*Perception & Psychophysics*
*1974, Vol. 16 (2), 276-282*

# Scale-free tests of an additive model for the size-weight illusion*

MICHAEL H. BIRNBAUM†
*Kansas State University, Manhattan, Kansas 66506*

and

CLAIRICE T. VEIT
*University of California, Los Angeles, Los Angeles, California 90024*

In two experiments, Ss rated the difference in heaviness between two objects varying in both size and weight. Assumption of the subtractive model and the use of factorial designs allow separation of judgmental effects from psychophysical processes. Difference ratings were rescaled by monotone transformation to fit the subtractive model, yielding scale-free values for the size-weight combinations. The subtractive model provided a good description of the difference ratings, but critical violations of the additive model for the size-weight illusion were obtained. The experiments illustrate how ordinal information can be used to differentiate additive from multiplicative processes.

"Size-weight illusion" refers to the fact that the subjective heaviness of an object depends upon its size as well as its weight. The smaller the object, holding physical weight constant, the heavier it feels when lifted. Although both *weigh* the same, a pound of lead does *feel* heavier than a pound of feathers.

## Theories of the Size-Weight Illusion

Two general theoretical interpretations have led to two alternative models of the illusion. The *expectancy* interpretation (e.g., Anderson, 1970) assumes that judgments of heaviness reflect the contrast between felt weight and the expected weight based on size. Since larger objects would be expected to be heavier, they seem lighter.

The expectancy theory led to an *additive* model, which can be written:

$$h = s - s^*, \qquad (1)$$

where h is the heaviness of the object, s reflects the subjective heaviness due to physical weight apart from the effect of size, and s* represents the effect of size on heaviness. According to this model, the magnitude of the illusion (the effect of size) is independent of weight.

The alternative interpretation is that subjective heaviness is a judgment of *density*, i.e., the ratio of perceived weight to size of an object (e.g., Huang, 1945). The larger an object (holding weight constant), the less the density; hence, the lighter it feels.

The *ratio* model (Sjöberg, 1969) predicts that the illusion (the effect of size) should be directly proportional to weight:

$$h = s/s^*, \qquad (2)$$

where h, s, and s* are defined as above.

## Scale-Dependent Tests

Although the models give quite different quantitative accounts of the illusions, both models have received some support from experimental studies. Anderson (1970, 1972) observed that ratings of heaviness appear to satisfy the parallelism prediction of the additive model. J. C. Stevens and Rubin (1970) and Sjöberg (1969), using magnitude estimations, found that the apparent effect of size increased as weight increased, possibly indicative of a ratio model.

These seemingly contradictory results may be due to the nonlinear relation between ratings and magnitude estimations. If magnitude estimations and ratings are exponentially related, and if $R(h) = s - s^*$, then $M(h) = \exp(s - s^*) = e^s/e^{s^*}$, where R(h) represents the rating of heaviness and M(h) represents the magnitude estimate. Thus, if ratings satisfy an additive model, magnitude estimations would be

*The first author received support from a National Institute of Mental Health postdoctoral fellowship at the Center for Human Information Processing, University of California, San Diego, which provided additional assistance through Grant MH 15828. Further support was provided by National Science Foundation Grant GB-21028; computing funds were supplied by Campus Computing Network, University of California, Los Angeles. We thank James C. Shanteau for helpful comments on the manuscript, and we thank Dwight R. Riskey and Valerie P. Hans for assistance.

†Requests for reprints should be addressed to Michael H. Birnbaum, who is now at the Department of Psychology, University of Illinois at Urbana-Champaign, Champaign, Illinois 61820.

expected to fit a multiplicative model (Birnbaum & Veit, 1974). Anderson (1972) assumed the additive model, and concluded that magnitude estimations "must be biased and invalid," since magnitude estimation data show a divergent interaction. If one were to assume the ratio model, however, then the parallelism obtained by Anderson (1970, 1972) might be taken as evidence against the rating scale.

The basic problem is that unless one response procedure—ratings or magnitude estimations—is assumed to be valid, the previous research cannot test between Eqs. 1 and 2. Unless one can assume one model or the other, neither response procedure can be validated. The situation is circular. Because the ordinal predictions of additive and multiplicative models are equivalent (Krantz, Luce, Suppes, & Tversky, 1971), Eqs. 1 and 2 cannot in principle be differentiated by the ordinal information in a simple experiment without some further constraint. The previous tests can be called "scale-dependent," since the conclusions depend upon the choice of the dependent variable.

### Scale-Free Tests

This paper applies new scale-free techniques that have recently been applied in person perception to separate cognitive processes from judgmental effects (Birnbaum, 1972, 1974). This approach can separate multiplicative from additive models, for example, without having to assume that the dependent variable is an interval measure. The key idea is to gain greater ordinal constraint by embedding the process to be studied within a simpler process.

In the present experiment, Ss judge the *difference* in heaviness between two objects varying in both size and weight. This experiment requires only the minimal assumption that judged differences are an ordinal scale of subjective differences:

$$R_{ij} = J(h_i - h_j), \tag{3}$$

where $R_{ij}$ is the overt response, $h_i$ and $h_j$ are the heavinesses of the objects, and J is the monotone function relating subjective differences to overt responses. The assumption of Eq. 3 allows separate estimation of J and of the heavinesses, $h_i$, unique to an interval scale. Once the scale-free values of h are determined, it is a simple matter to test whether they satisfy Eq. 1 or Eq. 2.

Two important ordinal implications of the additive model of the size-weight illusion can be tested through the assumption of Eq. 3. The first implication of the additive model is that differences in heaviness between two same-sized objects should depend only on their weights, independent of size. This follows from Eqs. 1 and 3, since the constant size, s*, drops out of the equation; i.e., $R = J[(s_1 - s^*) - (s_2 - s^*)] = J[s_1 - s_2]$. The ratio model predicts that these differences should be inversely related to size since $R = J[s_1/s^* - s_2/s^*] = J[(s_1 - s_2)/s^*]$. A second implication of the additive model is that differences betwen two equal-weight objects should depend only on their sizes, independent of weight. The ratio model predicts these differences should be a monotonically increasing function of weight.

## METHOD

### Experiment 1

The Ss were run individually, seated at a table, separated from E by a screen. On each trial, E placed two objects before S, who lifted them simultaneously, one in each hand, and rated the difference in heaviness. Ratings were made on a 9-point scale from 1 (right-hand object is very, very much lighter) to 9 (right-hand object is very, very much heavier), with 5 designated "no difference."

**Stimuli.** The stimuli were black, plastic blocks, approximately cubical in shape. They were weighted inside with lead and clay. An 8-mm hook was mounted on the top for lifting. Nine different objects were constructed from a 3 by 3, Size by Weight, factorial design. The three levels of weight were 100, 150, and 200 g. The block dimensions for the three levels of size were 50 x 50 x 55, 62 x 62 x 70, and 78 x 78 x 87 mm in width, length, and height, respectively.

**Design.** The objects presented to the two hands formed a 9 by 9, Left Hand by Right Hand, symmetric factorial design, in which the objects in either hand could be one of the 9 (i.e., 3 by 3) size-weight combinations. The overall design can thus be expressed as a (3 by 3) by (3 by 3), Left Hand (Weight by Size) by Right Hand (Weight by Size), factorial design.

**Procedure.** Following a warm-up of representative practice trials, the 81 presentations were presented in random order, with a different random order for the second replicate.

**Subjects.** The Ss were 20 University of California, Los Angeles undergraduates fulfilling a requirement in introductory psychology.

### Experiment 2

As in Experiment 1, Ss lifted two blocks simultaneously, one in each hand, and rated the difference in heaviness. The chief difference between the experiments was the design and choice of stimulus levels. Experiment 2 used a -9-to-+9 scale in which the even-numbered points were given category labels varying from -8 (right-hand object is very, very much lighter) to +8 (right-hand object is very, very much heavier) centered at 0 (equal in heaviness).

**Stimuli.** The basic set of blocks were constructed from a 7 by 3, Weight by Size, factorial design in which the seven levels of weight were 50, 75, 100, 150, 200, 300, and 400 g. The three levels of size and other aspects of the stimuli were identical to those of Experiment 1.

**Design.** The basic design was a 4 by 21, Standard by Comparison, factorial design, in which the S judged the difference in heaviness between the Standard and Comparison objects presented to the two hands. The standard object could be 50 g in the smallest size, 100 g in the medium size, 200 g in the medium size, or 400 g in the largest size block. The 21 comparison objects were generated from the 7 by 3, Weight by Size, factorial design described above. The design was enlarged by allowing the standard object to be in the left and right hands; in addition, there were 2 replications of this entire design. The complete design can thus be conceptualized as a 2 by 2 by 4 by (7 by 3), Replicate by Hand by Standard by Comparison (Weight by Size) factorial design.

An additional set of eight trials per hand-replicate were produced by pairing standards of 50 g in the largest size and 400 g in the smallest size with comparisons of 50 and 400 g in the smallest size and 50 and 400 g in the largest size.

**Procedure.** The 184 trials of each replicate were separately randomized for each S and session by shuffling a deck of cards that represented the trials. Following a warm-up of 20 representative trials, the two sessions were run with a 10-min rest between sessions, requiring about 3 h to complete.

**Subjects.** The Ss were 16 University of California, San Diego undergraduates, half of either sex, who were paid $1.88/h for their services.

## RESULTS

### Experiment 1

**Subtractive Model for Difference Judgment.** Figure 1 plots the mean judgment of difference as a function of the marginal means for the right-hand object. Separate curves are for different left-hand objects. Since the difference ratings represent right minus left, the curves have positive slope.

If the subtractive model (Eq. 3) is valid *and if* the judgment function, J, is linear, the curves should be parallel except for error. As can be seen in the figure, the curves are very nearly parallel; i.e., the slopes of the curves are nearly equal.

Analysis of variance measures the nonparallelism of the Right by Left interaction, which was small in magnitude and nonsignificant, $F(64,1216) = 1.31$. The nonsignificance of this interaction, based on many degrees of freedom, and the apparent parallelism of the curves can be taken as support for the subtractive model with an interval response measure.

**Size-Weight Illusion.** Figure 2 plots the heaviness scale values for the 9 right-hand objects as a function of size, with a separate curve for each weight. Since the subtractive model fits, the subjective values are the marginal means, each point representing the mean of 360 judgments, averaged over 20 Ss, 2 replicates, and 9 left-hand objects. If there were no illusion, the curves in Fig. 2 would be horizontal. The figure illustrates the usual illusion—the smaller the size, the greater the heaviness. The large effects of size are of course statistically significant, $F(2,38) = 110.40$; they are also significant for left-hand objects, $F(2,38) = 142.88$.

The additive model for the size-weight illusion predicts that the slopes of the curves in Fig. 2 (the illusion) should be independent of weight; that is, the curves should be parallel. The curves show a small but very regular divergence—the effect of size is greater for the 200-g weights than for the 100-g weights. The interaction is statistically significant, $F(4,76) = 2.57$; it is also significant for left-hand objects, $F(4,76) = 3.34$.

**Scale-Free Tests of Additivity.** These interactions cannot be attributed to nonlinearity of the rating scale. Critical difference ratings illustrate ordinal violations of the additive model. According to Eq. 3, assuming only that J is monotone, the rating of difference in heaviness between a 200-g object and a 100-g object of the same size should be independent of the size under the additive model, since $R = J[(s_{200} - s^*) - (s_{100} - s^*)] = J[s_{200} - s_{100}]$. The open circles in Fig. 1 connected by dashed lines illustrate the type of ordinal violation of the additive model that would be predicted by the ratio model. The ratio model implies that these ratings be inversely related to size, since $R = J[s_{200}/s^* - s_{100}/s^*] = J[(s_{200} - s_{100})/s^*]$. The dashed curve shows that ratings of the difference in heaviness between 200- and 100-g objects of the same size increase with decreasing size.

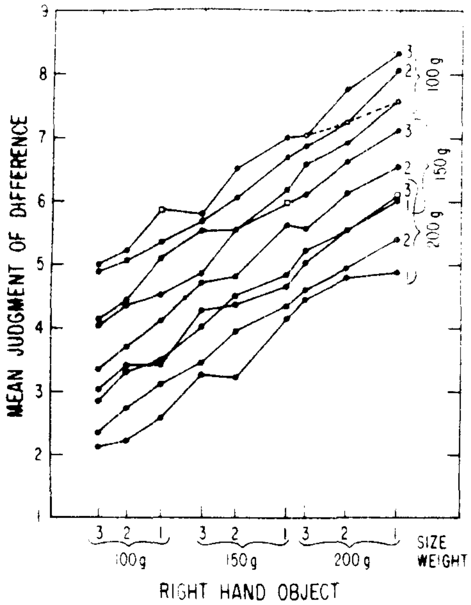


**Fig. 1. Mean ratings of difference in heaviness, plotted as a function of the heaviness of the right-hand object. Each curve represents a different size-weight object in the left hand. Abscissa values for the right-hand objects are spaced according to the marginal means. Open circles and dashed curve illustrate critical comparisons discussed in text (Experiment 1).**

To assess these ordinal violations, a 3 by 3, Size by Weight-Difference, subdesign was assessed by analysis of variance (after reversing the scale for left-hand objects). According to the additive model, there should be no main effect of size. Instead, the effect of size was highly significant, $F(2,38) = 10.91$. Consistent with the divergent interaction shown in Fig. 2, the greater the size, the less the effect of the weight difference.

A similar 3 by 3, Weight by Size Difference, subdesign was also analyzed in similar fashion. The additive model predicts that the effect of size

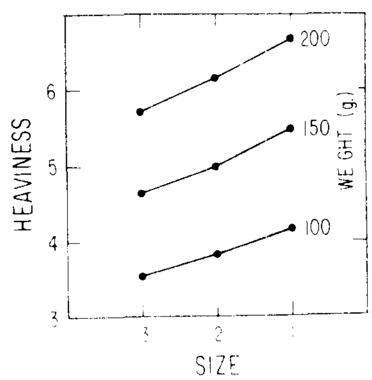


**Fig. 2. Heaviness values (marginal means) for the size-weight illusion, plotted as a function of size, with a separate curve for each level of weight (Experiment 1).**

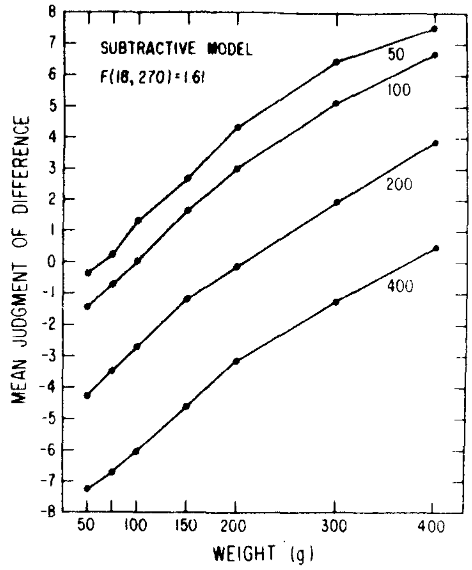


**Fig. 3. Mean ratings of difference in heaviness, plotted as a function of the weight of the comparison with a separate curve for each standard. Parallelism of the curves supports the subtractive model of difference judgment (Experiment 2).**

difference should be independent of weight, whereas the ratio model would predict that it should be directly proportional to weight. An example of this comparison is illustrated by the three open squares in Fig. 1, which show that the difference in heaviness between the smallest and largest size is slightly greater for 200 g than for 100 g. The results for these comparisons were not as clear, however, and the effects of weight were nonsignificant, $F(2,38) = 1.62$. This statistical nonsignificance may be due to the fact that the variation in weight was only 100 g in Experiment 1.

In summary, Experiment 1 provides support for the subtractive model with a linear judgment function. Although the traditional size-weight illusion was observed, it did not conform to the additive model. Instead, the illusion indicates a small, but regular,

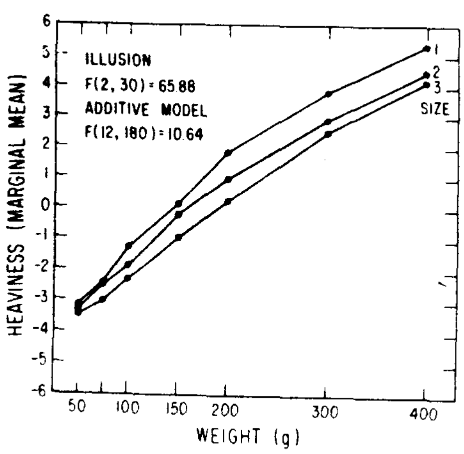


**Fig. 4. Heaviness values for the size-weight combinations, as a function of weight with a separate curve for each size. Nonparallelism violates the additive model of the size-weight illusion (Experiment 2).**

divergent interaction in the direction that would be predicted by the ratio model. The Size by Weight interaction cannot be attributed to the rating scale because it is confirmed by critical ordinal violations of additivity.

**Experiment 2**

**Subtractive Model for Difference Judgment.** Figure 3 plots the mean judgment of difference in heaviness as a function of the weight of the comparison object with a separate curve for each standard object. Each curve has been averaged across the three sizes for the comparison object. The scale was reversed by reflecting the signs of the responses for that part of the design in which the standard object was in the right hand. The corresponding figures, drawn separately for each hand, were very similar to those of Fig. 3.

According to the subtractive model, with the assumption of a linear judgment function, the curves in Fig. 3 should be parallel. The curves appear nearly parallel, and the analysis of variance test for interaction was nonsignificant, $F(18,270) = 1.61$. However, this is only part of the test of the subtractive model. The test of the entire 4 by 21, Standard by Comparison, design indicated small, but statistically reliable, deviations, $F(60,900) = 2.84$. Part of the problem appears to be due to a slight nonlinearity in the J function, possibly reflecting a slight tendency for some Ss to minimize small differences, as though they would respond "0" to difficult comparisons. Overall, the subtractive model appears to provide a reasonable account of the mean ratings.

**Size-Weight Illusion.** Figure 4 shows the heaviness values for the 21 size-weight comparison objects. Each point is the mean of 256 judgments, averaged over 16 Ss, 4 standards, 2 hands, and 2 replications. The heaviness values are plotted as a function of weight, with a separate curve for each size. The distances between the curves represent the illusion, which is very large and, of course, statistically significant, $F(2,30) = 65.88$.

The additive model of the size-weight illusion predicts parallel curves for Fig. 4. The curves are clearly nonparallel, showing a marked divergence from 50 to 200 g. The interaction is statistically significant, $F(12,180) = 10.64$. The data, plotted separately for each hand, for each standard, and for each replicate, showed similar divergent interactions. Figure 4 supports the findings of Experiment 1, providing further evidence against the additive model.

The data were plotted as in Fig. 4 separately for each S. A separate analysis of variance was also performed for each S. Of the 16 Ss, 15 showed significant effects of size, 13 showed divergent size-weight interactions that were similar in form and magnitude to Fig. 4.

The curves for the single-S counterparts of Fig. 3 were negatively accelerated for 13 of 16 Ss, and they were approximately parallel for 10 of 16 Ss. Three Ss showed a slight tendency to respond "0" to small differences, causing the curves to be flatter near an ordinate value of 0. An equal number of Ss showed slightly steeper slopes near the ordinate value of 0, apparently reflecting the opposite response bias, to exaggerate small differences. These two analyses showed that the majority of single-S data very closely resembled the group averages.

**Scale-Free Tests**. To remove any residual nonlinearity in the J function, each S's raw data was separately fit to the subtractive model (Eq. 3) using MONANOVA, a computer program (Kruskal & Carmone, 1969) based on Kruskal's (1965) monotone rescaling procedure. Since the subtractive model did quite well without transformation, it was not surprising that the J functions derived from MONANOVA were nearly linear for most Ss, with slight cubic (S-shaped) trends for a few of the Ss. The slight cubic transformations were interpreted as correction of response biases (to exaggerate or minimize small differences). From the fit of the 4 by 21 subtractive model, MONANOVA yielded scale-free heaviness values for the 21 (3 by 7) size-weight comparison objects for each S.

These scale-free values were then subjected to analysis of variance, giving results similar to those for the raw ratings. Figure 5 plots the mean of the scale-free heaviness values. The results are very similar to those of Fig. 4. The effect of size was again significant, $F(2,30) = 74.84$, as was the Size by Weight interaction, $F(12,180) = 8.40$. The fact that the Size by Weight interaction remained divergent and statistically significant shows that the violations of the additive model cannot be attributed to nonlinearity in the rating scale. Although this analysis gives the same conclusions as the analysis of the untransformed ratings, it illustrates how the subtractive model can be used as the basis for response rescaling, allowing a scale-free test of a substantive model.

Figures 4 and 5 show a mild divergent interaction that would be inconsistent with the additive model. Divergence would be consistent with the ratio model, but the curves also show a mild tendency to reconverge from 200 to 400 g, which would be inconsistent with the ratio model. The heaviness function for the smaller object appears to have more curvature than that for the larger object. This effect is analogous to Stevens and Rubin's (1970) finding of different power function exponents for objects of different size.

**Scales for Weight and Size**. Figure 6 shows the marginal effects of size and weight for each S derived from the scale-free MONANOVA values. The effects of weight are very similar (negatively accelerated), for all but three of the Ss. Under the assumption of either an additive or ratio model for the size-weight illusion, these marginal effects of weight would represent the psychophysical functions for heaviness. However, because of the shortcomings of both models, these

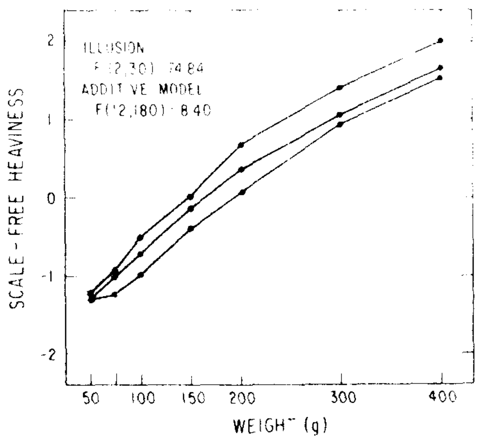


**Fig. 5. Scale-free values for the size-weight combinations, plotted as in Fig. 4, derived from fit of the subtractive model (Experiment 2).**

scales should be limited in generality to the sizes studied. Most single-Ss also show a moderate effect of size, as shown by the vertically spaced points.

In summary, Experiment 2 provides some highly consistent data that indicate a divergent size-weight interaction, contrary to the additive model. Evidence against the ratio model is provided by the finding that the heaviness functions are more concave downwards for smaller sized objects. The subtractive model for difference judgments appeared to provide a nice account of the difference ratings, in spite of small deviations that may be attributable to slight nonlinearity in the rating scale. When the data are transformed to fit the subtractive model, the size-weight interactions remain. Different Ss showed similar size-weight interactions with similar negatively accelerated scales for heaviness.

## DISCUSSION

These experiments suggest several conclusions: (1) The subtractive model appears to give a good fit to

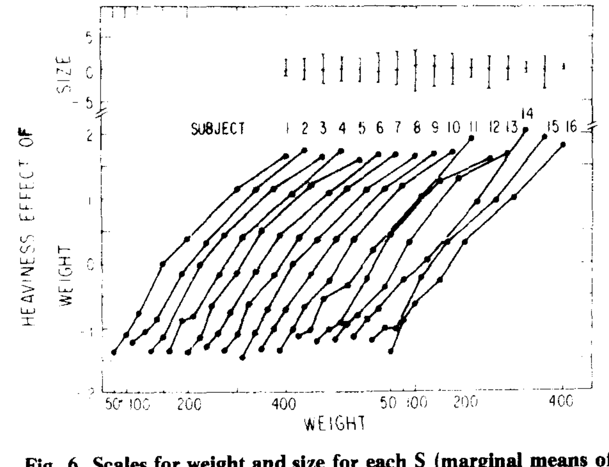


**Fig. 6. Scales for weight and size for each S (marginal means of transformed ratings). Abscissa values have been shifted for each S (Experiment 2).**

difference ratings; (2) this good fit can be taken as support for the linearity of category ratings; (3) the assumption of the subtractive model permits a scale-free test between additive and ratio formulations of the size-weight illusion; (4) the data contain ordinal violations of the additive model of the size-weight illusion—the illusion (the effect of size) appears to be smaller for lighter weights. The implications of these conclusions for measurement, model testing, and the perception of heaviness are discussed below.

**Subtractive Model**

The assumption of the subtractive model, Eq. 3, is the basis for the present scale-free tests of Eqs. 1 and 2. These experiments also provide tests of internal consistency for the subtrative model.[1] Although the present experiments only require ordinal consistency, the near parallelism of Figs. 1 and 3 shows that the data are also "metrically" consistent with Eq. 3 under the assumption that ratings are an interval response measure (i.e., that J in Eq. 3 is linear).

The subtractive model has also had a fair degree of success in previous research with ratings of differences in heaviness (Birnbaum & Veit, 1974), differences in likeableness (Birnbaum, 1974), and preferences for foods (Shanteau & Anderson, 1969). Curtis, Rule, and their associates have obtained consistent results for heaviness using the subtractive model as a scaling framework (summarized in Rule & Curtis, 1973).

A final assumption of the present scale-free approach deserves brief comment. It has been implicitly assumed that the procedure of lifting two blocks and judging the difference does not affect the process by which size and weight combine to form heaviness. This nonreactive assumption seems reasonable for the present situation and is supported by the overall success of the subtractive model and the comparability of the findings with the results for simple ratings of single blocks. However, there are probably situations in which the process of comparing multidimensional stimuli that are similar on one or more dimensions may induce cancellation or other short-cut strategies that would disrupt the information integration process (e.g., Tversky, 1969). Such effects would show up as violations of the subtractive model and would preclude use of the present approach.

**Size-Weight Illusion**

Contrary to the additive model of the size-weight illusion, the magnitude of the illusion depends on weight. The divergent interactions in Figs. 2, 4, and 5 show systematic deviations from additivity: Lighter weights are less affected by the variation in size. These violations of additivity cannot be attributed to nonlinearity in the rating scale, since Experiment 1 shows that direct ratings of heaviness differences depend on size, and Experiment 2 shows that even after ratings are transformed to fit the subtractive model of difference judgment, the Size by Weight interaction remains.

Although divergent interactions would be consistent with a ratio model, the data of Experiment 2 (Figs. 4 and 5) also contain some evidence against the ratio model, since the curve for the smaller block exhibits slightly more curvature. Both additive and ratio models imply that the psychophysical relationship between heaviness and weight be independent of size. The present results suggest that the psychophysical function may depend on size.

The relationship between heaviness and weight is negatively accelerated for all three sizes. Negatively accelerated psychophysical scales have been obtained in previous studies using additive and subtractive models (Anderson, 1972; Birnbaum & Veit, 1974; Rule & Curtis, 1973). However, magnitude estimations of heaviness have typically been fit with positively accelerated functions (Stevens & Galanter, 1957). Research based on additive and subtractive models has generally supported rating scales, indicating that magnitude estimation may induce a positively accelerated judgmental transformation that confounds the study of psychophysical processes (Birnbaum & Veit, 1974).

Anderson (1972) obtained ratings of size-weight combinations that appeared nearly parallel. Size and weight may be nearly additive for the range of sizes and weights employed by Anderson (1972); however, it is also possible that a slight nonlinearity in the rating scale, possibly produced by the extreme "anchor" stimuli used in that study, may have distorted the results. Birnbaum, Parducci, and Gifford (1971) have shown that nonlinearity in rating scales can be produced by varying the stimulus distribution. The advantage of the present approach is that it requires only that the dependent variable be a monotone function of subjective differences to test the model.[2]

Although these results pose problems for the additive model, they would not seem to seriously damage the expectancy theory of the illusion. As anyone who has lifted an empty milk carton thinking it full will testify, expectancy plays an important role in heaviness judgment. Research with an analogous size-numerosity illusion (Birnbaum & Veit, 1973; Birnbaum, Kobernick, & Veit, 1974) indicates that manipulation of the expectancies (by varying the size-numerosity correlation) can lead to reversal of the illusion. The violations of the additive model in the present study may indicate that other factors (in addition to expectancies) are also at work in the size-weight illusion, or that size-based expectancies do not combine with weight additively.

One possibility is that each weight is judged in part with respect to the expected distribution of weights for its size. Instead of contrasting felt weight with the *mean* expected weight, felt weight is compared with the entire distribution of expectancies. If this process of comparison follows Parducci's range-frequency theory (Birnbaum, 1974; Parducci, 1974), then the greater the *range* of expected values, the less the effect of felt weight on heaviness; and the greater the median expected value, the less the judged heaviness. If objects of all sizes were made of all substances, then the median weight would be greater for larger objects and the range of weights would also be larger. This explanation could thus give a qualitative account of the present interactions. It could be tested by using systextual design such as that employed by Birnbaum and Veit (1973). Systematically varying the range of expectancies should affect the interaction.

## Conclusions

This study illustrates how additive and multiplicative models can be differentiated using ordinal information. The subtractive model for difference judgment seems to work quite well in accomplishing this goal. Results indicate that category ratings are nearly linearly related to subjective differences. The large size-weight illusions did not appear to conform to the additive model. Critical ordinal tests revealed that differences in heaviness due to size or weight, *ceteris paribus*, were not independent of weight or size, in violation of the additive model. But the negatively accelerated functions for heaviness showed slightly more curvature for smaller sizes, apparently inconsistent with the ratio model. These experiments illustrate a basis for distinguishing between simple alternatives, but like so many critical tests in psychology, the data are not wholly consistent with either simple model.

## REFERENCES

ANDERSON, N. H. Averaging model applied to the size-weight illusion. *Perception & Psychophysics*, 1970, **8**, 1-4.

ANDERSON, N. H. Cross-task validation of functional measurement. *Perception & Psychophysics*, 1972, **12**, 389-395.

BIRNBAUM, M. H. The nonadditivity of personality impressions. *Journal of Experimental Psychology Monograph*, 1974, **102**, 543-561.

BIRNBAUM, M. H. Using contextual effects to derive psychophysical scales. *Perception & Psychophysics*, 1974, **15**, 89-96.

BIRNBAUM, M. H., KOBERNICK, M., & VEIT, C. T. Subjective correlation and the size-numerosity illusion. *Journal of Experimental Psychology*, 1974, **102**, 537-539.

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. Judgmental illusion produced by contrast with expectancy. *Perception & Psychophysics*, 1973, **13**, 149-152.

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.

HUANG, I. The size-weight illusion and the "weight density illusion." *Journal of General Psychology*, 1945, **33**, 65-84.

KRANTZ, D. H., LUCE, R. D., SUPPES, R., & TVERSKY, A. *Foundations of measurement*. New York: Academic Press, 1971.

KRUSKAL, J. B. Analysis of factorial experiments by estimating monotone transformations of the data. *Journal of the Royal Statistical Society* [*B*], 1965, **27**, 251-263.

KRUSKAL, J. B., & CARMONE, F. J. MONANOVA: A FORTRAN IV program for monotone analysis of variance. *Behavioral Science*, 1969, **14**, 165-166.

PARDUCCI, A. Contextual effects: A range-frequency analysis. In E. C. Carterette and M. P. Friedman (Eds.), *Handbook of perception*. Vol. II. New York: Academic Press, 1974.

RULE, S. J., & CURTIS, D. W. Conjoint scaling of subjective number and weight. *Journal of Experimental Psychology*, 1973, **97**, 305-309.

SHANTEAU, J. C., & ANDERSON, N. H. Test of a conflict model for preference judgment. *Journal of Mathematical Psychology*, 1969, **6**, 312-325.

SJÖBERG, L. Sensation scales in the size-weight illusion. *Scandinavian Journal of Psychology*, 1969, **10**, 109-112.

STEVENS, J. C., & RUBIN, L. L. Psychophysical scales of apparent heaviness and the size-weight illusion. *Perception & Psychophysics*, 1970, **8**, 225-230.

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

TVERSKY, A. Intransitivity of preferences. *Psychological Review*, 1969, *76*, 31-48.

## NOTES

1. If one were to assume instead that "difference" ratings represented subjective ratios, then the near-parallelism of Figs. 1 and 3 would be taken as evidence that ratings are an exactly logarithmic function of impressions (Birnbaum & Veit, 1974). But the data would still refute the additive model of the size-weight illusion since Figs. 2, 4, and 5 would represent the log of heaviness. Exponential transformation of the ordinates of Figs. 2, 4, and 5 would *increase* the divergent interaction, making the fit of the additive model even worse. Under the interpretation that Ss are rating ratios, the nonparallelism in Figs. 2, 4, and 5 also constitute evidence against the ratio model of the size-weight illusion, since it implies parallelism in a logarithmic plot. In general, to assume that difference ratings represent some other arbitrary function would require an explanation of the parallelism of Figs. 1 and 3. It would be necessary to postulate a judgmental transformation that exactly compensates for the nonadditive composition function. The most simple interpretation seems to be that Ss are forming differences and that ratings of differences are nearly linear functions of subjective differences.

2. A more general form of averaging model could describe the present size-weight interaction by allowing mathematical weights to vary with size and weight. Lighter weights would have greater importance relative to size. This approach seems unattractive, however, since the mathematical weights that would be inferred from judgments of average heaviness would vary directly with weight (Birnbaum & Veit, 1974).

(Received for publication December 7, 1973;
revision received June 4, 1974.)