Journal of Experimental Psychology:
Human Perception and Performance
1986, Vol. 12, No. 1, 108–109

0096-1523/86/$00.75

OBSERVATIONS

Model Testing Is Not Simple: Comments on Lane, Anderson, and Kellam

Colleen F. Surber
University of Wisconsin–Madison

Research by Lane, Anderson, and Kellam (1985) on covariation judgment raises important issues in testing models of human judgment. It is argued that any test of a model of human judgment must consider the possibility that the psychophysical functions for the stimulus variables are not the identity function, and the possibility that rating scale responses are not a linear function of the subject's impressions. Neither of these possibilities was considered by Lane et al.

Lane, Anderson, and Kellam's (1985) research on judgments of relatedness represents a new approach to this area and thus provides a valuable contribution. Instead of varying only the degree of relation between variables and studying the psychophysical function for judged relatedness, these authors have varied three components of Pearson's correlation and have studied judged relatedness as a function of those components. Other researchers could fruitfully follow their example, perhaps examining other possible parameterizations of the equation for Pearson's correlation or combinations of other variables that might influence perceived covariation. Although their article should stimulate new research on perceived covariation, my opinion is that the authors have not really succeeded in testing Pearson's equation [$r^2 = b^2S_x^2/(b^2S_x^2 + S_e^2)$] as a psychological model for judgments of relatedness. Below, I outline the difficulties in testing Pearson's equation as a psychological model of covariation judgment. Most of these are general issues in testing models of human judgment.

Lane et al. concluded that "the components do not influence covariation judgments in precisely the same manner as they influence the value of Pearson's r." (p. 649). This conclusion is correct if the phrase "precisely the same manner" is interpreted as meaning that (a) the manipulated values of the variables are perceived veridically; (b) the perceived values are combined by a psychological process that is isomorphic with Pearson's equation; and (c) the responses are a linear function of the subjects' impressions of covariation.

In general, any attempt to test an equation as a psychological model of judgment must consider the possibility that the manipulated values are not perceived veridically, that is, that the psychophysical function may not be the identity function. The data of Lane et al. cannot reject Pearson's equation as a model of covariation judgment unless the psychophysical functions are assumed to be the identity function. This is demonstrated by comparing the upper and center panels of Figure 1. The upper panel of Figure 1 plots the Pearson's correlation values for Lane et al.'s manipulated values of slope, X-variance, and error variance. The points surrounded by squares have $r = .78$, and those surrounded by circles have $r = .53$. According to Lane et al., the members of each set of stimuli (those highlighted by squares or those highlighted by circles) should result in equivalent covariation judgments if covariation judgments are based on Pearson's equation. The center panel of Figure 1 shows the Pearson correlation values, assuming that subjects "misperceive" the low error variance to be 600 rather than 1,000 but maintaining the assumption that the variables are combined by a psychological process that is isomorphic to Pearson's equation. Notice that in the center panel the three points surrounded by squares no longer have the same r value, and analogously for the three points surrounded by circles. Thus, unless the psychophysical function for the manipulated variables is the identity function, the nonequivalence of the judgments for the key stimuli in Figure 1 does not tell us whether Pearson's equation is viable model of covariation judgment.

The bottom panel of Figure 1 presents the values of r^2 for the same stimulus combinations that are presented in the center panel of the figure. This monotonic transformation of the ordinate is analogous to the possibility that ratings of covariation are not a linear function of the subject's impression of covariation. In the bottom panel, when the slope is equal to 2, the data show an interaction of X-variance and error variance, whereas in the center panel there is no interaction for a slope of 2. Thus, unless it is assumed that the response scale is used linearly, the X-Variance × Error Variance interaction cannot be used to test Pearson's equation. The X-Variance × Error Variance interactions in the data of Lane et al. can be predicted by Pearson's equation if response scale nonlinearity is allowed. In order to test mathematical models of human judgment, it is necessary either to make some simplifying assumptions about the response scale or to use experimental designs and methods that have the leverage to test models ordinally (Birnbaum, 1982a; Krantz & Tversky, 1971). The difficulty in testing models of human judgment is reflected in the fact that the field of psychophysics has a long history of debate over the interpretation of rating scale responses (Attneave, 1962; Birnbaum, 1982b; Krantz, 1974; Marks, 1974; Torgerson, 1961).

In sum, Lane et al. have ignored both the psychophysical function and the possibility of response scale nonlinearity in their initial attempt to test Pearson's correlation equation as a psychological model of covariation judgment. Although they conceptualize their experiment as examining the functional relation between the three cues and judgment of covariation, with-

Correspondence concerning this article should be addressed to Colleen F. Surber, Department of Psychology, University of Wisconsin, Madison, Wisconsin 53706.

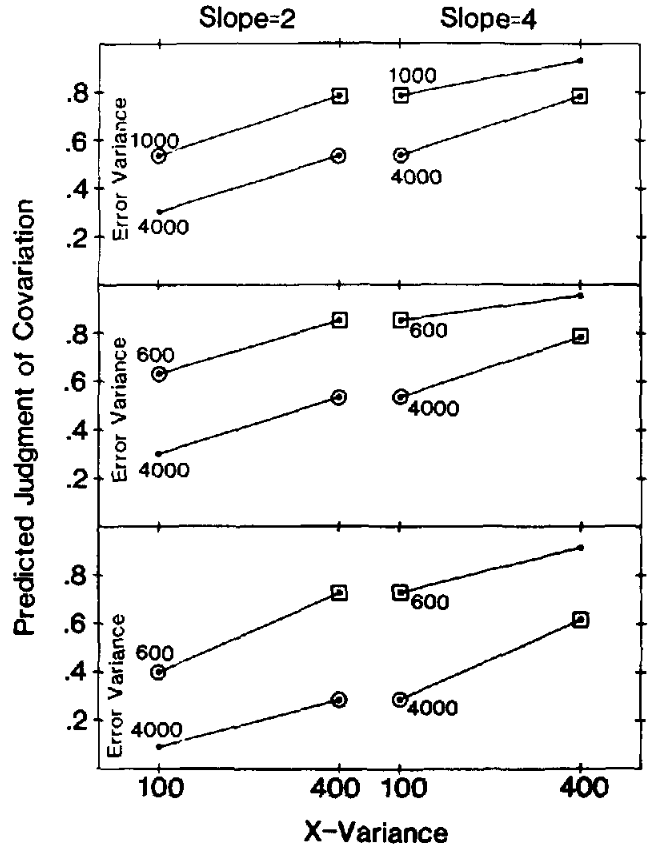

Figure 1. Predicted judged covariation under three different assumptions. (The top panels present predicted r assuming veridical perception of all stimulus variables and a linear response scale. The center panels present predicted r assuming tht low error variance is valued at 600 rather than 1,000. The bottom panels present predicted r^2 for the same stimulus values as shown in the center panels.)

out considering the problems of psychological scaling of the manipulated variables and response scale nonlinearity, it is not possible to articulate and test theories of the functional relation between cues and covariation judgment. These issues are important regardless of whether the stimulus dimensions are perceived integrally or separably.

Relative Importance of Cues

Lane et al. concluded that error variance had a larger influence on judged relatedness than did either X-variance or slope. The evidence presented for this conclusion was a set of Newman-Keuls comparisons showing that objectively equal stimulus changes do not produce equal changes in judged covariation. These conclusions regarding the relative importance of cues also depend on assuming the psychophysical function to be the identity function. The authors' particular findings can be predicted by proposing that the low error variance stimulus is perceived to be lower than its objective value, whereas the values of the other variables are perceived to be close to their objective values.

A more serious issue is the meaning of the term *importance* when applied to the effect of a stimulus variable in psychology. Measures of importance such as the proportion of variance accounted for, effect size, and omega squared depend critically on what other variables are present in the experiment and the degree to which each variable in the experiment is manipulated. A variable that is manipulated over a wider psychological range will have a larger effect than a variable manipulated over a narrower subjective range. To draw a general conclusion that one variable is more important in judgment than others, much more is required than a demonstration that stimulus changes that are objectively equal do not have equivalent effects on judgment. The definitions of the psychological importance of variables that are implicit in different theoretical approaches to human judgment have been articulated by Hammond, McClelland, and Mumpower (1980, pp. 213–218) and by Shanteau (1980). For some approaches, psychological importance is represented in part in scaling constants, whereas for other approaches psychological importance is represented in weights that are theoretically independent of scale value. Although there is no method for measuring the psychological importance of variables that is generally agreed upon by researchers of human judgment, Lane et al. have not attempted to use any of the extant approaches and have not articulated their own approach to the measurement of the psychological importance of variables in judgment.

References

Attneave, F. (1962). Perception and related areas. In S. Koch (Ed.), *Psychology: A study of a science* (Vol. 4, pp. 619–659). New York: McGraw-Hill.

Birnbaum, M. H. (1982a). Controversies in psychological measurement. In B. Wegener (Ed.), *Social attitudes and psychophysical measurement* (pp. 401–485). Hillsdale, NJ: Erlbaum.

Birnbaum, M. H. (1982b). On rescaling data to fit the model and concluding that the model fits: A note on monotone transformation. *Perception & Psychophysics, 32,* 293–296.

Hammond, K. R., McClelland, G. H., & Mumpower, J. (1980). *Human judgment and decision making: Theories, methods and procedures.* New York: Praeger.

Krantz, D. H. (1974). Measurement theory and qualitative laws in psychophysics. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.), *Measurement, psychophysics and neural information processing* (pp. 160–199). San Francisco: Freeman.

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

Lane, D. M., Anderson, C. A., & Kellam, K. L. (1985). Judging the relatedness of variables: The psychophysics of covariation detection. *Journal of Experimental Psychology: Human Perception and Performance, 11,* 640–649.

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

Shanteau, J. (1980). *The concept of weight in judgment and decision making: A review and some unifying proposals* (Report No. 228). Boulder, CO: University of Colorado, Center for Research on Judgment and Policy.

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

Received June 14, 1985
Revision received September 29, 1985 ■