Swanson, R. A. (1970). Auditory automotive mechanics diagnostic achievement test. Journal of Industrial Teacher Education, 1(5), 32-36.

AUDITORY AUTOMATIC MECHANICS DIAGNOSTIC ACHIEVEMENT TEST

by Richard A. Swanson[1]

For the sake of expedience, many instructional programs are evaluated with readily available, and often inadequate, criteria. Careful analysis of what the student is supposed to be able to do as a result of his training, stated in behavorial terms, is often neglected.

Readily available and more easily developed criteria instruments are usually of the paper and pencil type. In the practical arts areas of education these instruments fail to measure the gamut of behavioral changes expected of students. Little has been done in the way of developing performance tests to remedy this deficiency.

The question of the importance of performance testing is a logical one. If a test is to represent the ultimate criterion, the dimensions of the criterion must be analyzed and tested accordingly. Numerous studies have shown that the criterion of success in most cases is multidimensional. Because a person has a comprehensive understanding of automotive operational theory does not signify that he is an outstanding manipulator of automotive tools and equipment. For a truly representative test, each dimension of the criterion needs to be assessed, particularly where the several dimensions are uncorrelated. If an automotive tool and equipment manipulative test were developed, it would probably not replace the paper and pencil test of "understanding". Instead, it would add another dimension to the assessment of an automotive mechanic's ability. More important, it would give the mechanic credit for that important dimension of the criterion which he has learned.

An attempt was therefore made to develop and validate a test to measure a dimension of performance in automotive mechancis that is not now being measured by the more conventional paper and pencil tests. Specifically, the test attempts to measure the ability to diagnose malfunctions in automobiles through the auditory sense. The test is referred to as the Auditory Automotive Mechanics Diagnostic Achievement Test (AAMDAT).

Procedure

The potential sounds or test items for the AAMDAT were compiled by a consulting staff of auto mechanics and auto mechanics instructors. Those malfunctions selected for inclusion in the test were built into an otherwise perfectly running automobile and recorded with a high quality binaural recording-reproducing system. The final recordings, along with recorded instructions, made up the forty-five question, one hour long test.

The Fleishman and Friedman Auditory Aptitude Test Battery was specially shortened for use in this study.

Four populations were sampled. The first was a no-experience group that consisted of high school seniors (N=67). The second group consisted of first year post-high school auto mechanics students (N=91), and the third group were graduated auto mechanics students (N=67). The fourth group con-

[1]Dr. Swanson is Associate Professor, Department of Industrial Education, Bowling Green State University.

sisted of auto mechanics (N=44) who had anaverage of over twelve years on-the-job experience. A total of 269 people were tested.

The recorded AAMDAT and Aptitude Battery were transmitted through stereo headsets. Item alternatives were presented in a test booklet as well as on the recording.

The no-experience group and the auto mechanics received the AAMDAT. The first year auto mechanics students were administered the AAMDAT and the Aptitude Battery. The graduating auto mechanics students received the AAMDAT, the Aptitude Battery, and the Paper and Pencil Auto Mechanics Achievement Test.

Test Reliability

The item analysis pointed out ineffective item distractors, item to total test score correlations, and other detailed information that will be of great assistance in test revision.

AAMDAT reliability was estimated by the Kuder-Richardson Formula 14. The results are shown in table 1.

The low correlations found within the selected samples may be partially explained by the inclusion of ineffective test items and the limited range of scores. In an attempt to improve the reliability, the items that appeared least efficient (according to the point biserial correlations) were eliminated. The odd-even reliabilities were recalculated to obtain an appreciation of the effect on the reliability from this minimal test revision. A consistent increase in the reliability is evident. An even greater increase can be expected from the recomputation of the KR-14 reliability, since it more closely represents the data.

TABLE 1
AAMDAT RELIABILITY

Groups	N	Total Test KR-14	Total Test Odd-Even	Shortened Test* Odd-Even
No experience subjects	67	.335	.204	.230
First year auto mechanics students	91	.433	.448	.599
Graduating auto mechanics students	67	.436	.235	.333
Auto mechanics	44	.252	.268	.278
Total	269	.710	.674	.738

*Items 5, 7, 12, 18, 25, 33, 38, and 40 eliminated.

The AAMDAT reliabilities within the sub-groups are presently not high enough to warrant making fine distinctions in test achievement among those subjects.

Construct Validity

In order to test the theoretical construct that auditory diagnostic ability increases with auditory diagnostic experience, a comparison of the achievement of the subjects of the four different experience groups was made. Table 2 presents the mean scores and standard deviations of the four groups at the AAMDAT. Table 3 shows the results of a simple one-way analysis of variance.

TABLE 2
AAMDAT ACHIEVEMENT SCORES OF GROUPS

Group	X	S.D.
Automotive Mechanics	26.57	3.32
Graduating Automotive Mechanic Students	23.21	3.87
First Year Automotive Mechanic Students	21.64	3.92
No Exerience Group	14.49	3.66

TABLE 3
ANALYSIS OF VARIANCE OF AAMDAT ACHIEVEMENT SCORES OF GROUPS HAVING VARYING DEGREES OF AUTOMOTIVE EXPERIENCE

Source of Variation	Sum of Square	df	Mean Square
Between groups	4,565	3	1,521.66
Within groups	3,750	265	14.15
Total	8,315	268	
$df = 3$ $F = 107.54$ $(p < .0001)$			

The F value was found to be significant beyond the .0001 level. The theoretical construct was supported.

Concurrent Validity

The peer nomination technique was chosen as the second measure of auditory diagnostic ability because of the lack of a comparable achievement test in this area.

TABLE 4
PEER NOMINATION TO AAMDAT ACHIEVEMENT CORRELATION COEFFICIENTS

Group	N	Correlation
First year auto mechanic students	76	.256*
Graduating auto mechanic students	58	.327**
Auto mechanics	32	.161

* $p < .05$
** $p < .01$

As shown in Table 4, the first year students had a .256 ($p < .05$) correlation between AAMDAT test scores and peer nominations. The graduating students had a .327 ($p < .01$) correlation coefficient and the auto mechanics had a .161 correlation.

Although two of the three correlation coefficients were found to be statistically significant they are not of the magnitude to lend full support to the concurrent validity of the test. The low AAMDAT group reliabilities and limited ranges of test scores can account partially for these low correlations. A compounding factor is the questionable validity of peer nominations as a measure of auditory diagnostic ability. To date, the peer nomination technique has been used primarily for evaluating more complex factors, such as leadership ability. Only further study will validate this technique as a good intermediate measure of specific dimensions of sensory achievement.

Dimensions of Audition

A rotated six factor analysis on the AAMDAT results was performed.

The stems, distractors, and recordings for the items loading high in factor one were carefully studied by this investigator and several people from other fields. The conclusion was that the common element in these items was that their malfunctions made very definite sounds that were separable from the other simultaneous sounds of the running automobile. Factor one was considered a basic experience factor.

The concensus on the identification of factor two was the hypothesis that these malfunctions produced more complex sounds that were imbedded in the total automobile noise. In an attempt to psychologically describe the auditory discrimination needed for these more complex items, it was suggested that one of the auditory aptitude sub-test results would correlate highly with them. A correlation of .288 for the Tonal Memory to factor two questions was computed for the automotive mechanics students. Factor two was therefore tentatively identified as a subtle tonal memory factor.

The third factor consisted of items that were believed to contain a rhythm factor. The correlation of the results on items in factor three with the results on the Rhythm Discrimination Test was computed. The resulting coefficient of .344 bave some tentative support to the concensus that factor three was indeed a rhythm factor.

A meaningful psychological interpretation of the remaining three factors was not possible. No concensus on the underlying dimension could be found. It should be noted also that the interpretation of the first three factors is quite tentative. Only with replication can confidence be placed in these psychological interpretations.

Relationship to Auditory Aptitude

The correlations between performance on the nine auditory aptitude subtests and the AAMDAT achievement for the first year and graduating auto mechanics students were computed. These correlation coefficients, shown in Table 5, range from .052 to .323.

Relationship to Pencil and Paper Test Achievement

The test results of those graduating auto mechanics students who took both the AAMDAT and the paper and pencil Auto Mechanics Achievement

TABLE 5
CORRELATIONS OF FIRST YEAR AND GRADUATING AUTO MECHANICS STUDENTS PERFORMANCE ON AUDITORY APTITUDE SUB-TESTS TO AAMDAT ACHIEVEMENT

Sub-Tests	Correlation	N
Pitch discrimination	.055	157
Loudness discrimination	.052	157
Rhythm discrimination	.121	157
Time discrimination	.074	154
Timre discrimination	.066	154
Tonal memory	.279	153
Copying behind test	.102	144
Dot perception test	.142	145
Hidden tunes test	.323*	128

* $p < .01$

The multiple correlation of the AAMDAT with the Auditory Aptitude Battery, without any corrections, was .41 ($p < .001$).
Test were correlated. A correlation coefficient of .249 was obtained. With a correction for range restriction on the AAMDAT the correlation between the AAMDAT and the paper and pencil test would be expected to increase from .249 to .346. Even with this increase in the AAMDAT range, it is obvious from the relatively low initial correlation that the AAMDAT is at least partially measuring a dimension not being measured by the paper and pencil test. The results lend support to the need for refining these independent measures of sensory performance.

Implications for Further Research

1. The most obvious implication for further research is to use the available item analysis data to revise the AAMDAT. A much larger sample should then be tested to establish norms.

2. The support this study gave to the probable independence of the performance and cognitive test suggests a greater need for multi-dimensional performance testing. This can be done primarily in areas where sensory discriminations are an important part of the criterion.

3. The validity of the peer nomination technique and other intermediate criteria of success in sensory achievement would assist in validating newly developed performance tests.

4. Further research on the relationship of auditory aptitude and auditory achievement as it relates to occupational success is needed.

5. Many of the sensory dimensions that are difficult to test and yet are an important part of an ultimate criterion, are also difficult dimensions to teach systematically. For the most part, auditory diagnostic ability is taught on a hit-or-miss basis in automotive mechanics programs. In this effort to evaluate achievement in this dimension of auto mechanics, a high quality binaural test tape was developed. This same tape could easily be converted into a training instrument. Thus, a student in training could receive the equivalence of five years on-the-job auditory diagnostic experience in a few hours through headsets and a carefully prepared training tape. It is possible that research concerned with developing instructional materials in these more difficult areas could do well by taking an evaluation approach to help provide instructional answers.