“Ratios” and “differences” in perceived sweetness intensity

CEES DE GRAAF and JAN E. R. FRIJTERS
Agricultural University, Wageningen, The Netherlands

For a number of perceptual continua, it has been shown in previous studies that subjects use only one quantitative comparison between two sensory impressions of a pair of stimuli, irrespective of whether they are instructed to judge “ratios” or “differences.” This comparison can be described by algebraic subtraction. The present study was designed to investigate whether this one-operation theory for psychophysical judgment also applied to the sensory continuum of sweetness. Subjects were presented with pairs of fructose solutions, and judged “ratios” of, or “differences” in, perceived sweetness intensities. The pairs were constructed on the basis of a factorial judgment design. The results showed that the reported “differences” could be adequately described by a difference response model, and that the reported “ratios” could be adequately described by a ratio response model. However, the reported “ratios” and reported “differences” were monotonically related, and the marginal means of the log-transformed response matrix of “ratios” were a linear function of the marginal means of the response matrix of “differences.” These results are incompatible with the notion that subjects judged differences when instructed to judge “differences,” but ratios when instructed to judge “ratios.” The consistency of the ratio response model with “ratio” judgments is probably caused by a comparative operation based on “differences” in combination with an exponential response output function. It may be concluded that subjects judge only “differences,” and not “ratios,” between perceived sweetness intensities.

For several perceptual continua, such as heaviness (Birnbau & Veit, 1974; Mellers, Davis, & Birnbau, 1984), loudness (Birnbau & Elmasian, 1977; Schneider, Parker, Kanow, & Farell, 1976), pitch (Schneider, Parker, & Upenieks, 1982), darkness of dot patterns (Birnbau, 1978; Mellers & Birnbau, 1982), and of gray papers (Veit, 1978), it has been shown that the way in which subjects compare the absolute magnitudes of two sensations elicited by a pair of stimuli is independent of the instructions given to the subject. The type of comparison is the same whether the subject is instructed to judge “differences” or to judge “ratios.” This cognitive operation can be described by algebraic subtraction. Birnbau and his colleagues (e.g., Birnbau, 1982; Mellers et al., 1984) have argued that under proper experimental conditions, “difference” instructions (e.g., category scales) induce a linear response output function (i.e., the function that relates subjective intensity “differences” to observable “difference” responses is linear). However, “ratio” instructions (e.g., magnitude estimation), in which the numerical examples are geometrically spaced (e.g., the standard is 10; if a stimulus is 10 times stronger than the standard, one is to assign the number 100; if a stimulus is $\frac{1}{10}$ as strong as the standard, one is to assign the number 1) are supposed to induce an approximately exponential response function (Birnbau, 1978, 1980; Mellers et al., 1984). An exponential response output

function transforms perceived “differences” into numerical ratings that are consistent with a ratio response model.

The empirical evidence in favor of Birnbau’s one-operation theory has evolved from experiments in which the functional measurement approach, in combination with a two-stimulus procedure, was applied (e.g., Anderson, 1974, 1981; Birnbau, 1982). In these experiments, pairs of stimuli were presented along with the instruction to judge either the “ratio” or the “difference” between the subjective intensities of each of a number of row stimuli, and each of a number of column stimuli. The results of these experiments showed that the reported “ratios” were monotonic with, and an approximately exponential function of, reported “differences.” Such monotonicity is incompatible with the notion that subjects report differences when instructed to judge “differences,” and that they report ratios when instructed to judge “ratios.” The rank order in magnitude of reported ratios must be different from the rank order in magnitude of reported differences, if each kind of instruction generates the comparative operation as intended (e.g., Birnbau, 1982; Veit, 1978).

Monotonicity between “difference” and “ratio” judgments is an indication that subjects use one comparative operation only, given the condition that the psychological range of the sensory impressions of the experimental stimuli is sufficiently large. When the psychological range is small, “difference” and “ratio” judgments will have the same rank order, even if subjects distinguish between psychological differences and ratios (Rule & Curtis, 1980; Schneider et al., 1982). Therefore, rank order between “ratio” and “difference” judgments must be evaluated

Correspondence should be addressed to Jan E. R. Frijters, Department of Food Science, Agricultural University, De Dreijen 12, 6703 BC Wageningen, The Netherlands.

in combination with the psychological range of sensory impressions.

The present study was designed to investigate whether or not the one-operation theory of comparative judgment holds for judgments of perceived sweetness intensity. The design of the study was similar to that of the studies previously cited, where there were two experimental conditions, one for each type of instruction. In one condition, subjects were instructed to judge the "difference" in perceived sweetness intensity of two fructose solutions, and in the other, to judge the "ratio."

METHODS AND MATERIALS

Subjects

Forty paid volunteers were used as subjects. All were undergraduate students from the Agricultural University, whose ages ranged from 18 to 29 years. None of the subjects had had prior experience in psychophysical scaling tasks. They were given no information about the aim of the experiment or the nature of the substances used.

Stimuli

The stimuli were solutions of fructose (Merck 5321) dissolved in demineralized water. The concentrations of the six fructose solutions were 0.125, 0.218, 0.379, 0.660, 1.149, and 2.00 M. A pair consisting of 0.125 M fructose as the first stimulus and 2.50 M fructose as the second stimulus served as the reference pair. The solutions were prepared at least 24 h prior to tasting, and were stored at 4°C for no longer than 7 days.

Design

The subjects were assigned randomly to each of two groups of 20 subjects. One group was instructed to judge the "difference" in perceived sweetness intensity between the two stimuli of each pair, and the other group was instructed to judge the "ratio" of the intensities.

Both conditions involved a 6 (number of different stimuli tasted first) $\times$ 6 (number of different stimuli tasted second) symmetrical factorial judgment design (De Graaf, Frijters, & Van Trijp, 1987). Thus, each of the 6 fructose stimuli served as both first and second stimulus in all possible pairs of stimuli (i.e., 36 in total).

Procedure

In both the "difference" and the "ratio" condition, the instructions emphasized that only the sweetness intensity was to be judged, and that the hedonic value and potential side tastes were to be disregarded. The subjects were requested to rinse their mouths thoroughly with demineralized water, both within and between pairs. The stimuli were presented at room temperature, in polystyrene medicine cups, each containing about 10 ml of solution. The 36 pairs in the factorial judgment design were presented in a random sequence and in a different order for each subject. The reference pair (0.125 M fructose as first and 2.5 M fructose as second stimulus) was presented at the beginning of each session, and again after the 12th and 24th pair in each series of 36. The "difference" between, or "ratio" of, perceived sweetness intensities in this pair was expected to be larger than the "difference" in, or "ratio" of, any other pair to be judged.

The time interval between the two stimuli within a pair, as well as the time interval between pairs, was 50 sec. Each subject tasted a series of 36 pairs three times, one in each of three sessions.

"Difference" task. The subjects were instructed to judge the magnitude of the "difference" in perceived sweetness intensity between the first and second stimulus of each pair. The ratings had to be expressed by a slash mark on a 250-mm visual analogue scale (see De Graaf et al., 1987), of which the midpoint was defined as "the

first and second stimulus are equal in sweetness intensity." If the first stimulus was perceived as sweeter than the second, the subject put a mark on the left side of the scale. Similarly, the subject put a mark on the right side of the scale when the second stimulus was perceived as being sweeter. The larger the "difference" between stimuli, the larger the distance of the mark from the middle of the scale. The left and the right anchors of the scale were labeled "maximum difference." In the instructions, "maximum difference" was defined as the "difference" in sweetness intensity between the stimuli of the reference pair.

For data analysis, the ratings were measured as the distance in millimeters from the left anchor of the scale. Thus, the numerical value 125 meant no "difference," values below 125 indicated that the first stimulus was perceived as sweeter than the second, and values above 125 indicated that the second stimulus was perceived as sweeter.

"Ratio" task. The subjects were instructed to judge the magnitude of the "ratio" of the perceived sweetness intensity of the first and the perceived sweetness intensity of the second stimulus of each pair. They had to first identify which stimulus of each pair was the sweetest, and subsequently to assign a number reflecting the "ratio" of the perceived sweetness intensity of the sweetest stimulus to the intensity of the least sweet stimulus. If the first and second stimuli were equal in perceived sweetness intensity, the subjects had to assign the number 1. The maximum "ratio," being the "ratio" of the perceived sweetness intensities of the stimuli of the reference pair, was declared in the instructions to be 25.

RESULTS

In Figure 1, panel A, the arithmetic means of "difference" judgments are plotted as a function of the marginal mean of the "difference" responses to the second (column) stimulus, with a separate curve for each level of the first stimulus. A subtractive comparative operation in combination with a linear judgment function implies that this factorial plot must exhibit six straight lines, which run parallel to each other. This is confirmed by analysis of variance (ANOVA), showing that the row (first) $\times$ column (second) interaction is statistically not significant [$F(25,475) = .93, p > .5$]. The row $\times$ column interaction accounts for 0.16% of the systematic variance (throughout this paper, systematic variance is the variance due to the first stimulus, the second stimulus, and the interaction between the first and second stimulus).

The "ratio" responses consisted of numbers between 1 and 25, where about half of the responses referred to pairs in which the first stimulus was perceived as the sweetest, and the other half referred to pairs in which the second stimulus was perceived as the sweetest. The "ratio" response matrix consisted of the numerical values reflecting how many times the second stimulus was rated as sweeter than the first (sweetness second/sweetness first). So, if it was reported that the second stimulus was five times sweeter than the first stimulus, the numerical value in the response matrix is 5. Similarly, if it was reported that the first stimulus tasted five times sweeter than the second stimulus, the value in the response matrix is 0.2 (1/5).

If the ratio response model applies, then the matrix of the natural logarithms of reported "ratios" can be described by a subtractive model. In Figure 1, panel B,

Figure 1. Panel A: Arithmetic means of reported "differences," averaged over subjects and replicates, as a function of scale values of responses to the second stimulus, with a separate curve for each concentration of the first stimulus. A scale value of 125 implies no difference; scale values above 125 imply that the second stimulus of the pair was perceived as being sweeter; scale values below 125 indicate that the first stimulus of the pair was perceived sweeter than the second. Panel B: The arithmetic means of the log-transformed reported "ratios" averaged over subjects and replicates as a function of the scale values of the log-transformed responses to the second stimulus, with a separate curve for each concentration of the first stimulus.

the arithmetic means of the log-transformed "ratio" responses are plotted as a function of the marginal means of the log-transformed "ratio" responses of the second (column) stimulus, with a separate curve for each level of the first (row) stimulus. If the ratio response model is valid, then the lines within this panel should run parallel. However, parallelism is not a sufficient condition for the conclusion that reported "ratios" are a linear representation of judged ratios, as will be shown below. Visual inspection shows deviations from parallelism. This is confirmed by ANOVA, showing that the deviations from parallelism are statistically significant [$F(25,475) = 13.10, p < .0001$]. This row $\times$ column interaction accounts for only 2.4% of the total systematic variance. Thus, although the deviations from the ratio model are statistically significant, these deviations are not substantive.

In summary, the difference response model is confirmed by the results, and the ratio response model provides a good description of the data, although minor (but statistically significant) deviations occur. Thus, the results of the experimental condition in which subjects were instructed to judge "differences" suggest that subjects have indeed judged differences. Similarly, the results of the "ratio" condition suggest that subjects have judged ratios. On the basis of these results, one might be inclined to conclude that subjects judge differences when instructed to judge "differences" and judge ratios when instructed to judge "ratios."

However, if both models are valid at the same time, then the rank orders of reported "differences" and of

reported "ratios" must be different (Birnbau, 1978, 1980). Additional analysis (Figure 2) shows that reported "ratios" are a monotonic function of reported "differences." The Spearman rank correlation coefficient between the geometric means of reported "ratios" and of the arithmetic means of reported "differences" is 0.996.

Figure 2. Geometric means of reported "ratios" as a function of arithmetic means of reported "differences."

This observation indicates that subjects must have used the same comparative operation in both judgmental tasks (Birnbbaum, 1982). According to the one-operation theory, subjects judge only differences and not ratios. If the "ratio" instructions induce an exponential response output function, the scale values (i.e., the marginal means) of the log-transformed responses of the "ratio" response matrix must be a linear function of the scale values of the "difference" response matrix (Birnbbaum, 1980). Figure 3 shows that is confirmed by the data. The R^2 value of the function fitted by orthogonal linear regression (Hampton, 1983) is 0.998. This result also indicates that the deviations from parallelism in the log-transformed "ratio" response matrix (earlier found to be 2.4% of the systematic variance; see Figure 1, panel B) are not due to deviations from judged "ratios," but result from deviations from an exponential response output function of judged "differences." Apparently, "ratio" instructions induce a response output function that transforms judged "differences" into responses spuriously consistent with a ratio model.

Figure 4 shows, for the "difference" response matrix, the relationship between the fructose concentration and the mean scale value for each stimulus averaged over the row marginal mean and the column marginal mean. This function is an empirical estimation of the psychophysical function of fructose.

The magnitude of the psychological range of the sensory impressions in the present study can be estimated by using the data from other studies on the sweetness intensity of fructose. De Graaf et al. (1987) and De Graaf and Frijters (1988) suggested that they used a scaling method that yielded a ratio scale of perceived sweetness intensity. In the study of De Graaf et al. (1987), 0.125 M

Figure 3. Relationship between the scale values of the "difference" response matrix and the scale values of the log-transformed "ratio" response matrix.

Figure 4. Scale values of the perceived sweetness intensity plotted as a function of the concentration of the fructose stimuli, averaged over the row and column scale values. The scale value of 0.125 M fructose was set equal to zero, and the scale values for the other stimuli were calculated as the absolute difference from this scale value. Thus, the origin of this scale is arbitrary.

fructose had a sweetness intensity of 14 units and 2.00 M fructose had a sweetness intensity of 115 units. The calculated ratio between these two sensory impressions is about 8. De Graaf and Frijters obtained scale values for 0.125 M and 2.00 M fructose. In that study, 0.125 M fructose had a scale value of about 8 and 2.00 M fructose had a scale value of about 80, a calculated ratio of about 10. Thus, the actual ratio between 2.00 M fructose and 0.125 M fructose is about 8–10, provided that the scales developed by De Graaf et al. (1987) and De Graaf and Frijters (1988) are genuine ratio scales as they contend. Schneider et al. (1982) have shown that a range of this magnitude is sufficiently large to produce nonmonotonicity between judged ratios and judged differences, if subjects distinguish between psychological differences and psychological ratios. In addition, when subjects distinguished between differences and ratios, the relationship shown in Figure 3 should be clearly nonlinear (see Schneider et al., 1982, Figure 9), which is not the case.

DISCUSSION

The results of this study concur with those of similarly designed experiments that have investigated other perceptual continua. If the results of both experimental conditions had been analyzed separately, it could have been concluded that both the ratio and the difference response models were valid. The psychological range of the sensory impressions in this study was large enough to produce a nonmonotonic relationship between "difference" and "ratio" judgments, had subjects actually distinguished between differences and ratios of perceived sweetness intensity. However, the monotonic relationship between the

arithmetic means of judged "differences" and the geometric means of judged "ratios" shows that either one of the two response models must be invalid. Apparently, subjects compared the perceived sweetness intensities of each pair in the same way, irrespective of instructions. It can therefore be concluded that Birnbaum's one-operation theory also applies to judgments of perceived sweetness intensity. The comparison between two subjective impressions is amenable to description by algebraic subtraction.

The results from this study could also be explained by assuming that subjects judged ratios of perceived sweetness intensities (Birnbaum & Veit, 1974) instead of differences in both "ratio" and "difference" conditions. However, there is sufficient experimental evidence from studies on other perceptual continua to conclude that the comparative judgmental operation relating two subjective impressions is a subtractive and not a ratio operation (Birnbaum & Mellers, 1978; Hagerty & Birnbaum, 1978; Mellers & Birnbaum, 1982; Mellers et al., 1984; Schneider et al., 1976; Schneider et al., 1982; Veit, 1978).

Rule, Curtis, and their colleagues (e.g., Rule & Curtis, 1980, 1982; Rule, Curtis, & Mullin, 1981) argued that subjects are able to judge either differences or ratios when instructed to do so. According to these authors, category ratings are linear with subjective intensity, and magnitude-estimation instructions induce a response output function that has the form of a power function. This two-operation theory predicts that the geometric means of reported "ratios" are a nonmonotonic function of the arithmetic means of reported "differences" (Rule & Curtis, 1980). This theory also predicts that the relationship between the marginal means of the log-transformed "ratio" response matrix (S_r) and the marginal means of the "difference" response matrix (S_d) has the form $S_r = a \log(S_d + b) + c$, where a , b , and c are constants (Birnbaum, 1980). Figures 2 and 3 show that these data are not in line with either prediction, and hence do not support the two-operation theory.

The one-operation theory assumes that psychophysical judgment encompasses two stages, a psychophysical stage, relating physical intensity to perceived intensity, and a judgmental stage, relating perceived intensity to observable response (e.g., Attneave, 1962; Torgerson, 1961; Treisman, 1964). Anderson's functional measurement approach, in combination with a two-stimulus procedure, provides a means of separating these processes. The parallelism in Figure 1, panel A, indicates that the judgmental stage can be described by a linear (difference) function. This implies that the function plotted in Figure 4 is an empirically estimated psychophysical function of fructose, of which the scale values represent perceived sweetness intensities on an interval scale.

This function has a sigmoid form when plotted on a log-linear plot. This finding is in line with the results of De Graaf et al. (1987) and De Graaf and Frijters (1988), who used a similar methodology. The sigmoid shape of

the psychophysical function of fructose on a log-linear plot also implies that the psychophysical function of fructose as assessed by the present procedure is not a power function.

The two-stage interpretation of psychophysical judgment (the S-O-R paradigm) conflicts with the S-R approach of Stevens (1975) and Moskowitz (e.g., 1970, 1971), who believed that sensory responses obtained by using magnitude estimation instructions have a one-to-one relationship with perceived taste intensities. Following the two-stage interpretation, we hold the view that the observable response is a behavioral parameter, which is not necessarily a linear representation of subjective intensity (Frijters & De Graaf, 1987; Frijters & Oude Ophuis, 1983). Investigators who have adopted the two-stage interpretation of psychophysical judgment have shown that the responses obtained by using magnitude estimation are a nonlinear, positively accelerating function of subjective intensity (e.g., Rule & Curtis, 1982; Veit, 1978; Weiss, 1972). This implies that the power functions, which Stevens (1969) and Moskowitz (1970, 1971) obtained by magnitude estimation, do not reflect the psychophysical functions as they intended. These (S-R) functions comprise two functions, the psychophysical function (S-O) and the response output function (O-R). As the response output function depends on several experimental factors (e.g., Baird & Noma, 1978), it is not surprising, for example, that the exponents of fitted power functions for substances such as sucrose and NaCl (Meiselman, 1972) show a large variability.

REFERENCES

- ANDERSON, N. H. (1974). Algebraic rules in perception. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (Vol. 2, pp. 215-298). New York: Academic Press.
- ANDERSON, N. H. (1981). *Foundations of information integration theory*. New York: Academic Press.
- ATTNEAVE, F. (1962). Perception and related areas. In S. Koch (Ed.), *Psychology: A study of science* (Volume IV, Study II, pp. 619-659). New York: McGraw-Hill.
- BAIRD, J. C. & NOMA, E. C. (1978). *Fundamentals of scaling and psychophysics*. New York: Wiley.
- BIRNBAUM, M. H. (1978). Differences and ratios in psychological measurement. In N. J. Castellan & F. Restle (Eds.), *Cognitive theory* (Vol. 3, pp. 33-74). Hillsdale, NJ: Erlbaum.
- BIRNBAUM, M. H. (1980). Comparison of two theories of "ratio" and "difference" judgments. *Journal of Experimental Psychology: General*, **109**, 304-319.
- BIRNBAUM, M. H. (1982). Controversies in psychological measurement. In B. Wegener (Ed.), *Social attitudes and psychophysical measurement* (pp. 401-485). Hillsdale, NJ: Erlbaum.
- BIRNBAUM, M. H., & ELMASIAN, R. (1977). Loudness "ratios" and "differences" involve the same psychophysical operation. *Perception & Psychophysics*, **22**, 383-391.
- BIRNBAUM, M. H., & MELLERS, B. A. (1978). Measurement and the mental map *Perception & Psychophysics*, **23**, 403-408.
- BIRNBAUM, M. H., & VEIT, C. T. (1974). Scale convergence as a criterion for rescaling: Information integration with difference, ratio, and averaging tasks. *Perception & Psychophysics*, **15**, 7-15.
- DE GRAAF, C., & FRIJTERS, J. E. R. (1988). Assessment of the taste interaction between two qualitatively similar tasting substances: A comparison between comparison rules. *Journal of Experimental Psychology: Human Perception & Performance*, **14**, 526-538.

- DE GRAAF, C., FRIJTERS, J. E. R., & VAN TRIJP, H. C. M. (1987). Taste interaction between glucose and fructose assessed by functional measurement. *Perception and Psychophysics*, **41**, 383-392.
- FRIJTERS, J. E. R., & DE GRAAF, C. (1987). The equiratio taste mixture model successfully predicts the sensory response to the sweetness intensity of complex mixtures of sugars and sugar-alcohols. *Perception*, **16**, 615-628.
- FRIJTERS, J. E. R., & OUDE OPHUIS, P. A. M. (1983). The construction and prediction of psychophysical power functions for the sweetness of equiratio sugar mixtures. *Perception*, **12**, 753-767.
- HAGERTY, M., & BIRNBAUM, M. H. (1978). Nonmetric tests of ratios vs. subtractive theories of stimulus comparison. *Perception & Psychophysics*, **24**, 121-129.
- HAMPTON, J. A. (1983). A note on describing the linear relationship between a pair of correlated dependent variables. *Bulletin of the British Psychological Society*, **36**, 408-409.
- MEISELMAN, H. L. (1972). Human taste perception. *Critical Reviews in Food Technology*, **3**, 89-117.
- MELLERS, B. A., & BIRNBAUM, M. H. (1982). Loci of contextual effects in judgment. *Journal of Experimental Psychology: Human Perception and Performance*, **8**, 582-601.
- 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.
- MOSKOWITZ, H. R. (1970). Ratio scales of sugar sweetness. *Perception & Psychophysics*, **7**, 315-320.
- MOSKOWITZ, H. R. (1971). Ratio scales of acid sourness. *Perception & Psychophysics*, **9**, 371-374.
- RULE, S. J., & CURTIS, D. W. (1980). Ordinal properties of subjective ratios and differences: Comment on Veit. *Journal of Experimental Psychology: General*, **109**, 296-300.
- RULE, S. J., & CURTIS, D. W. (1982). Levels of judgmental processing: Strategy for the evaluation of a model. In B. Wegener (Ed.), *Social attitudes and psychophysical measurement* (pp. 107-122). Hillsdale, NJ: Erlbaum.
- RULE, S. J., CURTIS, D. W., & MULLIN, L. C. (1981). Subjective ratios and differences in perceived heaviness. *Journal of Experimental Psychology: Human Perception & Performance*, **7**, 459-466.
- SCHNEIDER, B., PARKER, S., KANOW, G., & FARELL, G. (1976). The perceptual basis of loudness ratio judgments. *Perception & Psychophysics*, **19**, 309-320.
- 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.
- STEVENS, S. S. (1969). Sensory scales of taste intensity. *Perception & Psychophysics*, **6**, 302-307.
- STEVENS, S. S. (1975). *Psychophysics*. New York: Wiley.
- TORGERSON, W. S. (1961). Distance and ratios in psychophysical scaling. *Acta Psychologica*, **19**, 201-205.
- TREISMAN, M. (1964). Sensory scaling and the psychophysical law. *Quarterly Journal of Experimental Psychology*, **16**, 11-22.
- VEIT, C. T. (1978). Ratio and subtractive processes in psychophysical judgment. *Journal of Experimental Psychology: General*, **107**, 81-107.
- WEISS, D. J. (1972). Averaging: An empirical validity criterion for magnitude estimation. *Perception & Psychophysics*, **12**, 385-388.

(Manuscript received June 19, 1987;
revision accepted for publication April 25, 1988.)