Predicting with intuitive correlations¹

CLIFTON W. GRAY, DIVISION OF THE SOCIAL SCIENCES,
UNIVERSITY OF MINNESOTA AT MORRIS, 56267

An experiment in the context of man as intuitive statistician tested ability to produce an intuitive equivalent of simple linear-regression predictions as a function of the correlation between two scaled variables. Matching, rather than optimizing, characterized the resulting prediction behavior. Overmatching was most prominent when cue validity was low.

This study was done in the context of what Peterson & Beach (1967) have called "man as an intuitive statistician." The question was whether people can produce an intuitive analog to regression equations.

Method. The paradigm was to present to the S two scaled variables, X and Y, in pairs requiring the S to predict Y using X as a cue. The XY pairs were taken in random order from some controlling XY scatterplot, where X and Y were the integers 0 to 10, and where the correlation r_{xy} was of major interest.

The setting was a simulated gambling casino. The values of X and Y were integers printed on decks of playing cards dealt by E, who represented the dealer for the house.

Under an "active" payoff condition, E dealt a card showing X, the S predicted Y, then E turned the card over to show both the X and Y values together. A second payoff condition, called "active tallied" was similar and in addition S was staked with a supply of poker chips, from which he was required to surrender back to the house a number of chips, zero to nine, according to how far off his guesses were from Y. Under a "passive" condition, the S merely observed first X on the front side, then XY together on the reverse side of the card, with instructions to study the XY relationship in preparation for a later test in which only X would appear and he would be asked to predict Y.

Each of 70 cards in a shuffled deck represented one point on an XY scatterplot. Eight different scatterplots were used, representing four correlations and two shapes. The four correlations were approximately .90, .60, .30, and zero, and are called "high," "medium," "low," and "zero." Each plot had seven different X-values, with a column of 10 Y-values for each X (nine only under high correlation), making 70 points in all (63 under high). The scatterplots were constructed so as to be as close to mathematically ideal as the small number of points permitted. Each Y-column above each X was symmetrical about the regression-line value, Y-hat ($\hat{Y}$).

There were two shapes, unimodal and rectangular, equated within each correlation level on r_{xy} and the regression-line values Y-hat. The Y-arrays in unimodal plots were peaked high at each regression-line point, but with wide scatter about $Y_{\hat{x}}$. Rectangular plots distributed their Y-arrays flatly, with no peak anywhere, but compactly about $Y_{\hat{x}}$.

An S had only one of the eight plots, and he went through it four times, with careful and ostentatious shuffling by the dealer in between. Each of

these four pairing blocks was followed by a testing block. In testing blocks the S was told to instruct a proxy to make bets for him during his brief absence from the casino. The proxy must have no discretion, he would always make the same Y-guess whenever a particular X-value appeared. Thus the proxy-bets represented the S's idea of the single best Y-guess from each X.

The Ss, 96 undergraduates, competed for cash prizes ranging from \$3 to \$10 that were based upon how close their guesses were to Y. For active-tallied Ss, this meant how many poker chips they retained at the end of the game.

Results. Pairing trials and testing trials were analyzed separately. The Ss' guesses are called "G." There are four basic quantities to look at: X, Y, G, and the statistically-optimum value, Y-hat.

Predictions made by a statistically ideal S would be completely under cue control, i.e., r_{xg} would be unity regardless of the stimulus r_{xy} except where that correlation was zero. On the other hand, matching behavior, i.e., feeding back the same scatterplots that were shown in the stimuli, would make r_{xg} correspond to r_{xy} .

Figure 1 compares stimulus scatterplots with the intuitive scatterplots that they evoked, correlation by correlation. Those on top were the XY inputs, those below were the corresponding XG outputs under unimodal conditions in the final pairing trial. Figure 2 makes a similar comparison with rectangular plots. Corresponding data from the final testing trials look very similar.

The output scatterplots are group consolidations, showing all the data from all Ss under each labelled condition. Group data are reported because they were highly congruent with and presumably more stable than that of the individual median S (Table A2), whose data also support the findings which follow.

The numerical entries in Figs. 1 and 2 show behavior intermediate between optimizing and matching with substantial improvement over matching at the lower correlations. This result confirms an earlier finding (Gray, Barnes, & Wilkinson, 1965) with a similar paradigm and different stimuli. The scatterplots themselves are perhaps more indicative than this particular numerical summary of them. What is most striking is how qualitatively close the input scatterplots are mirrored by the output, both in their overall contour and in gross internal detail. Rectangular inputs make rectangular looking outputs, and unimodal inputs make unimodal outputs, demonstrating matching behavior.

Whatever may be the proportion of variance of guesses that the X-cues controlled as shown by r_{xg} , S had room for considerable variation in how he weighted his cues, as shown by the regression coefficient, b_{xg} , of the line that best fits his G-values. In a statistically ideal S, as well as in a matching S, the value would correspond to the stimulus regression line slope, b_{xy} . Results (Table B2) were that the subjective regression line slopes were too steep, though almost exactly correct with high and with zero

Fig. 1. Input (top row) vs. output with unimodal scatter plots

Fig. 2. Input (top row) vs. output with rectangular scatter plots

stimulus correlations. This result can be seen by close inspection of how points cluster in Figs. 1 and 2, and it means that most Ss did not learn during the experiment to regress toward the mean as X-cues became extreme. Inquiry after the experiment confirmed this finding, which also appeared in the earlier study (Gray et al. 1965).

Treatment effects other than correlation level were not of great interest and so are presented in Table C² and summarized briefly: correlation level, also trials (1 through 4): $p < .001$ in pairing and in testing trials. Payoff: equivocal; poker chips perhaps helped, i.e., elicited more accurate predictions, $p < .01$ in pairing but $p > .05$ in testing. Scatterplot shape: also equivocal; unimodal plots were perhaps better, $p < .05$ in pairing, $p < .01$ in testing³.

Conclusions: (1) Following limited exposure to correlated variables, Ss succeeded in producing a rudimentary scheme of simple linear-regression prediction. Matching, rather than maximizing, best describes this behavior. (2) Manipulation of the correlation between paired stimulus variables produced corresponding differences in output. (3) Confronted with correlations so low as to be almost useless, Ss did not degenerate to random responding; they overmatched. (4) Predictions with intuitive

correlations behave with sufficient regularity to warrant further study.

REFERENCES

- GRAY, C. W., BARNES, C. B., & WILKINSON, E. F. The process of prediction as a function of the correlation between two scaled variables. *Psychon. Sci.*, 1965, 3, 231-232.
PETERSON, C. R., & BEACH, L. R. Man as an intuitive statistician. *Psychol. Bull.*, 1967, 68, 29-46.

NOTES

1. The author is indebted to Julie Kraabel for research assistance in all phases of this experiment.
2. One-page tables containing (a) comparisons of pooled data with median Ss, (b) comparisons of input with output regression slopes, and (c) an ANOVA summary have been deposited with the office of this journal. Copies may be obtained without charge by writing the journal sending a self-addressed stamped envelope.
3. In a subsequent experiment, designed to replicate this particular finding about shape using sequential analysis, the null hypothesis was accepted, after testing 16 pairs of Ss, with a Beta risk of .05 against the alternative hypothesis of a difference as great as that observed in the present experiment.