

P.
Patricia Anderson

A Comparison of Listening Tests Found Within
Sequential Tests of Educational Progress (1979)
and the Metropolitan Achievement Tests (1978)

This session reports an examination of the relationship between students' scores from the Sequential Tests of Educational Progress (STEP) Listening Test (Educational Testing Service, 1979) and scores from the Metropolitan Achievement Test (MAT) Listening Comprehension Test (Balow, Hogan, Farr, and Prescott, 1978). It was hypothesized that the students' scores on the STEP Listening Test would not be related to students' scores on the MAT Listening Comprehension Test.

The STEP is a comprehensive assessment program of educational achievement for kindergarten through twelfth grade students. Its authors purport that the battery's listening test measures students' information-processing abilities through its 20 items. Level F (Intermediate), suitable for grades 4.5 through 5.5, was used for this study. That testing component is divided into two sections, Following Directions and Listening Comprehension, each using a different testing format.

The MAT is a comprehensive program for assessing the achievement in skill and content areas of the school curriculum. Its authors claim that the MAT Listening Comprehension Test, one component of the battery, measures in its 21 items the acquisition of information through attention to aural stimuli. The Elementary Level for grades 3.5 through 4.9 of the Instructional Battery, Form J1, was used in this study. The two sections labeled by the test authors are Comprehension of Meaning and Comprehension of Syntactic Structure; both sections use the same format for evaluation within the test.

Fourth grade students from Greenville, Mississippi, participated in the eight group administration of the two listening tests. During each testing session, the order of the administration of the listening tests was varied to eliminate any

*Feb 1983 4th Annual IZLA convention
Miss MN*

fatigue or practice effect. Test procedures as recommended by the test authors were followed; a single examiner administered all tests without the use of audio equipment.

After scoring, it was found that the correlation of scores on the two tests was .411 ($p < .001$). That relationship was lower than would be expected for two tests claiming to measure the same behavior. A higher correlation between the two standardized listening tests would have provided better indices of the validity of each separate test.

Although the correlation was lower than expected between the STEP Listening Test and the MAT Listening Comprehension Test, there was a moderate percent (17%) of the variability in one test that could be explained by knowing subjects' scores on the other test. It can be concluded that the tests did indeed measure, at least in some part, some similar components of listening behavior.

The tests were further examined according to the two sub-sections described by the test publishers. Based on the descriptions of the sections of each of the two listening tests, those examining general aspects of listening behavior (the subtests on Listening Comprehension and Comprehension of Meaning) were informally hypothesized as most likely to be most highly correlated. Yet, it was found that the STEP Listening Test's section on Listening Comprehension and the MAT Listening Comprehension Test's section on Comprehension of Meaning had only a small, positive correlation ($r = .377$). When examining the other sections of both tests (Following Directions of the STEP and Comprehension of Syntactic Structures of the MAT), it was found that each of the correlations between sections were small (from $r = .243$ to $r = .377$), except for the moderate correlation ($r = .472$) between the two sections of the STEP Listening Test. The correlation between the two sections of the MAT Listening Comprehension Test ($r = .368$) would indicate that the two sections of that test examine somewhat different aspects of the behavior.

Since all correlation coefficients reported were small or moderate, it appeared that the two tests were influenced by intervening variables not yet explained by the test authors. It can be concluded that while the two listening tests appeared to measure some similar components of the listening process, they did not measure other components of that process. Prospective users of either test should consider carefully the intent of the test administration and the aspects of listening to be measured.

REFERENCES

Balow, I. H., Hogan, T. P., Farr, R. D., & Prescott, G. A. Metropolitan Achievement Tests. New York: The Psychological Corporation, 1978.

Educational Testing Service. Sequential Tests of Educational Progress. Menlo Park, California: Addison-Wesley Publishing Company, 1979.