

The judgment operation of ratio production

SERGIO CESARE MASIN

Università di Padova, Padua, Italy

The judgment operation that individuals perform when they set a sensory intensity in a predefined ratio to a standard sensory intensity is unknown. The present study was conducted to determine this operation by functional measurement. Participants first produced a sensory intensity for each combination of various predefined ratios and standard sensory intensities. Subsequently, they rated these produced intensities. Factorial graphs derived from these ratings were essentially parallel. This result shows that participants produced each sensory intensity by operating as if they were adding an amount of sensory intensity to the respective standard sensory intensity. It was found that this added amount varied nonlinearly with the predefined ratio, implying that ratio productions are nonlinear.

For each number n taken from a predefined set of numbers, the method of ratio production requires participants to produce a sensory intensity, Ψ_p , equal to n times a fixed standard sensory intensity, Ψ_s (Marks, 1974, p. 14; Stevens, 1975, p. 20). The measure of Ψ_s is an arbitrarily specified number M_s (the modulus), and the measure of each Ψ_p is

$$M_p = n \cdot M_s. \quad (1)$$

Garner (1954a) suggested that when participants produce Ψ_p , they could be behaving as if they were using some unknown number v in place of n such that

$$\Psi_p = v \cdot \Psi_s. \quad (2)$$

The following six different tests indicate that $n \neq v$.

Test 1. Garner (1954a) and Engen and Tulunay (1957) had different groups of participants judge on each trial whether the second of two successive sensory intensities was less than or more than half as large as the first sensory intensity. For each group, the first of these intensities was produced by the same standard stimulus whereas the second was produced by a variable stimulus. Each group had a different range of variable stimuli. These ranges were not overlapping. The mean variable stimulus that was found to have the sensory intensity that was judged to be half as large as the standard sensory intensity ($n = 1/2$) was, for each group, in the middle of the respective range of variable stimuli. Thus, for at least two of the groups, the stimulus that produced the sensory intensity that was half as large as the standard sensory intensity ($v = 1/2$) must have been outside the range of variable stimuli. Consequently, $v \neq 1/2$ for at least two of the groups.

Test 2. Consider the sensory ratios $v_1 = \Psi_a/\Psi_b$ and $v_2 = \Psi_b/\Psi_a$ and their numerical estimates n_1 and n_2 , respectively. Since $v_1 \cdot v_2 = 1$, $n_1 \cdot n_2 = 1$ when $n_i = v_i$ (where $i = 1$ or 2). Svenson and Åkesson (1966a, 1966b) found that $n_1 \cdot n_2 \neq 1$ and, thus, that $n_i \neq v_i$.

Test 3. Consider the sensory ratios $v_1 = \Psi_a/\Psi_b$, $v_2 = \Psi_b/\Psi_c$, and $v_3 = \Psi_a/\Psi_c$ and their numerical estimates n_1 , n_2 , and n_3 , respectively. Since $v_1 \cdot v_2 = v_3$, the equality $n_1 \cdot n_2 = n_3$ occurs when $n_i = v_i$ (where $i = 1, 2$, or 3). Fagot and Stewart (1969) found that $n_1 \cdot n_2 \neq n_3$. Consequently, $n_i \neq v_i$.

Test 4. Masin (1983) had participants produce the sensory intensity $2\Psi_s$ and subsequently the intensity Ψ_b that bisected the interval between Ψ_s and $2\Psi_s$. If $v = 2$, then $\Psi_b = 1.5\Psi_s$. It turned out that $\Psi_b > 1.5\Psi_s$ and, thus, that $v \neq 2$.

Test 5. Masin and Tommasi (2000) found that participants, without noticing any contradiction, easily fulfilled the task of producing the sensory intensity $2\Psi_s$ when Ψ_s was so close to one end of the sensory continuum that $2\Psi_s$ was outside the continuum. In this case, the production of $2\Psi_s$ necessarily implies that $v \neq 2$.

Test 6. Consider the sensory ratios $v_1 = \Psi_1/\Psi_s$, $v_2 = \Psi_2/\Psi_1$, and $v_3 = \Psi_3/\Psi_s$ and their respective numerical estimates n_1 , n_2 , and n_3 . The equality $v_1 \cdot v_2 = v_3$ occurs when $\Psi_2 = \Psi_3$. Thus, when $n_1 \cdot n_2 = n_3$, the equality $\Psi_2 = \Psi_3$ occurs when $n_i = v_i$. The condition that $\Psi_2 = \Psi_3$ when $n_1 \cdot n_2 = n_3$ is called the *multiplicativity axiom* (Narens, 1996). Ellermeier and Faulhammer (2000) and Zimmer (2005) found that this axiom was violated. For example, the production of $6\Psi_s$ differed from the production of 2 times the production of $3\Psi_s$. These results further demonstrate that $n_i \neq v_i$.

The repeated finding that $n \neq v$ is neutral regarding the possibility of participants' ability to estimate sensory ratios. It means only that if participants are able to estimate sensory ratios, they do so imperfectly. However, the following discussion shows that participants may not be able to estimate sensory ratios.

The Problem of the Judgment Operation

Given the sensory intensities $\Psi_1 < \Psi_2 < \Psi_3$ with Ψ_1 and Ψ_3 fixed, Garner (1954b) had participants set Ψ_2 so that $\Psi_2/\Psi_1 = \Psi_3/\Psi_2$ on one occasion and $\Psi_2 - \Psi_1 = \Psi_3 - \Psi_2$

on another occasion. He found that the produced Ψ_2 was the same on these two occasions. Garner (1954b, pp. 79–80) concluded that participants judged a single quantitative relation whether they judged the difference or the ratio between two sensory intensities. This conclusion was endorsed by Torgenson (1961) and confirmed experimentally by many investigators using essentially the following test: Participants were asked to estimate differences or ratios in pairs of sensory intensities. If estimated differences differ from estimated ratios, the rank order of sensory intensities obtained from difference estimates must differ from that obtained from ratio estimates. These orders were repeatedly found to be equal, implying that participants judged the same quantitative relation whether they judged sensory differences or sensory ratios (Birbaum & Elmasian, 1977; Hardin & Birbaum, 1990; Mellers, Davis, & Birbaum, 1984; Parker & Schneider, 1974; Schneider, 1982; Schneider, Parker, Farrell, & Kanow, 1976).¹ Recently, Ellermeier, Narens, and Dielmann (2003) confirmed these results by another test. In their study, participants produced a sensory intensity $b \cdot \Psi_s$ and subsequently added an amount c of sensory intensity to this produced intensity, or they added amount c of sensory intensity to Ψ_s and subsequently produced sensory intensity $b \cdot (c + \Psi_s)$. They found that sensory intensity $c + b \cdot \Psi_s$ equaled sensory intensity $b \cdot (c + \Psi_s)$. This result shows that the participants judged the same quantitative relation both when they were asked to judge a sensory ratio and when they were asked to add an amount of sensory intensity.

Let participants produce the sensory intensities $b \cdot \Psi_s$ and $c \cdot (b \cdot \Psi_s)$ and subsequently produce the sensory intensities $c \cdot \Psi_s$ and $b \cdot (c \cdot \Psi_s)$. The condition $b \cdot (c \cdot \Psi_s) = c \cdot (b \cdot \Psi_s)$ is termed the *commutativity axiom* (Narens, 1996). By having participants produce multiples or fractions of sensory intensities, Ellermeier and Faulhammer (2000) and Zimmer (2005) have found that the commutativity axiom was satisfied—for example, the production of two thirds of the production of one fourth of Ψ_s matched the production of one fourth of the production of two thirds of Ψ_s . Since $n \neq v$, at first one could interpret this result as showing that participants were able to produce ratios of sensory intensities, the only drawback being that these ratios were imperfect. However, one should consider that the commutativity axiom is satisfied both in the event that sensory intensities are produced on the basis of sensory ratios and in the event that sensory intensities are produced by adding some amount of sensory intensity to Ψ_s (Narens, 1996). That is, when participants are asked to produce the sensory intensities $b \cdot \Psi_s$ and $c \cdot (b \cdot \Psi_s)$, they might instead produce the sensory intensities $b + \Psi_s$ and $c + b + \Psi_s$; when participants are asked to produce the sensory intensities $c \cdot \Psi_s$ and $b \cdot (c \cdot \Psi_s)$, they might instead produce the sensory intensities $c + \Psi_s$ and $b + c + \Psi_s$. Thus, the finding that the commutativity axiom is satisfied could equally mean that $b \cdot (c \cdot \Psi_s) = c \cdot (b \cdot \Psi_s)$ or that $b + c + \Psi_s = c + b + \Psi_s$. That is, the test of the commutativity axiom cannot distinguish Equation 2 from the equation

$$\Psi_p = \delta + \Psi_s, \quad (3)$$

where δ is some amount of sensory intensity.

We are in the position of not being able to decide which operation participants perform in ratio production. Is the operation multiplicative (Equation 2) or additive (Equation 3)? Functional measurement can provide an answer to this question (Anderson, 1981, 1982; McBride & Anderson, 1991). Functional measurement is obtained through the joint use of a factorial experimental design and some method of self-estimation involving a linear response function. There is a large amount of evidence showing that the rating method provides such a response function—that is, that ratings of sensory intensity are related linearly to corresponding sensory intensities (Anderson, 1996, pp. 94–96; Curtis, 1970; Curtis & Fox, 1969; Masin, 2005; Weiss, 1972). The linearity of ratings may be approximate due to context effects. However, it is now recognized (Marks, 1996) that context effects are greatly reduced or practically eliminated when many response alternatives are available to the participants (as in graphic rating or when integer numerals in the range of 0–100 are employed—Parducci, 1982; Parducci & Wedell, 1986) and two anchor stimuli, one somewhat smaller than the smallest experimental stimulus and one somewhat larger than the largest experimental stimulus, are used to delimit some portion of the sensory continuum (Anderson, 1982). For each fixed value of n , Equation 2 and the linearity of the rating scale imply that the curves relating the rating of Ψ_p to Ψ_s are straight lines that diverge as Ψ_s increases, whereas Equation 3 and the linearity of the rating scale imply that these straight lines are parallel.

These two possibilities were tested in Experiments 1 and 2. The linearity of the rating scale was tested in Experiment 3.

EXPERIMENTS 1 AND 2

In Experiment 1, for ns from 1 to 4, productions of luminous intensity were obtained by the method of adjustment. Only ascending trials were used, to avoid the possibility that starting stimulus values set arbitrarily higher than the highest value of the standard stimulus would constrain participants not to select stimulus values higher than the starting stimulus values, as could be the case during descending trials. The results of Experiment 1 showed that this possibility was reasonably excluded if ns from 1 to 2.5 were used. In Experiment 2, which was a repetition of Experiment 1 using only these ns , ratio productions were obtained by the method of adjustment with both ascending and descending trials.

Method

Participants

Fifteen university students (3 males and 12 females, mean age = 21.4 years, $SD = 1.7$) participated in Experiment 1 in exchange for payment. A different group of 17 unpaid university students (6 males and 11 females, mean age = 24.7 years, $SD = 2.7$) participated in Experiment 2.

Stimuli for Experiment 1

In a dark room, with viewing distance held at 70 cm by a head- and chinrest, stimuli appeared in the middle of a 32×17 cm rectangular achromatic area of the frontal parallel screen of a display (Apple

Multiple Scan 1705) controlled by a computer (Power Macintosh 7200/90). The remaining area of the screen was covered with black cardboard. The luminance of the screen was 0.2 cd/m^2 .

Throughout the experiment, one horizontal $0.5 \times 186 \text{ mm}$ line and two $13 \times 13 \text{ mm}$ anchor squares were displayed constantly in the middle of the visible part of the screen. The reason for this constant presentation is explained at the end of this section. The center of the right vertical side of one anchor square and the center of the left vertical side of the other anchor square were positioned 1 cm from the left and right ends of the line, respectively. The luminance of the line was 13 cd/m^2 , and the luminances of the left and right anchor squares were 0.5 and 110 cd/m^2 , respectively.

The stimuli were two achromatic $13 \times 13 \text{ mm}$ squares. The centers of these squares and the center of the horizontal line were vertically aligned. The gap between each of these squares and the horizontal line was 35 mm.

These were two parts of Experiment 1, as described below.

First part. In the first part of the experiment, the participants were asked to ignore the anchor stimuli and were given no information about these stimuli. Using the adjustment method, each participant varied the luminance of the lower square by pressing one of two juxtaposed keys aligned horizontally. The starting luminance of this square was always the luminance of the screen. The luminance of the lower square decreased when the left key was held down and increased when the right key was held down. This luminance varied in discrete steps, the size of which progressively increased from a minimum of 0.1 cd/m^2 to a maximum of 2 cd/m^2 . The relation between the number of these steps (counted from the minimum luminance) and the luminance of the lower square was, essentially, a power function with an exponent of 2. The number of steps needed to pass from the minimum to the maximum luminance was 86.

The participants selected a luminance for the lower square such that the predefined ratio n of the luminous intensity of the lower square to the luminous intensity of the upper square was 1, 1.25, 1.5, 1.75, 2, 2.5, 3, or 4. Three complete $8(n) \times 4$ (standard luminance) factorial designs were prepared, which shall be called Designs 1, 2, and 3. The standard luminances for Design 1 were 1.5, 2.5, 3.5, and 4.5 cd/m^2 ; for Design 2, they were 1.5, 3.5, 5.5, and 7.5 cd/m^2 ; and for Design 3, they were 1.5, 4.5, 7.5, and 10.5 cd/m^2 . The designs resulting from combining n s with standard luminances involved 96 combinations of n and standard luminance. Of these combinations, 56 differed from each other and 40 matched one or more of these 56 combinations. Only these 56 different combinations were presented to each participant, in random order and without replacement, two times consecutively.

For each stimulus, the numeral representing n was displayed as an achromatic 12-point character with a luminance of 13 cd/m^2 at the right bottom corner of the visible part of the screen. It was read aloud by the experimenter.

Second part. In the second part of Experiment 1, the upper and lower squares were presented singly. The luminances of the upper square were those used in the first part of the experiment. For each participant, the luminances of the lower square were the means of the two luminances that the participant had produced in the first part of the experiment for each of the 56 different combinations of n and standard luminance. The criterion used was that these mean luminances were not presented if one of the luminances making up the mean was greater than or equal to 110 cd/m^2 . Only 1 participant met this criterion, and he did so four times. The luminances of lower and upper squares were presented to each participant in random order, without replacement, two times consecutively.

The participants rated the luminances of the upper or lower squares by positioning a cursor on the horizontal line. This cursor was a $0.3 \times 2.5 \text{ mm}$, 13-cd/m^2 vertical achromatic line with its center on the horizontal line. The participants moved the cursor left or right by sliding the computer mouse left or right, respectively. The position of the cursor varied in steps of 0.3 mm .

The anchor stimuli were presented throughout the experiment so that any possible configural or perceptual effect they may have had on produced luminous intensities was the same for intensities

produced in the first part of the experiment and those subsequently presented in the second part of the experiment (Diamond, 1962).

Stimuli for Experiment 2

Stimuli were presented as in Experiment 1. Experiment 2 was also divided into two parts, for the first of which the n s were 1, 1.5, 2, and 2.5 and the standard luminances of the upper square were 1.5, 3.3, 5.5, and 9.2 cd/m^2 . These luminances corresponded to equally spaced least squares estimates of rated luminous intensity computed from the data obtained in Experiment 1. The starting luminance of the lower square was either 0.2 or 110 cd/m^2 . Each of the 32 different combinations of n , standard luminance, and starting luminance were presented to each participant in random order, without replacement, two times consecutively.

For the second part of Experiment 2, the luminances of the upper square were those used in the first part of this experiment. For each participant, the luminances of the lower square were the means of the four luminances produced by the participant in the first part of Experiment 2 for each combination of n and standard luminance. The luminances of the lower and upper squares were presented to each participant in random order, without replacement, four times consecutively.

Procedure for Experiment 1

In the first part of the experiment, the method of adjustment was used. The participants were asked to set the perceived luminous intensity of the lower square such that the ratio of this intensity to the luminous intensity of the upper square equaled the n read aloud by the experimenter. All the participants confirmed that they understood that the instructions implied that an n of 1 meant that the luminous intensity of the lower square had to be set equal to the luminous intensity of the upper square. The participants were allowed to arrive at their settings by bracketing.

In the second part of the experiment, the method of graphic rating was used. The participants were asked to set the position of the cursor on the horizontal line such that this position represented the luminous intensity of the displayed upper or lower square. The left and right ends of the horizontal line were defined to represent the luminous intensities of the left and right anchors, respectively. To minimize the ceiling effect due to the vicinity of individual mean productions of the luminous intensity of the lower square to that of the right anchor square, the criterion was adopted that mean produced luminances of the lower square above 84 cd/m^2 were not presented for rating. Four participants met this criterion only for the luminance of the lower square that in the first part of the experiment had been set at four times the maximum luminance of the upper square.

Immediately before each part of the experiment was initiated, 10 stimuli selected at random were presented to familiarize the participant with the respective procedure.

Procedure for Experiment 2

The procedures for the first and second parts of Experiment 2 were the same as the respective procedures of Experiment 1 except that the participants had no preexperimental familiarization with the stimuli.

Results

Ratings of the luminous intensities produced for each combination of n and a given standard luminance of the upper square were linearly transformed to numbers in the range of 0–100. For each n , mean transformed ratings were plotted as a function of standard luminance.

The results of Experiment 1 are shown in Figure 1. Factorial graphs slightly converge or are approximately parallel. These results allow us to reject Equation 2, which predicts divergence of factorial graphs. A $8(n) \times 4$ (standard luminance) ANOVA with two replicates carried out

for each of Designs 1–3 showed that the interaction was significant for Design 3 [$F(21,294) = 1.9, p < .05$] and not significant for Designs 1 and 2 [$F_s(21,294) = 1.4$ and 1.3 , respectively], indicating, respectively, a slight convergence and a slight parallelism of factorial graphs.

The results of Experiment 2 are shown in Figure 2. These results show that factorial graphs were essentially parallel also when descending trials were used. The mean ratings of produced luminous intensity of the lower square calculated for each standard luminance across values of n are linear estimates, called *functional estimates*, of Ψ_s (Equation 3; see Anderson, 1982, p. 58). In conformity with the prediction of Equation 3, the slight curvature of the factorial graphs in Figure 2 disappears when mean ratings are plotted against the functional estimates of Ψ_s . These results further contradict Equation 2, which predicts that the factorial graphs should diverge. A $4(n) \times 4$ (standard luminance) ANOVA with four replicates confirmed that the interaction was not significant [$F(9,144) = 0.98$]. The linear-linear component of the interaction was significant [$F(1,16) = 6.9, p < .05$], whereas the residual components were not significant [$F_s(1,16) = 0.06$ to 1.5], indicating that factorial graphs were essentially parallel with a barely noticeable convergence.

The mean ratings of produced luminous intensity of the lower square calculated for each n across values of standard luminance are functional estimates of δ (Equation 3; see Anderson, 1982, p. 58). For each of Designs 1–3, Figure 3 shows these estimates as a function of n . The data are fitted by least squares logarithmic functions, showing that n is related nonlinearly to δ . Curves move increasingly upward in passing from Design 1 to Design 3 due to the obvious fact that the functional estimate of δ increases with the maximum standard luminance in passing from Design 1 to Design 3. For each of Designs 1–3, a $4(n) \times 4$ (standard luminance) ANOVA carried out for the equally

spaced predefined ratios of 1, 2, 3, and 4 confirmed that the quadratic trend of n was significant [$F_s(1,14) = 38, 41, \text{ and } 32$ for Designs 1, 2, and 3, respectively; $p < .0005$]. In Experiment 2, the quadratic trend of n was also significant [$F(1,16) = 19, p < .0005$].

Discussion

We have seen that Equation 3 and the linearity of the rating scale imply that the factorial graphs one obtains must be parallel. Conversely, if the rating scale is linear and the obtained factorial graphs are parallel, the judgment operation of ratio production must by necessity be additive, as described by Equation 3. Thus, for $n > 1$, the present finding that factorial graphs were essentially parallel confirms Equation 3 as a model for the judgment operation of ratio production.

This conclusion is based on previous empirical findings that the rating scale is linear. However, if the rating R_p of Ψ_p should instead be related linearly to the logarithm of Ψ_p (i.e., if $R_p = c_0 \log \Psi_p + c_1$, with c_0 and c_1 constant), then Equation 2 would imply that

$$R_p = c_0 (\log v + \log \Psi_s) + c_1. \quad (4)$$

Equation 4 shows that if the rating scale is logarithmic and factorial graphs are parallel, the operation of ratio production must be multiplicative, as described by Equation 2, rather than additive, as described by Equation 3. It is thus important to test whether or not the rating scale is logarithmic. In Experiment 3, this test was performed by asking each participant to rate the total amount ε of light emitted by a surface toward the participant. Phenomenally, achromatic surfaces have a *whiteness*, and this whiteness has an *insistency* (Koffka, 1935, p. 113). The insistency of the light of a surface should be proportional to the total amount of light emitted by the surface. This total amount

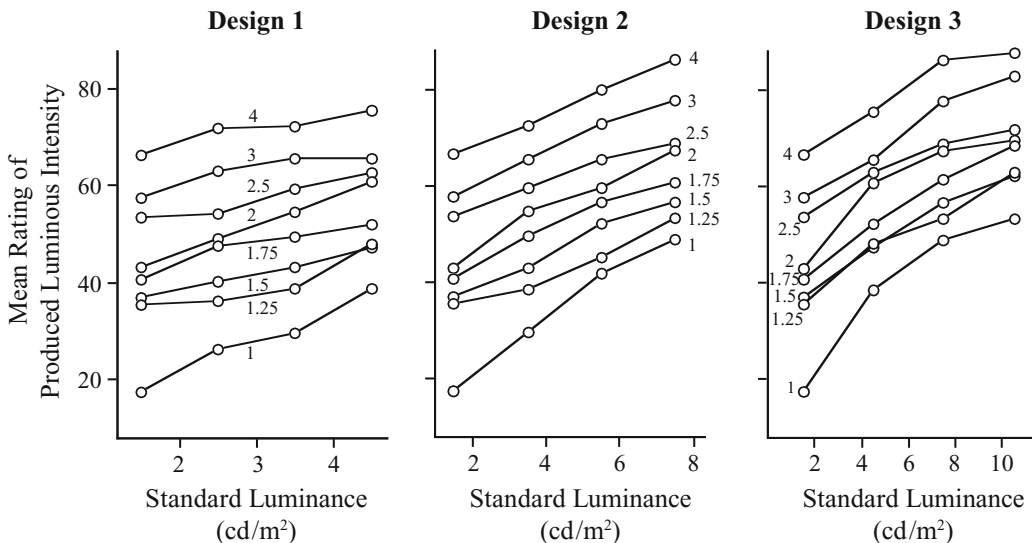


Figure 1. Results of Experiment 1. Participants produced a luminous intensity such that the ratio of this intensity to that of a standard stimulus equaled a predefined ratio. The panels show the mean ratings of this produced luminous intensity for Designs 1, 2, and 3 separately. Each line corresponds to a different predefined ratio.

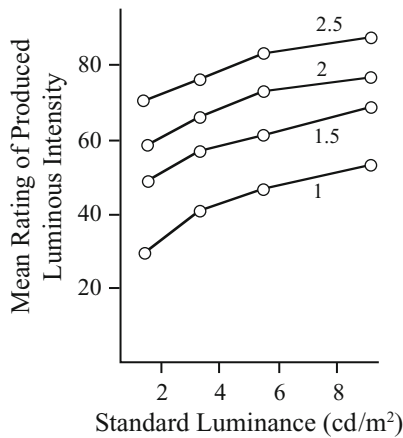


Figure 2. Mean rating of produced luminous intensity as a function of the luminance of the standard stimulus in Experiment 2. Each line corresponds to a different predefined ratio.

is the product of perceived area α of the surface by perceived luminous intensity λ of the same surface. That is,

$$\varepsilon = \alpha \cdot \lambda. \quad (5)$$

For each α , Equation 5 predicts that the curves relating the rating of ε to λ are straight lines that diverge as λ increases if the rating of ε is linearly related to ε , and are parallel logarithmic curves if the rating of ε is linearly related to the logarithm of ε .

EXPERIMENT 3

Method

Participants. Fifteen unpaid university students (1 male and 14 females, mean age = 23.9 years, $SD = 3.3$) participated in Experiment 3.

Stimuli. The viewing conditions were the same as those in Experiments 1 and 2. The same horizontal line and cursor used in Experiments 1 and 2, displaced downward 30 mm to make room for the stimuli, and two uniform achromatic disks used as anchors were constantly presented on the screen. These disks were horizontally aligned, one on the left and one on the right of the horizontal line, with 1-cm gaps between disks and line. These disks had the same area (1,925 mm²) but different luminances (0.3 and 90 cd/m², respectively). Stimuli were one of 16 uniform achromatic disks, each with a different combination of area and luminance. These disks had areas of 500, 1,000, 1,500, or 2,000 mm² and luminances of 5, 10, 20, or 40 cd/m². The centers of the stimulus and of the horizontal line were 65 cm from each other, vertically aligned.

Procedure. The participants were asked to judge the total amount of light emitted toward the participant by all points of the stimulus. To exemplify the meaning of “total amount of emitted light,” a stimulus with constant luminance and progressively increasing area was shown to the participants. They agreed that the total amount of light emitted by the disk increased with the area of the disk. The participants were then shown a stimulus with constant area and progressively increasing luminance. They agreed that the total amount of light emitted by the disk increased with the luminance of the disk. Using the rating method, the participants were asked to set the position of the cursor on the horizontal line in such a way that this position represented the total amount of light emitted by all points of the stimulus. The left and right ends of the horizontal line were defined to represent the total amount of light emitted by all points of the left and right anchors, respectively.

Results

Each rating of total amount of emitted light was transformed linearly to a number in the range of 0–100. For each stimulus area, Figure 4 shows mean transformed ratings of total amount of emitted light as a function of stimulus luminance. Factorial graphs diverge as luminance increases, showing that the rating scale is not logarithmic. The curvature of these graphs disappears completely when mean ratings are plotted against the functional estimate of λ , in agreement with the prediction of Equation 5 for the case in which the rating of ε is linearly related to ε . These results confirm the aforesaid findings of the previous literature that the rating scale is linear.

A 4 (area) \times 4 (luminance) ANOVA with four replicates confirmed the divergence of factorial graphs. The interaction [$F(9,126) = 4.2, p < .0005$] was significant. The linear–linear [$F(1,14) = 13, p < .005$] and quadratic–linear [$F(1,14) = 9.1, p < .01$] components of the interaction were significant, whereas the other components were not significant [$F_s(1,16) = 0.13\text{--}0.84$].

GENERAL DISCUSSION

The task of ratio production required the participants to set a luminous intensity such that the ratio between this intensity and a standard luminous intensity equaled a predefined ratio. The present results show that, to fulfill this task, the participants operated as if they were adding a luminous intensity to the standard luminous intensity.

For $n > 1$, the results presented in Figure 3 indicate empirically that

$$\delta = p \cdot \log n + q, \quad (6)$$

with $p \neq 0$ and $q > 0$ constant.

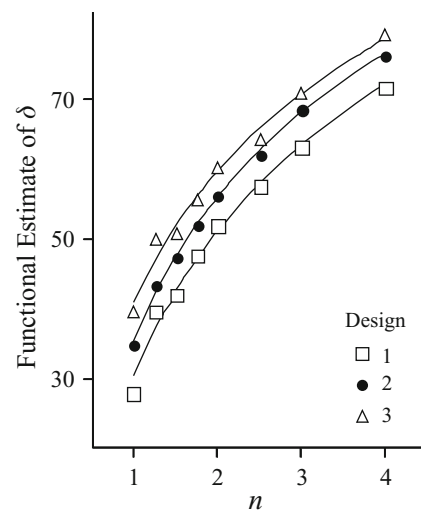


Figure 3. Functional estimates of δ as a function of the predefined ratio n for each of Designs 1–3 of Experiment 1. For each predefined ratio, participants produced luminous intensities by performing an operation equivalent to adding an amount δ of luminous intensity to the respective standard luminous intensity, rather than by estimating the predefined ratios.

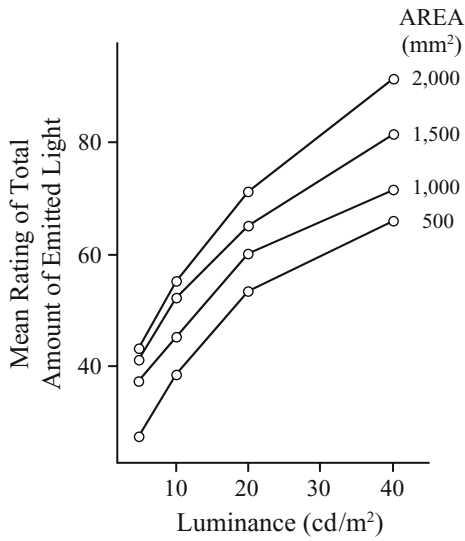


Figure 4. Experiment 3: Mean rating of total amount of light emitted by disks of different areas toward the eyes of the participant as a function of disk luminance for each disk area.

Equation 6 implies that Equation 3 satisfies the commutativity axiom but violates the multiplicativity axiom, in agreement with Ellermeier and Faulhammer's (2000) results. Let us consider first the multiplicativity axiom. With a , b , and c being numbers such that $a = b \cdot c$, when participants produced the sensory intensities $c \cdot \Psi_s$, $b \cdot (c \cdot \Psi_s)$, and $a \cdot \Psi_s$, Ellermeier and Faulhammer found empirically that $a \cdot \Psi_s < b \cdot (c \cdot \Psi_s)$. Equations 3 and 6 predict this result. If participants produce the sensory intensities $p \cdot \log c + q + \Psi_s$ (instead of $c \cdot \Psi_s$), $p \cdot \log b + q + (p \cdot \log c + q + \Psi_s)$ [instead of $b \cdot (c \cdot \Psi_s)$], and $p \cdot \log a + q + \Psi_s$ (instead of $a \cdot \Psi_s$), it may be easily shown that $p \cdot \log a + q + \Psi_s < p \cdot \log b + q + (p \cdot \log c + q + \Psi_s)$. Let us now consider the commutativity axiom. When participants produce the sensory intensities $b \cdot \Psi_s$ and $c \cdot (b \cdot \Psi_s)$ and subsequently produce the sensory intensities $c \cdot \Psi_s$ and $b \cdot (c \cdot \Psi_s)$, Ellermeier and Faulhammer found empirically that $b \cdot (c \cdot \Psi_s) = c \cdot (b \cdot \Psi_s)$. In agreement with Equations 3 and 6, it may be easily shown that $p \cdot \log b + q + (p \cdot \log c + q + \Psi_s) = p \cdot \log c + q + (p \cdot \log b + q + \Psi_s)$.

Equation 6 implies that ratio productions are nonlinear. Indeed, ratio productions are linear only if δ and n are related linearly. To see this, let us assume that

$$n = a \cdot \delta + b, \quad (7)$$

with a and b constant and $a \neq 0$. Equations 1, 3, and 7 yield

$$M_p = M_s \cdot a \cdot (\Psi_p - \Psi_s) + M_s \cdot b. \quad (8)$$

Ratio productions are linear when M_p is related linearly to Ψ_p . Since Ψ_s and M_s are constant, Equation 8 implies that M_p is related linearly to Ψ_p only if a and b are constant. Hence, ratio productions are linear only if δ and n are related linearly.

The finding that ratio productions are nonlinear—that is, that M_p is a nonlinear function of Ψ_p —agrees with results of previous studies on magnitude estimation. Curtis, Attneave, and Harrington (1968) proposed that (1) sensory intensity, Ψ , is related to corresponding physical intensity, ϕ , by the psychophysical function $\Psi = k_1 \cdot \phi^\beta$, and that (2) the magnitude estimate J of Ψ is related to Ψ by the response function $J = k_2 \cdot \Psi^\gamma$, with k_1 and k_2 being proportionality constants. These functions imply that magnitude estimate J_{ij} of the difference of the sensory intensities corresponding to physical intensities ϕ_i and ϕ_j is $J_{ij} = k_3 (\phi_i^\beta - \phi_j^\beta)^\gamma$, with k_3 being a proportionality constant. In a factorial experiment, Curtis et al. had participants estimate magnitudes of differences of sensory intensities. By fitting their model to these estimates, they found that $\gamma > 1$. That is, J was related nonlinearly to Ψ . Weiss (1972) confirmed this finding. In a factorial experiment, he had participants estimate the magnitude of the average intensity of pairs of sensory intensities. Since magnitude estimates are nonlinear (Curtis et al., 1968), the additivity of this prescribed operation involves nonparallel factorial graphs for average sensory intensity. The factorial graphs obtained by Weiss were clearly nonparallel, confirming that magnitude estimates were nonlinear.

The conclusions reached in the present study depend on the possibility that the rating scale is linear. The results of Experiment 2 for luminous intensity show that factorial graphs were essentially parallel, in accordance with the possibility that the rating scale for luminous intensity is linear. The results of Experiment 3 further support this possibility. These results confirm the following previous results obtained specifically for luminous intensity. Using Curtis et al.'s (1968) procedure, Curtis (1970) had participants rate magnitudes of differences of luminous intensities. Curtis found that J was a linear function of Ψ , confirming that ratings were linear. Weiss (1972) had participants rate the average luminous intensity of pairs of Munsell chips. The additivity of this prescribed operation and the linearity of the rating scale implied that mean ratings of mean luminous intensities formed parallel factorial graphs. The factorial graphs obtained by Weiss were indeed parallel, confirming that ratings were linear.

Many authors have warned that the numerals (n) used in the magnitude estimation of sensory intensities differ from the numbers (ν) that represent the sensory intensities on a ratio scale (Attneave, 1962; Garner, 1954a; Garner, Hake, & Eriksen, 1956; Graham, 1958; Graham & Ratoosh, 1962; Luce, 2002; McGill, 1960; Narens, 1996; Shepard, 1981). Garner's (1954a) empirical finding that $n \neq \nu$ involves the possibility for magnitude estimation that $\nu = f(n)$, with $f(n)$ a monotonic function other than the identity function. Recently, there have been attempts to determine $f(n)$. As Attneave originally proposed, empirical results indicate that $f(n)$ could be approximately a power function (Steingrimsson & Luce, 2007; Zimmer, 2005; Zimmer & Baumann, 2003). Should $f(n)$ be correctly identified, $f(n)$ would permit one to obtain transformed magnitude estimates of sensory intensity that could be used in place of ratings for functional measurement.

AUTHOR NOTE

I thank Norman Henry Anderson and three anonymous reviewers for useful comments about the manuscript. Correspondence concerning this article should be addressed to S. C. Masin, Department of General Psychology, Università di Padova, Via Venezia 8, 35131 Padova, Italy (e-mail: scm@unipd.it).

REFERENCES

- ANDERSON, N. H. (1981). *Foundations of information integration theory*. New York: Academic Press.
- ANDERSON, N. H. (1982). *Methods of information integration theory*. New York: Academic Press.
- ANDERSON, N. H. (1996). *A functional theory of cognition*. Mahwah, NJ: Erlbaum.
- ATTNEAVE, F. (1962). Perception and related areas. In S. Koch (Ed.), *Psychology: A study of a science. Study II: Empirical substructure and relations with other sciences. Vol. 4. Biological oriented fields: Their place in psychology and in biological science* (pp. 619-659). New York: McGraw-Hill.
- BIRNBAUM, M. H., & ELMASIAN, R. (1977). Loudness "ratios" and "differences" involve the same psychophysical operation. *Perception & Psychophysics*, **22**, 383-391.
- CURTIS, D. W. (1970). Magnitude estimations and category judgments of brightness and brightness intervals: A two-stage interpretation. *Journal of Experimental Psychology*, **83**, 201-208.
- CURTIS, D. W., ATTNEAVE, F., & HARRINGTON, T. L. (1968). A test of a two-stage model of magnitude judgment. *Perception & Psychophysics*, **3**, 25-31.
- CURTIS, D. W., & FOX, B. E. (1969). Direct quantitative judgments of sums and a two-stage model for psychophysical judgments. *Perception & Psychophysics*, **5**, 89-93.
- DIAMOND, A. L. (1962). Simultaneous contrast as a function of test-field area. *Journal of Experimental Psychology*, **64**, 336-345.
- ELLERMEIER, W., & FAULHAMMER, G. (2000). Empirical evaluation of axioms fundamental to Stevens's ratio-scaling approach: I. Loudness production. *Perception & Psychophysics*, **62**, 1505-1511.
- ELLERMEIER, W., NARENS, L., & DIELMANN, B. (2003). Perceptual ratios, differences, and the underlying scale. In B. Berglund & E. Borg (Eds.), *Proceedings of the 19th Annual Meeting of the International Society for Psychophysics* (pp. 71-76). Stockholm: International Society for Psychophysics.
- ENGEN, T., & TULUNAY, Ü. (1957). Some sources of error in half-heaviness judgments. *Journal of Experimental Psychology*, **54**, 208-212.
- FAGOT, R. F., & STEWART, M. (1969). Tests of product and additive scaling axioms. *Perception & Psychophysics*, **5**, 117-123.
- GARNER, W. R. (1954a). Context effects and the validity of the loudness scales. *Journal of Experimental Psychology*, **48**, 218-224.
- GARNER, W. R. (1954b). A technique and a scale for loudness measurement. *Journal of the Acoustical Society of America*, **26**, 73-88.
- GARNER, W. R., HAKE, H. W., & ERIKSEN, C. W. (1956). Operationism and the concept of perception. *Psychological Review*, **63**, 149-159.
- GRAHAM, C. H. (1958). Sensation and perception in an objective psychology. *Psychological Review*, **65**, 65-76.
- GRAHAM, C. H., & RATOOSH, P. (1962). Notes on some interrelations of sensory psychology, perception, and behavior. In S. Koch (Ed.), *Psychology: A study of a science. Study II: Empirical substructure and relations with other sciences. Vol. 4. Biological oriented fields: Their place in psychology and in biological science* (pp. 483-514). New York: McGraw-Hill.
- HARDIN, C., & BIRNBAUM, M. H. (1990). Malleability of "ratio" judgments of occupational prestige. *American Journal of Psychology*, **103**, 1-20.
- KOFFKA, K. (1935). *Principles of Gestalt psychology*. New York: Harcourt, Brace, & World.
- LUCE, R. D. (2002). A psychophysical theory of intensity proportions, joint presentations, and matches. *Psychological Review*, **109**, 520-532.
- MARKS, L. E. (1974). *Sensory processes*. New York: Academic Press.
- MARKS, L. E. (1996). Psychophysics in the scientific market-place: Peer review of grant applications. In S. C. Masin (Ed.), *Proceedings of the 12th Annual Meeting of the International Society for Psychophysics* (pp. 329-334). Padua: International Society for Psychophysics.
- MASIN, S. C. (1983). An experimental check on Stevens's explanation of the partition paradox by successive doubling and bisecting. *Perception & Psychophysics*, **34**, 294-296.
- MASIN, S. C. (2001). Study of the bisection operation. In E. Sommerfeld, R. Kompass, & T. Lachmann (Eds.), *Proceedings of the 17th Annual Meeting of the International Society for Psychophysics* (pp. 517-522). Berlin: Pabst.
- MASIN, S. C. (2005). Test of linearity of the response function for ratings of perceived area. In J. S. Monahan, S. M. Sheffert, & J. T. Townsend (Eds.), *Proceedings of the 21st Annual Meeting of the International Society for Psychophysics* (pp. 199-204). Mt. Pleasant, MI: International Society for Psychophysics.
- MASIN, S. C., & TOMMASI, M. (2000). Test of the validity of the judged sensory ratio of 1: 2. *Perceptual & Motor Skills*, **90**, 988-994.
- MCBRIDE, R. L., & ANDERSON, N. H. (1991). Integration psychophysics in the chemical senses. In N. H. Anderson (Ed.), *Contributions to information integration theory* (pp. 295-319). Hillsdale, NJ: Erlbaum.
- MCGILL, W. (1960). The slope of the loudness function: A puzzle. In H. Gulliksen & S. Messick (Eds.), *Psychological scaling: Theory and applications* (pp. 67-81). New York: Wiley.
- MELLERS, B. A., DAVIS, D. M., & BIRNBAUM, M. H. (1984). Weight of evidence supports one operation for "ratios" and "differences" of heaviness. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 216-230.
- NARENS, L. (1996). A theory of ratio magnitude estimation. *Journal of Mathematical Psychology*, **40**, 109-129.
- PARDUCCI, A. (1982). Category ratings: Still more contextual effects! In B. Wegener (Ed.), *Social attitudes and psychophysical measurement* (pp. 89-105). Hillsdale, NJ: Erlbaum.
- PARDUCCI, A., & WEDELL, D. H. (1986). The category effect with rating scales: Number of categories, number of stimuli, and method of presentation. *Journal of Experimental Psychology: Human Perception & Performance*, **12**, 496-516.
- PARKER, S., & SCHNEIDER, B. (1974). Nonmetric scaling of loudness and pitch using similarity and difference estimates. *Perception & Psychophysics*, **15**, 238-242.
- PARKER, S., SCHNEIDER, B., & KANOW, G. (1975). Ratio scale measurement of the perceived lengths of lines. *Journal of Experimental Psychology: Human Perception & Performance*, **1**, 195-204.
- SCHNEIDER, B. (1982). The nonmetric analysis of difference judgments in social psychophysics: Scale validity and dimensionality. In B. Wegener (Ed.), *Social attitudes and psychophysical measurement* (pp. 317-337). Hillsdale, NJ: Erlbaum.
- SCHNEIDER, B., & BISSETT, R. (1988). "Ratio" and "difference" judgments for length, area, and volume: Are there two classes of sensory continua? *Journal of Experimental Psychology: Human Perception & Performance*, **14**, 503-512.
- SCHNEIDER, B., PARKER, S., FARRELL, G., & KANOW, G. (1976). The perceptual basis of loudness ratio judgments. *Perception & Psychophysics*, **19**, 309-320.
- SHEPARD, R. N. (1981). Psychological relations and psychophysical scales: On the status of "direct" psychophysical measurement. *Journal of Mathematical Psychology*, **24**, 21-57.
- 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.
- STEVENS, S. S. (1975). *Psychophysics*. New York: Wiley.
- SVENSON, O., & ÅKESSON, C. Å. (1966a). *Fractional and multiple estimates in ratio scaling* (Report No. 202). Stockholm: University of Stockholm, Psychological Laboratories.
- SVENSON, O., & ÅKESSON, C. Å. (1966b). *On the relation between fractional and multiple estimates in ratio scaling* (Report No. 220). Stockholm: University of Stockholm, Psychological Laboratories.
- TORGENSEN, W. S. (1961). Distances and ratios in psychological scaling. *Acta Psychologica*, **19**, 201-205.
- WEISS, D. J. (1972). Averaging: An empirical validity criterion for magnitude estimation. *Perception & Psychophysics*, **12**, 385-388.
- ZIMMER, K. (2005). Examining the validity of numerical ratios in loudness fractionation. *Perception & Psychophysics*, **67**, 569-579.
- ZIMMER, K., & BAUMANN, O. (2003). Validity of Luce's transformation function relating numerals and numbers in loudness fractionation. In

B. Berglund & E. Borg (Eds.), *Proceedings of the 19th Annual Meeting of the International Society for Psychophysics* (pp. 343-348). Stockholm: International Society for Psychophysics.

NOTE

1. This finding applies to sensory intensities, such as heaviness or loudness, but not to sensory extensions, such as apparent length or area. For example, Parker, Schneider, and Kanow (1975) asked participants to estimate line length differences and line length ratios. They found that the rank order of line pairs obtained from difference estimates differed from the rank order of line pairs obtained from ratio estimates. Schneider and Bissett (1988) extended this finding to area and volume. In these cases, it might seem on first thought that one could conclude that difference and ratio instructions involved different judgment operations. However, consider the following empirical results. Given the perceived

areas $a_1 < a_2 < a_3$, with a_1 and a_3 being fixed, Masin (2001) had participants set a_2 so that $a_2/a_1 = a_3/a_2$. It was found that a_2 bisected the interval between a_1 and a_x , with a_x being the value that bisected the interval between a_1 and a_3 . That is, the participants fulfilled the task of equalizing sensory ratios by successively performing two sensory bisections. This finding indicates that the participants arrived at their responses by estimating sensory differences (two bisections) under ratio instructions. Differences in rank order of stimulus pairs obtained for sensory extensions under difference instructions and ratio instructions may thus be due to a single operation of estimation of sensory distance applied once under difference instructions and more than once under ratio instructions.

(Manuscript received August 2, 2005;
revision accepted for publication January 20, 2006.)