

ITEM DIFFICULTY SEQUENCING AND RESPONSE STYLE: A FOLLOW-UP ANALYSIS

ALBERT D. SMOUSE AND DAVID C. MUNZ
University of Oklahoma

STUDIES have shown the Achievement Anxiety Test (AAT) to predict aptitude and achievement performance (Alpert and Haber, 1960; Dember, Nairne, and Miller, 1962; Jewell and Carrier, 1965; Milholland, 1964; Pervin, 1967), and a recent study by the writers (Munz and Smouse, 1968) has shown that individuals stereotyped on the basis of the AAT perform differently on academic achievement tests depending on whether the items are sequenced easy-to-hard (*E-H*), hard-to-easy (*H-E*), or at random (*R*). In that study it appeared from the pattern of interactions that the AAT explained more variance in a distribution resulting from achievement test items sequenced *R* than when sequenced in some other way. One of the implications suggested by these data was that when one is attempting to assess academic achievement, *H-E* sequencing should be used since it seemed to provide least variance attributable to test-taking personality factors. Random sequencing, on the other hand, appeared to yield relatively more variance which could be attributed to the AAT. If a statistical comparison showed this to be true, then it would follow that criteria, differently sequenced, are measuring different things, to say the least. Inasmuch as the above observations were based on a method whereby the above stereotypes were operationally constructed by selecting extreme scores on the AAT distributions (see the original study for the complete method), the use of all the data, including the mid-ranges of the AAT distribution, would constitute a more conservative test of this notion. Also it would make possible an extension of the initial findings to practical situations, such as the classroom, where decisions are needed on every member of the group.

Method. The *Ss* were 181 beginning psychology students at the University of Oklahoma on which AAT+ (facilitation scale) and AAT- (debilitation scale) scores had been obtained. Each subject had been assigned one of three forms of the course final examination, the forms differing only in the difficulty sequencing of the items (*R*, *E-H*, or *H-E*). Tests of the original hypothesis called for an analysis of variance design, but for the re-analysis, the two AAT scores were combined by means of multiple correlation to predict the final examination score. Thus, the same predictors were used to predict three criteria differing only as to item difficulty sequence. Analyses of variance of the regression data from the respective criterion groups permitted a comparison as to the relative amount of variance in each criterion group that could be attributed to the AAT (the regression line). These comparisons were made by means of Ryan's "method of adjusted significance levels" (1960). In addition, the criterion measures themselves were compared for homogeneity of variance.

Results. Results of the multiple correlations and the analysis of variance for the respective group regressions are shown in Table 1. Also shown in Table 1 is a comparison of the criterion variances using Cochran's test, which indicated that they were homogeneous ($C = .358$; $df = 3, 71$; $p > .05$). Table 2 shows intergroup comparisons, as to variance attributable to the AAT scales, using Ryan's method.

Discussion and conclusions. It can readily be seen from Tables 1 and 2 that the AAT explains more of the criterion variance when the items are sequenced *R* than when sequenced *E-H*. Similarly, more variance is accounted for by the AAT in the case of

TABLE 1
Summary Table for Analysis of Variance for the Regression of AAT Scales
on Three Forms of the Criterion, Showing Multiple *R* Coefficients
and Standard Deviations for Each Criterion Group

Sequence Group	Criterion SD	<i>N</i>	Multiple <i>R</i>	Mean Sq. For Regression ($df = 2$)	Mean Sq. For Deviations From Regression ($df = N - 3$)	<i>F</i>
Random	11.51	55	.515**	948.80	101.12	9.38**
Easy-Hard	10.52	54	.382**	427.56	98.15	4.36*
Hard-Easy	11.28	72	.221	219.84	124.62	1.76

* $p < .05$.

** $p < .01$.

TABLE 2

Multiple Comparison of Variance Explained (Attributable to Regression) by the Combined AAT Scores Under Three Item Difficulty Sequences Using Ryan's Method

Samples				
Degrees of Freedom	<i>R</i>	<i>E-H</i>	<i>H-E</i>	
Variance	54	53	71	
	948.799	427.564	219.843	
Summary of Computations				
Comparison Groups	<i>df</i>	Adjusted Alpha Level	<i>F</i>	<i>p</i>
<i>R</i> vs. <i>H-E</i>	54, 71	.0084	4.32	.05
<i>R</i> vs. <i>E-H</i>	54, 53	.0166	2.22	.05
<i>E-H</i> vs. <i>H-E</i>	53, 71	.0166	1.94	.05

E-H sequencing than with *H-E*. It is not surprising that there are significant correlations inasmuch as the AAT was validated against academic criteria. But the interesting part is that the amount of explained variance changes systematically across the *R*, *E-H*, and *H-E* forms. The possibility of these results being a statistical artifact due to differences in variances of the respective criterion distributions is ruled out by the fact that the criterion variances were homogeneous.

From an applied standpoint, the above results have significance in that they indicate that the predictive power of the AAT varies with the item difficulty sequence of the criterion even when scores from the middle ranges of the AAT scales are included. From a theoretical standpoint, the results offer evidence that item sequencing does affect content validity. Hence, for assessment testing, item difficulty sequencing appears to produce the noncontent-determined variance which Cronbach (1946, 1950) says should be eliminated or controlled.

While one may point to the superiority of the *R* and *E-H* criterion sequences when using the AAT as the predictive instrument, precaution should be taken to avoid the generalization that the *H-E* sequence is the ideal format for avoiding all test-taking variables which might interfere with assessment purposes. For although *H-E* sequencing minimizes variance attributable to the personality variables as measured by the AAT, it is possible that such sequencing introduces other test-taking responses, not measured by the AAT, which may contaminate content validity to an even greater extent than that indicated by the AAT scales. Thus, whether

the *H-E* sequencing reduces variance attributable to test-taking personality factors or simply substitutes one response style for another is still very much open to investigation.

Summary. Following the report of a study which suggested that the Achievement Anxiety Test (AAT) explains more achievement test variance when the criterion items are arranged randomly (*R*) than when sequenced easy-to-hard (*E-H*) or hard-to-easy (*H-E*), data already reported on plus additional data, eliminated by the original design, were re-analyzed to test the suggested hypothesis. *Ss* were 181 beginning psychology students whose AAT scores were used to predict their final examination scores, based on an exam sequenced either *R*, *E-H*, or *H-E*. Analyses of variance of the regression data from the three criterion groups permitted a direct comparison of the amount of variance attributed to the respective regression lines. The hypothesis was confirmed and interpreted in terms of response styles.

REFERENCES

- Alpert, R. and Haber, R. N. Anxiety in Academic Achievement Situations. *Journal of Abnormal and Social Psychology*, 1960, 61, 207-215.
- Cronbach, L. J. Response Sets and Test Validity. *EDUCATIONAL AND PSYCHOLOGICAL MEASUREMENT*, 1946, 6, 475-494.
- Cronbach, L. J. Further Evidence on Response Sets and Test Design. *EDUCATIONAL AND PSYCHOLOGICAL MEASUREMENT*, 1950, 10, 3-31.
- Dember, W. N., Nairne, F., and Miller, F. J. Further Validation of the Alpert-Haber Achievement Anxiety Test. *Journal of Abnormal and Social Psychology*, 1962, 65, 427-428.
- Jewell, D. O. and Carrier, N. A. Anxiety in the Measurement of Academic Performance. Unpublished paper read at American Psychological Association, Chicago in September, 1965.
- Milholland, J. E. Note on the Further Validation of the Alpert-Haber Achievement Anxiety Test. *Journal of Abnormal and Social Psychology*, 1964, 69, 236.
- Munz, D. C. and Smouse, A. D. The Interaction Effects of Item Difficulty Sequence and Achievement Anxiety Reaction on Academic Performance. *Journal of Educational Psychology*, 1968, 59, 370-374.
- Pervin, L. A. Aptitude, Anxiety, and Academic Performance: A Moderator Variable Analysis. *Psychological Reports*, 1967, 20, 215-221.
- Ryan, T. A. Significance Tests for Multiple Comparison of Proportions, Variances, and Other Statistics. *Psychological Bulletin*, 1960, 57, 318-328.