



Intuitive Numerical Prediction

Michael H. Birnbaum

The American Journal of Psychology, Vol. 89, No. 3. (Sep., 1976), pp. 417-429.

Stable URL:

<http://links.jstor.org/sici?sici=0002-9556%28197609%2989%3A3%3C417%3AINP%3E2.0.CO%3B2-V>

The American Journal of Psychology is currently published by University of Illinois Press.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/illinois.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

The JSTOR Archive is a trusted digital repository providing for long-term preservation and access to leading academic journals and scholarly literature from around the world. The Archive is supported by libraries, scholarly societies, publishers, and foundations. It is an initiative of JSTOR, a not-for-profit organization with a mission to help the scholarly community take advantage of advances in technology. For more information regarding JSTOR, please contact support@jstor.org.

Intuitive numerical prediction

Michael H. Birnbaum

University of Illinois at Urbana-Champaign

Subjects were trained, with feedback, to predict a numerical criterion from each of two separate cues and then asked, without feedback, to predict it from a pair of independent cues or a single cue. Their intuitive predictions were qualitatively inconsistent with an additive model, since the effect of one cue varied inversely with the number of cues available, and with a constant-weight averaging model, since the effect of one cue varied inversely with the validity of the other cue. The data were consistent with a relative-weight averaging model, which assumes that subjective cue values are averaged using weights that depend on cue validities. Normative and descriptive theories of intuitive prediction are compared.

Early research on intuitive statistics focused on the *normative accuracy* of subjective statistical estimates, predictions, or decisions. This research is discussed from different points of view in review articles by Peterson and Beach (1967), Slovic and Lichtenstein (1971), Hammond and Summers (1972), Bremer (1974), Anderson (1974), and Levin (1975). One issue in subjective statistics that received considerable attention is multiple-cue prediction. In the early work, it was concluded that intuitive predictions could be well approximated by an additive model of multiple linear regression, based on what seemed to be large multiple correlations between the theory and the data.

Recent research departs from the earlier work in two ways. First, since correlations of fit can be misleading (Anderson, 1972; Birnbaum, 1973, 1974b), it is difficult to draw theoretical conclusions from much of the previous research on this topic, which used the additive model both as a theory and a statistical device. Therefore, more recent work uses improved techniques to separate model testing and measurement. Second, the new concern is with testing *descriptive psychological theory* rather than assessing normative accuracy, or 'performance.' Recent work has

viewed intuitive statistics as a problem in information integration, using descriptive theories of information processing to account for intuitive statistical judgments (Anderson, 1974; Birnbaum, Wong, and Wong, 1976; Kahneman and Tversky, 1973; Levin, 1975; Lichtenstein, Earle and Slovic, 1975). The present experiment uses new approaches to differentiate three types of models describing intuitive numerical prediction.

For simplicity, the models will be described in the special forms applicable to the present experiment. Generalization to other cases is straightforward. In the present experiment, the subjects were trained, with feedback, to predict a criterion, Y , from each of two separate cues, X s. Then they were asked to predict the criterion, without feedback, from single cues or from pairs of independent (uncorrelated) cues. The first cue was always of low validity (correlating .447 with the criterion), whereas the second cue was of either low or high validity (either .447 or .894). In all of the following models, it is assumed that subjective scale value (s) depends only on the objective cue value (X) and that weight depends only on cue validity, which the subject learns in the training phase of the experiment.

Additive models

An additive model of multiple linear regression for two independent cues can be written

$$Y^* = s_0 + w_1s_1 + w_2s_2, \quad [\text{Equation 1}]$$

where Y^* is the intuitive prediction, s_0 is the additive constant, w_1 is the weight of the first cue, s_1 and s_2 are the scale values of the first and second cues respectively, and w_2 is the weight reflecting the validity of the second cue. If only a single cue is presented, the weight of the omitted cue is set to zero. Since w is presumed to depend only on the validity of a cue and s only on its value, Equation 1 predicts a bilinear interaction between the validity and value of a cue. It predicts that the effect of the first cue will be independent of the presence of the second cue or its validity. Equation 1 also predicts no interaction between the values of the cues.

Equation 1 is a more general statement of the *optimal* multiple linear-regression model. The *least-squares* equation is the special case of Equation 1 in which s_0 equals the mean of Y , the values of s are the deviations of the cues about their means, and the weights are $\rho_{iY}\sigma_Y/\sigma_{X_i}$, where ρ_{iY} is the cue/criterion correlation and σ_Y and σ_{X_i} are the standard deviations of criterion and cue respectively. The *linear* model is also a special case of Equation 1 in which the values of s are assumed to be equal to the objective cue values, X , but the other parameters are unconstrained.

Constant-weight averaging models

The constant-weight averaging model (Birnbbaum, Wong, and Wong, 1976) can be written

$$Y^* = (as_0 + bw_1s_1 + cw_2s_2)/(a + b + c), \quad [\text{Equation 2}]$$

where Y^* , s , and w are defined as above and a , b , and c are constants that are independent of cue validity. When only the first cue is presented, c is set to zero. Like Equation 1, this model predicts a bilinear interaction between the validity and value of a cue and no interaction between cues. Unlike the additive model, however, Equation 2 predicts that the effect of each cue will be inversely related to the total number of available cues.

The constant-weight model has formal similarities to the impression-formation model of Rosenbaum and Levin (1968, 1969), since it predicts that the effect of a cue is independent of the validity of the other cue. The *average-of-regressed-values* model (Lichtenstein et al., 1975) is a special case of Equation 2 in which $a = 0$ and $c = b = 1$. Under these restrictions, the model could not accommodate an effect of set size: predictions based on two independent, extreme values could be no more extreme than those based on one cue of the same value. By allowing a to attain positive values, however, Equation 2 can accommodate such an effect.

Relative-weight averaging model

The relative-weight averaging model (Anderson, 1971; Birnbbaum et al., 1976) can be written

$$Y^* = (w_0s_0 + w_1s_1 + w_2s_2)/(w_0 + w_1 + w_2), \quad [\text{Equation 3}]$$

where Y^* , w , and s are defined as above. When a single cue is presented, the weight of the omitted cue is zero. It is useful to conceptualize the value of s_0 as the prediction based on no information and the value of w_0 as a type of predicted or 'regressed' value based on one cue, apart from all other effects. It is then instructive to rewrite Equation 3 in the form

$$Y^* = s_0 + [w_1(s_1 - s_0) + w_2(s_2 - s_0)]/(w_0 + w_1 + w_2). \quad [\text{Equation 4}]$$

Written in this form, it can be seen that without information, the response will be s_0 , and that with information, the response will deviate from s_0 in proportion to an average of deviations. The relative weight of a cue will be directly proportional to its validity and inversely proportional to the validity of other cues. For example, the relative weight of the first cue

will be $w_1 / (w_0 + w_1 + w_2)$; therefore, the greater the validity of the second cue (the larger the value of w_2), the less the relative effect of the first cue. Similarly, the effect of the first cue will be greater when other cues are not presented, since the relative weight will be $w_1 / (w_0 + w_1)$.

Previous tests of the models

Birnbaum et al. (1976) reported two experiments that tested the additive, constant-weight averaging, and relative-weight averaging models. In one experiment, subjects estimated the value of used cars based on Blue Book value and an estimate provided by one of three friends who examined the car. The friends differed in mechanical expertise. In the second experiment, subjects rated the likableness of a person described by two sources, who varied in their length of acquaintance with the person to be rated. In both experiments, both the additive model and the constant-weight averaging model could be rejected, since the effect of information provided by one source was inversely related to the expertise or credibility of the other source. The relative-weight model can account for these effects. These experiments tested sources whose credibilities are symbolically defined. Presumably, credibility/outcome relationships are learned from everyday experience outside the laboratory. Thus, the symbolically manipulated credibilities should be analogous to learned cue/criterion validities. This analogy predicts that if the cue/criterion relationships are learned within an experiment, the relative-weight model should apply.

Lichtenstein et al. (1975) trained subjects to predict a criterion from individual cues, then tested them with combinations of cues. The additive model was inconsistent with the data, since intuitive predictions became less extreme as moderate cue values were added to extreme values. Lichtenstein et al. (1975) advanced an average-of-regressed-values model in which the hypothesized response is a simple average of the regressed values of the cues. As noted above, this model is a special case of the constant-weight averaging model considered by Birnbaum et al. (1976); see Equation 2 above. Unfortunately, experiments like those of Lichtenstein et al. (1975) cannot in principle distinguish the constant-weight from the relative-weight averaging model because cue validities are not independently varied. The present study was a test of the additive models favored by early research on this topic, the average-of-regressed-values model of Lichtenstein et al. (1975), and the relative-weight averaging model. This was accomplished by manipulating the validity of one cue independent of the validity of another cue.

METHOD

The general procedure of the experiment was similar to that of Lichtenstein et al. (1975). Subjects were first trained to predict the criterion (Y) without any information, then with single cues (X). During training, subjects were given feedback so that they could learn the distributions of Y and X and the individual cue/criterion relationships. Finally, they were tested without feedback, predicting Y from pairs of independent cues or from single cues.

Instructions

Printed instructions explained the nature of the task and stated that the criterion, Y , had a mean of 20. In the first phase of the training, subjects were shown a deck of plain white cards and asked to guess the value of Y that would occur on the back of the first card. After writing down their guess, Y^* , the card was turned over to read, for example, ' $Y = 18$.' Then the subjects were trained to compute their residual, $Y - Y^*$, and to square it, $(Y - Y^*)^2$. They made several guesses without information in this fashion, computing each time their 'squared mistakes.' They were told that their goal was to make predictions as close to the actual values as possible, in the sense that the squared deviations would be minimal.

In the second phase of the training, the subjects were shown a deck of cards and told that one side of the card contained a 'hint,' a cue, X , that could be used to predict the value of the criterion, Y , printed on the other side of the card. It was explained that Y could not be perfectly predicted from X but that the average squared mistake would be reduced if the cue/criterion relationship was used.

The subjects were told that all of the predictors, X s, had means of 100 and identical distributions. They were informed that the first cue was low in validity ($\rho = .447$) and that the second cue could be either low ($\rho = .447$) or high ($\rho = .894$) in validity. The two cues would be independent, uncorrelated cues. One cue could not be predicted from the other, but both could be used to predict the criterion.

Individual relationships were then learned in the following fashion. A deck of 25 cards was shuffled, the cue was identified as of low or high validity, and the subjects predicted Y from X . One side of the card might read ' $X_1 = 130$,' the subject would write down his prediction, then the card would be turned over to reveal, for example, ' $Y = 24$.' Figure 1 shows the relationships between the low-validity and high-validity cues and the criterion. The low-validity regression equation is given by the formula $Y = (X_1 - 100)/15 + E$, where $X_1 = 70, 85, 100, 115, \text{ or } 130$, and $E = -4, -2, 0, 2, \text{ or } 4$. The 25 cards factorially combined the five values of X_1 with the five values of E . The low-validity second cue, X_2 , had the same characteristics as X_1 . The criterion could be predicted from the high-validity cue from the equation $Y = 2(X_3 - 100)/15 + E$, where values of $X_3 = 70, 85, 100, 115, \text{ and } 130$ were factorially combined with $E = -2, -1, 0, 1, \text{ or } 2$.

After two blocks of training trials with each cue, the nature of the task was again discussed, the independence of the cues was emphasized, and another block of training trials with each type of cue was presented to clarify the individual relationships between each type of cue and the criterion. Then the sub-

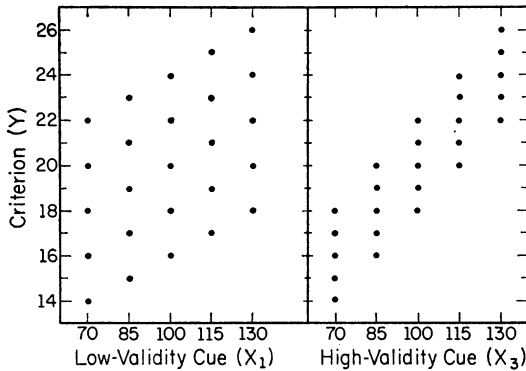


Figure 1. Individual cue/criterion relationships used for training. At the left, the relationship between the low-validity cue, X_1 , and criterion, Y ; the correlation for X_2 was the same. At the right, the relationship between the high-validity cue, X_3 , and criterion, Y

jects were told to make the best predictions they could, based on either one cue or two independent cues.

Design

The testing trials were constructed from a 2 (validity of the second cue) \times 5 (value of the second cue) \times 5 (value of the first cue) factorial design, in which the first cue (X_1) was always designated as low in validity ($\rho_{1Y} = .447$), the second cue (X_2 or X_3) was either low ($\rho_{2Y} = .447$) or high ($\rho_{3Y} = .894$) in validity, and each cue could take on five values: 70, 85, 100, 115, and 130. Fifteen additional test trials were produced by the inclusion of single-cue values for X_1 , X_2 , and X_3 .

Subjects

The subjects were 53 undergraduates at the University of Illinois at Urbana-Champaign. Of these 53, a subset of 20 were enrolled in a course in psychological statistics; the rest were students in introductory psychology who received extra course credit for participating in the experiment. A small number of additional undergraduates were tested, but failed to follow instructions.

RESULTS

The mean intuitive predictions are plotted as solid points in Figure 2 as a function of the value of the second cue, with separate curves for different values of the first cue, X_1 . The curves at the left show the results when the second cue was low in validity (X_2); the curves at the right show the results when the second cue was high in validity (X_3). The

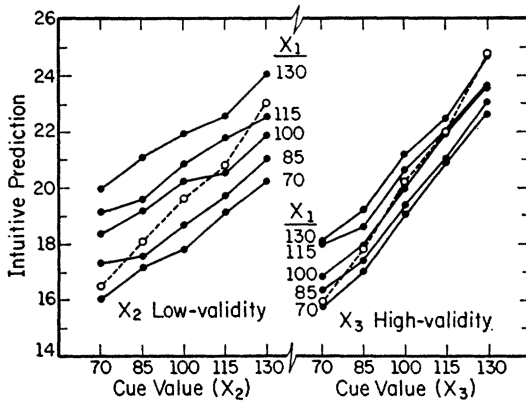


Figure 2. Mean intuitive predictions as a function of the second cue, with a separate curve for each level of the first cue, X_1 . At the left, the results for the low-validity second cue, X_2 . At the right, the results for the high-validity second cue, X_3 . Open circles and dashed lines plot intuitive predictions based on a single cue of low (at the left) or high (at the right) validity

slopes of the curves represent the effects of the second cue; the vertical separations represent the effects of the first cue. Consistent with all of the models, and with the fact that the independent cues were all positively correlated with the criterion, it can be seen that the predictions were monotonically increasing functions of each cue's value.

Figure 2 also illustrates that a given variation for the high-validity cue had a greater effect on the prediction than the same variation for a low-validity cue: the slopes at the right are much greater than the slopes at the left. The interaction between cue value and cue validity was statistically significant [$F(4, 208) = 52.39$]. This change in slope is predicted by the multiplicative relation between cue value (s) and cue validity (w) posited by all of the models. Of the 53 subjects, 47 showed this pattern, indicating that they detected and utilized the cue/criterion validities.

The open circles connected by dashed curves represent intuitive predictions based on a single low-validity cue (at the left) or a single high-validity cue (at the right). The additive model of Equation 1 can be rejected, since it predicts that the dashed curves should parallel the solid curves — that the effect of each cue should be independent of the presence of the other cue. Instead, the effect of each cue was greater when presented alone than in combination with another cue. The fact that the dashed curves cross the solid curves rules out Equation 1.

The constant-weight model of Equation 2 can also be rejected on the

basis of these data, since it predicts that the effect of each cue will be independent of the validity of the other cue. Instead, the effect of X_1 (the spread of the curves) was less when combined with the high-validity cue (at the right of Figure 2) than when combined with the low-validity cue (at the left). This interaction was significant [$F(4, 208) = 24.28$]. Single-subject analyses revealed that 47 of the 53 subjects showed this crossover interaction, including 18 of the 20 statistics students. Therefore, the constant-weight model can be rejected. This result also eliminates the average-of-regressed-values model (Lichtenstein et al., 1975), since it is a special case of Equation 2.

These findings, which rule out Equations 1 and 2, are predicted by Equation 3. Since the effect of the low-validity cue is proportional to its relative weight, it should be clear that when only a single cue is presented, the effect of the cue will be proportional to $w_1/(w_0 + w_1)$; when an additional cue is presented, the effect of the first cue will be reduced, $w_1/(w_0 + w_1 + w_2)$. Furthermore, the greater the weight of the second cue (w_2), the *less* the relative weight of the first cue. Therefore, the relative-weight model of Equation 3 gives a simple account of these qualitative effects.

Equation 3 was fit to the data using STEPIT, a parameter-estimation subroutine (Chandler, 1969), programmed to minimize the squared deviations from the model. Since the distributions of the predictors were identical, the scale values were assumed to be the same for X_1 , X_2 , and X_3 . Similarly, the weight of X_1 was assumed to be the same as the weight of X_2 , since the validities were the same. The best-fit estimates of weights (with w_0 arbitrarily set to 1.0) were $w_{low} = 1.64$ and $w_{high} = 5.21$; scale values were $s_0 = 18.67$, $s_{70} = 15.62$; $s_{85} = 17.47$, $s_{100} = 20.36$, $s_{115} = 22.60$, and $s_{130} = 25.60$. The mean squared discrepancy for the 60 means was .044, apparently a close fit. Inspection of the residuals did not reveal any systematic pattern of errors. All of the models predict that the solid curves in Figure 2 should be parallel. The curves do appear approximately parallel; however, the small deviations from parallelism were of borderline statistical significance [$F(16, 832) = 2.04$].

Figure 3 plots mean judgments for the two-cue data (solid points in Figure 2) as a function of the cues' scale values (derived from the fit of Equation 3). The left of Figure 3 plots mean judgments averaged over values of X_1 as a function of the scale value of the second cue. The two different slopes represent the low- and high-validity second cues. The model predicts that the interaction between cue value and cue validity should be bilinear, with slopes proportional to $w_{low}/(w_0 + 2w_{low})$ and $w_{high}/(w_0 + w_{low} + w_{high})$. The empirical points fall close to the pre-

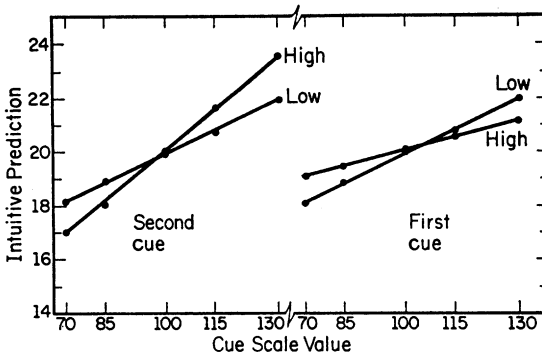


Figure 3. Mean intuitive predictions based on two cues as a function of the cues' scale value. Lines are theoretical predictions based on the relative-weight averaging model, points are empirical. Curves at the left plot predictions as a function of the second cue, with a separate curve for each level of its validity. Curves at the right show the effect of the first cue, with a separate curve for each level of validity of the second cue. The interaction of the curves at the right rules out the additive and the constant-weight averaging models

dicted lines, showing that the weights derived by STEPIT give a good fit to this relationship.

The right of Figure 3 plots an important test of the models. Mean judgments are plotted as a function of the scale value of the first cue, with a separate curve for each level of validity of the second cue. The additive and constant-weight averaging models predict no interaction. There is no necessary connection between the interactions at the left and right of Figure 3. Therefore, the relative-weight averaging model receives some measure of support from the fact that it correctly predicts the slopes. The same weights that increase the slopes at the left decrease the slopes at the right, since they appear in the denominators of the expressions for the relative weights, $w_{low}/(w_0 + 2w_{low})$ and $w_{low}/(w_0 + w_{low} + w_{high})$. Again, the data appear to be well fit by the predicted lines. In brief, in spite of some small deviations, the relative-weight averaging model appears to give a good account of the important features of the data.

Correlational analyses

To illustrate the difference between the present methods of analysis and the correlational techniques that have been utilized in previous research on this topic, mean intuitive predictions were fit by multiple regression as linear functions of the objective values of the cues.

The linear model correlated .984 with either set of two-cue data (solid points), at the left and right of Figure 2, analyzed separately. Furthermore the *optimal* regression equation (optimal weights) correlated .979 with all of the two-cue data. These correlations are considered 'high' by contemporary standards, yet the model they represent cannot account for the crossover interactions in Figure 2 and at the right of Figure 3. Such crossing curves, although inconsistent with additivity, can achieve high correlations with additive models (Birnbaum, 1973, 1974b); therefore, large correlations should not be interpreted to mean that data are theoretically compatible with the model. Unless extraneous factors are unconfounded, correlations can even be higher for less appropriate models than for correct models (Birnbaum, 1973, 1974b). In the present case, the correlation between predictions from the relative-weight averaging model of Equation 3 is higher (.995) than for the other models, but its superiority should be judged on the basis of its ability to account for qualitative features of the data rather than for its higher correlation.

It is interesting to compare the linear 'model of the subjects' with the optimal regression equation. The optimal weights for low- and high-validity cues are .067 and .133 respectively. For the $X_1 \times X_2$ subdesign, the predictions were fit by the equation

$$Y^* = 7.32 + .0637X_1 + .0625X_2, \quad [\text{Equation 5}]$$

where Y^* is the subject's prediction. The $X_1 \times X_3$ combinations were fit by the equation

$$Y^* = 5.62 + .0353X_1 + .1095X_3. \quad [\text{Equation 6}]$$

By comparing the two equations, it can be seen that the regression coefficient for X_1 changes depending on whether X_2 or X_3 is paired with it. Had only the $X_1 \times X_2$ design been used, one might be tempted to conclude that subjects are not only like multiple-regression equations but also discover the appropriate coefficients! Had only the $X_1 \times X_3$ design been used, one might tend to conclude that subjects are 'conservative,' since the weights were less than optimal. In this language, the single-cue results could be called 'counterconservative.' From the viewpoint of normative theory, then, the present experiment finds evidence of 'optimal' behavior, 'conservative' behavior, and 'counterconservative' behavior. It should be clear that these terms are descriptions, not explanations. All of these effects — the finding that the weights vary directly with the cue validities, vary inversely with the number of cues available, and vary inversely with the sum of the validities of all of the cues — that would be so puzzling without a proper descriptive theory can be explained by Equation 3.

DISCUSSION

The intuitive numerical predictions refute both the additive model of Equation 1 and the constant-weight averaging model of Equation 2. Much of the early research in intuitive prediction was interpreted as support of the additive model, but it reported only correlations between predicted and obtained responses. Since these correlations are not useful for purposes of theoretical interpretation (Anderson, 1972; Birnbaum, 1973, 1974b; Lichtenstein et al., 1975), it is difficult to know how to interpret early research on this problem. The inverse crossover interaction in Figure 3 demonstrates that the effect of the first cue is inversely related to the validity of the second cue. This finding rules out the additive model and the constant-weight averaging model, including the average-of-regressed-values model of Lichtenstein et al. (1975). Instead, the data are consistent with the predictions of the relative-weight averaging model.

In addition to functional measurement (Anderson, 1971, 1972, 1974; Birnbaum, 1973, 1974a, 1974b; Levin, 1975), the design of the present experiment permits differentiation among the models under investigation. A single two-factor design cannot in principle differentiate the additive from the averaging models (Anderson, 1971). But by varying the number of cues, the additive can be differentiated from the averaging models (Anderson, 1974; Lichtenstein et al., 1975). And by varying the validity of one cue independent of another, it is possible to differentiate relative-weight from constant-weight averaging models (Birnbaum et al., 1976).

The present results are consistent with those obtained by Birnbaum et al. (1976) for estimates of the values of used cars based on Blue Book value and friends' estimates and those for ratings of the likableness of persons described by sources of different credibility. Consistent with the relative-weight averaging formulation of Equation 3, both experiments found that the greater the credibility of one source, the less the effect of information provided by another source. Thus, source credibility in impression formation can be modeled in the same way as cue validity in multiple-cue prediction.

Kahnemann and Tversky (1973) have argued that intuitive statistics follow a heuristic of 'representativeness': intuitive predictions are supposed to be representative of the evidence. While Kahnemann and Tversky (1973) did not discuss the situation of the present experiment, 'averaging' seems consistent with the notion of representativeness, since subjects would probably judge an average to be representative of a set of scores. However, the models considered here are more precise than the notion of representativeness, which appears to attempt to explain one dimension,

intuitive prediction, in terms of other psychological dimensions, 'similarity' and 'representativeness,' that themselves need explanation.

The qualitative results implied by the averaging model of Equation 3 can also be interpreted in terms of the statistical regression equation if the 'independent' cues are intuitively correlated. The regression coefficient of a cue, X_1 , will be proportional to β_1 ; that is,

$$\beta_1 = (\rho_{1Y} - \rho_{12} \rho_{2Y}) / (1 - \rho_{12}^2), \quad [\text{Equation 7}]$$

where ρ_{1Y} is the validity of the cue, ρ_{2Y} is the validity of the other cue, and ρ_{12} is the cue intercorrelation. Assuming all of the correlations are positive, Equation 7 implies that the effect of a cue should be inversely related to the validity of another cue. This could be taken to suggest that subjects do not understand independence but act as if they thought that 'independent' estimates of the same quantity should be correlated.

This suggestion loses some of its appeal when one considers that similar results were found for six graduate students at the University of Illinois. Three were majoring in quantitative psychology and had extensive experience in multivariate statistics. These students should understand the independence of the cues, which was made clear by a diagram of the factorial combinations of predictors. One student, majoring in social psychology, compulsively computed the optimal values. The other five graduate students, including the three in quantitative psychology, performed the task intuitively, with results nearly identical to those shown in Figures 2 and 3. When the results were explained to these last five students, some reported that their intuitive predictions had 'felt right' and that the experiment and its explanation gave them new insights into both multiple regression and human judgment.

In summary, intuitive predictions are inconsistent with the additive model and the constant-weight averaging model, including the simple average-of-regressed-values model. Instead, intuitive predictions appear compatible with the relative-weight averaging model in which regressed subjective values are averaged using weights that depend on their validities.

Notes

This research was facilitated by a summer research fellowship from the University of Illinois at Urbana-Champaign. The author thanks Barbara Mellers for a critical reading of the manuscript. Address requests for offprints to the author, Department of Psychology, University of Illinois, Champaign, Illinois 61820. Received for publication August 6, 1975.

References

- Anderson, N. H. 1971. Integration theory and attitude change. *Psychological Review* 78:171-206.
- Anderson, N. H. 1972. Looking for configularity in clinical judgment. *Psychological Bulletin* 78:93-102.
- Anderson, N. H. 1974. Information integration theory: A brief survey. In *Contemporary developments in mathematical psychology*, vol. 2, ed. D. H. Krantz, R. C. Atkinson, R. D. Luce, and D. Suppes. San Francisco: W. H. Freeman.
- Birbaum, M. H. 1973. The devil rides again: Correlation as an index of fit. *Psychological Bulletin* 79:239-242.
- Birbaum, M. H. 1974a. The nonadditivity of personality impressions. *Journal of Experimental Psychology* 102:543-561 (monograph).
- Birbaum, M. H. 1974b. Reply to the devil's advocates: Don't confound model testing and measurement. *Psychological Bulletin* 81:854-859.
- Birbaum, M. H., Wong, R., and Wong, L. 1976. Combining information from sources that vary in credibility. *Memory and Cognition* 4:330-336.
- Bremer, B. 1974. Inductive inferences from uncertain information. Um a Psychological Reports no. 78.
- Chandler, J. P. 1969. STEPIT — Finds local minima of a smooth function of several parameters. *Behavioral Science* 14:81-82.
- Hammond, K. R., and Summers, D. A. 1972. Cognitive control. *Psychological Review* 79:58-67.
- Kahneman, D., and Tversky, A. 1973. On the psychology of prediction. *Psychological Review* 80:237-251.
- Levin, I. P. 1975. Information integration in numerical judgments. *Journal of Experimental Psychology: General* 104:39-53.
- Lichtenstein, S., Earle, T. C., and Slovic, P. 1975. Cue-utilization in a numerical prediction task. *Journal of Experimental Psychology: Human Perception and Performance* 104:77-85.
- Peterson, C. R., and Beach, L. R. 1967. Man as an intuitive statistician. *Psychological Bulletin* 68:29-46.
- Rosenbaum, M. E., and Levin, I. P. 1968. Impression formation as a function of source credibility and order of presentation of contradictory information. *Journal of Personality and Social Psychology* 10:167-174.
- Rosenbaum, M. E., and Levin, I. P. 1969. Impression formation as a function of source credibility and the polarity of information. *Journal of Personality and Social Psychology* 12:34-37.
- Slovic, P., and Lichtenstein, S. 1971. Comparison of Bayesian and regression approaches to the study of information processing in judgment. *Organizational Behavior and Human Performance* 6:649-744.