Perception & Psychophysics
1979, Vol. 25 (2), 150-156

Notes and Comment

Reply to Eisler: On the subtractive theory of stimulus comparison

MICHAEL H. BIRNBAUM
University of Illinois, Urbana-Champaign, Illinois 61820

A series of experiments found that when subjects are instructed to judge "differences" or "ratios," the rank order of the responses is virtually identical.[1] Since actual ratios and differences of the same numbers are not monotonically related in suitable (factorial) designs, this finding has been interpreted as consistent with Torgerson's (1961) theory that subjects use one comparison operation for both tasks (Birnbaum, 1978; Birnbaum & Elmasian, 1977; Birnbaum & Mellers, 1978; Birnbaum & Veit, 1974a; Hagerty & Birnbaum, 1978; Rose & Birnbaum, 1975; Schneider, Parker, Farrel, & Kanow, 1976; Veit, 1978). Birnbaum (1978) has considered five theories of stimulus comparison and has shown that with four-stimulus tasks, one can distinguish theories that would be indistinguishable with two-stimulus experiments.

Birnbaum (1978) concluded that the Subtractive Theory gave the most coherent account of the data for six scaling tasks studied by Hagerty and Birnbaum (1978) and Veit (1978). Eisler (1978) challenged this conclusion and suggested two theories that permit one to retain the ratio model for "ratio" judgments by postulating an internal transformation that occurs only for "difference" judgments. Both theories succeed in explaining the results for five of the six tasks, but both make an incorrect prediction for one task. To make the Transformation Theories equivalent to the Subtractive Theory, Eisler proposes that subjects "reinterpret" the instructions for one task. The present article contends that, since the Subtractive Theory accounts for all of the results without such additional post hoc assumptions, it seems the preferred interpretation.

Figure 1 shows theoretical outlines of two- and four-stimulus judgment tasks. In the outline, Φ_i, Φ_j, Φ_k, and Φ_l, represent physical values of the stimuli to be compared or combined, s_i, s_j, s_k, and s_l represent the scale values of the stimuli, the psychophysical function, H, relates subjective to objective stimulus values. The function, $\Psi_{ij} = C(s_i,s_j)$, represents the process of combination or comparison, which relates the combined subjective impression (Ψ_{ij}) to the component scale values. In the four-stimulus tasks, δ_{ijkl} represents the comparison of two different stimulus combinations or comparisons; the function, $\delta = G(\Psi_{ij},\Psi_{kl})$, represents the process by which Ψ_{ij} and Ψ_{kl} are compared. The judgment function, J, represents the monotonic relationship between overt numerical responses and subjective impressions.

The principle of scale convergence (Birnbaum, 1974a, 1974b; Birnbaum & Veit, 1974a) assumes that the scale values, s_i, are independent of tasks to judge "ratios" or "differences," etc. The G function represents the process by which subjects compare two *comparisons* or *combinations* (e.g., "ratio of two differences"), whereas the C function represents the process by which subjects compare or combine two *stimuli*. It is nontrivial if the G function can be represented with the same models as those for C.

The judgment functions are assumed to be strictly monotonic, and they are assumed to depend lawfully upon such factors as the procedure for responding

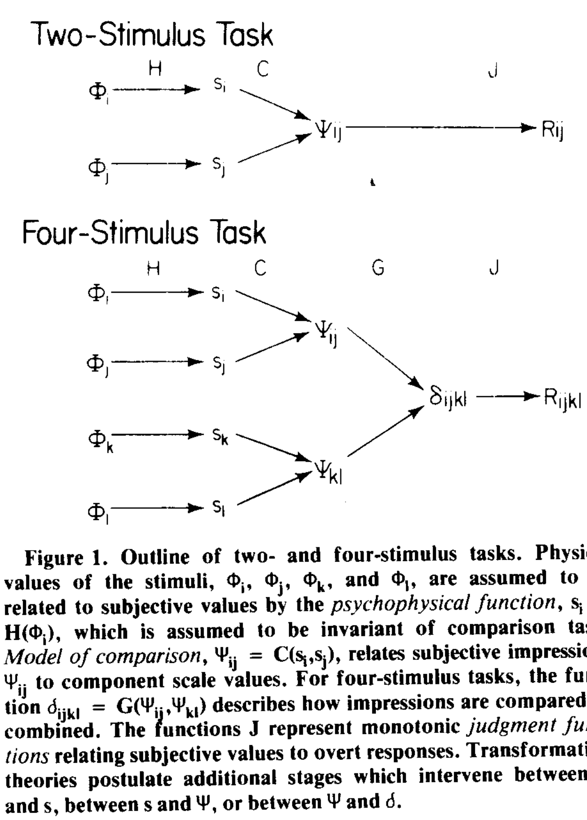

Figure 1. Outline of two- and four-stimulus tasks. Physical values of the stimuli, Φ_i, Φ_j, Φ_k, and Φ_l, are assumed to be related to subjective values by the *psychophysical function*, $s_i = H(\Phi_i)$, which is assumed to be invariant of comparison task. *Model of comparison*, $\Psi_{ij} = C(s_i,s_j)$, relates subjective impressions Ψ_{ij} to component scale values. For four-stimulus tasks, the function $\delta_{ijkl} = G(\Psi_{ij},\Psi_{kl})$ describes how impressions are compared or combined. The functions J represent monotonic *judgment functions* relating subjective values to overt responses. Transformation theories postulate additional stages which intervene between Φ and s, between s and Ψ, or between Ψ and δ.

1 thank Hannes Eisler for helpful comments on the manuscript and a stimulating exchange of ideas. Requests for reprints should be addressed to Michael H.Birnbaum, Department of Psychology, University of Illinois, Champaign, Illinois 61820.

0031-5117/79/020150-07$00.95/0

Table 1
Theories Discussed by Birnbaum (1978)

Task	Theory: Model = Task	Subtractive	Ratio	Indeterminate	Two Worlds
"Ratios"	A/B	A−B	A/B	A−B	a/b
"Differences"	A−B	A−B	A/B	A−B	A−B
"Ratios of Ratios"	(A/B)/(C/D)	(A−B)−(C−D)	(A/B)/(C/D)	(A−B)−(C−D)	(a/b)/(c/d)
"Differences of Ratios"	(A/B)−(C/D)	(A−B)−(C−D)	(A/B)−(C/D)	(A−B)−(C−D)	(a/b)−(c/d)
"Ratios of Differences"	(A−B)/(C−D)	(A−B)/(C−D)	(A/B)/(C/D)	(A−B)−(C−D)	(A−B)/(C−D)
"Differences of Differences"	(A−B)−(C−D)	(A−B)−(C−D)	(A/B)/(C/D)	(A−B)−(C−D)	(A−B)−(C−D)

Note—A, B, C, D refer to s_i, s_j, s_k, s_l in Figure 1, respectively. Each entry represents the model for each task predicted by each theory. Judgment functions are omitted for simplicity. For the Two Worlds theory, a = exp(A), b = exp(B), etc.

(e.g., category rating vs. magnitude estimation), stimulus distribution, individual differences, etc. Changes in the modulus of a magnitude estimation task, the number of categories in a rating task, the range or spacing of the stimuli will all affect the numbers reported to a given stimulus. Assuming that J depends on context helps to explain how a subject can convert internal values into numerical values. Birnbaum, Parducci, and Gifford (1971) have provided evidence that the stimulus distribution affects the J function.

The judgment function in Figure 1 seems the best way to represent the differences among category rating, magnitude estimation, production, and other methods for obtaining numerical responses from the judge. Experiments by Sarris and Heineken (1976) and by Weiss (1972) show that when ratings fit an additive model, magnitude estimations fit a multiplicative model. These findings are consistent with the theory that the magnitude estimation response procedure produces an exponential transformation (Birnbaum, 1978; Birnbaum & Veit, 1974a, 1974b).

Theories Considered by Birnbaum (1978)

Table 1 summarizes five theories considered by Birnbaum (1978) for a set of experiments in which subjects perform six scaling tasks. Each theory implies a set of models for the six tasks. Subjects can be asked to judge "ratios" and "differences" of two stimuli, and can also be given four tasks involving four stimuli: "ratios of ratios" [(A/B)/(C/D)], "differences of ratios" [(A/B) − (C/D)], "ratios of differences" [(A − B)/(C − D)], and "differences of differences [(A − B) − (C − D)]. For example, for the "ratio of differences" task, the judges were asked to judge the "ratio of two differences." If one assumes $G(\Psi_{ij},\Psi_{kl}) = \Psi_{ij}/\Psi_{kl}$ and $C(s_i,s_j) = s_i - s_j$; then $\delta_{ijkl} = (s_i - s_j)/(s_k - s_l)$. Letting $A = s_i$, $B = s_j$, $C = s_k$, and $D = s_l$, the first column of Table 1 shows that if the model is equivalent to the task, then "ratios of differences" can be represented by the ratio of differences model, (A − B)/(C − D).

If the model were the same as the task, as in the first column of Table 1, then the rank order of "differences" would be distinct from the rank order of "ratios," but the two rank orders would be related by the same scale values. The four rank orders for the four-stimulus tasks would also be distinct; each rank order should be consistent with its model. Birnbaum (1978) has noted that the ratio of differences model is a distributive polynomial, the difference of ratios model is dual distributive, and the others are additive. Therefore, nonmetric analyses (Krantz & Tversky, 1971) can be used to distinguish the models. Metric analyses can also be employed to test the theories, and scale convergence adds additional leverage to the comparisons among the theories (Birnbaum, 1978).

The Subtractive Theory, Ratio Theory, and Indeterminacy (one operation) Theory all conform to the assumption of scale convergence. The Subtractive Theory predicts that "ratios of ratios," "differences of ratios," and "differences of differences" should all have the same rank order, consistent with a difference of differences model. The Subtractive Theory can be distinguished from Ratio Theory and the Indeterminacy Theory, in that it predicts that "ratios of differences" should have a distinct rank order, consistent with a ratio of differences model. A "ratio of differences" task was studied by Hagerty and Birnbaum (1978) and by Veit (1978), who found evidence consistent with the Subtractive Theory.

The Subtractive Theory predicts that subjects can judge both differences and ratios *of stimulus differences*. A stimulus difference will have a well-defined zero point even when the stimuli are no more than an interval scale. Thus, the question, "What is the ratio of the distance from San Francisco to Philadelphia relative to the distance from San Francisco to Denver?" has a meaningful answer even though the positions of cities do not constitute a ratio scale. Similarly, the ratio of two temperature differences is independent of whether one uses Celsius or Farenheit scales, although ratios of temperatures on these scales would not be meaningful.

The Two Worlds Theory violates scale convergence by allowing two scales for "ratios" and "differences." However, scale convergence is assumed within each set involving a given task (and assumed model) for C in Figure 1. The theory predicts three distinct rank orders for the four-stimulus tasks; "differences of ratios" should not be monotonically related to "differences of differences," but should conform instead to a differences of ratios model. Only one experiment (Hagerty & Birnbaum, 1978) has investigated all four four-stimulus tasks.[2] Hagerty and Birnbaum (1978) found that "differences of differences" and "differences of ratios" were nearly identical and could both be fit to the difference of differences model, yielding very similar scale values. For this reason, Birnbaum (1978) noted that the complicated Two Worlds Theory was not required by the data of Hagerty and Birnbaum.

In sum, Table 1 presents five theories that make different ordinal predictions for six sets of data, if the values of A, B, C, and D are independently manipulated with a sufficient number of levels. Veit (1978) employed three tasks, "ratios," "differences," and "ratios of differences," and found data that would be compatible with either the Subtractive Theory or the Two Worlds Theory. Her "ratios of differences" data fit the ratio of differences model and could not be rescaled to fit another simple polynomial. Hagerty and Birnbaum (1978) employed all six tasks and found a set of results that were consistent with Subtractive Theory.

Eisler's Transformation Theories

Eisler (1978) has proposed that new stages be added to Figure 1 to allow additional transformations to intervene before C or G, in order to salvage the ratio model for "ratio" judgments. Marks (1978) has independently proposed a stage theory of auditory stimuli which can also be interpreted as a transformation theory.

Eisler (1978) discussed two theories in which "ratio" judgments can be represented by a ratio model. Neither theory is equivalent to the Subtractive Theory, as noted by the question mark in Figure 1 and Table 1 of Eisler (1978) and as discussed below.

The first theory assumes that there is only one comparison operation, ⊙, for "ratio" and "difference" tasks, and a monotonic transformation, T, which follows the operation only when the subject is instructed to judge "differences." The first column of Table 2 shows this representation. The theory is consistent with the Subtractive Theory (Table 1) except for the "differences of differences" task. Without specifying the operation or the transformation, it should be clear from the last two lines of Table 2 that the transformation theory predicts that "ratios of differences" and "differences of differences" should be monotonically related. In particular, if ⊙ is represented by division and T by the logarithmic function (assume $A > B$, $C > D$, and $A/B > C/D$), then the theory predicts that both "ratios of differences" and "differences of differences" can be represented by a distributive, ratio of differences model. The data of Hagerty and Birnbaum (1978) do not support these predictions.

In order to make the otherwise distinct theories (Subtractive Theory and the Transformation Theory) equivalent, Eisler (1978) argued that although a ratio of differences model will fit the "ratio of differences" task, subjects will "reinterpret" under "difference of differences" instructions. But, if the theory (Table 2) predicts that the rank orders for the two instructions are identical, why should the subject need to "reinterpret" under one instruction but not the other? Eisler (1978) notes that under the *ratio model* specification of the general theory, the argument of the logarithm may be zero or negative, and therefore would be undefined. To solve this problem, he proposed that "differences of differences" are reinterpreted as "differences of ratios." One can reasonably ask why the subject doesn't follow the theory and reinterpret the task as a "ratio of differences," as predicted by the theory? Even so, the reinterpretation argument seems a complex explanation for the fact that "difference of differences" judgments actually fit the difference of differences model.

Eisler concedes that the "reinterpretation" argument seems post hoc, but suggests that perhaps some evidence for it could be found by comparing the

Table 2
Transformation Theories

Task	Theory: One Operation (⊙); T follows ⊙ for "differences"	Two Operations (θ,ϕ); T precedes θ for "differences"
"Ratios"	A⊙B	$A\phi B$
"Differences"	T(A⊙B)	$T(A)\theta T(B)$
"Ratios of Ratios"	(A⊙B)⊙(C⊙D)	$(A\phi B)\phi(C\phi D)$
"Differences of Ratios"	T[(A⊙B)⊙(C⊙D)]	$T(A\phi B)\theta T(C\phi D)$
"Ratios of Differences"	T(A⊙B)⊙T(C⊙D)	$[T(A)\theta T(B)]\phi[T(C)\theta T(D)]$
"Differences of Differences"	T[T(A⊙B) ⊙ T(C⊙D)]	$T[T(A)\theta T(B)]\theta T[T(C)\theta T(D)]$

Note–The operation θ corresponds to the instruction to judge "differences." The operation ϕ corresponds to instruction to judge "ratios." The transformation, T, has the property that $A\phi B$ is monotonically related to $T(A)\theta T(B)$.

error terms for the "difference of differences" task and the "difference of ratios" task. If subjects reinterpret the "difference of differences" task, according to Eisler, the error variance for this task should be greater than that for the "difference of ratios" task. To check this possibility, Michael Hagerty conducted two separate ANOVAs on the data plotted in Figures 4 and 5 of Hagerty and Birnbaum (1978). Each analysis was a factorial, Replications (subjects and repetitions) by 48 Treatments (stimulus combinations), design. For the "difference of ratios" task, the sum of squares for the main effects of Treatments, Replications, and the Replications by Treatments interaction were 6,528, 2,348, and 8,458, respectively. For the "difference of differences" task, they were quite similar: 7,027, 2,244, and 8,840, respectively. Standard deviations for individual cells were comparable and did not not appear to differ systematically between the two tasks. Thus, these data show that the two tasks have comparable error variances and thus provide no evidence for the "reinterpretation" idea required to remove the question mark from Figure 1 and Table 1 of Eisler (1978) (see Footnote 2).

The second theory (Column 2 of Table 2) is that there are two operations corresponding to the two tasks and a monotonic transformation, T, which *precedes* the "difference" operation. This theory is similar to ideas proposed by Marks (1974, 1978). Although this theory has two scales and two operations in common with the two worlds view, it is distinct from the Two Worlds Theory (see Table 1). Since "ratios" and "differences" are assumed to be monotonically related, the operations $\emptyset$ (for "ratios") and θ (for "differences") are assumed to be interlocked by the transformation T, as follows:

$$x \emptyset y = M[T(x) \theta T(y)], \qquad (1)$$

where M is a monotone transformation. It follows that "ratios of differences" (RD) and "differences of differences" (DD) should be monotonically related. [Proof: Let $x = T(A) \theta T(B)$ and $y = T(C) \theta T(D)$; then $RD = x \emptyset y$ and $DD = T(x) \theta T(y)$; therefore, by Equation 1, $RD = M(DD)$].

In sum, both transformation theories in Table 2 can account for five of the six tasks but incorrectly predict that "differences of differences" and "ratios of differences" should be monotonically related, contrary to data (Hagerty & Birnbaum, 1978; Veit, 1978). Both transformation theories predict that "differences of differences" and "differences of ratios" would *not* in general by monotonically related. The data of Hagerty and Birnbaum (1978) suggest, instead, that they are linearly related with comparable error terms.

Two Worlds Revisited

Eisler (1978) has suggested that a case can be made for a modification of the Two Worlds Theory if it is allowed that the "difference" scale is a discriminability scale of its arguments. If it is allowed that the standard deviation of the "ratio" scale varies in proportion to s_i, and the standard deviation of $\Psi_{ij} = s_i/s_j$ in Figure 1 varies in proportion to Ψ_{ij} (Eisler, 1962a, 1962b, 1963, 1965), then the "difference" scale should be a log scale of "ratios" and "differences of ratios" can be represented by $\log(a/b) - \log(c/d)$, or $(A - B) - (C - D)$.

However, this theory predicts that when the standard deviation of sensation is independent of value (as it is postulated to be for so-called "metathetic" scales), then "differences" and "ratios" should produce two rank orders, consistent with two operations on a single scale. Birnbaum and Mellers (1978) showed that "ratios" and "differences" of easterliness and westerliness of the positions of U.S. cities are monotonically related. If position is supposed to be a "metathetic" continuum, in which discriminability is constant along the scale (Stevens & Galanter, 1957), the results of Birnbaum and Mellers (1978) do not support the modified Two Worlds Theory of Eisler (1978).

Subtractive Theory Explains Eisler's Evidence

Eisler (1978) cited a number of results in support of the ratio model. Subtractive Theory provides at least as good an account of these findings as the ratio model.

(1) Eisler and Ekman (1959) and Eisler (1960) asked subjects to judge "similarities" of stimuli and fit the model:

$$S_{ij} = 2 \frac{s_i}{s_i + s_j}, \qquad (2)$$

where S_{ij} is the judged "similarity" of the two stimuli, and s_i and s_j are the subjective scale values. If "ratio" judgments, R_{ij}, are assumed to be ratios, $R_{ij} = s_i/s_j$, then "similarities" ought to be a monotonic function of "ratios," but they should be *nonlinearly* related according to the equation:

$$S_{ij} = 2 \frac{R_{ij}}{1 + R_{ij}}. \qquad (3)$$

Data of Eisler (1960) show that "ratios" and "similarities" are nearly linearly related, contrary to the theory. Sjöberg (1971) has also shown, in a series of experiments, that the two kinds of judgments are approximately linearly related. Schneider, Parker, Valenti, Farrell, and Kanow (1978) obtained category ratings and magnitude estimations of stimulus "differences" and "similarities." For both loudness and pitch, they concluded that all four tasks can be represented by the subtractive model using a single set of scale values. In sum, the data of several experiments appear consistent with the hypothesis that "ratios," "differences," and "similarities" are governed by the same comparison operation, which can be represented by subtraction.

(2) Eisler (1962a, 1962b, 1963) has shown that if magnitude estimations are rescaled to have approximately equal variances, the rescaled values are more nearly linearly related to category ratings. Furthermore, such rescaled magnitude estimations of loudness will be more nearly linearly related to rescaled magnitude estimations of softness. Eisler has shown that if the standard deviation is a linear function of the mean magnitude estimation, then the log of a linear function of magnitude estimation will render the transformed standard deviations more nearly equal.

Although Eisler (1963) interprets the above relationships in terms of a subjective Weber's law (integrating the Weber function of magnitude estimations should yield the category scale), these findings can be reinterpreted as follows: (a) Both magnitude estimation and category ratings are types of category judgment tasks. (b) A variation of Thurstone's laws apply to both category judgment and magnitude estimation, with the same scale values and constant standard deviations for stimuli under both procedures.

Let P_{ik} be the cumulative probability that stimulus i is judged greater than or equal to category k in a category judgment experiment. Let Q_{ik} be the cumulative probability that stimulus i is judged greater than response value x_k in a magnitude estimation experiment. A simple possibility is as follows:

$$P_{ik} = F(s_i - t_k) \quad (4)$$

$$Q_{ik} = F(s_i - u_k), \quad (5)$$

where F is a distribution function (e.g., cumulative normal), s_i is the scale value of stimulus i, t_k is the boundary value for category limen k, and u_k is the subjective value of magnitude estimation x_k. Thus, the theory assumes that only the spacing of the response values differs for the two procedures (see Birnbaum & Veit, 1974a; Birnbaum, 1978).

For judgments of combinations, such as "differences" and "ratios" of pairs of stimuli, s_i could be replaced by Ψ_{ij} in Equations 4 and 5; however, it may prove necessary in this case to allow variation in the discriminal dispersions.

(3) Eisler (1978) referred to the use of an additive model to decide between Case V and Case VI of Thurstone's law of comparative judgment (Bock & Jones, 1968). The additive model could be better fit when using Case VI than Case V. However, serious doubts can be raised about the validity of the additive model for those studies (Shanteau, Note 1; Birnbaum, 1974).

Birnbaum (1974) found that ratings of "differences" in likeableness between two adjectives could be fit to a subtractive model, but judgments of the likeableness of a person described by the combination of the same two adjectives could not be fit to an additive model using the same scale values. The two rank orders for "combinations" and "differences" were each compatible with additive and subtractive representations; however, the two were not simultaneously compatible with these models using the same scale values. Birnbaum (1974, Experiment 4) introduced the scale-free approach to test the additive (or equivalent constant-weight averaging) model of impression formation (see also Birnbaum & Veit, 1974b). Subjects judged "differences between combinations." Suppose "differences" can be represented by subtraction, i.e., $DC_{ijkl} = J(\Psi_{ij} - \Psi_{kl})$, where DC_{ijkl} is the "difference in likeableness between a person described by the combination of adjectives i and j and another described by the combination of k and l." Suppose $\Psi_{ij} = s_i + s_j$ [or in this case, equivalently, $\Psi_{ij} = (w_0 s_0 + w_1 s_i + w_2 s_j)/(w_0 + w_1 + w_2)$], then $DC_{ijil} = J(s_i + s_j - s_i - s_l) = J(s_j - s_l)$, where stimulus i is in both combinations. Thus, "differences" should be a function of stimuli j and l only, and should be independent of stimulus i, a prediction that does not require the assumption that J is linear. Birnbaum (1974) found a systematic violation: if one adjective is low in likeableness, the difference due to the other adjective is less. For example, most subjects rate the difference in likeableness between Loyal and Understanding – Loyal and Obnoxious to be greater than the difference in likeableness between Malicious and Understanding – Malicious and Obnoxious, although the additive model predicts that both differences should be equal.

Suppose, however, that the "difference" is represented by a ratio model. Would this assumption allow one to represent impression formation with addition? The answer is negative, since if $a > (b,c) > d$ and $a/b > c/d$, then $a - b > c - d$. Thus, representation of "differences" with a ratio model cannot save the additive model for either impression formation (Birnbaum, 1974) or for the size-weight illusion (Birnbaum & Veit, 1974b).

Can one save the additive model by means of the present transformation theories? The answer is that, to explain the results of Birnbaum (1974), one must postulate a *positively accelerated* transformation which precedes category rating or subtraction, *contrary* to the logarithmic T postulated by Eisler for the same process.

It seems reasonable to ask whether subjects in psychophysical tasks can judge both differences and ratios of "total intensities," i.e., $(A + B) - (C + D)$ and $(A + B)/(C + D)$. If so, it may be possible to recover a ratio scale of "totals," and if the "totals" obey an additive model, a ratio scale of subjective value can be derived. On the other hand, if the "totals" do not obey the consistency required by the additive model (as in Birnbaum, 1974), then the additive model of combination cannot be used as a lever for distinguishing ratio and subtractive models of stimulus comparison.

Concluding Comments

In summary, the transformation theories suggested by Eisler (1978) are not equivalent to the Subtractive Theory without complex "reinterpretation" arguments. The reinterpretation argument is deemed complex because it is not dictated by the theory (Table 2) but is added post hoc to change one of the theory's predictions. Of course, with enough reinterpretations and postulated internal transformations, any two theories can be made equivalent. Choice among theories would then become a matter of convenience, or simplicity. Some philosophers hold the view that one can never truly refute a theory, but rather render it so complicated that it loses all attraction.

In order to account for data currently available, the ratio model must be complicated indeed.

(1) The ratio model requires easterliness to be the reciprocal of westerliness and leads to two cognitive maps. The subtractive model yields a single cognitive map independent of the task and the direction (Birnbaum & Mellers, 1978).

(2) The ratio model implies a psychophysical function for numerals that is positively accelerating, contrary to scales derived by range-frequency theory and other methods. The subtractive model implies a negatively accelerated psychophysical function for numerals that agrees with previous findings (Rose & Birnbaum, 1975).

(3) The Ratio Theory (Table 1) does not give a coherent account of the results of the six tasks studied by Hagerty and Birnbaum (1978) and Veit (1978). Complex reinterpretation arguments are required to save the Transformation Theories (Table 2) that attempt to represent "ratio" judgments with the ratio model.

In sum, the theory that "ratio" judgments should be represented by a ratio model has reached the state of complexity for the results cited above where it seems preferable to represent "ratio" judgments by subtraction.

REFERENCE NOTE

1. Shanteau, J., & Phelps, R. *Things just don't add up: A reexamination of additive utility theory.* Human Information Processing Institute (KSU-HIPI) Report No. 75-9, Kansas State University, 1975.

REFERENCES

Birnbaum, M. H. The nonadditivity of personality impressions. *Journal of Experimental Psychology Monograph*, 1974, **102**, 543-561. (a)

Birnbaum, M. H. Using contextual effects to derive psychophysical scales. *Perception & Psychophysics*, 1974, **15**, 69-96. (b)

Birnbaum, M. H. Differences and ratios in psychological measurement. In N. J. Castellan & F. Restle (Eds.), *Cognitive theory* (Vol. 3). Hillsdale, N.J: Erlbaum, 1978.

Birnbaum, M. H., & Elmasian, R. Loudness "ratios" and "differences" involve the same psychophysical operation. *Perception & Psychophysics*, 1977, **22**, 383-391.

Birnbaum, M. H., & Mellers, B. A. Measurement and the mental map. *Perception & Psychophysics*, 1978, **23**, 403-408.

Birnbaum, M. H., Parducci, A., & Gifford, R. K. Contextual effects in information integration. *Journal of Experimental Psychology*, 1971, **88**, 158-170.

Birnbaum, M. H., & Veit, C. T. Scale convergence as a criterion for rescaling: Information integration with difference, ratio, and averaging tasks. *Perception & Psychophysics*, 1974, **15**, 7-15. (a)

Birnbaum, M. H., & Veit, C. T. Scale-free tests of an additive model for the size-weight illusion. *Perception & Psychophysics*, 1974, **16**, 276-282. (b)

Bock, R. D., & Jones, L. V. *The measurement and prediction of judgment and choice.* San Francisco: Holden-Day, 1968.

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

Eisler, H. On the problem of category scales in psychophysics. *Scandinavian Journal of Psychology*, 1962, **3**, 81-87. (a)

Eisler, H. Empirical test of a model relating magnitude and category scales. *Scandinavian Journal of Psychology*, 1962, **3**, 88-96. (b)

Eisler, H. Magnitude scales, category scales, and Fechnerian integration. *Psychological Review*, 1963, **70**, 243-253.

Eisler, H. The connection between magnitude and discrimination scales and direct and indirect scaling methods. *Psychometrika*, 1965, **30**, 271-289.

Eisler, H. On the ability to estimate differences. A note on Birnbaum's subtractive model. *Perception & Psychophysics*, 1978, **24**, 185-189.

Eisler, H., & Ekman, G. A mechanism of subjective similarity. *Acta Psychologica*, 1959, **16**, 1-10.

Hagerty, M., & Birnbaum, M. H. Nonmetric tests of ratio vs. subtractive theories of stimulus comparison. *Perception & Psychophysics*, 1978, **24**, 121-129.

Krantz, D. H., & Tversky, A. Conjoint measurement analysis of composition rules in psychology. *Psychological Review*, 1971, **78**, 151-169.

Marks, L. E. On scales of sensation: Prolegomena to any future psychophysics that will be able to come forth as a science. *Perception & Psychophysics*, 1974, **16**, 358-376.

Marks, L. E. PHONION: Translation and annotations concerning loudness scales and the processing of auditory intensity. In N. J. Castellan, Jr., & F. Restle (Eds.), *Cognitive theory* (Vol. 3). Hillsdale, N.J: Erlbaum, 1978.

Rose, B. J., & Birnbaum, M. H. Judgments of differences and ratios of numerals. *Perception & Psychophysics*, 1975, **18**, 194-200.

Sarris, V., & Heineken, E. An experimental test of two mathematical models applied to the size-weight illusion. *Journal of Experimental Psychology: Perception and Performance*, 1976, **2**, 295-298.

Schneider, B., Parker, S., Farrell, G., & Kanow, G. The perceptual basis of loudness ratio judgments. *Perception & Psychophysics*, 1976, **19**, 309-320.

Schneider, B., Parker, S., Valenti, M., Farrell, G., & Kanow, G. Response bias in category and magnitude estimation of difference and similarity for loudness and pitch. *Journal of Experimental Psychology: Human Perception and Performance*, 1978, **4**, 483-496.

Sjöberg, L. Three models for the analysis of subjective ratios. *Scandinavian Journal of Psychology*, 1971, **12**, 217-240.

Stevens, S. S., & Galanter, E. H. Ratio scales and category scales for a dozen perceptual continua. *Journal of Experimental Psychology*, 1957, **54**, 377-411.

Torgerson, W. S. Distances and ratios in psychological scaling. *Acta Psychologica*, 1961, **19**, 201-205.

VEIT, C. T. Ratio and subtractive processes in psychophysical judgment. *Journal of Experimental Psychology: General*, 1978, **107**, 81-107.

WEISS, D. J. Averaging: An empirical validity criterion for magnitude estimation. *Perception & Psychophysics*, 1972, **12**, 385-388.

NOTES

1. Quotation marks are used to denote instructions to judge "ratios" or "differences" or for numbers obtained with these instructions. Quotations are not used for ratio or difference models or for theoretical statements about actual ratios or differences. Theories that make predictions for a set of experiments (Subtractive Theory, Two Worlds Theory) are capitalized; models of a single set of data are not.

2. In this discussion, great weight is placed on the results of Hagerty and Birnbaum (1978) and Veit (1978), who employed four-stimulus tasks. It would be highly desirable to see these experiments replicated and extended with a greater number of levels of the stimuli, with other continua, and in other laboratories.

(Received for publication and accepted November 7, 1978.)