

Judgmental illusion produced by contrast with expectancy*

MICHAEL H. BIRNBAUM† and CLAIRICE T. VEIT
University of California, Los Angeles, Los Angeles, California 90024

Judgments of the numerosness of dots vary inversely with the size of the background on which the dots are presented. An expectancy-contrast model is proposed and tested in an experiment with this size-numerosity illusion. According to the model, the expectancy for numerosness depends upon the subjective correlation between size and numerosity; judgments of numerosness are assumed to reflect an additive contrast with the expectancy. In agreement with the model, the data demonstrate that the effects of size depend upon the correlation between size and numerosity; when this correlation is negative, the illusion is reversed. The model received tentative support from quantitative tests of fit, based on functional measurement procedures.

Certain perceptual illusions have been attributed to the contrast between the judged stimulus and the expectancy produced by unjudged background cues (e.g., Brunswik & Herma, 1951). The size-weight illusion, in which the same weight seems lighter when presented in a larger package, is a familiar example. The expectancy interpretation assumes that everyday experience establishes a positive subjective correlation between size and weight; the expectancy for heaviness is consequently greater for the larger package, and by contrast we judge it lighter.

The expectancy interpretation can be tested by experimental manipulation of the correlation between the judged and unjudged variables. In the case of the size-weight illusion, this manipulation would have to overcome powerful effects of past experience; however, the correlations presumably underlying other illusions might be less firmly fixed by past experience. The present tests are therefore directed to the investigation of a size-numerosity illusion, in which the judged numerosness of a pattern of dots varies inversely with the size of the background on which the dots are presented (see, e.g., Bevan & Turner, 1964). This illusion is assumed to depend upon the learned correlation between numerosity and background size.

THE MODEL

A formal representation of the expectancy interpretation includes two parts. The first, an additive-contrast model proposed by Anderson (1970b), describes how the expectancy produces judgmental contrast. For a size-numerosity application, this model can be written as follows:

$$J = N + a(N - N^*), \quad (1)$$

where J is the judged numerosness of dots, N is what the judgment would be apart from the background effect, a is a weighting factor expression the magnitude of the illusion, and N^* is the background-based "expectancy." Since N depends only on numerosity and N^* presumably depends only on size, Eq. 1 predicts no interaction between these two factors in an appropriate Numerosity by Size factorial test. This part of the model received tentative support in experimental tests of the size-weight illusion (Anderson, 1970b).

The second part of the model describes how the expectancy, N^* , is formed. The present proposal is that it depends upon the subjective correlation between numerosity and background size:

$$N^* = R_{NB}B, \quad (2)$$

where B is the psychological size and R_{NB} is the psychological correlation between numerosness and size. This subjective correlation is presumed to vary directly with the actual correlation between numerosity and size established by experimental manipulation. Because of its multiplicative form, Eq. 2 predicts an interaction between the effects of experimental correlation and the size of background; it also predicts that this interaction should be located entirely in the bilinear component (Anderson, 1970a).

METHOD

Undergraduates rated the numerosness of patterns of dots on a 9-point scale, with 1 representing "very very few dots" and 9 representing "very very many dots."

Stimuli

The stimuli were patterns of solid black dots, 2 mm in diam, arranged randomly on white, square, cardboard backgrounds of varying size. Since the dots were distributed approximately uniformly over each background, density (dots/area) varied inversely with size.¹ The stimuli were displayed for

*The authors thank Allen Parducci and Norman H. Anderson for their helpful comments on earlier versions of the manuscript. Computing assistance was received from Campus Computing Network, UCLA.

†Requests for reprints should be sent to Michael H. Birnbaum, Department of Psychology, University of California, San Diego, La Jolla, California 92037.

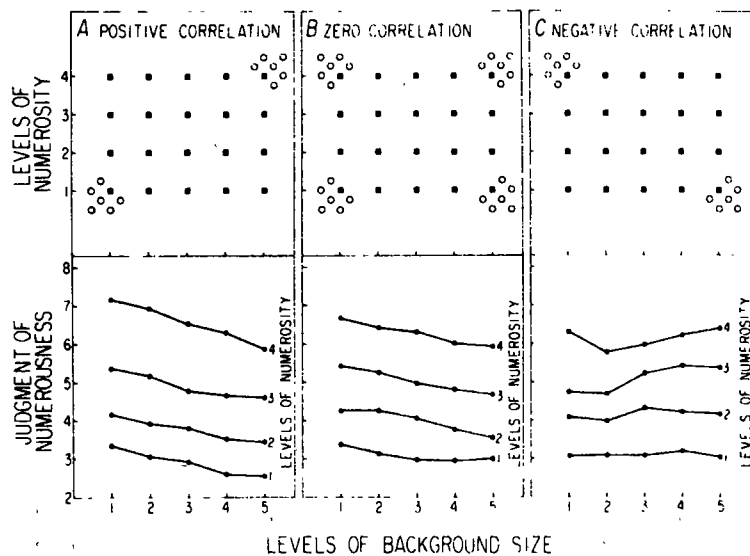


Fig. 1. (Upper panels) Schematic representation of the experimental design. Solid squares represent test stimuli, which were the same for all three conditions; open circles represent contextual stimuli for each correlation condition. (Lower panels) Mean judgments of test stimuli in the corresponding conditions, plotted as a function of size with a separate curve for each level of numerosity.

approximately .5 sec each to nine groups of 10 Ss seated 2-5 m from the E.

Subjects

The Ss were 90 UCLA undergraduates, fulfilling a requirement in introductory psychology. Thirty Ss served in each of three conditions, with a different correlation between size and numerosity for each condition.

Design

The upper panels of Fig. 1 give a schematic representation of the experimental design used to manipulate the correlation between variables. Each symbol (solid square or open circle) represents a different stimulus presented to the Ss.

As shown in Fig. 1, all Ss judged the same test stimuli, but the three groups judged different contextual stimuli. Figure 1 shows how each of the 20 test stimuli represents a different cell of a Numerosity by Size factorial design. The four levels of numerosity were 15, 30, 60, and 120 dots. The five levels of background size were 6.0, 8.5, 12.0, 17.0, and 24.0 cm in side length.

The contextual stimuli were presented to manipulate the overall correlation between size and numerosity while holding constant the range of each, as shown in Fig. 1. For example, the open circles in Fig. 1A indicate that the contextual stimuli for the positive correlation consist of large numbers of dots on large backgrounds and small numbers of dots on small backgrounds. The Ss in this condition presumably learn to expect more dots on larger backgrounds. On the other hand, the contextual stimuli for the negative correlation condition (Fig. 1C) consist of large numbers of dots on small backgrounds and small numbers on large backgrounds; therefore, Ss in the negative condition should learn to expect fewer dots on the larger backgrounds.

The first 20 trials of the experiment consisted entirely of randomly ordered presentations of contextual stimuli. The next 80 trials included three replicates of the 20 test stimuli (randomly ordered) with 20 additional contextual trials interspersed randomly. The overall correlations were .57, .00, and -.57 for the positive, zero, and negative correlation conditions, respectively.

RESULTS

The vertical separations between the curves in each of

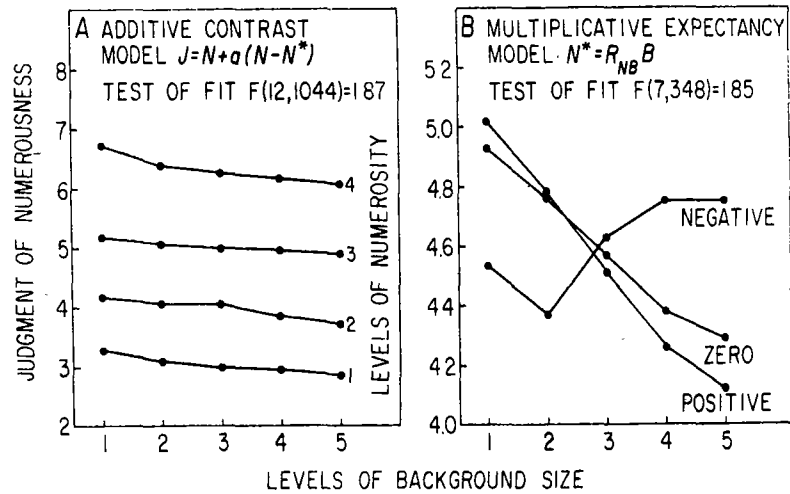
the lower panels of Fig. 1 represent the effects of numerosity upon judged numerosity of the test stimuli. The slopes of the curves represent the effects of size. When the correlation between numerosity and size is positive, judgments vary inversely with size; however, judgments vary directly with size when the contextual correlation is negative. The nature of this interaction is shown in Fig. 2B, where mean judgments of the test stimuli are plotted against size for each level of correlation. The order of the slopes in Fig. 2B is predicted by the expectancy hypothesis: Ss in the positive correlation condition expect more dots on larger backgrounds, and by contrast, judge there to be fewer. Ss in the negative correlation condition judge there to be more dots on the larger backgrounds. By reversing the expectancy, the experimental manipulation was able to reverse the direction of the illusion.

The statistical support for the expectancy hypothesis (that the slopes in Figs. 1 and 2B vary with the contextual correlation) lies in the highly significant interaction between experimental correlation and size, $F(8,348) = 10.41, p < .001$. The linear trend of the slope for the negative correlation condition is significantly positive, $F(1,29) = 5.21, p < .05$, indicating that the reversal was statistically significant.

The size of the effect can be inferred by comparing the scale of Fig. 2B with Fig. 2A. Figure 2A shows that the difference between Categories 4 and 5 corresponds to a change of about 30 dots; this is the scale of Fig. 2B. Thus, the illusion in the positive condition produces a highly significant *decrease* corresponding to about 25 dots, whereas the illusion in the negative condition is significantly reversed, producing an *increase* corresponding to about 9 dots.

These results are in qualitative agreement with theoretical predictions. Because the experimental method employs factorial designs, it is possible to employ powerful quantitative tests of fit to evaluate

Fig. 2. Model evaluation. (A) Mean judgments of test stimuli averaged across correlation conditions. (B) Mean judgments of test stimuli, averaged across levels of numerosity and plotted as a function of size for each correlation condition.



each part of the model. These tests employ Anderson's (1970a) functional measurement approach, estimating stimulus parameters from the data rather than requiring a priori estimates.

The additive-contrast model (Eq. 1) predicts parallel curves for each of the lower panels of Fig. 1, since the differences between the curves are a function of numerosity only. Figure 2A shows the overall Numerosity by Size interaction (averaged across levels of correlation). The curves are nearly parallel; the mean absolute discrepancy from the best-fit parallel curves is only .04, indicating an overall fit of the model. Statistically, the parallelism prediction of Eq. 1 is supported by a diminutive overall Numerosity by Size interaction, $F(12,1044) = 1.84, p < .05$. Although of borderline statistical significance, this result is encouraging, considering the fact that this test is powerful enough to detect an average discrepancy of less than 1/20th of one category.

The Numerosity by Size interaction was analyzed for each correlation condition separately, and a significant discrepancy was obtained for the negative condition only, $F(12,348) = 5.05, p < .05$. A separate analysis of the data for the third replicate indicated a nonsignificant interaction, $F(12,348) = 1.51$. The discrepancy appears to be due to a single point (largest numerosity and smallest size) in the negative condition, and to disappear by the third replicate. The same discrepancy also occurred and also disappeared in a pilot study.² It may thus be a real, but unimportant and perhaps transitory, effect.

According to the multiplicative-expectancy model (Eq. 2), the three functions in Fig. 2B should be linear, intersecting at a common point. This is very nearly the case, as can be seen. Statistically, Eq. 2 implies that the highly significant interaction between correlation and size be concentrated entirely in a single degree of freedom—the bilinear component. After the variance from this component is removed, the residual is nonsignificant, $F(7,348) = 1.85, p > .05$, in agreement

with the model. Thus, the manipulation of the correlation produced the specific effect predicted by the theory.

DISCUSSION

The fact that the size-numerosity illusion depends on the correlation between a background cue and the judged dimension gives further support to an expectancy interpretation of illusions. An alternative interpretation that Ss are partly responding to density implies nonparallel curves for Fig. 2A (Anderson, 1970b). Anderson's findings for the size-weight illusion were also inconsistent with this prediction of the density interpretation. In addition, the density interpretation would not have predicted that the illusion depends upon the correlation between size and numerosity.

It should be noted that the present model is based upon *subjective* correlation rather than upon *actual* correlation. The subjective correlations are estimated from the slopes of Fig. 2B; of course, the signs would be reversed. A slope of zero would imply that $R_{NB} = 0$, but the unit appears to be arbitrary.³ The present model assumes only that subjective correlation is monotonically related to the actual correlation in the experiment; it would be an interesting matter for future investigation to describe how this correlation depends upon actual correlation, number of trials, and prior experience. When the expectancy has not been established experimentally, it is assumed to depend upon prior experience. Since the judgments varied inversely with size when the contextual correlation was zero, it appears that Ss' prior experience established a positive correlation between size and numerosity.

An important distinction is made by separating the two parts of the model: the *expectancy* varies *directly* with the contextual correlation and the *judgment* varies *inversely*. Therefore, one should distinguish situations which require the S to report his expectancy from those which require the S to judge the situation *following* the

formation of an expectancy. For example, a S's judgment of heaviness would vary directly with the size of the object if he did not lift it, and inversely with size if he did.

It is interesting that the correlational basis for expectancy is revealed within the framework of systematic design. It has been assumed that factorial designs, by combining each level of each variable with every level of every other variable, *must* obscure the intercorrelations among the independent variables (Brunswik, 1955, 1956; Hammond, 1966; Hoffman, Slovic, & Rorer, 1968; Postman, 1955). Brunswik (1955, 1956) and Hammond (1966) have warned that if experimental results depended upon these correlations, then factorial designs would lead to ungeneralizable results. They recommended the use of representative design; however, representative design cannot assess the effects of these correlations. The present experiment represents a different approach—that of systextual design (Birnbaum, 1972), in which the contextual features of the experimental design are systematically manipulated. The present experiment demonstrates that the correlations among variables can be manipulated, while maintaining a factorial design.

One aim of research with psychophysical stimuli is to uncover principles of judgment which may have generality to our everyday judgments (Parducci, 1968). The common notions of "disappointment" and "pleasant surprise" suggest that the present model may provide insight into everyday affective judgments. For example, our enjoyment of a movie may reflect not only the quality of the movie itself, but also its relation to our expectancy based on reviews, advertisements, friends' reports, or having read the book. The present model provides a formal framework for discussion of social, clinical, and affective judgment.

REFERENCES

- Anderson, N. H. Functional measurement and psychophysical judgment. *Psychological Review*, 1970a, 77, 153-170.
- Anderson, N. H. Averaging model applied to the size-weight illusion. *Perception & Psychophysics*, 1970b, 8, 1-4.
- Bevan, W., & Turner, E. D. Assimilation and contrast in the estimation of number. *Journal of Experimental Psychology*, 1964, 67, 458-462.
- Birnbaum, M. H. Systematic vs representative design: Systextual design for stimulus variation and covariation. Paper presented at Conference on Human Judgment and Social Interaction, Boulder, Colorado, March 1972.
- Brunswik, E. Representative design and probabilistic theory in a functional psychology. *Psychological Review*, 1955, 62, 193-217.
- Brunswik, E. *Perception and the representative design of experiments*. Berkeley: University of California Press, 1956.
- Brunswik, E., & Herma, H. Probability learning of perceptual cues in the establishment of a weight illusion. *Journal of Experimental Psychology*, 1951, 41, 281-290.
- Hammond, K. R. Probabilistic functionalism: Egon Brunswik's integration of the history, theory and method of psychology. In K. R. Hammond (Ed.), *The psychology of Egon Brunswik*. New York: Holt, Rinehart, & Winston, 1966.
- Hoffman, P. J., Solvic, P., & Rorer, L. G. An analysis of variance model for the assessment of configural cue utilization in clinical judgment. *Psychological Bulletin*, 1968, 69, 338-349.
- Krueger, L. E. Perceived numerosity. *Perception & Psychophysics*, 1972, 11, 5-9.
- Parducci, A. The relativism of absolute judgment. *Scientific American*, 1968, 219, 84-90.
- Postman, L. The probability approach to nomothetic theory. *Psychological Review*, 1955, 62, 218-225.

NOTES

1. An important difference in method between the present study and that of Krueger (1972), deserves emphasis: in the present study, the size of the background is varied; Krueger (1972) varied the area in which the dots appeared, holding background size constant. He found a slight tendency for dots to appear less numerous when bunched together than when spread out; we have also obtained this result in preliminary work with similar stimuli. Of course, when the size of the background is held constant, it cannot be used by the Ss to form expectancies; in addition, this procedure introduces a white surround of varying size, which may affect the subjective size and density of the dot pattern.

2. In the pilot study, 14 Ss served in each correlation condition. The stimuli and general procedure were almost identical to those reported here, except that the 20 contextual stimuli were all presented before the 60 trials of test stimuli. The results were nearly identical to those of the present study, except that the effects of correlation decayed as a function of replicates. The pilot study also contained a group who judged the test stimuli only; the results for this group paralleled those for the zero correlation group reported here.

3. The present model for the formation of expectancy (Eq. 2) is analogous to the best-fit linear regression equation: $\hat{z}_Y = r_{XY}z_X$, where \hat{z}_Y is the predicted z score of Y, r_{XY} is the correlation coefficient, and z_X is the z score of X. An anonymous reviewer has suggested that the term "subjective regression coefficient" would be preferable to "subjective correlation" for R_{NB} in Eq. 2. This would be the case if N and B were interpreted as subjective deviation scores rather than as subjective z scores. However, since the overt response is only assumed to be a linear function of the impressions of numerosity (J in Eq. 1), it would not seem possible to differentiate these interpretations.

(Received for publication February 28, 1972;
final revision received October 26, 1972.)