DOCUMENT RESUME

ED 233 062 TM 830 510

AUTHOR Day, H. D.; Marshall, David
TITLE A Note on the Incremental Validity of Aggregate Predictors.
PUB DATE [82]
NOTE 14p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Behavioral Science Research; *Correlation; *Personality Studies; *Predictive Validity; *Research Problems; Simulation; Statistical Analysis
IDENTIFIERS *Aggregate Model

ABSTRACT

Three computer simulations were conducted to show that very high aggregate predictive validity coefficients can occur when the across-case variability in absolute score stability occurring in both the predictor and criterion matrices is quite small. In light of the increase in internal consistency reliability achieved by the method of aggregation and consequent increase in predictive validity, such high aggregate validity coefficients are expected. The particular value of the demonstrations reported here is that they show that in some applications of the method of aggregation, high aggregate validity coefficients can occur even though the actual variability in the predicted aggregate score is trivially small; in other words, little may be gained in the explanation of the variability in the criterion measure by predicting the aggregate criterion score since this predicted criterion score is so close to the overall mean of the criterion scores. The routine reporting of the standard deviation of the predicted aggregate score was recommended for all applications of the aggregation method in the study of personality. (PN)

A Note on the Incremental Validity of Aggregate Predictors

H. D. Day

Department of Psychology and Philosophy

and

David Marshall

Department of Mathematics, Computer Science, and Physics

Texas Woman's University

Mail proofs to Jim Day, Department of Psychology and Philosophy, Box 23029, Texas Woman's University, Denton, Texas 76204

Running head: Aggregate Predictive Validity Coefficients

Abstract

Three computer simulations were conducted to show that even though high aggregate predictive validity coefficients can occur, the incremental validity of the method of aggregation may be trivially small. The routine reporting of the standard deviation of the predicted aggregate score was recommended for all applications of the aggregation method in the study of personality.

A Note on the Incremental Validity of Aggregate Predictors

Although the existence of stable behavioral dispositions is crucial for theories of personality, the actual occurrence of stable behaviors across time and situations has not been demonstrated in a generally satisfactory manner (Mischel, 1969). In response to Mischel's criticism of the trait position, Epstein (1979) has argued that the stability of a behavior can be demonstrated by correlating two sets of aggregates, or averages, of the behavior in contrast to the usual practice of correlating just two observations. Such aggregations, according to Epstein, reveal behavioral stability as a result of the averaging out of the within-case error variance.

In addition to the empirical fact that recent applications of the method of aggregation have not been entirely successful in demonstrating behavioral stability (Mischel & Peake, in press a; Moskowitz & Schwartz, 1982), the method of aggregation has been criticized on both conceptual and methodological grounds (Allen, Note 1; Day, Marshall, Day & Christy, Note 2; Day, Marshall, Hamilton, & Christy, in

press; Lutsky, Note 3; Mischel & Peake, in press b; Nisbett, 1980; Peake, Note 4). The particular utility of the method of aggregation in the study of personality is the enhancement of the predictive validity of a trait measurement resulting from the aggregation of repeated measurements of the trait (Epstein, 1979). In light of the fact that the method of aggregation simply results in relatively high internal consistency reliability following an increase in "test" length (Day et al., in press), given even trivially low day-to-day consistencies, an increase in the correlation between two aggregations would be expected.

The adoption of the method of aggregation implies that the prediction of an aggregate of behavior, i.e., a general behavioral trend, is of more interest than predicting day-to-day behavior (Allen, Note 1). Aggregate scores, however, may behave statistically in ways that are unique to the behavior of the sets of day-to-day measurements contributing to the aggregate. For example, the predicted aggregate scores of all but a very few of the cases may regress to the mean aggregate criterion score (Peake, Note 4); therefore, although the aggregate predictive validity coefficient may be high, accuracy in predicting general

behavior would be enhanced for only a few cases. In other words, only a small percentage of the cases may exhibit aggregate stability (cf. Bem & Allen, 1974). The purpose of the studies reported here is to demonstrate via computer simulations the danger of this approach-to-the-mean phenomenon in the application of the method of aggregation to the study of personality.

Method and Results

Data for 200 cases were generated according to the following formula

$$f(X) = \alpha \sin(\beta X - \gamma) + \delta$$

where X was increased from $\pi/8$ to 6π in increments of $\pi/8$ ($\pi = 3.1415926$), resulting in 48 scores for each case.[1] Absolute instability occurred across these measurements as a result of the adherence of the generated scores to the sine wave form, and exaggerated absolute instability was induced by selecting an arbitrarily large amplitude ($\alpha = 3$) for the generation of these scores for each case. β was set to 1 for each case producing a constant period of 2π across cases. Since $\beta = 1$, the phase shift is given by γ, and the value of γ for a particular case was randomly

selected from the range 0 to 2π; therefore, the generated scores for the 200 cases were randomly out of phase. δ in the above function gives the vertical displacement of the baseline of the generated curve from the X axis. Relative stability was induced into the score matrix by letting δ randomly vary across cases within arbitrarily small limits. In Study 1, δ ranged from -.1 to .1; δ ranged from -.01 to .01 in Study 2 and from -.001 to .001 in Study 3. Therefore, the variability in the matrix due to aggregate stability was relatively small and increasingly small across the three studies.

Two 200 x 48 score matrices were generated in each study according to the above formula in order that the demonstration would be analogous to the situation where one aggregate is used to predict another aggregate. Relative stability was maintained across the two measurements by holding δ constant for each case to produce the two matrices.

An aggregate score for each case in each matrix was subsequently computed by taking the mean of the 48 within-case scores for each matrix. The aggregate predictive validity coefficient was then the Pearson product-moment correlation between the two sets of 200 aggregate scores.

Since the aggregate score for each case is the assigned δ value for that case and δ was held constant across the two matrices, the aggregate validity coefficient is 1.0 for each study. The standard deviation of the predicted aggregate scores (equal to the standard deviation of δ here) was .061, .0061, and .00061 for Studies 1 through 3, respectively. These standard deviations of the predicted aggregate scores are in contrast to the standard deviation of each of the 48 predictor and 48 criterion trials which ranged from 2.05 to 2.20 across the three studies.

Discussion

These demonstrations show that very high aggregate predictive validity coefficients can occur when the across-case variability in absolute score stability occurring in both the predictor and criterion matrices is quite small. In light of the increase in internal consistency reliability achieved by the method of aggregation and consequent increase in predictive validity, such high aggregate validity coefficients are expected (Day et al., in press). The particular value of the demonstrations reported here is that they show that in some applications of the method of

aggregation, high aggregate validity coefficients can occur even though the actual variability in the predicted aggregate score is trivially small; in other words, little may be gained in the explanation of the variability in the criterion measure by predicting the aggregate criterion score since this predicted criterion score is so close to the overall mean of the criterion scores. It could be said that the incremental validity (Sechrest, 1963) of the predictor aggregate is quite low.

It is, of course, not possible to know how closely the highly contrived conditions of these demonstrations may be approximated by real data; however, that such a phenomenon may occur to some degree is suggested by Epstein's (1979) report of large decreases in the standard deviation of aggregate scores as the number of aggregated trials increases. Since only the mean aggregate standard deviation for up to 12 variables occurring in eight categories are reported in Epstein's 1979 paper, it seems likely that the aggregate standard deviation for some of the variables in each category did approach zero.

This approach of the predicted aggregate scores to the mean of the aggregate criterion scores may be a serious

limitation of the application of the method of aggregation to the search for lawful trait relationships (Peake, Note 4). As a check for the occurrence of this approach-to-the-mean phenomenon, we suggest that the standard deviation of the predicted aggregate score be routinely reported and contrasted with the day-to-day or situation-to-situation trial standard deviations.

Reference Notes

1. Allen, B. P. Intraindividual variability is alive and well: Comments on Epstein's paper. Unpublished manuscript. Available from Bem P. Allen, Department of Psychology, Western Illinois University, Macomb, IL 61455.

2. Day, H. D., Marshall, D., Day, K. C., & Christy, J. The construct validity of some different conceptualizations of behavioral stability. In review.

3. Lutsky, N. Possibilities of personality: Self-reported behavior and perception of personality. In N. Lutsky (Chair), From behavior to personality: Conceptual and methodological transformations. Symposium presented at the 90th annual convention of the American Psychological Association, Washington, D.C., 1982.

4. Peake, P. K. Predicting "sums" of the person "sums" of the time. In N. Lutsky (Chair), From behavior to personality: Conceptual and methodological transformations. Symposium presented at the 90th annual convention of the American Psychological Association, Washington, D.C., 1982.

References

Bem, D. J. & Allen, A. On predicting some of the people some of the time: The search for cross-situational consistencies in behavior. Psychological Review, 1974, 81, 506-520.

Day, H. D., Marshall, D., Hamilton, B., & Christy, J. Some cautionary notes regarding the use of aggregated scores as a measure of behavioral stability. Journal of Research in Personality, in press.

Epstein, S. The stability of behavior: I. On predicting most of the people much of the time. Journal of Personality and Social Psychology, 1979, 37, 1097-1126.

Mischel, W. Continuity and change in personality. American Psychologist, 1969, 24, 1012-1018.

Mischel, W. & Peake, P. K. Analyzing the construction of consistency in personality. In M. M. Page and R. A. Dienstbier (Eds.). Nebraska symposium on motivation, 1982: Personality--current theory and research. Lincoln: University of Nebraska Press, in press. (a)

Mischel, W. & Peake, P. K. Beyond deja vu in the search for cross-situational consistency. Psychological Review, in press. (b)

Moskowitz, D. S. & Schwarz, J. C. Validity comparison of behavior counts and ratings by knowledgeable informants. Journal of Personality and Social Psychology, 1982, 42, 518-528.

Nisbett, R. E. The trait construct in lay and professional psychology. In L. Festinger (Ed.). Retrospections on social psychology. New York: Oxford University Press, 1980.

Sechrest, L. Incremental validity. In D. N. Jackson & S. Messick (Eds.). Problems in human assessment. New York: McGraw-Hill, 1967.

Footnote

[1]Forty-eight trials were used in order that the sine function would be symmetrical about some X value.