.2.4 of Krantz, Luce, Suppes, and Tversky (1971) and in Section 19.6 of Luce, Krantz, Suppes, and Tversky (1990).

One can show under weak assumptions (see Proposition 1 of Luce, 2002) that $\succsim$ is a weak order (that is, transitive and connected), that (x, u) is strictly increasing in each variable, and that 0 is a right identity of $\oplus_l$, that is,

$$x = x \oplus_l 0, \tag{6}$$

and 0 is a left identity of $\oplus_r$, that is,

$$u = 0 \oplus_r u, \tag{7}$$

whereas 0 is not an identity of $\oplus_s$ at all. However, the symmetric operation is idempotent in the sense that

$$x \oplus_s x = x. \tag{8}$$

These properties play important roles in some of the proofs.

We assume that the function $\Psi(x, u)$ is decomposable in the sense that it depends just on its components $\Psi(x, 0)$ and $\Psi(0, u)$,

$$\Psi(x, u) = F[\Psi(x, 0), \Psi(0, u)]. \tag{9}$$

One natural question is the following: What is the nature of that dependence, that is, what is the mathematical form of F? A second natural question is the following: How do $\Psi(x, 0)$ and $\Psi(0, u)$ depend on the physical intensities x and u, respectively? These are ancient problems with very large literatures which we make no attempt to summarize here. Some references appear later. Neither question, it should be mentioned, is resolved in any fully satisfactory manner if one restricts attention just to the primitive ordering $\succsim$ of the conjoint structure of intensities, $\langle \mathbb{R}_+ \times \mathbb{R}_+, \succsim \rangle$. To have a well constrained theory that arrives at specific answers seems to require some structure beyond the ordering so far introduced. Later, in Section 4, we encounter two examples of such additional linking structures, which, in these cases, are two forms of a distribution law. This important point, which is familiar from physics, does not seem to have been as widely recognized by psychologists as we think that it should be.

Two points should be stressed. The first is that the theory is not domain specific, which means that it has many potential interpretations in addition to our auditory one. For example, also in audition, Karin Zimmer and

Wolfgang Ellermeier,[5] interpreted (x, u) to mean a brief signal of intensity x followed almost immediately by a another brief signal of intensity u. Other interpretations, using visual stimuli, are brightness summation of hemifields or across the two eyes. Each conceivable interpretation will, of course, require separate experimental verification, although drawing on our experience with the two ear experiments should be beneficial.

The second point is that the approach taken here is entirely behavioral and so is independent of any particular biological account of the behavior. Consequently, we do not attempt to draw any such conclusions from our results.

1.2. Ratio Productions

To the ordering of signal pairs, we add the independent structure of a generalized form of ratio production. This entails the presentation to a respondent of a positive number p and the stimuli (x, x) and (y, y), where $y < x$, and asking the respondent to produce the stimulus (z, z) for which the loudness "interval" from (y, y) to (z, z) is perceived to stand in the ratio p to the "interval" from (y, y) to (x, x). Because the z chosen by the respondent is a function of p, x, and y, we may again represent that functional dependence in the operational form

$$(x, x) \circ_p (y, y) := (z, z). \tag{10}$$

This operation, which we call (subjective) ratio production, is somewhat like Stevens's magnitude production[6] (for a summary, see Stevens, 1975) which is usually described as finding the signal (z, z) that stands in proportion p to stimulus (x, x). Thus, his method is the special case of ours but with $(y, y) = (0, 0)$—the threshold intensity or below. Thus, $(x, x) \circ_p (0, 0) = (z, z)$.

We assume two things about $\circ_p$: (a) it is strictly increasing in the first variable and nonconstant and continuous in the second one, and (b) that Ψ over $\circ_p$ is also decomposable in the sense that

$$\Psi[(x, x) \circ_p (y, y)] = G_p[\Psi(x, x), \Psi(y, y)]. \tag{11}$$

[5] As reported at the 2001 meeting of the European Mathematical Psychology Group in Lisbon, Portugal.

[6] In a generalized ratio estimation, the respondent is presented with two pairs of stimuli, (y, y) to (z, z) and (y, y) to (x, x), where $y < x, z$, and is asked to state the ratio $p = p(x, y, z)$ of the interval between the first two to the interval between the second two. This is discussed in SL-III and is summarized later in Sec. 6.1. This procedure is, of course, conceptually related to S. S. Stevens's magnitude estimation where no standard is provided [see after (10)].

1.3. The Representations

Building on the assumptions given earlier, Luce (2002, 2004) presented necessary and sufficient qualitative conditions the following representations, which are discussed later in Sections 3 and 4.[7]

$$\Psi(x,u) = \Psi(x,0) + \Psi(0,u) + \delta\Psi(x,0)\Psi(0,u) \quad (\delta \geq 0), \tag{12}$$

$$W(p) = \frac{\Psi\left[(x,x) \circ_p (y,y)\right] - \Psi(y,y)}{\Psi(x,x) - \Psi(y,y)} \quad (x > y \geq 0), \tag{13}$$

where δ is a (non-negative) constant and the function $W : [0,\infty[\overset{onto}{\longrightarrow} [0,\infty[$ is strictly increasing. The "summation" formula (12) has been dubbed p-additive because it is the unique polynomial function of $\Psi(x,0)$ and $\Psi(0,u)$ with $\Psi(0,0) = 0$ that can be transformed into additive form (see Section 3.2).

Under certain assumptions, one can also show that, for some $\gamma > 0$,

$$\Psi(x,0) = \gamma\Psi(0,x), \tag{14}$$

which we call constant bias; however, for other assumptions, constant bias is not forced. More specifically, if the properties stated later in Sections 3 and 4 hold for asymmetric matches, then constant bias, (14), holds in addition to the two representations (12) and (13) (Luce, 2002, 2004). In contrast, if the properties hold using symmetric matches, then one can prove that (12) holds with $\delta = 0$, that (13) holds, but that constant bias, (14), need not hold. Because constant bias seems intuitively unlikely—the ears often do not seem to be identical—we are probably going to be best off with the symmetric theory. We discuss data on whether our young respondents satisfy the assumption of having symmetric ears above threshold. We also investigate empirically whether $\delta = 0$ (Section 3.3), which has to hold if symmetric matches satisfy the conditions. If $\delta = 0$, empirical testing of the theory is simplified considerably.

Nothing in the theory giving rise to (12) and (13) dictates explicit mathematical forms for $\Psi(x,0)$ as a function of the physical intensity x, for $\Psi(0,u)$ as a function of u, or for $W(p)$ as a function of p. One attempt to work out the form of Ψ based just on the summation operation is summarized later in Section 5.4. It leads to a sum of power functions. Another condition, also leading to a power function form, which is based on ratio

[7] In Luce (2002), all of the results are presented in terms of psychophysical functions on each signal dimension, as was also the first submitted version of Luce (2004). As a reviewer, Ehtibar Dzhafarov saw how they could be neatly brought together as a psychophysical function over the signal pairs, and Luce adopted that formulation.

productions, is provided in Section 6.3. The experimental data make clear that our endeavors are incomplete. Our attempts to find out something about W, which currently also are incomplete, are summarized in Section 6.

Where do the representations (12) and (13) come from, and how do we test them? Various testable conditions that are necessary and sufficient for the representations are outlined and the results of several experimental tests are summarized. Note that we make no attempt to fit the representations themselves directly to data. In particular, no parametric assumptions are imposed about the nature of the functions Ψ and W. Later, in Section 5.4, we see how to test for the power function form of Ψ using parameter-free properties, and then in Section 6, again using parameter-free properties, we arrive at two possible forms for W.

2. Methods of Testing

The many experiments discussed employ empirical interpretations of the two operations. One is $x \oplus_i u := z_i \quad (i = l, r, s)$, (4), which involves estimating a value z_i that is experienced as equal in loudness to the joint-presentation (x, u). The other is $(x, x) \circ_p (y, y) := (z, z) \quad (y < x)$, (10), which involves estimating a value z that makes the loudness "interval" between (y, y) and (x, x) be a proportion p of the interval between (y, y) and (z, z). The first procedure is referred to as matching and the second as ratio production.

The stimuli used were, in all cases, 1,000 Hz, in phase pure tones of 100 ms duration that included 10 ms on and off ramps. Throughout, signals are described as dB relative to sound pressure level (SPL).

2.1. Matching Procedure

To describe the testing, we employ the notation $\langle A, B\rangle$ to mean the presentation of stimulus A followed after 450 ms by stimulus B, where A and B are joint presentations. Then the three matches of (3) are obtained using whichever is relevant of the following three trial forms:

$$\langle (x, u), (z_l, 0)\rangle, \tag{15}$$

$$\langle (x, u), (0, z_r)\rangle, \tag{16}$$

$$\langle (x, u), (z_s, z_s)\rangle. \tag{17}$$

In practice, respondents heard a stimulus followed 450 ms later by another tone pair in the left, right, or both ears, as the case might be. Respondents used key presses either to adjust the sound pressure level of z_i, $i = l, r, s$ (one of four differently sized steps), to repeat the previous trial, or to indicate satisfaction with the loudness match. Following each adjustment, the altered tone sequence was played. This process was repeated until respondents were satisfied that the second tone matched the first tone in loudness.

2.2. Ratio Production Procedures

The basic trial form is $\langle\langle A, B\rangle, \langle A, C\rangle\rangle$ where $\langle A, B\rangle$ and $\langle A, C\rangle$ represent the first and the second intensity interval respectively. The $\langle A, B\rangle$ and $\langle A, C\rangle$ were separated by 750 ms, and between A and B (and A and C), the delay was 450 ms.

An estimate of $x \circ_{p,i} y = v_i$, in the case of $i = s$, was obtained using the trial type

$$\langle\langle(y, y), (x, x)\rangle, \langle(y, y), (v_s, v_s)\rangle\rangle, \tag{18}$$

where the value of v_s was under the respondents' control. In practice, respondents heard two tones separated by 450 ms (the first interval) then 750 ms later, another such set of tones was heard (the second interval). The tone in the first pair in both intervals is the same and less intense than the second tone. Respondents continued to alter the sound pressure level of v_s until they experienced the second loudness interval as being a proportion p of the first one.

As mentioned earlier, the special case of $y = 0$ is an operation akin to Stevens's magnitude production, which involves finding the signal (z, z) that stands in proportion p to stimulus (x, x). With $i = s$, this was estimated using the trial type

$$\langle(x, x), (v_s, v_s)\rangle. \tag{19}$$

In practice, respondents heard two tones, separated by 450 ms, and they adjusted the second tone to be a proportion p of the first tone.

Trial forms in the case of $i = r, l$ are constructed in a manner analogous to (18) and (19).

2.3. Statistics

The four SL articles examined parameter-free null hypotheses of the form $L = R$, where L means the signal equivalent on the left side of the condition and R is the parallel equivalent for the right side. Not having any a priori idea concerning the distribution of empirical estimates, we used the nonparametric Mann-Whitney U test at the 0.05 level. To improve our statistical evaluation, we checked, using Monte Carlo simulations based on the bootstrap technique (Efron & Tibshirani, 1993), whether L and R could, at the 0.05 level, be argued to come from the same underlying distribution. This was our criterion for accepting the null hypothesis as supporting the behavioral property.

Recently there has been a flurry of activity concerned with Bayesian approaches to axiom testing. The first published reference is Karabatsos (2005). These methods have not been applied to our data.

2.4. Additional Methodological Observations

During the course of doing these studies, we encountered and overcame or attenuated some methodological issues (details in SL-I).

The well-known time-order error—that is, the impact of (x, u) depends on whether it occurs before or after (y, v)—means that it is important to use some counterbalancing of stimulus presentations or to ensure that the errors are balanced on the two sides of a behavioral indifference.

Some experiments require us to use an estimate from one step as input to a second one. If a median or other average from the first step is used as the input in a second step, then whatever error it contains is necessarily carried over into that second one, but all information about the variance is lost. After some experience we concluded that the results are more satisfactory if we used each individual estimate from the first step as an input to the second one. Then the errors of the first estimate are carried into the second step and average out there, while preserving the variance information.

Variability for ratio productions tends to be higher than for matching. This fact means that, in evaluating our conclusions, some attention needs to be paid to the number of observations made by each respondent.

3. Summations and Productions Separately

Much of the mathematical formulation of the theory was first developed for utility theory (summarized[8] in Luce, 2000) under the assumption that the following property, called joint-presentation symmetry (jp-symmetry), holds:

$$(x, u) \sim (u, x). \tag{20}$$

Under the current psychophysical interpretation, this means that the ears are identical in dealing with intensities above their respective thresholds. We know this need not always hold (e.g., single-ear deafness resulting from exposure of one ear, usually the left, to percussive rifle shots), but at first we thought that it might be approximately true for young people with no known hearing defects. Note that jp-symmetry, (20), is equivalent to $\oplus_l \equiv \oplus_r$ and $\oplus_s$, all being commutative operators, that is,

$$x \oplus_i y = y \oplus_i x \quad (i = l, r, s).$$

[8] Errata: see Luce's web page at http:www.socsci.uci.edu.

3.1. Evidence Against Symmetric Hearing Using Symmetric Matching

Using symmetric matching, we obtained

$$z = x \oplus_s y \text{ and } z' = y \oplus_s x,$$

using the trial form (17). Each respondent made from 34-50 matches per stimulus. We used tones with intensities $a = 58$ dB, $b = 64$ dB, $c = 70$ dB SPL, which gave rise to six ordered stimulus pairs: (a, b), (a, c), (b, c) and (b, a), (c, a), (b, c). For each (x, u) pair, we tested statistically whether the null hypothesis $z = z'$ held. With 15 respondents there were 45 tests of which 23 rejected the null hypothesis. The pattern of results suggests that jp-symmetry fails for at least 12 of the 15 respondents.

The negative outcome of this experiment motivated the developments in Luce (2004) where jp-symmetry is not assumed to hold.

Later, in Section 5.1, we turn to the use of asymmetric matches to study the properties underlying the representation. They sometimes exhibit an undesirable phenomenon for which we suggest an explanation in terms of filtering, and after the fact, show that the properties, described later using asymmetric matches, are unaffected by the filter. However, some of the arguments rest on a property that corresponds to the psychophysical function being sums of power functions, and that may not always be sustained.

3.2. Thomsen Condition

The representation (12) with $\delta = 0$,

$$\Psi(x, u) = \Psi(x, 0) + \Psi(0, u), \tag{21}$$

is nothing but an additive conjoint representation (Krantz et al., 1971, Ch. 6). And, for $\delta > 0$, the p-additive representation, (12), can be rewritten as

$$1 + \delta\Psi(x, u) = [1 + \delta\Psi(x, 0)]\,[1 + \delta\Psi(0, u)]\,,$$

so under the transformation

$$\Theta(x, u) = \ln\left[1 + \delta\Psi(x, u)\right], \tag{22}$$

the conjoint structure again has an additive representation. So data bearing on the existence of an additive presentation is of interest whether or not $\delta = 0$.

With our background assumptions—weak ordering, strict monotonicity, solvability, and that intensity changes in either ear affect loudness—we can formulate a property that is analogous to the numerical Archimedean property that for any two positive numbers a and b, one can find an integer n such that $na > b$. Thus, by Krantz et al. (1971) we need only the following condition, called the Thomsen condition, in order to construct an additive representation Θ.

$$\left.\begin{array}{r}(x,t) \sim (z,v)\\(z,u) \sim (y,t)\end{array}\right\} \Longrightarrow (x,u) \sim (y,v) \tag{23}$$

This notation is used in the conjoint measurement literature. It can, of course, be rewritten in operator form, but that is both less familiar and appears to be more complex. If all of the $\sim$ are replaced by $\succsim$, the resulting condition is called double cancellation. The reason for that term is that the condition can be paraphrased as involving the two "cancellations" t and z, each of which appears on each side of the hypotheses, to arrive at the conclusion.

We know of no empirical literature in audition, other than our study described later, that tests the Thomsen condition, per se. What has been published concerning conjoint additivity all examined double cancellation, which we feel is a somewhat less sensitive challenge than is the Thomsen condition. Of the double cancellation studies, three support it: Falmagne, Iverson, and Marcovici (1979), Levelt, Riemersma, and Bunt (1972), and Schneider (1988), where the latter differed from the other studies in having frequencies varying by more than a critical band in the two ears. Rejecting it were Falmagne (1976) with but one respondent, and Gigerenzer and Strube (1983) with 12 respondents. Because of this inconsistency, we felt it necessary to test the Thomsen condition within our own experimental context. Our experimental design was closest to that of Gigerenzer and Strube (1983).

The Thomsen condition was tested by successively obtaining the estimates, z', y', and y'',

$$\begin{aligned}(x,t) &\sim (z',v)\\(z',u) &\sim (y',t)\\(x,u) &\sim (y'',v)\end{aligned}$$

using the trial form in (17), where the first of the two tones in the second joint-presentation is varied. The property is said to hold if we do not reject the hypothesis that the observations y' and y'' all come from a single distribution.

We used two stimulus sets, **A** and **B**, in our test of the Thomsen condition:

$$\mathbf{A}: x = 66, t = 62, v = 58, \text{ and } u = 70 \text{ dB},$$
$$\mathbf{B}: x = 62, t = 59, v = 47, \text{ and } u = 74 \text{ dB}.$$

Stimulus set **B** consisted of stimuli having the same relative intensity relations as those used by Gigerenzer and Strube (1983), although we used 1,000 Hz whereas they used both 200 Hz and 2,000 Hz, a difference that may be relevant.

We initially ran the respondents on **A**, after which we decided to add **B** to have a more direct comparison with their study.

With 12 respondents, there were 24 tests of which 5 rejected the null hypothesis. Of the five failures, four occurred in set **A** and one in **B**. This fact suggests that a good deal of practice may regularize the behavior. (See SL-I for details.) In summary, we feel that the Thomsen condition has been adequately sustained.

3.3. Bisymmetry

On the assumption that we have a p-additive representation, (12), we next turn to the question of whether $\delta = 0$. All of the experimental testing is a good deal simpler when $\delta = 0$ than it would be otherwise—an example is the testing of the property called joint-presentation decomposition (Section 4.1).

Given the p-additive representation, one can show (Luce, 2004, Corollary 2 to Theorem 1, p. 450) that for a person who violates jp-symmetry, (20), $\delta = 0$ is equivalent to the following property, called bisymmetry:

$$(x \oplus_i y) \oplus_i (u \oplus_i v) = (x \oplus_i u) \oplus_i (y \oplus_i v) \quad (i = l, r, s). \tag{24}$$

Note that the two sides of bisymmetry simply involve the interchange of y and u. Bisymmetry is not predicted when $\delta \neq 0$ except for constant bias with $\gamma = 1$. Because we have considerable evidence against $\gamma = 1$ (Section 3.1), we know that bisymmetry holds if, and only if, $\delta = 0$.

Testing involved obtaining the estimates

$$w_i = x \oplus_i y \text{ and } w'_i = u \oplus_i v, \quad \text{[right side of (24)]},$$
$$z_i = x \oplus_i u \text{ and } z'_i = y \oplus_i v \quad \text{[left side of (24)]},$$

and then in a second step obtaining

$$t_i = w_i \oplus_i w'_i \text{ and } t'_i = z_i \oplus_i z'_i.$$

The property is said to hold if t_i and t'_i are found to be statistically equivalent. The property was tested, for both symmetric and left-ear matches, using trials of the form (17) and (15), respectively, and intensities $x = 58$ dB, $y = 64$ dB, $u = 70$ dB, and $v = 76$ dB. With six respondents there were no rejections of bisymmetry. So we assume $\delta = 0$ in what follows (SL-I).

3.4. Production commutativity

If we rewrite (13)[9] as

$$\Psi\left[(x,x)\circ_p(y,y)\right]=W(p)[\Psi(x,x)-\Psi(y,y)]+\Psi(y,y),$$

then by direct substitution the following behavioral property, called production commutativity, readily follows. For $p>0,q>0$,

$$[(x,x)\circ_p(y,y)]\circ_q(y,y)\sim[(x,x)\circ_q(y,y)]\circ_p(y,y). \tag{25}$$

Observe that the two sides differ only in the order of applying p and q, which is the reason for the term commutativity. This property also arose in Narens's (1996) theory of magnitude estimation. Ellermeier and Faulhammer (2000) tested that prediction in the special case where $y=0$ for $p,q>1$ and Zimmer (2005) did so for $p,q<1$. Both studies found it sustained. The general form of production commutativity has yet to be tested with $p<1<q$.

In the presence of our other assumptions, production commutativity turns out to be sufficient as well as necessary for (13) to hold.

Production commutativity was tested using symmetric ratio productions requiring four estimates in two steps. The first involved obtaining estimates of v and w satisfying

$$(x,x)\circ_p(y,y)\sim(v,v),$$
$$(v,v)\circ_q(y,y)\sim(w,w),$$

and the second of obtaining estimates of v' and w' satisfying

$$(x,x)\circ_q(y,y)\sim(v',v'),$$
$$(v',v')\circ_p(y,y)\sim(w',w').$$

The property is considered to hold if w and w' are found to be statistically equivalent. Trials were of the form in (18). The intensities used were $x=64$ dB and $u=70$ dB and the proportions used were $p=2$ and $q=3$, giving rise to four trial conditions in each step. Four respondents yielded four tests, and the null hypothesis of production commutativity was not rejected in any of them (SL-I).

3.5. Discussion

The results of the experiments on the Thomsen condition and on production commutativity support the existence of a $\Psi_{\oplus}$ as in (12) and $\Psi_{\circ_p}$ as in (13)

[9] To those familiar with utility theory, the following form is basically subjective weighted utility (Luce & Marley, 2005).

separately. However, from these data alone we cannot conclude that the same function Ψ applies both to summations and productions, that is, $\Psi_{\oplus} = \Psi_{\circ_p}$. Although we have no evidence at this point to assume that, we do know that both are strictly increasing with $\succsim$, and so there is a strictly increasing, real-valued function connecting them: $\Psi_{\circ_p}(x,u) = f(\Psi_{\oplus}(x,u))$.

So our next task is to ask for conditions necessary and sufficient for the function f to be the identity function. Such conditions involve some interlocking of the two structures $\langle \mathbb{R}_+ \times \mathbb{R}_+, \succsim \rangle$ and $\langle \mathbb{R}_+ \times \mathbb{R}_+, \succsim, \circ_p \rangle$, which can be reduced to the one dimensional structures of the form $\langle \mathbb{R}_+, \geq, \oplus_i \rangle$ and $\langle \mathbb{R}_+, \geq, \circ_{p,i} \rangle$, respectively. We turn to that interlocking issue.

4. Links Between Summation and Production

It turns out that two necessary properties of the representations establish the needed interlock or linkage between the primitives, and these properties along with those discussed earlier are sufficient to yield a common representation, $\Psi = \Psi_{\oplus} = \Psi_{\circ_p}$ (Theorem 2 of Luce, 2004). In a sense, the novelty of this theory lies in formulating their interlock purely behaviorally.

The links that we impose are analogous to the familiar "distribution" properties such as those in set theory, namely,

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C), \tag{26}$$

$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C). \tag{27}$$

If we replace $\cup$ by $\oplus$ and $\cap$ by $\circ_p$ we get, respectively, what are called later simple joint-presentation decomposition, (Section 4.1), and segregation, (4.2). Some of the significance of such an interlock is discussed by Luce (2005).

To help formulate these properties, we define the following induced production operators $\circ_{p,i}$, $i = l, r, s$, which are special cases of the general operation $\circ_p$ defined by (10):

$$(x \circ_{p,l} y, 0) := (x,0) \circ_p (y,0), \tag{28}$$

$$(0, u \circ_{p,r} v) := (0,u) \circ_p (0,v), \tag{29}$$

$$(x \circ_{p,s} y, x \circ_{p,s} y) := (x,x) \circ_p (y,y). \tag{30}$$

4.1. Simple Joint-Presentation Decomposition

As suggested above, the analogue of (26), linking the two operations $\oplus_i$ and $\circ_{p,i}$ is a property that is called simple joint-presentation (SJP-) decomposition: For all signals x, u and any number $p > 0$,

$$(x \oplus_i u) \circ_{p,i} 0 = (x \circ_{p,i} 0) \oplus_i (u \circ_{p,i} 0) \quad (i = l, r, s). \tag{31}$$

When $\delta \neq 0$, the corresponding property becomes vastly more difficult to test because the term $u \circ_{p,i} 0$ is replaced by $u \circ_{q,i} 0$ where $q = q(x, p)$. Thus, one must first determine $q(x, p)$ empirically and then check the condition corresponding to (31) with q replacing the second p on right.

SJP-decomposition has two levels of estimation which were done in two steps. First, the estimates

$$\begin{aligned} t_s &= (x \oplus_s u) \circ_{p,s} 0, \\ w_s &= x \circ_{p,s} 0, \\ s_s &= u \circ_{p,s} 0, \end{aligned}$$

were obtained using trials of the form in (18). We computed the means[10] of the empirical estimates of w_s and s_s and used them in the second step, which consisted of the match

$$t'_s = w_s \oplus_s s_s,$$

using trials of the form in (16). The property is considered to hold if t_s and t'_s are found to be statistically equivalent. We used one pair of intensities, $x = 64$ dB and $u = 70$ dB and the two values of p, namely $p = 2/3$ and $p = 2$. With four respondents there were eight tests and SJP-decomposition was not rejected in six of the eight (SL-II).

4.2. Segregation

The second property linking the two operations, the analogue of (27) but taking into account the noncommutativity of $\oplus_i$, is what is called segregation:

For all $x, u, p \in \mathbb{R}_+$, left segregation holds if

$$u \oplus_i (x \circ_{p,i} 0) \sim (u \oplus_i x) \circ_{p,i} (u \oplus_i 0) \quad (i = l, r, s). \tag{32}$$

And right segregation holds if

$$(x \circ_{p,i} 0) \oplus_i u \sim (x \oplus_i u) \circ_{p,i} (0 \oplus_i u) \quad (i = l, r, s). \tag{33}$$

If jp-symmetry, (20), holds, then right and left segregation are equivalent. Otherwise they are distinct.

[10] At the time this experiment was run, we did not fully understand the advantages of using sequential estimates.

Note that because 0 is a right identity of $\oplus_l$, (7), testing left segregation is easier for $i = l$, and similarly, right segregation is easier for $i = r$. For $i = s$ both need to be tested.

For each respondent (except one), we studied only one form of segregation, either left or right (see SL-II for details).

In the case of right segregation, four estimates must be made:

$$\begin{aligned} w_r &= x \circ_{p,r} 0, \\ t_r &= w_r \oplus_r u, \\ z_r &= x \oplus_r u, \\ t'_r &= z_r \circ_{p,r} u. \end{aligned}$$

The property is said to hold if t_r and t'_r are not found to be statistically different.

Note that the intensities w_r and z_r are first estimated in the right ear but then they must be presented in the left ear for the case $(w_r, u) \sim (0, w_r \oplus_r u)$. The converse is true for left segregation. The trials used for matching were of the forms in (15) to (17) depending on the matching ear; the ratio productions used were (18) and (19), or their equivalents for asymmetric productions.

We used one intensity pair, $x = 72$ dB and $u = 68$ dB, except for one respondent where each was decreased by 4 dB to avoid productions limited by a 85 dB safety bound. A theoretical predication is that the property holds for both $p < 1$ and $p \geq 1$, hence $p = 2/3$ and $p = 2$ were used.

Four respondents produced 10 tests and the null hypothesis was accepted in eight of them (SL-II).

4.3. Discussion

Given the complexity of testing these two properties of the model and given the potential for artifacts, we feel that the support found for the model leading to the additive representation, (12) with $\delta = 0$, and the subjective proportion representation, (13), is not too bad.

Assuming that there is a common Ψ underlying the representation, an interesting theoretical challenge exists. Taken by itself, the p-additive representation of $\oplus_i$ could have $\delta \lesseqgtr 0$. For $\delta < 0$, it is not difficult to see that Ψ is bounded by $1/|\delta|$ and so $\Psi : \mathbb{R}_+ \times \mathbb{R}_+ \xrightarrow{onto} I = [0, 1/|\delta|]$. Of course, we have given data on bisymmetry that suggests $\delta = 0$. However, we cannot really rule out that δ may be slightly different from 0, as we discuss later in Section 5.8. On the face of it, boundedness seems quite plausible. Psychophysical scales of intensity seem to have upper bounds tied in with potential sensory damage and so infinite ones are decidedly an idealization.

However, within the theory as currently formulated, bounded Ψ is definitely not possible[11] because one can iterate the operator as, for example, in the second step: $(x \circ_{p,i} 0) \circ_{p,i} 0$, and this forces Ψ to be unbounded.[12] So the challenge is to discover a suitable modification of (13) that is bounded and work out its properties.

5. Sensory Filtering, Multiplicative Invariance, and Forms for Ψ

5.1. Asymmetric matching and jp-symmetry

Earlier (see the beginning of Section 3) we discussed the use of symmetric matches in checking jp-symmetry. Now we discuss the use of asymmetric left and right matches. The latter sometimes exhibited the following phenomenon which at first seemed disturbing but, in fact, seems to have rather mild consequences. Consider the following asymmetric matches:

$$(x, u) \sim (x \oplus_l u, 0),\ (u, x) \sim (u \oplus_l x, 0),$$
$$(x, u) \sim (0, x \oplus_r u),\ (u, x) \sim (0, u \oplus_r x).$$

Suppose that jp-symmetry fails, as it often seems to, and suppose that $(x, u) \succ (u, x)$. Then one expects to observe that both $x \oplus_l u > u \oplus_l x$ and $x \oplus_r u > u \oplus_r x$—that is, that left and right matches will agree in what they say about jp-symmetry. We carried out such an experiment, obtaining the asymmetric matches above using the trials forms in (15) and (16), and the same stimuli as used earlier (Section 3.1). Although the expected agreement held for four respondents, it did not hold for two, even after considerable experience in the experimental situation. Moreover, for those who were qualitatively consistent, the magnitude of the differences $x \oplus_l u - u \oplus_l x$ and $x \oplus_r u - u \oplus_r x$ varied considerably. Evidently, matching in a single ear had some significant impact. Of course, one impact is manifest in a sharp change of localization, which at first concerned some readers of our work.

[11] This fact was pointed out by Ehtibar Dzhafarov in a referee report, dated October 8, 2000, of Luce (2002).

[12] Set $y = 0$ in (13) and consider a sequence x_n with $x_0 > 0$ such that

$$W(p) = \Psi(x_n, x_n)/\Psi(x_{n-1}, x_{n-1}).$$

A simple induction yields

$$\Psi(x_n, x_n) = W(p)^n \Psi(x_0, x_0) > 0.$$

For $W(p) > 1$, this is unbounded.

But the inconsistency just described is more worrisome as it means that an experimental procedure that relies on the assumption of bias independent of the matching ear will not be reliable. In practice, this has not proven to be an obstacle in our other experiments. We offer a possible account of this effect in the next two subsections.

5.2. Sensory Filtering in the Asymmetric Cases

Suppose that asymmetric matching has the effect of either enhancing (in the auditory system) all signals in the matching ear or attenuating those in the other ear. If these effects entail a simple multiplicative factor on intensity, that is, a constant dB shift, then the two ideas are equivalent. If we assume that there is an attenuation or intensity filter factor η on the non-matched ear, then for the left matching case, the experimental stimulus (x, u) becomes, effectively, $(x, \eta u)$. And when matching in the right ear, (x, u) becomes effectively $(\eta x, u)$, where $0 < \eta \leq 1$. Thus, when we ask the respondents to solve the three indifferences of (3) what they actually do, according to this theory, is set

$$\begin{aligned}
z_l &= x \oplus_l \eta u \Leftrightarrow (z_l, 0) \sim (x, \eta u), \\
z_r &= \eta x \oplus_r u \Leftrightarrow (0, z_r) \sim (\eta x, u), \\
z_s &= x \oplus_s u \Leftrightarrow (z_s, z_s) \sim (x, u).
\end{aligned}$$

Note that the filter plays no role in the symmetric matches.

Under a further condition that is called multiplicative invariance (Section 5.3), which is equivalent to $\delta = 0$ and that $\Psi(x, 0)$ and $\Psi(0, x)$ are each a power function of x, but with different powers, one can show that the filtering concept does indeed accommodate the aforementioned phenomenon of asymmetric matching in connection with checking jp-symmetry.

5.3. Multiplicative invariance

Fortunately, we were able to show that filtering does not distort any of the experimental tests of the properties discussed earlier (Sections 3 and 4), where asymmetric matching is used, provided that the operations $\oplus_i$ have an additive representation shown earlier (Sections 3.2 and 3.3), and that the following property of $\sigma-$Multiplicative Invariance ($\sigma-$MI) holds: For all signals $x \geq 0, u \geq 0$, for any factor $\lambda \geq 0$, and for $\oplus_i$, $i = l, r$, defined by (3) and (4), there is some constant $\sigma > 0$ such that

$$\lambda x \oplus_l \lambda^\sigma u = \lambda(x \oplus_l u), \tag{34}$$

and

$$\lambda x \oplus_r \lambda^\sigma u = \lambda^\sigma(x \oplus_r u). \tag{35}$$

We observe that this property is, itself, invariant under sensory filtering because with filtering that expression becomes

$$\lambda x \oplus_l \eta\lambda^\sigma u = \lambda(x \oplus_l \eta u), \tag{36}$$

$$\eta\lambda x \oplus_r \lambda^\sigma u = \lambda^\sigma(\eta x \oplus_r u). \tag{37}$$

Because $\eta\lambda^\sigma = \lambda^\sigma\eta$, setting $v = \eta u$ in (36) shows that it is of the form (34) and setting $y = \eta x$ (37) shows it is of the form (35). Thus, filtering does not affect the use of $\sigma-$MI when discussing other properties.

Turning to our other necessary properties discussed earlier (Sections 3 and 4), elementary calculations show that they are invariant under filtering either with no further assumption or assuming multiplicative invariance (see Table 1).

Table 1: Effect of filtering on properties.

Property	Assumption None	Assumption $\sigma-$MI
Thomsen	X	
Bisymmetry		X
Production Commutativity	X	
SJP-Decomposition	X	
Segregation		X

SJP = Simple Joint-Presentation
MI = Multiplicative Invariance

We examine one important implication of $\sigma-$MI in the next subsection and report some relevant data.

5.4. Ψ a Sum of Power Functions

So far, we have arrived at a representation with two free parameters, δ and γ, and two free increasing functions, Ψ and W, and we have shown that, most likely, $\delta = 0$. It is clear that one further goal of our project is to develop behavioral characterizations under which each of the functions belong to a specific family with very few free parameters. In this section we take up one argument for the bivariate Ψ being a sum of power functions and later (Section 6.1) we give a different argument for the power function form of Ψ and also consider two possible forms for W, rejecting one and possibly keeping the other.

Assuming that the representation (12) holds (see Sections 1.3 and 3.2) and that $\delta = 0$ (see Section 3.3), then one can show that $\sigma-$MI is equivalent to Ψ being a sum of power functions, (51), with exponents β_l and β_r such that $\sigma = \beta_l/\beta_r$, that is,

$$\Psi(x,u) = \alpha_l x^{\beta_l} + \alpha_r u^{\beta_r} = \alpha_l x^{\beta_l} + \alpha_r u^{\beta_l/\sigma}. \tag{38}$$

The proof, which is in SL-III, is a minor modification of that given by Aczél, Falmagne, and Luce (2000) for $\sigma = 1$. Thus, $\sigma-$MI is a behavioral test for the power function form (38).

Note that

$$\frac{\Psi(x,0)}{\Psi(0,x)} = \frac{\alpha_l}{\alpha_r} x^{\beta_l - \beta_r}.$$

Thus the constant bias property (14) holds iff $\gamma = \frac{\alpha_l}{\alpha_r}$ and $\sigma = \frac{\beta_l}{\beta_r} = 1$.

Recall that x and u in (34) and (35) are intensity differences between the signal intensity actually presented and the threshold intensity for that ear. However, the experimental design and results are typically reported in dB terms. In the current situation, this practice represents a notational difficulty because, for example, λx in dB terms is

$$10\log(\lambda x) = 10\log\lambda + 10\log x.$$

Thus, the multiplicative factor becomes additive when written in dBs. In the following, the intensity notation will be maintained in equations but actual experimental quantities are reported in dBs SPL where $\lambda_{dB} = 10\log\lambda$ stands for the additive factor.

5.5. Tests of $1-$MI

We did this experiment before we had developed the general result about $\sigma-$MI. The test was carried out in two steps: The first is an experimental one in which the respondents estimate

$$t_i = (\lambda x) \oplus_i (\lambda u) \quad \text{and} \quad z_i = x \oplus_i u,$$

obtained using trial-form (15) or (16) as the case might be. This is followed by a purely "arithmetic" step in which the multiplication $t_i' = \lambda \times z_i$ is performed by the experimenter. $1-$MI is said to hold if the hypothesis $t_i = t_i'$ is not statistically rejected.

For the experiment, we used $x = 64$ dB and $u = 70$ dB and two values for λ_{dB}, 4 and -4 dB ($\lambda = 2.5$ and 0.4, respectively).

Of 22 respondents, 12 satisfied $1-$MI in both tests, three failed both, and seven failed one. So we have a crude estimate of about half of the respondents satisfying multiplicative invariance with $\sigma = 1$. The fact of so many failures led us to explore how to estimate σ and then to estimate whether or not the 1-MI results were likely to change by doing the $\sigma-$MI experiment using that estimate.

5.6. Estimating σ and η

To test multiplicative invariance, it is most desirable to estimate σ and not to have to run a parametric experiment. To that end, using the representation (38), one can show that there exist constants c_1 and c_2 such that

$$(0 \oplus_l x)_{dB} = c_1 + \frac{1}{\sigma} x_{dB}, \tag{39}$$

$$(x \oplus_r 0)_{dB} = c_2 + \sigma x_{dB}, \tag{40}$$

from which it follows that there is a constant c_3 such that

$$(x \oplus_r 0)_{dB} = c_3 + \sigma^2 (0 \oplus_l x)_{dB} \tag{41}$$

follows. One can regress as shown and also in the other direction. Each gives an estimate of σ and we used the geometric mean of the two estimates of σ. This appears to be a suitable way to estimate σ—suitable in the sense that if (38) holds, then this is what it must be.

In terms of the power function representation itself one can show that the constants c_1 in (39), c_2 in (40), and c_3 in (41) are explicit functions of γ and η and, solving for these parameters, one can show that

$$\log \eta = \frac{\sigma c_1 + c_2}{10(1+\sigma)}, \tag{42}$$

$$\log \gamma = \frac{\beta_r c_3}{10(1+\sigma)}. \tag{43}$$

5.7. Estimates of σ

For seven of the respondents for whom we tested $1-$MI, we also collected the estimates $z_r = (x \oplus_r 0)$ and $z_l = (0 \oplus_l x)$ using trial-forms (15) and (16), and the three instantiations of x, 58, 66, and 74 dB SPL. Then, we estimated σ using (41) and linear regression. The estimates were obtained by regressing both on $(0 \oplus_l x)_{dB}$ and $(x \oplus_r 0)_{dB}$, separately, and the final estimate was taken as the geometric mean of the two and we tested statistically whether $\sigma = 1$. These results, including the numerical direction of the estimated σs, are summarized in the left portion of Table 2. To evaluate whether or not it would be worthwhile to do the $\sigma-$MI experiment, we asked the following: In which direction would σ have to deviate from 1 in order to alter the

Table 2: Summary of numerical direction of σ needed to fit data and obtained estimates.

	Test of 1-MI				Estimates of σ		
1−MI	Total	Needed $\sigma < 1$	Needed $\sigma > 1$	Contradictory	Total	Numerical $\widehat{\sigma} < 1$	Numerical $\widehat{\sigma} > 1$
Passed	12	2	7	3	5	1	4
Failed	10	2	7	1	2	0	2
Total	22	4	14	4	7	1	6

Note: MI = Multiplicative Invariance.

previous data testing 1−MI (22 respondents) toward equality of the two sides? These results are summarized in the right portion of Table 2.

From the last row of Table 2, we see that for four of the 22 respondents we need a value of $\sigma < 1$ to fit the data, for 14 of them a value of $\sigma < 1$, and for four the data suggest contradictory directions. In the subset of seven respondents for whom we estimated σ (Left portion of Table 2), one respondent had an estimate of $\sigma > 1$, and for six an estimate of $\sigma > 1$ was obtained.

For these 7 respondents, the needed and obtained numerical direction of σ is the same for 4 and different for 2. In 1 case, the needed direction of σ is inconclusive, which is well reflected in the obtained σ being close to one. This means the pattern of results appears reasonable for 5 and inappropriate for 2 of 7 respondents.

For those who passed 1−MI, a sum of power functions is already a reasonable description of behavior. The interesting cases are for those who either failed 1−MI or yielded σ estimates suggesting material deviations from 1. Of the two who failed 1−MI, the direction of the estimated σ was the same as the expected for one, and for those who passed 1−MI, σ was estimated different from 1 for one respondent. In the former case, a correction factor would add 0.02 dB to δ and in the latter it is 0.70 dB. The smaller factor is insignificant but the latter could well affect the results. Based on this sample, running the σ−MI experiment appears only worthwhile for one respondent. Here a correction factor would add 1.1 dB to δ. Testing led to an estimate of σ insufficiently large for the respondent to pass σ−MI.

In conclusion, the σ estimates are reasonably in line with expectations but in this current sample not much seems to be gained from them. Specifically, the results of the σ estimation do not seem to provide a correction factor that explains the respondent's deviations from 1−MI. Thus, we have evidence that about half of the respondents are well described by the sum of power functions, but that we do not know what forms fit the other half.

5.8. Ψ a p-Additive Sum of Power Functions

There is another possible reason for failures of $\sigma-$MI. Recall that, based on the empirical fact that we did not reject the property of bisymmetry, (24), we concluded that $\delta = 0$ could not be rejected. Nonetheless, our results for 1$-$MI seem to be consistent with that property for only about 50% of the respondents and we concluded on the basis of our estimates of σ that going to $\sigma-$MI would not improve the picture. But we cannot ignore the possibility that our test of bisymmetry simply was not sufficiently sensitive to catch the fact that, really, for some respondents $\delta \neq 0$. Assuming that the function for each ear individually, $\Psi(x, 0)$ and $\Psi(0, u)$ are each a power function, as discussed later (Section 6.2), then this line of argument suggests that instead of Ψ being the sum of power functions, (38), it possibly is the more general p-additive form:

$$\begin{aligned} \Psi(x,u) &= \alpha_l x^{\beta_l} + \alpha_r u^{\beta_r} + \delta\alpha_l\alpha_r x^{\beta_l} u^{\beta_r} \\ &= \alpha_l x^{\beta_l} + \alpha_r u^{\beta_l/\sigma} + \delta\alpha_l\alpha_r x^{\beta_l} u^{\beta_l/\sigma}. \end{aligned} \tag{44}$$

Note that the formulas (39) and (40) are unchanged by this generalization because when one signal is 0, the δ term vanishes. Thus, the formulas for estimating σ and η, (41) and (42), are also unchanged. So, the important question becomes the following: What property replaces $\sigma-$MI in characterizing the p-additive form with $\delta \neq 0$, rather than additive sum, of power functions, (44)? This theoretical question has yet to be answered. If and when we find that property, clearly it should be tested empirically.

6. Ratio Estimation and the Forms for W

To those familiar with the empirical literature on "direct scaling" methods, our discussion may seem unusual because so far it has focused exclusively on ratio production and not at all on ratio estimation and its close relative magnitude estimation. Magnitude estimation is far more emphasized in the empirical and applications literatures than is magnitude production. We remedy this lacuna in the theory now.

Here it is useful to define the following:

$$\psi_l(x) := \Psi(x, 0), \tag{45}$$

$$\psi_r(u) := \Psi(0, u), \tag{46}$$

$$\psi_s(x) := \Psi(x, x). \tag{47}$$

We work with the generic ψ_i.

6.1. Ratio Estimation Interpreted Within This Theory

A fairly natural interpretation of ratio estimations can be given in terms of (13) with $y = 0$. Instead of producing $z_i(x,p) = x \circ_{p,i} 0$, $i = l, r, s$, such that $z_i(x,p)$ stands in the ratio p to x, the respondent is asked to state the value of p_i that corresponds to the subjective ratio of z to x. This value may be called the *perceived ratio* of intensity z to intensity x. If we change variables by setting $t = z/x$, then p_i is a function of both t and x, that is, $p_i = p_i(t,x)$. Note that p_i is a dimensionless number. According to (13) and using the definition of ψ_i,

$$W\left(p_i(t,x)\right) = \frac{\psi_i(tx)}{\psi_i(x)}. \tag{48}$$

This relation among the three unknown functions, ψ_i, p_i, W, is fundamental to what follows.

The empirical literature on magnitude estimates has sometimes involved giving a standard x and in other experiments it was left up to the respondent to set his or her own standard. Stevens (1975, pp. 26-30) argued for the latter procedure. From our perspective, this means that it is very unclear what a person is trying to do when responding—comparing the present stimulus to some fixed internal standard or to the previous signal or to what? And, therefore, it means that averaging over respondents, who may be doing different things, is even less satisfactory than it usually is.

The literature seems to have assumed implicitly that the ratio estimate $p_i(t,x)$ depends only on t, not on x, that is,

$$p_i(t,x) = p_i(t). \tag{49}$$

The only auditory data we have uncovered on this are Beck and Shaw (1965) and Hellman and Zwislocki (1961). The latter article had nine respondents provide ratio estimates to five different standard pairs $(x_0, 10)$ where $x_0 = 40, 60, 70, 80, 90$ dB SPL. The geometric-mean results for the respondents are shown in their Fig. 6. If one shifts the intensity scale (in dB) so that all the standard pairs are at the same point of the graph, we get the plot shown in Fig. 1a. For values above the standard, there does not seem to be any differences in the curves, in agreement with (49). But things are not so favorable for values below the standard. Of course, there are possible artifacts. Experience in this area suggests that many people are uneasy about the lower end of the numerical scale, especially below 1. They seem to feel "crowded" in the region of fractions, and such crowding should only increase as one lowers the standard. It therefore seems reasonable to do the study with moduli of, say, 100 or larger.

This is exactly what Beck and Shaw (1965) did: they collected magnitude estimates of loudness as a function of four standards, 25, 77, 81, and 101 dB SPL, and two moduli, 100 and 500 (incomplete factorial design), and reported the median magnitude estimates, shown in their Fig. 1. They collected data for both even and irregularly spaced stimuli, but concluded that the results were the same. Hence, we have averaged over the stimulus spacing conditions. In our Fig. 1b, we have replotted their data by shifting them to a common standard (s) and modulus (m) and on the same scale as those of Hellman and Zwislocki (1961). Note, only their 77/81 dB conditions extends both below and above the standard. Here we find, contrary to the data of Hellman and Zwislocki (1961), for values below the standard, there seems to be a very small if any difference in the slope of the curves, which agrees with (49). However, for at least two of the four graphs, the slope is shallower for values above as compared to below the standard. The shallower slopes above the standard are both for graphs generated by the lowest standard (25 dB), where as for the moderate standards $(77, 81$ dB), the slopes appear unchanged on either side of the standard. It almost appears that respondents had established an upper bound to the response scale, and so exhibited response attenuation to achieve that.

Also, in our theory one should treat the abscissa as the intensity less the threshold intensity, which these authors had no reason to do. This has the potential of changing the slopes closest to threshold, that is, for Hellman and Zwislocki's (1961) data below the standard ($p < 1$) but not for intensities well above threshold ($p > 1$). They reported an average threshold of 6 dB SPL, which is clearly too small to alter the results in a material way. Nevertheless, were these experiments repeated, we would favor the data be plotted in terms of the intensity less the threshold intensity for individual respondents.

We conducted an analysis of the apparent effect of standards and moduli on slope value and concluded, first, that the data are consistent with slopes below the standard decreasing with increasing standards and above it to increase with decrease in standard. Second, by overlaying the graphs of the two studies, the data are consistent with slopes both below and above the standard to increase with decreasing moduli.

Poulton (1968), who examined the same data sets, came to a conclusion similar to ours. He modeled these effects in his Fig. 1C, according to which there is a range of standards and moduli for which $p(z,t) = p(z)$ is true in magnitude estimation. Although we do not test this hypothesis, the assertion that pairs of standards and moduli can be chosen such that magnitude estimates above and below the standard are the same, does accord with the available data. That is, ratio independence, (49), is satisfied in at least some cases.

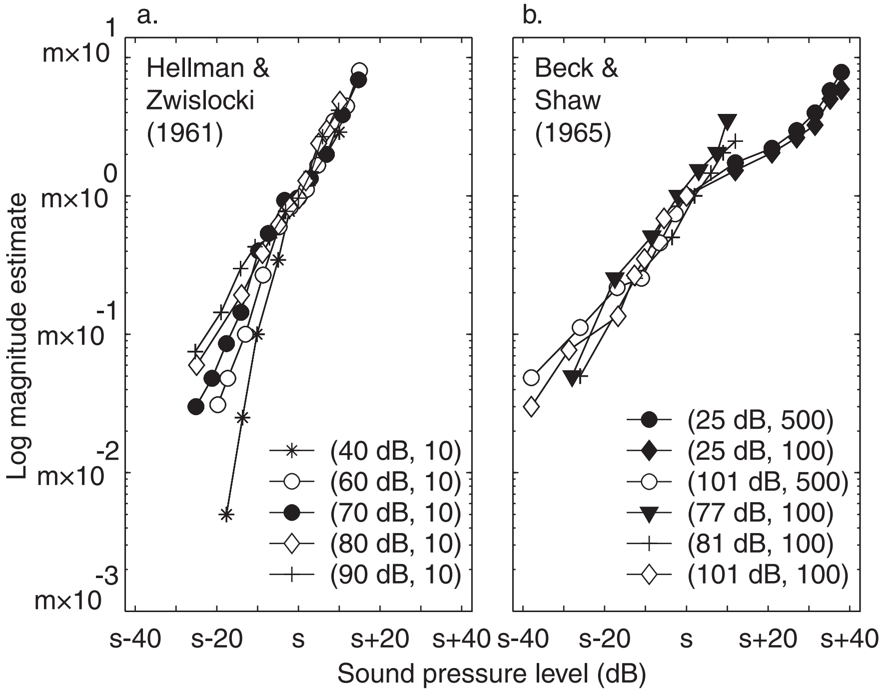

Fig. 1: Panel a. contains auditory data adapted from Fig. 6 of Hellman and Zwislocki (1961). Plotted in panel b. are data adapted from Fig. 1 of Beck and Shaw (1965). Each graph shows results of magnitude estimates as a function of stimuli in dB SPL and with respect to a common standard (s) and modulus (m), indicated as (s, m).

6.2. Psychophysical Power Functions

Anyhow, assuming that (48) holds, then (49) immediately yields

$$W(p_i(t)) = \frac{\psi_i(tx)}{\psi_i(x)}, \tag{50}$$

which is a Pexider functional equation (Aczél, 1966, p. 144) whose solutions with $\psi_i(0) = 0$ are, for some constants, $\alpha_i > 0, \beta_i > 0$,

$$\psi_i(t) = \alpha_i t^{\beta_i} \quad (t \geq 0), \tag{51}$$

$$W(p_i(t)) = t^{\beta_i} \quad (t \geq 0). \tag{52}$$

Recall that the ψ_i are the production psychophysical functions all defined in terms of Ψ by (45) for $i = l$ and by (46) for $i = r$. So (51) agrees with our earlier result about sums of power functions being implied when multiplicative invariance is satisfied (Section 5.4). And, of course, (49) holds if the psychophysical function is a power function.

6.3. Do Ratio Estimates Also Form Power Functions?

The conclusion (52) tells us that, when we observe empirically the estimation function $p_i(t)$, it is a power function, but it is seen through the distortion W^{-1}. Stevens (1975) claimed that the magnitude estimation psychophysical functions are, themselves, power functions, which was approximately true for geometric means over respondents; however this is not really the case for data collected on individuals (see Fig. 2). This fact is again a caution about averaging over respondents.

Moreover, Stevens (1975) attempted to defend the position that both the magnitude and production functions are power functions, although he was quite aware that empirically they do not prove to be simple inverses of one another (p. 31). Indeed, he spoke of an unexplained "regression" effect which has never really been fully illuminated (Stevens, 1975, p. 32).

So let us consider the possibility that, as Stevens claimed,

$$p_i(t) = \rho_i t^{\nu_i} \quad (t > 0, \rho_i > 0, \nu_i > 0). \tag{53}$$

Note that because p_i is dimensionless, the parameter ρ_i is a constant, not a free parameter. It is quite easy to see that if (52) holds, then p_i is a power function, (53), if, and only if, $W(p)$ is also a power function with exponent $\omega_i := \beta_i/\nu_i$, that is,

$$\begin{aligned} W_i(p) &= \left(\frac{p}{\rho_i}\right)^{\omega_i} \\ &= W_i(1)p^{\omega_i} \quad (p \geq 0). \end{aligned} \tag{54}$$

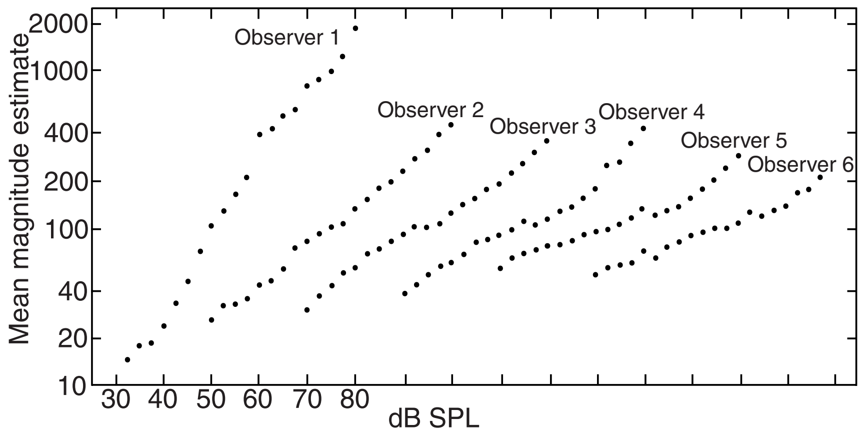

Fig. 2: Reproduction of Fig. 1 of Green and Luce (1974)

This form has different implications depending on whether the constant $\rho_i = 1$ or $\neq 1$. Note that $\rho_i = 1$ holds if, and only if, $W_i(1) = 1$. From here on we assume that W_i, being a cognitive function, is independent of $i = l, r$ and so can be denoted W. Both cases rest on an exploration of the property of threshold production commutativity:

$$(x \circ_{p,i} 0) \circ_{q,i} 0 = (x \circ_{q,i} 0) \circ_{p,i} 0 = x \circ_{t,i} 0, \tag{55}$$

which by (13), is equivalent to

$$W(p)W(q) = W(t). \tag{56}$$

To increase generality, we suppose that (56) holds for $p > 1, q > 1$, and separately, for $p < 1, q < 1$, but not necessarily for the crossed cases: $p > 1 > q$ or $q > 1 > p$. Assuming the continuity of $W(p)$ at $p = 1$, it is easy to show that if this obtains, then the following statements are equivalent: (1) There exist constants ω and ω^* such that

$$W(p) = W(1) \begin{cases} p^{\omega}, & p \geq 1 \\ p^{\omega^*}, & p < 1 \end{cases}. \tag{57}$$

(2) The relation among p, q, and t is given by:

$$t = kpq \text{ where } k = \begin{cases} W(1)^{1/\omega}, & p \geq 1 \\ W(1)^{1/\omega^*}, & p < 1 \end{cases}. \tag{58}$$

We call (58) $k-$multiplicative. If we also assume that (56) holds for $p > 1 > q$ or $p < 1 < q$, then $\omega = \omega^*$. Some pilot data we have collected strongly suggests that (58) does not hold for the crossed cases $p > 1 > q$ and $p < 1 < q$ and that $W(1) < 1$. Further empirical work is reported in SL-IV.

The only published data concerning (55), of which we are aware, are those of Ellermeier and Faulhammer (2000) and Zimmer (2005). They restricted their attention to the case of $\rho_i = 1$, which is equivalent to $W(1) = 1$, because Narens (1996) arrived at (58) with $k = 1$ as a consequence of his formalization of what he believed Stevens (1975) might have meant theoretically when invoking magnitude methods. Ellermeier and Faulhammer (2000) and Zimmer (2005) tested it experimentally and unambiguously rejected it. To our knowledge no one that we know, other than us, has attempted to collect sufficient data to see how well (58) fits the data with $\rho_i \neq 1$ in (53). Our preliminary data are promising, but incomplete.

So the answer to the question of the heading—"Do ratio estimates form power functions?"—is that at this point we do not know. The key prediction (58) has yet to be fully checked. If, however, the general power function form is rejected, then the task of finding the form of W remains open. We discuss next an interesting, but ultimately unsuccessful, attempt: the Prelec function.

6.4. If Ratio Estimation Is Not a Power Function, What Is W?

Prelec's Function

Within the context of utility theory for risky gambles and for $0 < p \leq 1$, a weighting function was proposed and axiomatized by Prelec (1998) that had the desirable feature that, depending on the combinations of the parameters, the function can be concave, convex, S-shaped, or inverse S-shaped. Empirical data on preferences among gambles seemed to suggest that the inverse S-shaped form holds (Luce, 2000, especially Fig. 3.10 on p. 99). The Prelec form for the weighting function, generalized from the unit interval to all positive numbers is

$$W(p) = \begin{cases} \exp\left[-\lambda\left(-\ln p\right)^{\mu}\right] & (0 < p \leq 1) \\ \exp\left[-\lambda'\left(\ln p\right)^{\mu}\right] & (1 < p) \end{cases}, \tag{59}$$

where $\lambda > 0$, $\lambda' > 0$, and $\mu > 0$. The special case of $\mu = 1$ is a power function with $W(1) = 1$, which we know is wrong.

Reduction Invariance: A Behavioral Equivalent of Prelec's W

Prelec gave one axiomatization of the form (59) and Luce (2001) gave the following simpler one, called reduction invariance, defined as follows: Suppose that positive $p, q, t = t(p, q)$ are such that (55) is satisfied for all $x > 0$. Then for any natural number N,

$$(x \circ_{p^N,i} 0) \circ_{q^N,i} 0 = x \circ_{t^N,i} 0. \tag{60}$$

In words, if the compounding of p and q in magnitude productions is the same as the single production of t, then the compounding of p^N and q^N is the same as the single production of t^N. On the assumptions that (56) holds for p^N, q^N, and t^N, and that W is strictly increasing function on the interval $]0, 1]$, Luce (2001) showed that reduction invariance, (60), is equivalent to the Prelec function (59) holding in the unit interval. Indeed, it turns out that its holding for two values of N such as $N = 2, 3$ are sufficient to get the result. Another pair that works equally well is $N = 2/3, 2$. One can also show that it works for N any positive real number; however, any two values without a common factor suffice. It is not difficult to see how to extend the proof to deal with the interval $]1, \infty[$

Zimmer (2005) was the first to test this hypothesis and she rejected it. Her method entailed working with bounds and showing that the observed data fall outside them. In SL-IV, we also tested it using our ratio-production procedure. We too found that it failed. The fact that W is a cognitive distortion of numbers may mean that it will also fail empirically in other domains, such as utility theory, when reduction invariance is studied directly.

Testing was done using two-ear ($i = s$) productions. First, the two successive estimates

$$v_s = x \circ_{p,s} 0, \tag{61}$$

$$t_s = z_s \circ_{p,s} 0, \tag{62}$$

were obtained. Then, using the simple Up-Down method (Levitt, 1971), a t was estimated such that $x \circ_{t,i} 0 \sim t_s$. With the estimate of t and our choice for N, the following estimates were obtained:

$$t'_s = (x \circ_{p^N,i} 0) \circ_{q^N,i} 0$$

$$w'_s = x \circ_{s^N,i} 0.$$

The property is said to hold if the hypothesis $t'_s = w'_s$ is not statistically rejected.

We used the two instantiations, $x = 64$ dB and $x = 70$ dB, and the proportions, presented as percentages, $p = 160\%$ and $q = 80\%$, except for one respondent where $q = 40\%$ and another where $p = 140\%$. The power

N was chosen as close to 2 as would provide numbers close to a multiple of five for each of p^N, q^N, and t^N.

The property was rejected for six of six respondents. For three, the failure was beyond much question. But taking into account the complexity of the testing procedure and the multiple levels of estimation, the failure for the other three was not dramatic. Indeed, had our data been as variable as Zimmer's (2005), we almost certainly would have accepted the property of reduction invariance in those three cases.

When we tested reduction invariance, we did not know about the potential problems outlined earlier (Section 6.3) of testing this property using the mixed case of $p > 1$, $q < 1$. Without further testing, the failure we observed is potentially related to this issue; however Zimmer's (2005) data are not based on mixed cases; she used $p < 1$, $q < 1$. This further suggests that the property should be tested with $p > 1$, $q > 1$; we aim to report such data in SL-IV.

No one has yet explored what happens to reduction invariance if it is assumed that the right side of (59) is multiplied by $W(1) \neq 1$.

6.5. Predictions About Covariation and Sequential Effects

When ψ_i is assumed to be a power function, we have the following inverse relations between ratio productions and ratio estimates:

$$r_i(p) = W(p)^{1/\beta_i} \quad (p \text{ given}), \tag{63}$$

$$p_i(r) = W^{-1}(r^{\beta_i}) \quad (r \text{ given}). \tag{64}$$

In the usual dB form in which data are plotted these are

$$r_{i,dB}(p) = \frac{1}{\beta_i} W_{dB}(p), \tag{65}$$

$$p_{i,dB}(r) = W_{dB}^{-1}(r^{\beta_i}) = W_{dB}^{-1}\left(\exp \frac{1}{10}\left(\beta_i r_{dB}\right)\right). \tag{66}$$

What Happens When W Is a Power Function?

If we suppose that W is a power function of the form (54), then a routine calculation yields

$$W^{-1}(r^{\beta_i}) = \rho_i r^{\nu_i},$$

and so

$$r_{i,dB}(p) = \frac{1}{\upsilon_i}\left(p_{dB} - \rho_{i,dB}\right),$$

$$p_{i,dB}(r) = \nu_i r_{dB} + \rho_{i,dB}.$$

In response to overwhelmingly clear empirical evidence, several authors have formulated sequential models in which the response in dB on trial n, $10 \log R_n$, depends linearly on the present signal in dB, $10 \log S_n$, the previous one, $10 \log S_{n-1}$, the previous response $10 \log R_{n-1}$, and in some cases, $10 \log S_{n-2}$ (DeCarlo, 2003; DeCarlo & Cross, 1990; Jesteadt, Luce, & Green, 1977; Lacouture, 1997; Lockhead, 2004; Luce & Steingrimsson, 2003; Marley & Cook, 1986; Mori, 1998; Petrov & Anderson, 2005)[13]. Both Lockhead (2004) and Petrov and Anderson (2005) provided many other references to the literature.

Setting

$$r_{s,n} = \frac{S_n}{S_{n-1}}, p_{s,n} = \frac{R_n}{R_{n-1}},$$

then each weighting function yields a sequential model for estimation. With symmetric stimuli (x, x), we see that for power functions

$$R_{n,dB} = R_{n-1,dB} + \nu_s \left(S_{n,dB} - S_{n-1,dB}\right) + \rho_{i,dB}.$$

What Happens When W Is a Prelec Function

If we assume that W is given by (59), then putting that form into the expressions for (65) and (66), doing a bit of algebra, and defining $\tau_i := \frac{1}{\beta_i}\left(\frac{\log 10}{10}\right)^{\mu-1}$, yields the following forms for $r_i(p)_{dB}$ and $p_i(r)_{dB}$, respectively:

$$r_i(p)_{dB} = \tau_i \begin{cases} -\lambda \left(-p_{dB}\right)^{\mu} & 0 < p \leq 1 \\ \lambda' \left(p_{dB}\right)^{\mu_i} & 1 < p \end{cases}, \tag{67}$$

$$p_i(r)_{dB} = \frac{1}{\tau_i^{1/\mu}} \begin{cases} -\left(-\frac{\beta_i}{\lambda} r_{dB}\right)^{1/\mu} & 0 < r \leq 1 \\ \left(\frac{\beta_i}{\lambda'} r_{dB}\right)^{1/\mu} & 1 < r \end{cases}. \tag{68}$$

When μ is approximately 1, then $r_i(p)$ and $p_i(r)$ are approximately power functions, that is, $r_i(p)_{dB}$ and $p_i(r)_{dB}$ are approximately linear, but with a change of the power at $p = 1$ and $r = 1$, respectively. Of course, it cannot be exactly a power function without contradicting the data of Ellermeier and Faulhammer (2000) and Zimmer (2005). Some examples of (67) are shown in Fig. 3. Such functions seem consistent with the data reported in Fig. 1.

[13] We thank A. A. J. Marley for supplying some of these references.

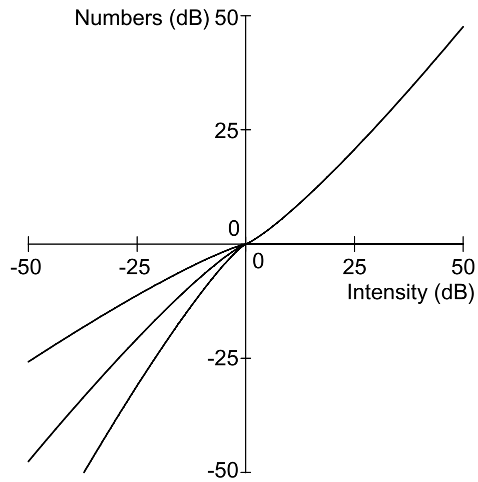

Fig. 3: $\psi(x) = \begin{cases} (.5x)^{1.2} & \text{if} \quad 0 \le x < 50 \\ -(-.cx)^{1.2} & \text{if} \quad -50 < x < 0, \end{cases}$
were $c = 0.3, 0.5, 0.7$ from top to bottom.

Proceeding as with the power function but using Prelec's function, we get the following predicted sequential effects:

$$R_{n,dB} = R_{n-1,dB} + \tau_s^{1/\mu} \begin{cases} -\left[-\frac{\beta_s}{\lambda}\left(S_{n,dB} - S_{n-1,dB}\right)\right]^{1/\mu} & S_{n,dB} \le S_{n-1,dB} \\ \left[\frac{\beta_s}{\lambda'}\left(S_{n,dB} - S_{n-1,dB}\right)\right]^{1/\mu} & S_{n,dB} > S_{n-1,dB} \end{cases}.$$

In commenting on an earlier draft of this chapter, A. A. J. Marley raised the following issue: " ... An important phenomenon related to sequential effects (especially in absolute identification) is *assimilation* of responses to the value of the immediately previous stimulus (with smaller *contrast* effects for earlier stimuli)." (Personal communication, December 11, 2004.) For W a power function with $W(1) = 1$, they are not predicted. No one yet has investigated these phenomena when $W(1) \neq 1$. What happens in the Prelec case is not yet clear. The rank order of signal intensities seems to matter substantially.

Some aspects of Stevens's magnitude estimation and production functions may be illuminated by these results. Let us assume that when the experimenter provides no reference signal x, each respondent selects his or her own. Thus, the usual data, which are averaged over the respondents, is the average of approximately piece-wise linear functions with the break occurring in different places. Although (63) and (64) are perfect inverses, it is no surprise that under such averaging, the results are not strict inverses of one another. Something like this may provide an account of Stevens's "regression" phenomenon.

7. Summary and Conclusions

7.1. Summary of the Theory

The theory has three primitives:

1. The (loudness) ordering $\succsim$ on $\mathbb{R}_+ \times \mathbb{R}_+$, where $\mathbb{R}_+$ is the set of non-negative numbers corresponding to signals which are intensities less threshold intensity (intensities less than the threshold are set to 0).
2. The presentation of signal pairs, (x, u), to (e.g., the two ears of) the respondent with the defined matching operations $\oplus_i$.
3. Judgments of "interval" proportions, $\circ_p$.

Within the fairly weak structural assumptions of the theory, necessary and sufficient properties were stated that yield the representations: There exist a constant $\delta \geq 0$ and two strictly increasing functions Ψ and W such that

$$\Psi(x,u) = \Psi(x,0) + \Psi(0,u) + \delta\Psi(x,0)\Psi(0,u) \quad (\delta \geq 0),$$

$$W(p) = \frac{\Psi\left[(x,x) \circ_p (y,y)\right] - \Psi(y,y)}{\Psi(x,x) - \Psi(y,y)} \quad (x > y \geq 0),$$

and, under some conditions, there is a constant $\gamma > 0$ such that

$$\Psi(x,0) = \gamma\Psi(0,x),$$

which is quite restrictive.

The property characterizing the form of $\Psi(x,u)$ is the Thomsen condition, (23). We showed next that for most people, the ears are not symmetric in the sense that $(x,u) \nsim (u,x)$, in which case $\delta = 0$ is equivalent to bisymmetry of the operation $\oplus_s$. The property underlying the second expression, the one involving $\circ_p$, is production commutativity, (25). Axiomatized by themselves, these representations really are $\Psi_\oplus$ and $\Psi_{\circ_p}$ and they are not automatically the same function. To establish that equality requires two linking expressions, SJP-decomposition, (31), and one of two forms of segregation, either (32) or (33). These are two types of distribution conditions.

Next we took up the form of $\Psi(x,u)$ in terms of the intensities x and u. The property of σ−MI, (34) and (35), turns out to be equivalent to Ψ

being a sum of two power functions with the ratio of the exponents being σ. A predicted linear regression permits one to estimate σ. We also explored a simple filtering model to allow one to account for the, to us, unexpected phenomena connected with asymmetric matching. If the filter takes the form of an attenuation factor η, one can show that none of the tests of properties that we used with asymmetric matching are invalidated by the filtering. We gave formulae for estimating η and γ, respectively (42) and (43).

Our final topic was the form of the ratio estimation predicted by the theory. The results depend heavily on the assumed form of $W(p)$ as a function of p. We explored two cases: one where ratio estimates are power functions and $W(1) \neq 1$; and one where $W(1) = 1$ and W is a Prelec function, which has the nice properties of being either concave, convex, S-shaped, or inverse S-shaped depending on the parameter pairs. Both cases offered accounts of magnitude methods without a standard and of the ubiquitous sequential effects. Just how viable they are relative to data remains to be examined.

The case when W is a power function and $W(1) \neq 1$ leads to a prediction that has not been explored. That of the Prelec function for W has been shown to be equivalent to a behavioral property called reduction invariance, (60), with two studies, one of them ours, that both show that this condition fails. Thus, the problem of the form of W remains open but with a clear experiment to test the power function assumption.

7.2. Summary of Experimental Results

The theory discussed implies that properties Thomsen, 2, proportion commutativity, 4, JP-decomposition, 5, and segregation, 6, in Table 3 should hold. This is summarized in the top portion of the flow diagram of Fig. 4. Although the results are not perfect, we are reasonably satisfied. Issues concerning the forms of Ψ and W as functions are summarized at the bottom of the Fig. 4; they are clearly in much less satisfactory form at this point.

Had property jp-symmetry, that is, $(x, u) \sim (u, x)$, been sustained, which it was not, we could have used a somewhat simpler theoretical development formulated for utility theory. Given that it was not sustained, we know that bisymmetry, Property 3, corresponds to $\delta = 0$ in the representation

$$\Psi(x, u) = \Psi(x, 0) + \Psi(0, u) + \delta\Psi(x, 0)\Psi(0, u).$$

The data sustained bisymmetry, so we accepted that $\delta = 0$. Three implications follow: First, the peculiarities that we observed with asymmetric matching are predicted by a simple filtering model. Second, if $\sigma-$MI holds, various properties, 2 to 6, are not altered by the filtering model (see Table 1). Third, $\sigma-$MI is equivalent to $\Psi(x, u)$ being a sum of power functions.

Table 3: Summary of experimental results.

Property	Number of Respondents	Number of Tests	Number of Failures
Joint–Presentation Symmetry	15	45	23
Thomsen	12	24	5
Bisymmetry	6	6	0
Production Commutativity	4	4	0
Joint–Presentation Decomposition	4	8	2
Segregation	4	10	2
1−Multiplicative Invariance	22	44	13[a]
Reduction Invariance	6	12	12

[a] 12 Respondents passed both tests.

The special case of $\sigma = 1$ has been tested and was sustained for about 50% of respondents. For $\sigma-$MI we have developed a regression model for estimating σ and for seven respondents from the 1−MI experiment the estimated σ moved things in the correct direction for five of them. Mostly, however, the correction does not seem to be sufficiently large to expect that $\sigma-$MI will improve matters much.

Given the potential for experimental artifacts, we conclude that sufficient initial support for the general theory has been received to warrant further investigation—both for auditory intensity and for other interpretations of the primitives. However, questions about the forms of Ψ as a function of physical intensity and about W as a function of its argument remain unsettled.

7.3. Conclusions

The studies summarized here seem to establish the following points:

1. As in classical physics, one does a lot better by having two or more interlocked primitive structures rather than just one in arriving at constrained representations. Our structures were $\langle \mathbb{R}_+ \times \mathbb{R}_+, \succsim \rangle$, which we reduced to the one dimensional structures $\langle \mathbb{R}_+, \geq, \oplus_i \rangle$, and $\langle \mathbb{R}_+ \times \mathbb{R}_+, \succsim, \circ_p \rangle$, which , in turn, we reduced to the one dimensional $\langle \mathbb{R}_+, \geq, \circ_{p,i} \rangle$.
2. The adequacy of such a representation theory that has both free functions and free parameters can be judged entirely in terms of parameter-free properties without, at any point, trying to fit the representations themselves to data. Again, this is familiar from classical physics.
3. As usual, more needs to be done. Among the most obvious things are:
   – Collect more data. Several specific experiments were mentioned.
   – Continue to try to improve the experimental methodology.

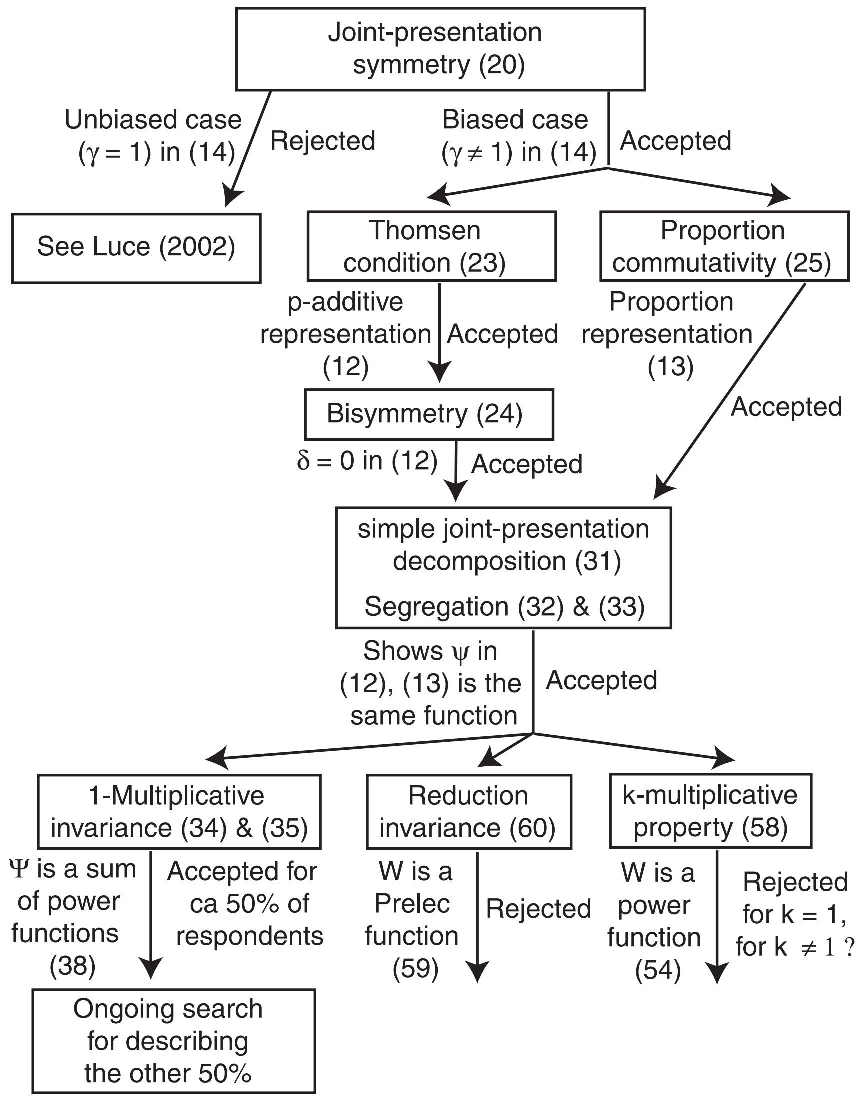

Fig. 4: The diagram shows the main testable properties discussed in this chapter and their inter-relation. A testable property is listed in each box. An arrow leads from that box to another listing the property whose testing logically follows. On the left side of the arrows is the main consequence of accepting the property and the testing result is indicated on the right side of the arrow.

- Statistical evaluation of behavioral indifferences is always an issue in testing theories of this type. We used the Mann-Whitney U test. Recently, an apparently very effective Bayesian method has been proposed for this task (Ho, Regenwetter, Niederée, & Heyer, 2005; Karabatsos, 2005; Myung, Karabatsos, & Iverson, 2005). It should be tried on our data.
- Work out the behavioral condition, presumably corresponding somewhat to $\sigma-$MI, that characterizes the p-additive form of power functions, and then test that empirically.
- Find a form for $W(p)$ as a function of p with $W(1) \neq 1$ that is characterized by a behavioral property that is sustained empirically. Open issues here are both an experiment about the form kpq, (58), and a behavioral condition corresponding to the Prelec function with $W(1) \neq 1$.
- Extend the theory to encompass auditory frequency as well as intensity.
- Study interpretations of the primitives other than auditory intensity. Currently the second author is collecting brightness data which, so far, seems comparable to the auditory data.

4. We are the first to admit, however, that the approach taken is no panacea:

   - We do not have the slightest idea how to axiomatize response times in a comparable fashion.
   - What about probabilistic versions of the theory? Everyone knows that when stimuli are close together, they are not perfectly discriminated and so not really algebraically ordered. Certainly this was true of our data, especially for our data involving ratio productions. Recognition of this fact has, over the years, led to probabilistic versions of various one dimensional ordered structures. But the important goal of blending probabilities with two interacting structures, $\oplus$ and $\circ_p$, in an interesting way has proved to be quite elusive.
   - Also we do not know how to extend the approach to dynamic processes that, at a minimum, seem to underlie both the learning that goes on in psychophysical experiments and the ever-present sequential effects. One thing to recall about dynamic processes in physics is that they are typically formulated in terms of conservation laws (mass, momentum, angular momentum, energy, spin, etc.) that state that certain quantities, definable within the dynamic system, remain invariant over time. Nothing really comparable seems to exist in psychology. Should we be seeking such invariants? We should mention that such invariants always correspond to a form of symmetry. In some systems, the symmetry is captured by the set of

automorphisms, and in others, by more general groups of transformations. For further detail, see Luce et al. (1990), Narens (2002), and Suppes (2002).

Acknowledgements: Many of the experiments discussed here were carried out in Dr. Bruce Berg's auditory laboratory at University of California, Irvine, and his guidance is appreciated. Others were conducted at New York University and we thank the Center for Neural Science and Dr. Malcolm Semple for making resources and laboratory space available to us. We appreciate detailed comments by A. A. J. Marley and J. Gobell on earlier drafts. Some of the experimental work was supported by University of California, Irvine, and some earlier National Science Foundation grants.

References

Aczél, J. (1966). *Lectures on functional equations and their applications.* New York: Academic.

Aczél, J., Falmagne, J.-C., & Luce, R. D. (2000). Functional equations in the behavioral sciences. *Mathematica Japonica*, *52*, 469-512.

Beck, J., & Shaw, W. A. (1965). Magnitude of the standard, numerical value of the standard, and stimulus spacing in the estimation of loudness. *Perceptual and Motor Skills, 21*, 151–156.

DeCarlo, L. T. (2003). An application of a dynamic model of judgment to magnitude estimation. *Perception & Psychophysics*, *65*, 152-162.

DeCarlo, L. T., & Cross, D. V. (1990). Sequential effects in magnitude scaling: Models and theory. *Journal of Experimental Psychology: General*, *119*, 375-396.

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap.* New York: Chapman and Hall.

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

Falmagne, J.-C. (1976). Random conjoint measurement and loudness summation. *Psychological Review, 83*, 65-79.

Falmagne, J.-C., Iverson, G., & Marcovici, S. (1979). Binaural "loudness" summation: Probabilistic theory and data, *Psychological Review, 86*, 25-43.

Gigerenzer, G., & Strube, G. (1983). Are there limits to binaural additivity of loudness? *Journal of Experimental Psychology: Human Perception and Performance, 9*, 126-136.

Green, D. M., & Luce, R. D. (1974). Variability of magnitude estimates: a timing theory analysis. *Perception & Psychophysics, 15,* 291-300.

Hellman, R. P., & Zwislocki, J. (1961). Some factors affecting the estimation of loudness. *Journal of the Acoustical Society of America*, 33, 687-694.

Ho, M.-H, Regenwetter, M, Niederée, R., & Heyer, D (2005). Observation: An alternative perspective on van Winterfeldt et al's consequence monotonicity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 365-373.

Jesteadt, W., Luce, R. D., & Green, D. M. (1977). Sequential effects in judgments of loudness. *Journal of Experimental Psychology: Human Perception and Performance, 3*, 92-104.

Karabatsos, G. (2005). An exchangeable multinomial model for testing deterministic axioms of decision and measurement, *Journal of Mathematical Psychology, 49,* 51-69.

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

Lacouture, Y. (1997). Bow, range, and sequential effects in absolute identification: A response-time analysis. *Psychological Research*, *60*, 121-133.

Levelt, W. J. M., Riemersma, J. B., & Bunt, A. A. (1972). Binaural additivity of loudness. *British Journal of Mathematical and Statistical Psychology, 25*, 51-68.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America, 49*, 467-477.

Lockhead, G. R. (2004). Absolute judgments are relative: A reinterpretation of some psychophysical data. *Review of General Psychology*, *8*, 265-272.

Luce, R. D. (2000). *Utility of gains and losses: Measurement-Theoretical and experimental approaches.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Luce, R. D. (2001). Reduction invariance and Prelec's weighting functions. *Journal of Mathematical Psychology, 45*, 167-179.

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. (2005). Measurement analogies: Comparisons of behavioral and physical measures. *Psychometrika, 70*, 227-251.

Luce, R. D., Krantz, D. H., Suppes, P., & Tversky, A. (1990). *Foundations of measurement, III.* San Diego, CA: Academic.

Luce, R. D., & Marley, A. A. J. (2005). Ranked additive utility representations of gambles: Old and new axiomatizations. *Journal of Risk and Uncertainty*, *30*, 21-62.

Luce, R. D., & Steingrimsson, R. (2003). A model of ratio production and estimation and some behavioral predictions. Berglund, B., & Borg, E. (Eds) *Fechner day 2003: Proceedings of the Nineteenth Annual Meeting of the International Society for Psychophysics,* Stockholm: International Society for Psychophysics (pp. 157-162).

Marley, A. A. J., & Cook, V. T. (1986). A limited capacity rehearsal model to psychophysical judgements applied to magnitude estimation. *Journal of Mathematical Psychology*, *30*, 339-390.

Mori, S. (1998). Effects of stimulus information and number of stimuli on sequential dependencies in absolute identification. *Canadian Journal of Psychology*, *52*, 72-83.

Myung, J. I., Karabatsos, G, & Iverson, G. J. (2005). A Bayesian approach to testing decision making axioms. *Journal of Mathematical Psychology, 49*, 205-225.

Narens, L. (1996). A theory of ratio magnitude estimation. *Journal of Mathematical Psychology*, *40*, 109-129.

Narens, L. (2002). *Theories of meaningfulness.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Petrov, A. A., & Anderson, J. R. (2005). The dynamics of scaling: A memory-based anchor model of category rating and absolute identification. *Psychological Review*, *112*, 383-416.

Poulton, E. C. (1968). The new psychophysics: six models for magnitude estimation. *Psychological Bulletin, 69,*1–19.

Prelec, D. (1998). The probability weighting function. *Econometrica*, 66, 497-527.

Ramsey, F. P. (1931). *The foundations of mathematics and other logical essays.* New York: Harcourt, Brace. (Ch. VII, reprinted in *Studies in Subjective Probability.* H. E. Kyburg, & H. E. Smokler, Eds., 1964. New York: Wiley, pp. 61-92.)

Schneider, B. (1988). The additivity of loudness across critical bands: A conjoint measurement approach. *Perception & Psychophysics*, *43*, 211-222.

Steingrimsson, R. (2002). *Contributions to measuring three psychophysical attributes: Testing behavioral axioms for loudness, response time as an independent variable, and attentional intensity.* Unpublished dissertation, University of California, Irvine.

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*, *49,* 290-307.

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*, *49,* 308-319.

Steingrimsson, R., & Luce, R. D. (in press). Empirical Evaluation of a Model of Global Psychophysical Judgments: III. A Form for the Psychophysical Function and Intensity Filtering. *Journal of Mathematical Psychology*

Steingrimsson, R., & Luce, R. D. (in preparation). Empirical evaluation of a model of global psychophysical judgments: IV. Forms for the weighting function.

Stevens, S. S. (1975). *Psychophysics: Introduction to its perceptual, neural, and social prospects.* New York: Wiley.

Suppes, P. (2002). *Representation and invariance of scientific structures.* Stanford, CA: CLSI Publications.

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