MASIN, S. C. (2022). Old and new views on ratio judgment. In G. R. Patching (Ed.), *Proceedings of the thirty-eight annual meeting of the International Society for Psychophysics*, Lund, Sweden: International Society for Psychophysics (pp. 61-66).

OLD AND NEW VIEWS ON RATIO JUDGMENT

Sergio Cesare Masin
University of Padova, Italy
<scm@unipd.it>

Abstract

Old and new psychophysical data suggest that judged sensory ratios fail to match the actual sensory ratios, and that this discrepancy is due to the incapacity of people to judge sensory ratios. Directions for future research are outlined.

To measure sensory magnitude, Merkel (1888) assumed that people are able to produce pairs of sensory magnitudes in the specific numerical ratio of 2:1, and Richardson (1929) more boldly assumed that people are able to estimate (that is, to express as a number) the ratio between any pair of sensory magnitudes. Stevens (1936) endorsed Richardson's assumption.

Productions and estimates of reciprocal sensory ratios

For each of different loudness levels of a standard sound *S*, Geiger and Firestone (1933) asked subjects to produce the loudness of a variable sound *V* that was, in turn, 1/2, 1/4, 1/10, 1/100, 2, 4, 10, and finally 100 times that of *S*. The authors exemplified their results by asserting that if subjects "are first asked to adjust the loudness of a sound to a value twice as loud, then are asked to adjust this sound to a value half as loud, they will not, on the average, arrive at the original sound, but to a value either louder than or less loud than the original sound" (p. 26). For any sensory magnitudes a and b, let J_1 and J_2 denote the estimates of the ratios a/b and b/a, respectively. These estimates are equivalent to the respective sensory ratios if $J_1 \cdot J_2 = 1$, given that $(a/b) \cdot (b/a) = 1$. Svenson & Åkesson (1966) found instead that $J_1 \cdot J_2 \neq 1$.

Test of Richardson's assumption

Engen and Levy (1955) and, perhaps surprisingly, Stevens (1956) himself found that estimation of sensory ratios yielded unreliable measures of sensory magnitude because the exponent of the psychophysical power function turned out to vary with the value of the standard stimulus used.

Arithmetical relationships between estimates of sensory ratios agree with this finding. For any sensory magnitudes a, b, and c, let J_1, J_2, and J_3 denote the estimates of the ratios a/b, a/c, and c/b, respectively. Because $a/b = (a/c) \cdot (c/b)$, these estimates yield reliable sensory measures only if $J_1 = J_2 \cdot J_3$. Eisler (1960) and Fagot and Stewart (1969) found instead that $J_1 \neq J_2 \cdot J_3$.

The Ross-Di Lollo hypothesis

The results of the above tests (and of a surplus of more recent similar tests not considered here) led to the belief that people misjudge sensory ratios (Luce, 2002; Shepard, 1978). Could this misjudgment indicate incapacity to produce or estimate sensory ratios? The empirical results of Ross and Di Lollo (1971) described below suggest a positive answer.

If subjects can estimate sensory ratios of variable stimuli *V* relative to a standard stimulus *S*, Richardson's assumption states that subjects produce the estimates

$$J = v/s \tag{1}$$

with v and s being the perceptual or memory magnitudes caused by V and S, respectively. If the measures of v and s are known, the functional relationship $J = v/s$ implies that plotting J in a Cartesian coordinate system against v separately for different values of s yields *factorial curves* diverging rightward. Unfortunately, v and s are unknown. However, considering that v increases monotonically with V, the relationship $J = v/s$ also implies that plotting J against V separately for each S yields factorial curves that diverge rightward.

With V and S known, a test of this implication serves as an indirect test of Richardson's assumption (Anderson, 1974). Since v and s are unknown, this test can only be an ordinal test: Richardson's assumption is rejected if the empirically-obtained factorial curves are parallel or converge rightward, and is supported—but not verified—if these curves diverge rightward. The test cannot verify Richardson's assumption because there are functional relationships other than $J = v/s$ that also imply a rightward divergence of factorial curves.

Ross and Di Lollo (1971) had three groups of subjects compare variable stimuli V relative to a single standard stimulus S. The range of V differed for each group, which involved a largely different remembered s for each group. Subjects were asked to estimate the ratio v/s. Plotting J against V separately for each s yielded nearly parallel factorial curves. Since estimates of $|v - s|$ imply parallel factorial curves, Ross and Di Lollo (1971) hypothesized that subjects manifested "interval scale behavior, that is, the direct representation of intervals as numbers by a process of comparison of intervals" (p. 520). That is, subjects did not estimate sensory ratios. Atkinson and Ward (1972), Fagot, Stewart, and Kleinknecht (1975), and Schneider, Parker, and Upenieks (1982) obtained data in line with this hypothesis.

Test of the Ross-Di Lollo hypothesis

On each trial of two experiments, Masin (2014) presented two successive stimuli: a surface of luminance V (8, 15, 26, or 49 cd/m^2) and a surface of luminance S (1, 2, or 4 cd/m^2) in one experiment, and an object of weight V (440, 535, 650, or 750 g) and an object of weight S (85, 175, or 260 g) in the other. For each combination of V and S, the subjects' task was to estimate the brightness ratio v/s in one experiment and the heaviness ratio v/s in the other. Instructions to make ratio estimates used no numerical examples to exclude spurious divergence of factorial curves (Guirao, 1987; Masin, Weiss, & Brancaccio, 2021; Robinson, 1976)

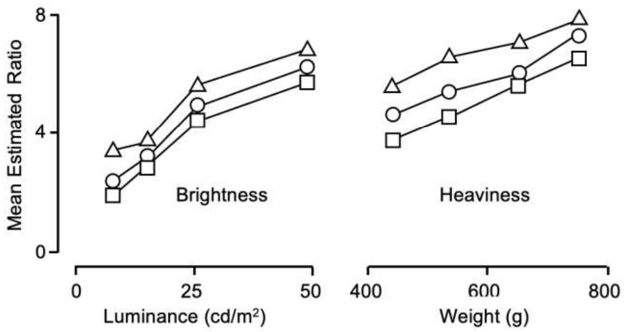

Fig. 1. Mean estimated brightness ratio or heaviness ratio as a function of luminance or weight, respectively. The curve parameter is the respective standard stimulus intensity. Triangles denote the smallest standard. Redrawn from data of Masin (2014, Figs. 1 and 2).

Fig. 1 shows the factorial curves obtained by plotting mean estimated brightness ratio and mean estimated heaviness ratio against the corresponding V, separately for each corresponding S. Triangles and squares denote the smallest and largest S, respectively. Factorial curves are nearly parallel. Because the relationship $J = |v - s|$ implies parallel factorial curves, the results agree with the Ross-Di Lollo hypothesis of "interval scale behavior." Ratio production yielded similar results (Masin, 2007, 2014).

Ratio judgment through counting

It may be that people easily believe that they can generally estimate sensory ratios because they often visually verify that they can somewhat accurately apprehend small ratios of visual extents. This automatic apprehension resembles subitizing, the ability to apprehend numerosities of 1 to 4 items at a glance (Taves, 1941; Balakrishnan & Ashby, 1991, 1992).

People subitize through an automatic counting process of which they are unaware. This counting process can, however, be demonstrated by measuring the mean time the subjects take to name the number of items presented in a visual display. The left diagram in Fig. 2 shows this mean response time plotted against item numerosity (Balakrishnan & Ashby, 1991). Mean response time increases with item numerosity indicating an automatic counting process.

People also use an automatic counting process to perform the task of estimating length ratios (Hartley, 1977). The central diagram in Fig. 2 shows the mean response time taken to name the apparent-length ratio between each of four variable lines of 2 to 5 cm and a standard line with mean length of 1 cm (Masin, 2013). Mean response time increases with the length of the variable line. This fact suggests that subjects automatically counted the number of times the standard line was contained in each variable line, rather than merely apprehending length ratios.

The Stevens-Moskowitz hypothesis

The following results indicate that the counting process used in the task of estimating length ratios might occur only for extensive sensations. The right diagram in Fig. 2 shows the mean time taken to name the brightness ratio between each of four variable luminances from 8 to 50 cd/m^2 and a standard mean luminance of 2.5 cd/m^2 (Masin, 2013). Mean response time was essentially invariant with luminance, indicating that subjects did not count brightness units. It might thus be that people cannot count units of intensive sensation.

If people fail to estimate brightness ratios (left diagram in Fig. 1) and to count brightness units (right diagram in Fig. 2), what other operation do they use to perform the task of estimating brightness ratios? Perhaps surprisingly, Stevens (1956) himself provided an answer, noting that

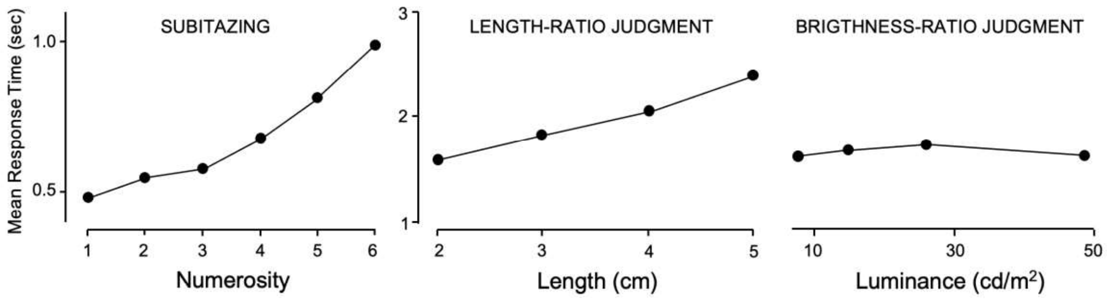

Fig. 2. Mean response time to report item numerosity (left) and to estimate length ratios (center) and brightness ratios (right) plotted against numerosity, length, and luminance, respectively. Redrawn from data of Balakrishnan and Ashby (1991, Fig. 1) and Masin (2013, Figs. 1 and 2).

"some [of the subjects instructed to estimate sensory ratios] seem to make their estimates on an interval-scale, or even on an ordinal scale, instead of on the ratio-scale we are trying to get them to use" (p. 23). Moskowitz (1977) also remarked that "[t]he range of numbers [ratio estimates] differs among individuals. Some people choose numbers between 1 and 10 and operate as if they were working on a category scale with equal intervals rather than equal ratios. Others choose a range between 0 and 100 and operate in the same way. Finally, others choose numbers between 0 and 1,000 (or some other range)" (p. 210).

Generality of the Stevens-Moskowitz hypothesis

The Stevens-Moskowitz hypothesis basically states that, if instructed to estimate sensory ratios, some subjects may instead rate sensory differences on a category scale. Masin (2022) tested the generality of this hypothesis, using instructions without numerical examples to prevent spurious effects possibly caused by such examples.

The experiment whose results are reported in the left diagram Fig. 1 was repeated using one range of *V* for one group of subjects, and about double this range for another group. If estimates of brightness ratios are proportional to brightness ratios, the height of factorial curves should increase with the range of *V*. If, instead, estimates of brightness ratios are ratings of brightness differences on a category scale, the height of factorial curves should be essentially the same for the two groups, considering that both groups would map brightness differences on scales with statistically the same endpoints. Fig. 3 illustrates the results. The patterns of factorial curves obtained by the two groups have the same height, indicating that subjects rated sensory differences on a category scale rather than estimating sensory ratios.

Different judgment operations

The results illustrated above suggest that people count units of apparent extent to fulfill the task of estimating ratios of extensive magnitudes (length) and that they rate differences of intensive magnitude to fulfill the task of estimating ratios of intensive magnitudes (brightness). It may be that people use various types of judgment operations (Masin, Brancaccio, & Tomassetti, 2019).

The results reported in the right diagram in Fig. 1 were obtained with subjects lifting pairs of weights unimanually. Masin and Brancaccio (2017) repeated this experiment with subjects now lifting pairs of weights bimanually. These pairs were factorial combinations of six variable weights *V* of 500 to 1000 g with four standard weights *S* of 100 to 400 g. Fig. 4 shows mean estimated heaviness ratio plotted against *V* separately for each *S*. The factorial curves diverge, thus indicating that subjects did not estimate heaviness differences. This divergence might seem

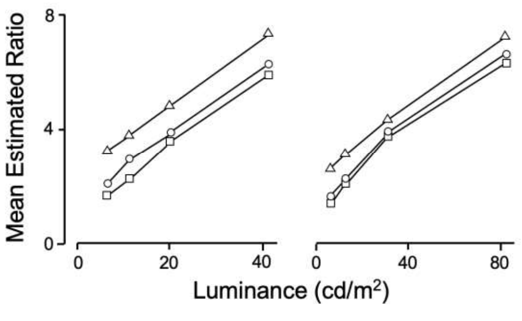

Fig. 3. Mean estimated brightness ratio plotted against variable stimulus luminance for each of three standard stimulus luminances (triangles denote the smallest standard) and for each of two different luminance ranges. Redrawn from data of Masin (2022, Fig. 2).

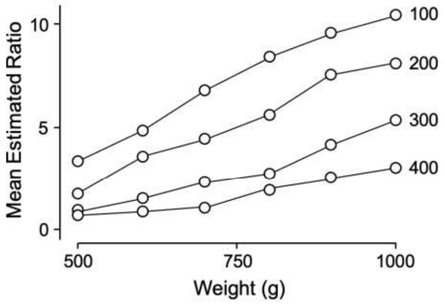

Fig. 4. Mean estimated heaviness ratio plotted against variable stimulus weight for each of four standard stimulus weights (grams). Redrawn from data of Masin and Brancaccio (2017, Fig. 3).

to support that subjects estimated heaviness ratios. However, the obtained factorial curves are essentially uniformly spaced, whereas the sensory relationships $J = v/s$ implies hyperbolically spaced factorial curves [the values used for S were equispaced and the psychophysical function for heaviness was essentially linear in the range of 100 to 400 g (J. C. Stevens & Rubin, 1970)]. These results thus indicate that subjects may have used a judgment operation based on a sensory relationship other than a sensory difference ($J = |v - s|$) or a sensory ratio ($J = v/s$).

Future research

The following may be relevant goals for future research: (i) determining which sensations allow for mental counting of sensory units, (ii) exploring comprehensively the sensory relationships people respond to while they fulfill the task of producing or estimating sensory ratios, and [since people can be conditioned to produce specific patterns of factorial curves (Price, Meyer, & Koh, 1992)] (iii) assessing new methods of constrained scaling where subjects learn to produce the patterns of factorial curves implied by the functional relationships $J = |v - s|$ or $J = v/s$.

References

Anderson, N. H. (1974). Algebraic models in perception. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (Vol. 2, pp. 215–298). New York: Academic Press.

Atkinson, W. H., & Ward, L. M. (1972). On the internal consistency of magnitude estimations. *Perceptual and Motor Skills*, 35, 903–913.

Balakrishnan, J. D., & Ashby, F. G. (1991). Is subitizing a unique numerical ability? *Perception & Psychophysics*, 50, 555–564.

Balakrishnan, J. D., & Ashby, F. G. (1992). Subitizing: Magical numbers or mere superstition?. *Psychological Research*, 54, 80–90.

Eisler, H. (1960). Similarity in the continuum of heaviness with some methodological and theoretical considerations. *Scandinavian Journal of Psychology*, 1, 69–81.

Engen, T., & Levy, N. (1955). The influence of standards on psychophysical judgments. *Perceptual and Motor Skills*, 5, 193–197.

Fagot, R. F., & Stewart, M. (1969). Tests of product and additive scaling axioms. *Perception & Psychophysics*, 5, 117–123.

Fagot, R. F., Stewart, M. R., & Kleinknecht, R. E. (1975). Representations for biased numerical judgments. *Perception & Psychophysics*, 17, 309–319.

Geiger, P. H., & Firestone, F. A. (1933). The estimation of fractional loudness. *Journal of the Acoustical Society of America*, 5, 25–30.

Guirao, M. (1987). Estimaciones numéricas de brillo y luminosidad con y sin rangos fijos. *Revista Latinoamericana de Psicología*, 19, 297–317.
Hartley, A. A. (1977). Mental measurement in the magnitude estimation of length. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 622–628.
Luce, R. D. (2002). A psychophysical theory of intensity proportions, joint presentations, and matches. *Psychological Review*, 109, 520–532.
Masin, S. C. (2007). The judgment operation of ratio production. *Perception & Psychophysics*, *69*, 1–8.
Masin, S. C. (2013). On the ability to directly evaluate sensory ratios. *Attention, Perception, & Psychophysics*, *75*, 194–204.
Masin, S. C. (2014). Test of the ratio judgment hypothesis. *Journal of General Psychology*, 141, 130–150.
Masin, S. C. (2022). Evidence that numerical estimates of subjective ratios may be numerical ratings of subjective differences. *Psicológica*, 43(1): e14568.
Masin, S. C., & Brancaccio, A. (2017). Judgments of differences and ratios of subjective heaviness. *Attention, Perception, & Psychophysics*, 79, 1860–1869.
Masin, S. C., Brancaccio, A., & Tomassetti, A. (2019). Tests of the abilities to judge ratios of extensive and intensive sensory magnitudes. *Attention, Perception, & Psychophysics*, 81, 2102–2119.
Masin, S. C., Weiss, D. J., Brancaccio, A. (2021). Further evidence of the invalidity of the magnitude estimation method. *European Review of Applied Psychology*, 71, 100649.
Merkel, J. (1888). Die Abhängigkeit zwischen Reiz und Empfindung, Erste Abtheilung. *Philosophische Studien*, 4, 541–594 [Zweite Abtheilung: 1889, 5, 245–291, 499–557].
Moskowitz, H. R. (1977). Magnitude estimation: Notes on what, how, when, and why to use it. *Journal of Food Quality*, 3, 195–227.
Price, P. C., Meyer, D. E., & Koh, K. (1992). Multivariable function learning: Applications of the adaptive regression model to intuitive physics. In J. K. Kruschke (Ed.), *Proceedings of the fourteenth annual conference of the Cognitive Science Society* (pp. 552–557). Hillsdale, NJ: Erlbaum.
Richardson, L. F. (1929). Quantitative mental estimates of light and colour. *British Journal of Psychology*, 20, 27–37.
Robinson, G. H. (1976). Biasing power law exponents by magnitude estimation instructions. *Perception & Psychophysics*, 19, 80–84.
Ross, J., & Di Lollo, V. (1971). Judgment and response in magnitude estimation. *Psychological Review*, 78, 515–527.
Schneider, B., Parker, S., & Upenieks, E. G. (1982). The perceptual basis of judgments of pitch differences and pitch ratios. *Canadian Journal of Psychology*, 36, 4–23.
Shepard, R. N. (1978). On the status of "direct" psychophysical measurement. In C. W. Savage (Ed.), *Minnesota studies in the philosophy of science* (Vol. 9, pp. 441–490). Minneapolis: University of Minnesota Press.
Stevens, J. C., & Rubin, L. L. (1970). Psychophysical scales of apparent heaviness and the size-weight illusion. *Perception & Psychophysics*, 8, 225–230.
Stevens, S. S. (1936). A scale for the measurement of a psychological magnitude: Loudness. *Psychological Review*, 43, 405–416.
Stevens, S. S. (1956). The direct estimation of sensory magnitudes—loudness. *American Journal of Psychology*, *69*, 1–25.
Svenson, O., & Åkesson, C. Å. (1966). *Fractional and multiple estimates in ratio scaling* (Report No. 202). Stockholm: University of Stockholm, Psychological Laboratories.
Taves, E. (1941). Two mechanisms for the perception of visual numerousness. *Archives of Psychology* (Columbia University), New York, No. 265.