Evaluating a Model of Global Psychophysical Judgments for Brightness: II. Behavioral Properties Linking Summations and Productions

Ragnar Steingrimsson

Department Cognitive Sciences

University of California, Irvine

Ragnar Steingrimsson

Department of Cognitive Sciences

University of California, Irvine

Irvine, CA, 92697-5100

e-mail: ragnar@uci.edu

Abstract

Steingrimsson (2009) outlined Luce's (2002, 2004) proposed psychophysical theory and tested, for brightness, behavioral properties that, separately, gave rise to two psychophysical functions, $\Psi_{\oplus}$ and $\Psi_{\circ_p}$. The function $\Psi_{\oplus}$ maps pairs of physical intensities onto the positive real numbers and represents subjective summation, and the function $\Psi_{\circ_p}$ represents a form of ratio production. This article tests the properties linking summation and production such that it forces $\Psi_{\circ_p} = \Psi_{\oplus} = \Psi$. The properties tested are a form of distributivity and in three experiments are subjected to an empirical evaluation. Considerable support is provided for the existence of a single function Ψ for both summation and ratio production. The scope of the this series of articles is to establish the theory as a descriptive model of binocular brightness perception.

Keywords: Binocular brightness, brightness summation, ratio production, magnitude production, magnitude estimation, psychophysics, matching, bisymmetry, joint-presentation decomposition, distributivity.

Evaluating a Model of Global Psychophysical Judgments for Brightness: II. Behavioral Properties Linking Summations and Productions

Luce's (2002, 2004, erratum, 2008) proposed a theory of global psychophysics.[1] The aim this series of papers is to apply this theory to binocular brightness perception and in so doing establish it as a descriptive model of that domain. In a parallel series of four papers, Steingrimsson and Luce (2005a, 2005b, 2006, 2007, erratum, 2008) evaluated the theory in loudness and provided broad support for the theory in that domain. Steingrimsson (2009) paralleled the work Steingrimsson and Luce (2005a) in loudness but for brightness. He provided evidence for a model of binocular brightness (cyclopic image) in which the percept of physical stimuli arriving at the two eyes could be described by the *summation representation* (3) respondents distortion of numbers in magnitude/ratio productions by the *proportion representation*(4).

A psychophysical function Ψ arises in both representations but as far as the results of Steingrimsson (2009) went the psychophysical functions arising in (3) and (4) are not guaranteed to be the same function. The topic of the current paper is to evaluate necessary and sufficient behavioral properties (axioms) to link the summation and production representations by showing that the same function arises in both representations. In form and content, the article parallels the loudness one of Steingrimsson and Luce (2005b).

Steingrimsson (2009) discusses the motivating background and the underlying theory in detail. To avoid unnecessary repetition, I will refer to the sections of that paper, as applicable, where the reader can seek a more in-depth treatment of relevant topics. However, for this paper to be self-contained, a brief summary of the relevant background and theory is provided.

As an article in a series, the results reported here deal with a portion of a larger whole. To see how the results reported here fit into the larger picture, it is instructive to recall that the psychophysical function is a formal model that maps physical stimuli into sensations. The linking of the two representations, (3) and (4), is equivalent to stating that this description, the psychophysical function, does not dependent on the operation the respondent is asked to perform. This is critical because a conceptual novelty of Luce's (2002, 2004) theory is the harnessing of the power of two operations to arrive at a complete model. This linking, typical in physics, has not been a feature of formal psychological modeling so the data and the question are novel.

As we shall see, the results reported favor the single psychophysical function and thereby confirm the model as applicable to binocular brightness. This result along with the auditory work (Steingrimsson & Luce, 2005a, 2005b) places the work into an even larger context in that the same model is found to capture behavior in two domains, which have traditionally mostly been modeled and studied separately. The unification of the two domains occurs on the level of description and as such suggests that many psychophysical phenomena can, as is the case in physics, can be described using a common language and on the basis of common primitives. Indeed, as is mentioned in the General Discussion, this fact has already shown itself useful for addressing several outstanding issues.

The paper is organized as follows:

- Relevant theory and its interpretation in brightness is summarized.
- Results from three experiments, which together link the two representations of the theory, are reported.
- A comprehensive summary of the current paper in context of the series' first paper is provided along with direction of subsequent research—see Figure 3 for a quick overview.

The theory, its interpretation in brightness, & background

The approach is that of axiomatic psychophysics. Steingrimsson (2009) provided an in-depth discussion of the approach (Appendices A-C), its interpretation in brightness (Introduction), and the current theory's relation to existing modeling efforts of binocular brightness perception (Appendix C).

The experimental setting involves a computer monitor, respondents seated in a dark room altering the luminance of achromatic stimuli (squares on a background), displayed on the monitor, according to instructions. In some cases, a stereoscope is used to generate the stimuli (see Figure 1 for a preview).

Luminance of the physical signal is taken to correspond to the sensation of brightness and a change in brightness is the sensation elicited by a change in this luminance, e.g., the sensation that results from varying the luminance setting on a computer monitor. This understanding accords with that of numerous papers (see the review of Grossberg & Kelley, 1999; Ding & Sperling, 2006).

Primitives

Luce's (2002, 2004) theory, as is typical of axiomatic theories, is non-domain specific, wherefore the theory is not one of binocular brightness, but rather the theory is applied to binocular brightness. The theory features three primitives, and applying the theory to brightness involves specifying and interpreting these in binocular brightness (See Appendix B, Steingrimsson, 2009, for details).

Joint presentations, ordering, and matching. Let x and u correspond to physical intensities, e.g., luminance of two achromatic squares on a monitor, then the ordered pair (x, u) is a stimulus that means that a luminance of intensity x and one of intensity u are presented jointly (simultaneously) such that x is presented to the left eye and u to the right eye. This *first primitive* is called a *joint presentation.*

The ordered pairs have the property that $(x, u) \succsim (y, v)$, which means that the stimulus (x, u) is judged to be at least as bright as (y, v). This *ordering*, $\succsim$, is the *second primitive*. The indifference relation $\sim$ is defined by: $(x, u) \sim (y, v)$ if and only if both $(x, u) \succsim (y, v)$ and $(y, v) \succsim (x, u)$. Note, $\sim$, $\succsim$, are used rather than $=$, $\geq$, because the latter refer to ordering of real numbers, and the former to psychological judgments, but the notational similarity refers to $\succsim$ behaving similarly to the ordering $\geq$ of the real numbers.

The ordering is assumed to agree with physical intensity, a feature shown to hold in all but the most extreme viewing conditions (produced under laboratory conditions)(See Steingrimsson, 2009, section on *Fechner's Paradox* for an in-depth discussion). Theoretically, testing is directed to the *symmetric* version of the theory (Luce, 2004), the case where some light is assumed to reach both eyes in binocular viewing conditions—rather reasonable considering the 70-80% overlap in the two eyes' visual fields.

Matching the perceived magnitude of a standard to another percept is a classic method in psychophysics (see Steingrimsson, 2009, Appendix A; e.g., Stevens, 1975, for a comprehensive discussion). Here this method is used and formalized as follows. Each joint presentation (x, u) can be matched perceptually by (z, z), i.e.,

$$(z, z) \sim (x, u). \tag{1}$$

This is referred to as a *symmetric match* but operationally it will be referred to as *brightness matching* or just *matching*. For notational convenience matching can be expressed using the operator notation

$$z =: x \oplus u, \tag{2}$$

where z is defined by 1 and the notation $A := B$ means A is defined by B. Technically $\oplus$ is a binary mathematical operator and is referred to as the *summation operator*.

Note that (x, u) refers to a joint presentation of a signal pair whereas, $x \oplus u$ refers to the cyclopic image that results. The cyclopic image/percept is a unitary percept (our everyday binocular view) that results from the combining of two separate input signals—hence referred to as subjective *summation*—of the two signals. Since the theory is descriptive, nothing is hypothesized about how this summation is biologically accomplished.

Ratio production. The *third primitive* is a generalization of magnitude production. Magnitude production was popularized by Stevens in the mid 20th century and subsequently became a standard psychophysical method (see Steingrimsson, 2009, Appendix A; Stevens, 1975, for a comprehensive discussion). In magnitude production the respondent produces a stimulus that is some proportion of a standard, e.g., two times as bright as a standard.

In ratio production, suppose that $x > y \geq 0$ are intensities and let $p > 0$ be a real number. Let (z, z) denote a signal that the respondent judges to make the brightness "interval"[2] from (y, y) to (z, z) stand in the ratio p to the brightness interval from (y, y) to (x, x). As in matching, (2), it is convenient to write the operation as a mathematical operator having the form: $(x, x) \circ_p (y, y) := (z, z)$, or $x \circ_{p_s} y = z$ in shorthand. The *generalization* part of the *ratio production* can be seen by the fact that it agrees with *magnitude production* when $(y, y) = (0, 0)$.

Notational convention. Let ϵ_l and ϵ_r denote thresholds for the left and the right eye respectively and let x' and u' be intensity arriving at the left and the right eye, respectively; then the notation is $x = x' - \epsilon_l$ and $u = u' - \epsilon_r$. Thus $x = 0$ denotes the threshold intensity (or less) of the left eye stimulus and $u = 0$ denotes the same for the right eye. For stimuli well above threshold, which are used, the difference $x - x'$ is negligible. Luminance intensities are reported in cd/m^2.

Representations of $\oplus$ *and* $\circ_p$

Luce (2002, 2004) formulated a set of necessary and sufficient (and testable) behavioral properties, formulated in terms of the presented primitives, that allowed him to construct a psychophysical function Ψ that is a strictly monotonic mapping of intensity pairs to the non-negative real numbers that preserves the order $\succsim$, and for which there exists a constant $\delta = 0$ or 1 such that

$$\Psi(x,u) = \Psi(x,0) + \Psi(0,u) + \delta\Psi(x,0)\Psi(0,u), \tag{3}$$

and a strictly increasing, numerical distortion function, $W(p)$, from a non-negative number p onto itself, such that

$$W(p) = \frac{\Psi\left[(x,u) \circ_p (y,v)\right] - \Psi(y,v)}{\Psi(x,u) - \Psi(y,v)},\ \left[(x,u) \succ (y,v) \succeq (0,0)\right]. \tag{4}$$

The representation (3) captures the combining of inputs to the left and right eye respectively, and is referred to as the *summation representation* (also known as a *p-additive representation* in the literature). The representation (4) describes the ratio production operation and is referred to as the *production representation*. Mapping subjective inputs into mathematical expressions allows us to bring to bear the toolbox of mathematics on unobservable subjective entities—see the General Discussion for examples of other recent applications.

The representations consist of two unspecified functions Ψ and W, and the constant δ. This allows for enormous freedom for capturing individual differences, as well as extending these representations to different domains. The key is that as long as certain parameter-free behavioral properties are satisfied, the functions and the representations are guaranteed to exist and have that status without the need for any fitting of data to functions (See Steingrimsson, 2009, Appendix B for details).

Steingrimsson and Luce (2006, 2007) addressed the question of the forms of the unknown functions in loudness and for the psychophysical, Ψ, and the weighting function, W, respectively, they found support for $\Psi(x,0)$ $\Psi(0,u)$ being power functions and W as a power or a Prelec function. Parallel work is ongoing for brightness, so far showing support for power functions for $\Psi(x,u)$.

What is known from Steingrimsson (2009)

Steingrimsson (2009) evaluated three behavioral axioms: Joint-Presentation Symmetry, which was rejected meaning that behaviorally the two eyes cannot be taken to be behaviorally identical; the Thomsen condition, which was supported and established additivity over the two eyes and along with certain reasonable background condition established the summation representations (3); and finally, Production Commutativity which established the production representation (4). Together, these results provided evidence that (3) and (4) hold, but separately. That is, in reality there are two distinct psychophysical functions, $\Psi_{\oplus}$ for summations and $\Psi_{\circ_p}$ for productions. At present, it is not known whether $\Psi_{\oplus} = \Psi_{\circ_p} = \Psi$. That question is the topic of the current paper.

Behavioral properties linking summations and productions

Bisymmetry. Luce (2004) initially held that under the assumptions of the theory, $\delta = 0$ was equivalent to the property of *bisymmetry:*

$$(x \oplus y) \oplus (u \oplus v) = (x \oplus u) \oplus (y \oplus v). \tag{5}$$

However, in a correction reported by Luce (2008), bisymmetry is shown to simply be a direct prediction of the summation representation (3). Therefore its holding pertains to that representation only and thus should logically have been evaluated by Steingrimsson (2009). However, this result became clear only well into the review process of that article so its evaluation is presented here.

Simple joint-presentation decomposition. One of the two key linking properties is *simple joint-presentation decomposition*: Using the shorthand $x \circ_{p_s} y = (x, x) \circ_p (y, y)$, then for positive real numbers p

$$(x \oplus u) \circ_{p_s} 0 = (x \circ_{p_s} 0) \oplus (u \circ_{p_s} 0). \quad (6)$$

Replacing $\oplus$ and $\circ_p$ with $+$ and $\times$, it is easy to see the arithmetic parallel to (6).

Distributivity. A second key property to linking the representations of (3) and (4), is the property of *Distributivity*. Distributivity comes in two forms:

Left distributivity:

$$z \oplus (x \circ_{p_s} u) = (z \oplus x) \circ_p (z \oplus u) \quad (u > 0). \quad (7)$$

Right distributivity:

$$(x \circ_{p_s} u) \oplus z = (x \oplus z) \circ_p (u \oplus z) \quad (u > 0). \quad (8)$$

The difference between the two forms of distributivity captures the fact that if the input order to the left and right eyes is switched on one side of the equality, it must also be switched on the other side. Because of the the failure of joint-presentation symmetry (Steingrimsson, 2009), left and right distributivity can not be assumed to be equivalent. [3]

As with (6), by replacing $\oplus$ and $\circ_p$ with $+$ and $\times$, it is easy to see that (7,8) are forms of distributivity of $\oplus$ over $\circ_p$ and if in (6), 0 is replaced by z the relation of (6) to (7,8) becomes evident.

Experiments

Three experiments are presented: The first is a test of bisymmetry (Exp. 1), which is an induced property of the summation representation, (3). The remaining two, the simple joint-presentation decomposition (Exp. 2), and distributivity (Exp. 3), are properties that link the summation and proportion representations and, if found to hold, support the hypothesis that $\Psi_{\oplus} = \Psi_{\circ_p} = \Psi$.

General Method

The testing strategies the experiments have in common are outlined in the following.

Respondents. A total of 12 students from New York University and University of California, Irvine, and the author[4] participated in the three experiments; for practical reasons not all respondents participated in all of the experiments. All respondents reported normal or corrected-to-normal vision and, except the author, received compensation of 10 per session. Every person provided written consent and was treated in accordance with the "Ethical Principles of Psychologists and Code of Conduct" (American Psychological Association, 2002). Consent forms and procedures were approved by the Institutional Review Boards of New York University and UC Irvine.

Stimuli . The stimuli consisted of squares, subtending 10 degrees of visual angle (Exp. 3 uses 5 degrees), of achromatic light (the RGB channels set to the same value) displayed on a computer monitor (e.g., Panel A, Figure 1).

Apparatus. Stimuli were presented using PsychToolbox extensions in MATLAB (Pelli, 1997; Brainard, 1997) generated with an Apple G4. At the University of California, Irvine, stimuli were presented on an 18" NEC Multisync FE 950+17" and at New York University, a ViewSonic P810 CRT, both at a resolution of 1024×768 pixels and refresh rate of 75 Hz. Experiments were conducted in a dark and light-insulated room.

Luminance calibration: Equipment calibration and background conditions were of two kinds. A photometer, PhotoResearch PR-650, was used to measure luminance. In both cases, a Gamma function was estimated by averaging 5 repeated measures of luminance at every 5th of the 255 RGB values (starting from 1).

NYU Calibration: Measures were taken from both the left and right side of the monitor, equidistant from the monitor's center. The luminance measures were fitted to a

Gamma function; the luminance disparity between sides was not appreciable.

UCI Calibration The NYU Calibration was improved with the aim of better counteracting any possible spatial luminance inhomogeneity of the monitor.[5] A Gamma function was determined for each stimulus location. At each location and using a reverse lookup procedure, the RGB value was determined that produced, as closely as possible, the desired luminance.

Data collected at UCI will be marked; otherwise, the data were collected at NYU.

Background and luminance range/steps: The monitors achieved an upper luminance of ∼100 cd/m^2 with the lowest stimulus level at ∼8 cd/m^2. Initially, stimuli were displayed on a no-luminance level (zeros for all RGB channels) background. In order to minimize the mixing of scotopic and photopic conditions, later experiments used 3.4 cd/m^2 background luminance, a level at which photopic vision is dominant (R. Blake, personal communication, September 12, 2007). This change largely coincides with the location change from NYU to UCI. To maximize available adjustment options, all available stimulus values were used.[6]

Stereoscope: A stereoscope (Panel B, Fig. 1) aided in the generation of some of the stimuli. A stereoscope is a mirror system that accomplishes the projection of left (right) half of the monitor to the left (right) eye. Thus a stimulus of intensity x (u) on the left (right) side, viewed through a stereoscope creates the cyclopic image of the stimulus primitive (x, u)—see joint presentations for details.

Procedure. An experimental session lasted no more than an hour. An initial session was used to obtain written consent, explain the task, answer questions, and run practice trials. When practiced, respondents typically completed around 60 estimates per session, organized into blocks of six or eight estimates. The block structure allowed for frequent rest periods which were encouraged but their frequency and duration were under respondents' control. Information about the current block and trial number were displayed in small letters in the upper left corner of the screen. Respondents received a minimum of 10 minutes of dark adaption prior to each session. Training was provided for each of the two tasks used (matching or magnitude/ratio production, as applicable).

Statistical method and presentation of results. We seek to evaluate parameter-free null hypotheses that have the generic form $L_{\text{side}} = R_{\text{side}}$. The statistical analysis is approached from three directions.

- As we have no theory that predicts the distributions of the estimates, a nonparametric test (Mann-Whitney U) is used for statistical evaluation, with a significance level of $.05$. This has become a common practice in similar studies (e.g., Ellermeier & Faulhammer, 2000; Zimmer, Luce, & Ellermeier, 2001; Ellermeier, Narens, & Dielmann, 2003; Zimmer, 2005; Steingrimsson & Luce 2005a, 2005b, 2006, 2007; Steingrimsson, 2009). Since intensity steps are discrete and estimates appear reasonably Gaussian, medians are known to be best estimated by the mean, and variability thus indicated by standard deviations. Hence these are the central tendency indicators reported.
- Since we have no a priori model of how individuals relate, all data analysis is done on individual data (e.g., Luce, 1995, p. 20).
- To evaluate whether the sample size is sufficiently large to detect a true failure of the null hypothesis, all statistical results were verified using Monte Carlo simulations (see Steingrimsson & Luce, 2005a, and Steingrimsson 2009, for details).
- No clear method exists for calculating the effect size for the Mann-Whitney U. Instead, I make use of the simple observation that should two medians (means) differ by less than Weber's fraction, they are arguably not noticeably different to an observer.[7] Teghtsoonian (1971) reports the mean Weber's fraction for brightness

from five, reportedly conservative, studies, to be 0.08.

The Mann-Whitney test, the Monte Carlo simulation, and the effect-size evaluation, together form the statistical criterion that will be required to be met for accepting the null hypothesis as supporting a given test—Steingrimsson (2009, Appendix A in particular) provides a more extensive discussion.[8]

Experiment 1: Bisymmetry

Method. Empirical testing of (5) requires obtaining several respondent-generated matches.

The summation operation, $\oplus$, and matches: The task is to find (z, z) that is perceive as equally bright as (x, u). Figure 1 describes the process: Panel A depicts what is displayed on the monitor, where the letters indicate stimulus intensity. Panel B depicts the stereoscope through which the respondents view the monitor. Panel C depicts what the subject sees. Since the stereoscope creates a cyclopic image, the percepts are those of $z \oplus z$ and $x \oplus u$.

To produce brightness matching, respondents adjust the intensity of z until they are satisfied that the two percepts—the upper and lower squares in Panel C—are experienced equal in brightness. Respondents used key presses to either adjust the luminance of z or to indicate satisfaction with the brightness match. Respondents could choose any of four luminance steps of 1, 2, 4, or 8 RGB values—equal change on the three channels—(described as extra-small, small, medium, large). After an adjustment, the screen was set to uniform background luminance for 100 ms and then the next stimulus was presented—subjectively, this was experienced as a blinking and signaled that the adjustment had been made. This process was repeated until respondents were satisfied with the match, which they indicated by a key-press, at which time the trial ended and z was recorded as the response. In verbal instructions to respondents, the task was explained as that of making the upper stimulus equal in brightness to the lower one.

Empirical testing of (5) requires obtaining six matches made in two steps. The left side of (5) is reduced to a single estimate through three (matching) estimates

$$w = x \oplus y,\ w' = u \oplus v, \text{ and then } t = w \oplus w'.$$

And similarly for the right side of (5)

$$z = x \oplus u,\ z' = y \oplus v, \text{ and then } t' = z \oplus z'.$$

The bisymmetry property is said to hold if t and t' are found to be statistically equivalent.

These six estimates were made within a block of trials.[9]

The intensities used were: $x = 13.01\ cd/m^2$, $y = 29.84\ cd/m^2$, $u = 53.78\ cd/m^2$, and $v = 84.93\ cd/m^2$.

Results. The data for seven respondents are presented in Table 1. Listed by respondents are the means and standard deviations for t and t', the number of observation, n, for each, and result of the hypothesis test $t \sim t'$, given as $p_{t \sim t'}$.

The bisymmetry property was not rejected for any of the seven respondents.

Discussion. The bisymmetry property is an induced property of the summation representation (3). Its being supported in 7/7 test provides favorable support for that representation.

Experiment 2: Simple joint-presentation decomposition

Method. The goal is to evaluate simple joint-presentation decomposition (6). This evaluation requires obtaining respondent-generated matches and several magnitude productions. The matching procedure was outlined in Experiment 1, here the magnitude production procedure is described.

Magnitude production: The task $(z', z') = (x', u') \circ_p (0, 0)$ is to produce a stimulus (z', z') that is perceived as a proportion p of the the standard (x', u'). Note, when $p = 1$

this task is in form equivalent to matching. Thus a magnitude production may be obtained using a procedure and stimulus identical to the matching task with only the addition of a proportion instruction (see method for Experiment 1 for details).

Recalling that $x \circ_{p_s} y = (x, x) \circ_p (y, y)$, (6) may be written as

$$((x \oplus u), (x \oplus u)) \circ_p (0, 0) = ((x, x) \circ_p (0, 0)) \oplus ((u, u) \circ_p (0, 0)). \tag{9}$$

Together, trial types A-B reduce the left side of (9) to a single estimate t, and in the subsequent trials C-E, the right side (9) is reduced to the estimate t'.

A: $(z, z) \sim (x, u) \circ_1 (0, 0)$

B: $(t, t) \sim (z, z) \circ_p (0, 0)$

C: $(v, v) \sim (x, x) \circ_p (0, 0)$

D: $(w, w) \sim (u, u) \circ_p (0, 0)$

E: $(t', t') \sim (v, w) \circ_1 (0, 0)$

(Note that in trial types A and E $p = 1$, which is equivalent to matching). The property is said to be supported if the hypothesis $t \sim t'$ is not rejected.

A theoretical prediction is that the property holds for both $p < 1$ and $p \geq 1$. Two production conditions meeting these constraints were chosen, namely $p = 2/3$ and $p = 2$. The luminance levels used in the case of $p = 2$ were $x = 12.98\ cd/m^2$ and $u = 29.80\ cd/m^2$ and for $p = 2/3$ the luminance values were $x = 29.80\ cd/m^2$ and $u = 53.74\ cd/m^2$.

With two proportion instructions, there were a total of 10 trial forms. These were run randomized[10] within a block of trials. The p-instruction was displayed in the upper left corner of the monitor.

To respondents, the task was described as making the upper square appear, e.g., twice $(p = 2)$ or two-thirds $(p = 2/3)$ that of the lower square. For the matching trials, the instruction was $p = 1$, and the task explained as equivalent to matching. Respondents were initially observed making the adjustments to help ensure complete understanding of the task.

Results. In Table 2, the means and standard deviations for t and t' are provided along with the proportion instruction p. In addition, the number of observations is indicated by n and the statistical test results given by $p_{t \sim t'}$.

The property held in both proportion conditions in 12/12 tests. Respondent R4 showed marginal $p_{t \sim t'}$ values but they were found to hold using the Monte Carlo evaluation of adequacy of the samples, but perhaps more convincingly, the obtained means were well within Weber's fraction so on all three components of the statistical criterion the evidence favored not rejecting the null hypothesis (see the General Method's statistical section for details).

Discussion. With the property not rejected for any of the respondents for either the $p < 1$ or the $p > 1$ conditions in 12 tests (all holding), the property is seen as having received reasonable initial support in brightness.

Experiment 3: Distributivity

Method. Distributivity comes in two forms, left (7) and right (8) distributivity. The method for testing left distributivity will be outlined; the method for right distributivity is analogous. The testing required matching and several ratio productions. The matching procedure was outlined in Experiment 1, here the ratio production procedure is described.

Ratio Production: The task $(t', t') = (z, x) \circ_p (z, u)$ is that of producing a stimulus (t', t') that makes the brightness "interval" from (z, u) to (z, x) be a proportion p of the "interval" from (z, u) to (t', t'). Figure 2 illustrates how this is accomplished. Panel A shows the stimuli as presented on the screen. Panel C illustrates the percept the respondent sees when the display in Panel A is viewed through the stereoscope of Panel

B. Recalling that the percept of any (z, x) is that of $(z \oplus x)$, then by a method analogous to that of matching (Fig. 1), the respondent need only adjust the luminance of t' to arrive at an estimate of the the ratio production.

The adjustment procedure is the same as for matching; the proportion p was displayed on the upper left side of the monitor (as noted earlier and to accommodate the number of stimuli on the monitor, each square subtended 5 degrees of visual angle).

Recalling that $x \circ_{p_s} y = (x, x) \circ_p (y, y)$, then (7) may be written as

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

The left side of (10) is reduced to a single estimate t using the trial types A-B; the right side is reduced to a single estimate t' using the trial type C.

A: $(v, v) \sim (x, x) \circ_p (u, u)$

B: $(t, t) \sim (z, v)$

C: $(t', t') \sim (z, x) \circ_p (z, u)$

Left distributivity is found to be supported if the hypothesis $t \sim t'$ is not rejected.

The property is predicted to hold for both $p < 1$ and $p \geq 1$, hence $p = 2$ and $p = 2/3$ were used. The RGB values for u and z were fixed in all conditions and thus the corresponding luminance values varied slightly between the two monitors used. At NYU (UCI), these were $u = 10.50$ (7.46) cd/m^2 and $z = 18.90$ (15.06) cd/m^2. Three values for x were used and mixed over proportion conditions; these are specified in Table 3.

With two proportion instructions, and three trial types, there was a total of six trial forms. The four trials (two each of type A and C) were randomized with a block and run within a session. A second session was comprised of the other two trials (type B) run in a single block consisting of each individual estimate of v (from A) matched with z. The p-instruction was displayed in the upper left corner of the monitor.

The task was described to respondents as adjusting the brightness of the upper right square such that the adjusted interval between it and the lower right square be twice ($p = 2$) or two-thirds ($p = 2/3$) that of the reference interval, the lower left to the upper left squares. Respondents were initially observed making the adjustments to help ensure complete understanding of the task.

Results. Six respondents provided data for one or more conditions. The results are reported in Table 3. Listed by respondents are the version of distributivity (left or right) and, by condition, the means and standard deviations for t and t', the luminance value of x, the proportion instruction p, the number of observations n, and the result of the hypothesis test $t \sim t'$, listed under $p_{t \sim t'}$.

The property failed outright in one (R4, left, $p = 2/3$) of 15 tests (on all the three components of the statistical criterion). For two tests (R4, right, $p = 2$, R12, left, $p = 2$), the $p_{t \sim t'}$-value was marginal. But of those, only one (R12, left, $p = 2$) deviated noticeably from the other results on the statistical criterion of effect size. Thus, the property is deemed supported in 13 of 15 tests.

Discussion. Evaluating Distributivity involves arguably the most complex psychophysical task and it is technically the most complex of the current three experiments, as well as of those evaluated by Steingrimsson (2009). Therefore, its being supported in 13 of 15 tests is a good initial support for the property.

General Discussion

Results of the tests of bisymmetry and the two linking properties are summarized in Table 4.

The topic has been the evaluation in brightness of three behavioral properties that arise in a theory of global psychophysical judgments that leads to the two representations:

$$\Psi(x,u) = \Psi(x,0) + \Psi(0,u) + \delta\Psi(x,0)\Psi(0,u) \quad (\delta = 0,1), \tag{3}$$

$$W(p) = \frac{\Psi\left[(x,u) \circ_p (y,v)\right] - \Psi(y,v)}{\Psi(x,u) - \Psi(y,v)} \quad [(x,u) \succ (y,v) \succsim (0,0)], \tag{4}$$

These representations have a number of necessary consequences (behavioral properties) that in turn are sufficient under certain structural conditions to give rise to the representations. Those that underlie (3) and (4) were examined separately and sustained in Steingrimsson (2009). Those that link the two representations and force a common psychophysical function were examined here.

The result of Steingrimsson (2009) and the current article are summarized in Figure 3, as well as how the two relate, and how they together form the conclusion that establishes the representations (3) and (4).

The main conclusion from this empirical evaluations is that the theory of global psychophysics (Luce, 2002, 2004, erratum, 2008) has received a reasonable support in the brightness domain for achromatic stimuli of intensity well above threshold.

These results readily open the way for a third article in the series, one whose topic is the functional form of Ψ and W. This work has been completed for loudness (Steingrimsson & Luce, 2006, 2007) and is in progress for brightness.

The representations (3,4) have been found to capture behavior in two separate domains, loudness and brightness, which to my knowledge no model has accomplished before. The results suggest not only exploring extension of the work to additional domains (work on perceived contrast is under way), but also to the level at which the description of behavior is unified. This question is the topic of Appendix C of Steingrimsson (2009).

In addition to a unification of description, the assumptions leading to (4) that were favorably evaluated by experiments in loudness ,

The results from loudness (Steingrimsson & Luce, 2005a, 2005b, 2006, 2007) and brightness (Steingrimsson, 2009, current) have already proved useful for asking and addressing several novel and outstanding questions. For example, Steingrimsson and Luce (under revision) used (4) to account for the time order error, Luce, Steingrimsson, and Narens (submitted) used the result to establish that the ratio scales of loudness at two frequencies can be interpreted as a common ratio scale of loudness over intensity and frequency pairs. As a direct consequence of the current results, comparable work is in progress on brightness over light intensity and wave length pairs, which, if successful, will lead to attempts to account for cross-modal matching as a single ratio scale of ratio scales of loudness and brightness. An explanation for the regression effect between magnitude production and magnitude estimation has been developed and written in draft form. Finally, work to place the Torgerson's (1961) conjecture—that respondents do not distinguish differences from ratios—on firm theoretical footing is on-going and suggests ways of understanding contradictory reports in the literature along with an experimental evaluation.

References

American Psychological Association (2002). Ethical principles of psychologists and code of conduct. *American Psychologist*, **57**, 1060–1073.

Ellermeier, W., Narens, L., & Dielmann, B. (2003). Perceptual ratios, differences, and the underlying scale. In B. Berglund and E. Borg (Eds.) *Fechner Day 2003. Proceedings of the 19th annual meeting of the International Society for Psychophysics*. Stockholm, Sweden: International Society for Psychophysics. Pp. 71–76.

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

Fechner, G. T. (1861). Über einige Verhältnisse des binocularen Sehens. *Abhandlungen der mathematisch-physischen Classe der königlich sächsischen Gesellschaft der Wissenschaften*, **5**, 337–564.

Grossberg, S., & Kelly, F. (1999). Neural dynamics of binocular brightness perception. *Vision Research*, **39**, 3796–3816.

Levelt, W. J. M. (1965). Binocular brightness averaging and contour information. *British Journal of Psychology*, **56**, 1–13.

Luce, R. D. (1995). Four tensions concerning mathematical modeling in psychology. *Annual Reviews of Psychology*, **46**, 1–26.

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. (2008). Correction to Luce (2004). *Psychological Review*, **115**, 601.

Luce, R. D., & Steingrimsson, R. (2008). Note on a changed empirical inference in several Steingrimsson and Luce articles due to C. T. Ng's correction of an error in Luce (2004). *Journal of Mathematical Psychology*, **52**, 263–264.

Luce, R. D, Steingrimsson, R., & Narens, L. (2010). *Are psychophysical scales of intensities the same or different when Stimuli wary on other dimensions? Theory with experiments varying loudness and pitch*. Manuscript submitted for publication.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, **10**, 437–442.

Steingrimsson, R. (2009). Evaluating a model of global psychophysical judgments in brightness: I. Behavioral properties of summations and productions. *Attention, Perception & Psychophysics*, **71**, 1916–1930.

Steingrimsson, R., & Luce, R. D. (2005a). Evaluating 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). Evaluating 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. (2006). Empirical evaluation of a model of global psychophysical judgments: III. A form for the psychophysical function and intensity filtering. *Journal of Mathematical Psychology*, **50**, 15–29.

Steingrimsson, R., & Luce, R. D. (2007). Empirical evaluation of a model of global psychophysical judgments: IV. Forms for the weighting function. *Journal of Mathematical Psychology*, **51**, 29–44.

Steingrimsson, R., & Luce, R. D. (2010). *Evaluation of time-order error predictions from a model of global psychophysics*. In review.

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

Teghtsoonian, R. (1971). On the exponents in Stevens' law and the constant in Ekman's law. *Psychological Review*, **78**, 71–80

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

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

Zimmer, K., Luce, R. D., & Ellermeier, W. (2001). Testing a new theory of psychophysical scaling: Temporal loudness integration. In G. Sommerfeld, R. Kompass, & T. Lachmann (Eds.), *Fechner Day 2001: Proceedings of the 17th Annual Meeting of the International Society for Psychophysics*. Lengerirch, Germany: Pabst.

Author Note

This research was supported in part by National Science Foundation grants SBR-9808057 & BCS-0720288 (R. D. Luce, PI) to the University of California, Irvine—any opinion, finding, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation. Additional financial support was provided by the School of Social Sciences and the Department of Cognitive Sciences at UC Irvine. I am especially grateful to Marisa Carrasco for unfettered access to her laboratory and equipment at New York University to carry out much of the research reported here. And I appreciate many helpful comments of Joetta Gobell and R. Duncan Luce on earlier versions.

Footnotes

[1]Global psychophysics deal with stimuli that can be discriminated correctly with near certainty.

[2]The term “interval” is being used figuratively to refer to the difference in brightness that respondents experience between two intensity pairs.

[3]Distributivity is a generalization of a the segregation property. The generalization is that it extends segregation to the case where $y > 0$, i.e., segregation is the special case of distributivity when $y = 0$ (Luce, 2004).

[4]This I judged acceptable because knowledge of the experimental design does not change the sensations on which the behavioral tasks of matching and magnitude/ratio productions are based, and the variability trial forms, conditions, and block structure do not readily suggest a biased behavior in favor of a particular result. The author is numbered R8.

[5]Particular thanks go to M. Rudd who, as a reviewer on Steingrimsson (2009), was responsible for this improvement in procedure.

[6]Some researchers use linearized luminance steps. Since I seek subjective judgments from respondents and it is not a linear function of luminance, physical liniarization does not clearly provide an advantage over finer adjustments. Subjective linearization is problematic due to individual differences.

[7]I thank J. Yellott for this simple and elegant observation.

[8]Statistical evaluations in which support for a null hypothesis is sought has a long and contentious history in psychology. Indeed the vast majority of psychological research rests on statistical inferences based on rejection of the null hypothesis. This practice is in contrast to the physical sciences where modeling centers generally on uncovering and verifying invariance relations. Steingrimsson (2009) discusses this issue at some length (Appendix A in particular) and that discussion is not repeated here. The statistical

approach used draws from the common practice in physics of establishing a criterion that empirical result must meet in order to be accepted as supportive of a given invariance. I and others seeking psychological invariances welcome a rigorous attention to both the statistical and philosophical issues as they manifest themselves in the psychological sciences.

[9]Steingrimsson and Luce (2005b) showed that in experiments requiring multiple estimation steps it was important that all estimates be collected within the same session (to avoid inter-session variability) and that individual estimates from one step be used as input into the following step (as opposed to using medians or averages of estimate) in order to preserve accumulation of variance in multi-step estimations.

[10]The randomization was constrained by, e.g., needing an estimate for z prior to estimating a t.

	Luminance Level					
	t		t'			
Respondent	M	SD	M	SD	n	$p_{t\sim t'}$
R3	61.21	7.10	61.53	8.12	30	1.000
R4	29.59	2.42	30.27	2.48	30	.285
R5	54.23	6.96	53.31	11.60	30	.414
R6	59.83	8.98	59.83	10.98	35	.925
R7	42.71	4.93	42.03	6.45	30	.620
R8	25.83	1.84	25.80	2.12	30	.847
R9	44.10	13.00	46.74	9.31	30	.241

Table 1

Results of Experiment 1: Bisymmetry. Listed by respondent are condition, his/her averaged estimates and their standard deviations, the number of observations, the results of the statistical testing.

	Luminance Level						
	t		t'				
Respondent	M	SD	M	SD	p	n	$p_{t\sim t'}$
R3	46.20	9.64	48.12	9.38	2	30	.155
	21.39	1.97	21.75	2.45	$2/3$	30	.941
R4	40.55	4.07	41.83	3.12	2	32	.075
	12.48	2.95	13.08	1.75	$2/3$	24	.279
R6†	63.23	10.01	64.15	9.89	2	30	.574
	36.41	3.51	35.57	3.34	$2/3$	30	.192
R8	33.4	9.39	32.36	8.08	2	30	.670
R9	57.96	10.34	55.90	11.47	2	30	.329
	28.44	4.86	28.16	5.18	$2/3$	30	.722
R15	78.93	9.03	78.06	5.90	2	30	.283
	16.14	2.91	15.66	2.78	$2/3$	30	.988

†R6 used the same stimulus value for $p = 2$ as for $p = 2/3$.

Table 2

Results of Experiment 6: Simple joint-presentation decomposition. Listed by respondent are conditions, his/her averaged estimates and their standard deviations, proportion instruction, the number of observations, the results of the statistical testing.

		Luminance Level							
	Distributivity		t		t'				
Respondent	Version	x	M	SD	M	SD	p	n	$p_{t\sim t'}$
R4	Left	25.87	38.26	1.64	38.04	1.44	2	30	.221
*		35.22	20.31	1.01	20.72	1.40	2/3	30	.536
*	Right	21.93	38.36	7.78	37.87	5.55	2	24	.090
*		65.34	29.05	6.14	32.61	2.17	2/3	24	$<$.001
R8	Left	25.87	37.45	4.35	37.22	4.72	2	30	.711
*		35.22	20.53	1.03	20.64	2.11	2/3	30	.112
*	Right	21.93	26.67	2.47	27.09	3.27	2	30	.894
*		35.22	21.57	1.26	21.34	1.78	2/3	30	.929
R10	Left	25.87	37.79	4.45	37.88	2.61	2	30	.976
		25.87	20.16	1.22	20.52	1.01	2/3	30	.225
R12	Left	38.51	41.70	3.53	43.49	3.41	2	30	.080
		38.51	22.17	1.20	22.70	1.07	2/3	30	.130
R13	Left	38.51	39.82	3.39	38.20	3.56	2	30	.131
		38.51	22.29	0.89	22.05	2.89	2/3	30	.422
R35	Left	21.93	37.78	4.51	39.09	2.65	2	30	.134

*Data collected at UCI.

Table 3

Results of Experiment 3: Distributivity. Listed by respondent are conditions, the form of Distributivity tested, his/her averaged estimates and their standard deviations, proportion instruction, the number of observations, the results of the statistical testing.

Property (Exp. #)	#Respondents	#Tests	#Fail
Bisymmetry (Exp. 1)	7	7	0
JP decomposition (Exp. 2)	6	12	0
Distributivity (Exp. 3)	6	15	2

Table 4

Summary of experimental results

Figure Captions

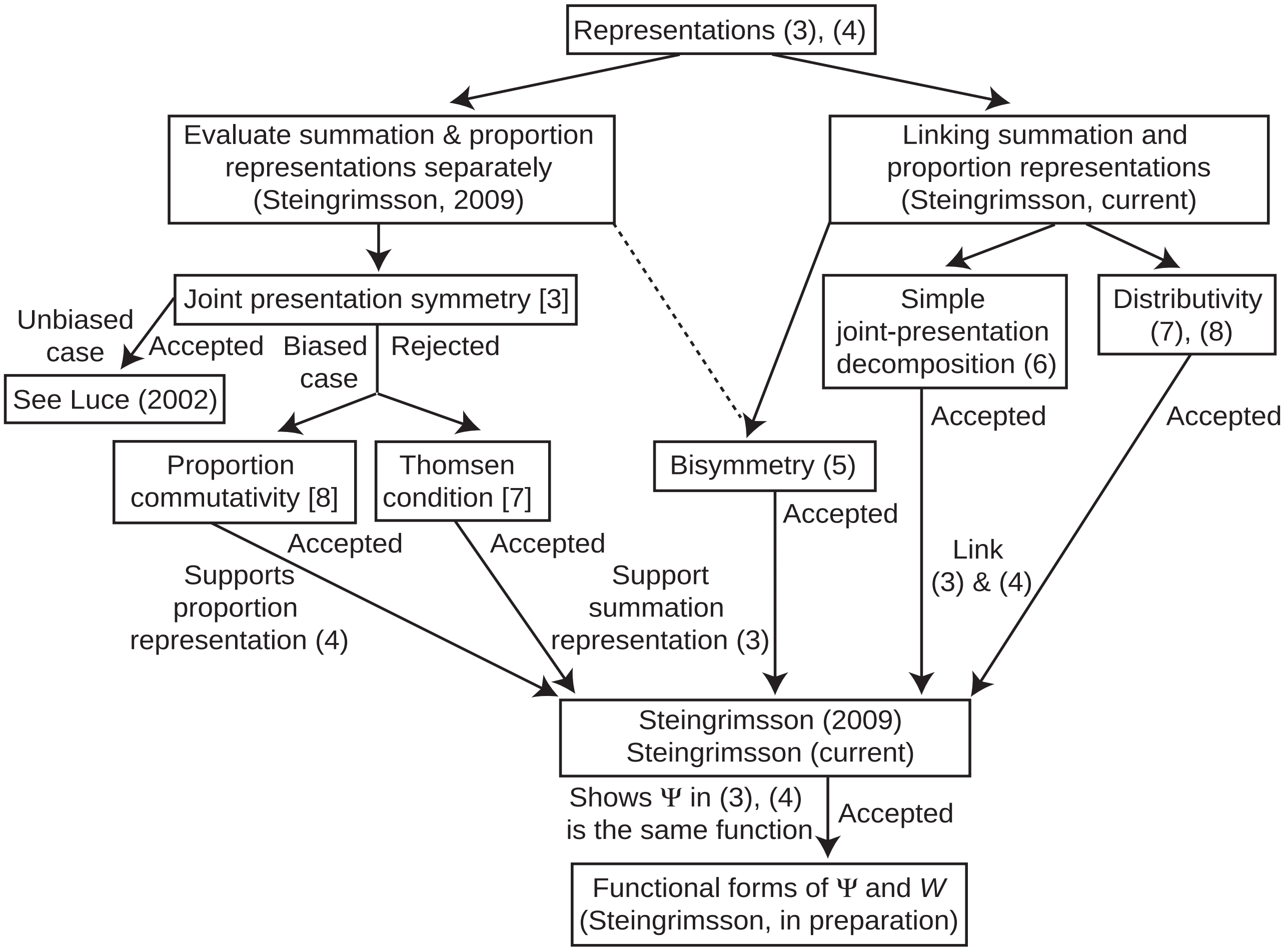
Figure 1. Stimuli displayed on a monitor (A) viewed through a stereoscope (B), produce the subjective percept seen by the respondents (C). The x, u, and z values are luminance.(Figure reprinted with permission from Steingrimsson, 2009).

Figure 2. Stimuli displayed on a monitor (A), are viewed through a stereoscope (B), produce the subjective percept seen by the respondents (C). The x, u, z, and t' values are luminance. The stimuli in C creates two stimulus "intervals", on the left, one from $z \oplus u$ to $z \oplus x$ and on the right side $z \oplus u$ to $t' \oplus t'$. Respondents adjust the luminance of t' until they are satisfied that the brightness interval between $z \oplus u$ and $z \oplus x$ is perceived as p time the brightness interval between $z \oplus u$ to $t' \oplus t'$. (Format and Panel B reprinted with permission from Steingrimsson, 2009).

Figure 3. The diagram shows, on the left, the properties tested by Steingrimsson (2009), on the right, those tested in this article, in the lower middle, their results, their implication and how they all come together to establish the representations (3) and (4). At the bottom, the topic of the third article (Steingrimsson, in preparation), the functional form of Ψ and W emerges as the next article in the series.


A
z
z
x
u
B
C
z⊕z
x⊕u


A
z
t’
x
t’
z
z
u
u
B
C
z⊕x
t’⊕t’
z⊕u
z⊕u


Representations (3), (4)
Evaluate summation & proportion representations separately (Steingrimsson, 2009)
Linking summation and proportion representations (Steingrimsson, current)
Joint presentation symmetry [3]
Unbiased case
Accepted
Biased case
Rejected
See Luce (2002)
Simple joint-presentation decomposition (6)
Distributivity (7), (8)
Accepted
Accepted
Proportion commutativity [8]
Thomsen condition [7]
Bisymmetry (5)
Accepted
Accepted
Accepted
Supports proportion representation (4)
Support summation representation (3)
Link (3) & (4)
Steingrimsson (2009)
Steingrimsson (current)
Shows Ψ in (3), (4) is the same function
Accepted
Functional forms of Ψ and W (Steingrimsson, in preparation)