

THE RELIABILITY AND VALIDITY OF BEM AND ALLEN'S MEASURE OF CROSS-SITUATIONAL CONSISTENCY¹

JACKIE L. GREANER
LOUIS A. PENNER
GTE Data Services, Inc.

The reliability and validity of Bem and Allen's (1974) technique of measuring cross-situational consistency was investigated. Subjects answered questions concerning consistency in academic behavior and general consistency and filled out Snyder's (1974) Self-monitoring Scale twice. Whereas the test-retest reliability of the Bem and Allen type questions was quite low, the reliability of the Self-monitoring Scale was relatively high. The correlation between self-reports of variability in academic behaviour and actual variability in grades was significant, while the correlation between self-reports of general consistency and self-monitoring scores was not. The problem of using a single-item technique to measure consistency was discussed.

Theorists who argue that personality traits are important determinants of social behavior have had to contend with one rather consistent empirical finding. It is that, typically, the correlation between trait-related behaviors across situations is 0.30 or less. As Epstein (1979) has noted, until trait theorists and interactionists can explain why this correlation is so small, their theoretical position is difficult to support.

One of the more promising explanations of the 0.30 correlation was put forth by Bem and Allen (1974). According to them, prior research failed to recognize that individuals may differ in the extent to which they are cross-situationally consistent in their trait-related behaviors. Some, people are quite consistent; others are not. If one examines the behavior of the former group, a cross-situational correlation much larger than 0.30 will be found. Thus, the 0.30 correlation may reflect the inconsistency of *some* individuals rather than a general lack of cross-situational consistency among all people.

To test this proposal, Bem and Allen asked subjects to indicate, first, how much of a certain trait they possessed. Then the subjects were asked how much their behavior related to this trait varied from situation to situation. The subjects were classified as consistent or inconsistent on the basis of their response to the second question and the actual cross-situational correlation of the two groups was determined. The clearest results were obtained for the trait, friendliness. Whereas the average cross-situational correlation for the inconsistent subjects was + 0.27; it was + 0.57 for the consistent subjects. As Bem and Allen described their results, "the magic + 0.30 barrier appears to have been penetrated" (p. 514).

Despite the intuitive appeal of this approach and its fairly widespread acceptance (cf. Snyder and Monson, 1975; Snyder, 1979; Zamma et al., 1980), there are several questions about it which remain unanswered. The purpose of this study was to address three of these questions.

¹ This article is based on part of the first author's doctoral dissertation.

The first concerned the test-retest reliability of this technique of classifying subjects. Although, as noted above, the Bem and Allen technique has been used in a number of studies, we were unable to find any data on its test-retest reliability. If a measure is to be used as a moderator variable (as Bem and Allen have proposed), the measure itself should be temporally stable. Measurement theory would suggest, however, that this measure may not be a very reliable one. Single-item tests typically contain large amounts of measurement error and, thus, are unreliable (Epstein, 1979; Kerlinger, 1973). Given this possibility, we decided to empirically determine the test-retest reliability of Bem and Allen's procedure.

The second question addressed the concurrent validity of this procedure. If the Bem and Allen technique provides a valid measure of variability in some behavior, then it should produce a substantial correlation with some objective, independent index of variability in this behavior. We examined this possibility by correlating subjects' self-reports of variability in academic behavior with the variability in their academic grade point average.

Finally, we examined the concurrent validity between Bem and Allen's procedure and another purported measure of cross-situational consistency – Snyder's (1974) Self-monitoring Scale. This scale measures a respondent's general propensity to use the behavior of others as a guideline for how s/he should behave in social situations. "High" self-monitors should display little cross-situational consistency, low self-monitors should display considerable consistency. Prior research on this question has produced mixed results. Snyder and Monson (1975) reported a substantial relationship between self-reports of variability on specific traits and self-monitoring, but others (Underwood & Moore, 1981; Zanna et al., 1980) found no correlation. It should be noted, however, that whereas Bem and Allen's procedure measures trait-specific consistency, Snyder's scale measures a *general* trait which is related to consistency. In the present study, we modified Bem and Allen's technique to provide a general measure of cross-situational consistency.

METHOD

SUBJECTS

The subjects were 467 undergraduates enrolled in introductory psychology. Subjects received extra credit class points in exchange for participation.

CONSISTENCY MEASURES

Two questions concerned self-reports of academic behavior. The first asked the subjects to indicate on a 7-point rating scale how concerned they were with doing well in school. Then they were asked to indicate on another 7-point scale how much they varied in this concern. General cross-situational consistency was assessed via another 7-point scale on which subjects were asked to indicate how much their behavior varied from day to day.

Self-monitoring was assessed by Snyder's (1974) Self-monitoring Scale. This is a 25-item true/false scale with a one-month test-retest reliability of + 0.83.

Variability in grades was obtained by computing the variance in subjects' grade point average for the three academic quarters immediately preceding the collection of the consistency measures.

PROCEDURE

Subjects signed an informed consent form in which they agreed to fill out the consistency measures and give the first author permission to access their student records. Subjects provided their social security numbers (these were used to identify students' records and to match responses from the two administrations) and filled out the measures. This procedure was repeated ten weeks later. At the conclusion of the study, all subjects were debriefed.

RESULTS

TEST-RETEST RELIABILITY

There were 362 subjects who answered the questions on: concern for doing well in school, variability in this concern, and general behavioral consistency at both administrations. The ten-week test-retest reliabilities for these questions were + 0.42, + 0.43, and + 0.38, respectively. These figures can be compared with the ten-week test-retest reliability of + 0.71 obtained for the 372 subjects who filled out the Self-monitoring Scale. All of these correlations were significant at beyond the 0.01 level.

CONCURRENT VALIDITY

It was possible to obtain data on variance in grade point average from 348 subjects who had answered the question about variability in their concern about doing well. The correlation between the self-respect of variability and actual variability was + 0.21, $p < 0.01$.

Finally, there were 467 subjects who answered the question on general consistency and filled out the Self-monitoring Scale. The correlation between these two measures was + 0.11, $p < 0.01$.

DISCUSSION

The theoretical rationale underlying the Bem and Allen procedure is that it would permit the identification of cross-situationally consistent and inconsistent individuals. The results of the present study suggest that this procedure may not provide a useful means for doing this.

The major problem appears to be the temporal stability of the scores produced by this approach. A measure which yields test-retest reliabilities only slightly in excess of + 0.40 is of questionable value in the classification of subjects. In passing, it should be noted that data were also collected on the five-week test-retest of this procedure. Over five weeks, none of the test-retest reliabilities was greater than + 0.57, a better but still questionable reliability figure.

It would be inappropriate to conclude, on the basis of this study alone, that the Bem and Allen procedure is without merit. We did not examine the traits used by Bem and Allen in the original study. One could argue that the trait we considered was appreciably more vague than those considered by Bem and Allen and this was the cause of the low test-retest reliability. Further, whereas we examined the reliability of a question about general consistency, they never proposed that this procedure could be used to measure general consistency.

The main conclusion from this study, therefore, is that there *may* be serious problems with the reliability of Bem and Allen's procedure. Future studies should examine the test-retest reliability of the specific questions used by Bem and Allen.

The size of the correlations concerned with the concurrent validity of the procedure was limited by its reliability. Thus, it is difficult to make any conclusive statements in this regard. There is some reason to believe that the self-report questions on variability in academic concern was a valid measure. The correlation between answers to this question and actual variability in grades, while small, was significant and in the right direction. Further, the size of this correlation was limited not only by the reliability of the measure, but also by a restriction of range in the variability of the subjects' grades.

The data on the correlation between a self-report measure of general consistency and scores on the Self-monitoring Scale did not suggest that the two techniques were tapping the same construct. Thus, the results of this study do not support Snyder's (1979) claim that his scale and Bem and Allen's procedure are measuring the same thing. Given the difference between Bem and Allen's conceptualization of cross-situational consistency (i.e., it is trait specific) and Snyder's (i.e., it is itself a trait), this result is not surprising.

REFERENCES

- Bem, D. J., Allen, A. (1974). On predicting some of the people some of the time: The search for cross-situational consistencies in behavior. *Psychological Review*, **81**: 506-20.
- Epstein, S. (1979). Stability of behavior: 1. On predicting most of the people much of the time. *Journal of Personality and Social Psychology*, **37**: 1097-1126.
- Kerlinger, F. (1973). *Foundations of behavioral research*. Holt, Rinehart & Winston, New York.
- Snyder, M. (1974). The self-monitoring of expressive behavior. *Journal of Personality and Social Psychology*, **30**: 526-37.
- Snyder, M. (1979). Self-monitoring processes. In L. Berkowitz (Ed.), *Advances in experimental social psychology*, Vol. 12. Academic Press, New York.
- Underwood, B., Moore, B. S. (1981). Sources of behavioral consistency. *Journal of Personality and Social Psychology*, **40**: 780-5.
- Zanna, M. P., Olson, J. M., Fazio, R. H. (1980). Attitude-behavior consistency: An individual difference perspective. *Journal of Personality and Social Psychology*, **38**: 432-41.

LOUIS A. PENNER, PhD,
Department of Psychology,
University of South Florida,
Tampa, Florida 33620,
U.S.A.

Reprints of this paper are available from Dr. Penner.