A STANDARDIZED ACHIEVEMENT TEST FOR HIGH SCHOOL HORTICULTURE STUDENTS

Henry G. Mityga
Department of Horticulture
Clifford L. Nelson
Ronald J. Seibel
Department of Agricultural and Extension Education
University of Maryland

Standardized achievement tests have become well accepted tools and are used extensively by professionals in the education field. Vocational agriculture has always recognized the need for suitable testing instruments. However, development of these tests has been rather limited. Recently, Seibel (1972) developed *The Plant and Soil Science Test* which was field tested in six states. A testing battery was developed at the University of Nebraska (Harvill, 1973) which provides scores for four areas: animal science, plant and soil science, mechanics, and management.

The use of standardized tests for research and evaluation in vocational agriculture has always been regarded as useful. The Vocational Education Act of 1963 and its amendments of 1968 strongly encouraged evaluation of federally funded vocational programs. The 1976 Education Amendments, Title II Vocational Education (P.L. 94-482), mandates evaluation of these programs.

The Problem

The problem consisted of the development of an instrument which was capable of measuring the achievement of students in the subject matter of horticulture.

The Procedure

The initial part of this study dealt with the development of a testing instrument based on item analysis of several tryout tests.

Letters were sent to the state supervisors of education in eight states asking for information about which high schools offer horticulture programs, the names of instructors concerned with these programs, and their addresses.

Journal of the American Association of Teacher Educators in Agriculture
Volume 19, Number 2, pp.11-16
DOI: 10.5032/jaatea.1978.02011

A subject matter outline for the test was developed, based on booklets and manuals that present suggested curriculums for development of horticulture programs at the high school level. Two subject matter specialists reviewed this outline. Table 1 presents the outline including the suggested changes and the mean weighted percentage for each content area.

Table 1

MAJOR AREAS FOR THE DEVELOPMENT OF AN ACHIEVEMENT TEST IN HORTICULTURE AND THE MEAN WEIGHTED PERCENTAGE ASSIGNED TO EACH AREA

Major Areas	Mean Weighted Percent of Test
1. Classification of plants	4
2. Plant structures and their function	3
3. Necessary nutrients for plant growth	4
4. Determination of fertilizer needs and correcting nutrient deficiences	10
5. Plant growth and development	3
6. Plant propagation	12
7. Turf management	12
8. Landscape and design principles	4
9. Floral design	4
10. Production of specific crops and cultural practices	16
11. Growing structures and other plant production equipment	10
12. Soil media and soil amendments	8
13. Identification and control of insects, diseases and weeds	10
TOTAL	100

For each of the major content areas of the subject matter outline, instructional objectives were written in the manner suggested by Mager (1962). A set of specifications for the development of an achievement test was drafted (Table 2). The relative weights assigned to each of the types of ability to be measured was based upon the proportions as suggested by Ebel (1956).

Content Areas Measured	Types of Ability Measured: Vocabulary and Factual Recall	Understanding and Generalizations	Application and Problem Solving	Total Per Content Area
1. Classification of plants	2%	2%	0%	4%
2. Plant structures and their functions	1%	2%	0%	3%
3. Necessary nutrients for plant growth	3%	1%	0%	4%
4. Determination of fertilizer needs and correcting nutrient deficiencies	2%	5%	3%	10%
5. Plant growth and development	2%	1%	0%	3%
6. Plant propagation	4%	6%	2%	12%
7. Turf management	5%	6%	1%	12%
8. Landscape and design principles	1%	3%	0%	4%
9. Floral design	1%	3%	0%	4%
10. Production of specific crops and cultural practices	5%	9%	2%	16%
11. Growing structures and other plant production equipment	4%	5%	1%	10%
12. Soil media and soil conditioners	5%	3%	0%	8%
13. Identification and control of insects, diseases, and weeds	5%	4%	1%	10%
TOTAL PERCENT	40%	50%	10%	100%

Over 200 multiple choice questions were written for an item pool. Each item offered four possible responses. A number of textbooks, teaching manuals, and publications dealing with horticultural practices were used as resource information in the construction of the items. Two 80-item tryout tests were constructed and a set of instructions for test administration was drafted. The items were reviewed by a subject matter specialist and changes were made on the basis of the recommendations received.

Nine high schools in the State of Maryland cooperated in the study by testing their students with the tryout tests. The authors personally administered the test in several schools in order to determine if there were any major problems in test administration or construction. Of the 328 returns, 168 examinees took test one and 159 took test two. In analyzing the data from the tryout tests, a discrimination index and a difficulty index was determined for each item and also for each alternative response to each item. A t test was performed on the correlation coefficient of each item to determine its level of significance.

The final test was constructed only of items which were significant at the 0.05 level or better. If several items fit into one of the categories within the table of specifications, then the one(s) with the best indices of effectiveness was (were) selected. The final test was composed of 80 items. Based on the tryout item analysis, the mean level of difficulty for the final test was 0.39 and item difficulties ranged from 0.23 to 0.86.

A letter was drafted and sent to horticulture instructors in 83 high schools in seven states asking them to administer the final horticulture test to their students. Of the 47 horticulture instructors who indicated they would be willing to participate in the study, 37 returned data. The states participating and the number of schools within each which returned data (in parentheses) are: Delaware (5), Illinois (6), Indiana (4), Maryland (7), New York (6), Ohio (6), and Virginia (3).

The final test was analyzed to determine the effectiveness of the items and to establish preliminary norms. For each item, a discrimination index and a difficulty index were computed. The same information was computed for each distractor in order to judge its effectiveness.

The Fog Index Formula (1968) was used to determine the lowest grade level capable of comprehending the test. This grade level was detemined as 8.1.

Findings

Test data were received from 1,485 horticulture students. Of the examinees, 785 had received one year of horticultural instruction, 198 two years, and 48 three years. When examinees were grouped according to grade level, there were 129 ninth graders, 397 tenth graders, 540 eleventh graders, and 406 twelfth graders. Norm tables for each of the above groups were established and presented data on percentile ranks, raw scores, *T* scores, mean, standard deviation, standard error of measurement, and standard error of mean.

An analysis of the final Horticulture Achievement Test, comparing desired and observed specifications, is presented in Table 3. Letters were sent to 83 high schools, which offered horticulture as part of the curriculum, asking them to participate in the testing program. It was hoped that a two-thirds response rate would provide an adequate sample assuming an average of 25 horticulture students per high school. Although only 57% of the schools responded as willing to participate, it was found the average number of horticulture students per school was 52.

Since the test was constructed according to a table of specifications, the desired specifications for relevance were compared with the observed specifications. Because of the close adherence of the observations to the desired specifications, the test is assumed to be high in content validity.

The range of item difficulty for the final test was 0.22 to 0.85 and the mean level of difficulty was 0.49. Wood (1960) has suggested that the difficulty of items in a test should range from 0.15 to 0.85 with an average difficulty of 0.50. The upper level was slightly greater than the recommended range. However, only one item had one index greater than 0.85.

The placement of discrimination indices into four levels of comparison showed that the observed values far exceeded the minimum-acceptable specifications. The item discrimination levels were rated according to the guidelines proposed by Ebel (1956).

The observed mean for the final test was found to be 39.91. This is very close to the ideal, or 40.00, since a test which has a mean one-half of the total raw score is expected to provide the best discrimination among examinees. The standard deviation was 15.35 and the range of scores was 2 to 79. The observed reliability coefficient was 0.93, while the minimum desired was 0.85. The minimum desired specification for the reliability coefficient was based on values reported for commercially available biology testing instruments which center

around 0.85. The standard error of measurement was 4.06, and the standard error of the mean was 0.40.

Table 3

ANALYSIS OF THE FINAL HORTICULTURE ACHIEVEMENT TEST

Characteristics		Specifications Desired	Observed
I.	Number of examinees	1390	1485
II.	Relevance		
	A. Vocabulary and factual recall	40%	40%
	B. Understanding and generalization	50%	50%
	C. Application and problem solving	10%	10%
III.	Item difficulty		
	A. Range	0.15-0.85	0.22-0.86
	B. Mean	0.50	0.49
IV.	Discrimination		
	A. Item		
	1. High (0.40 and above)	More than: 25%	46%
	2. Moderate (0.21-0.40)	More than: 25%	54%
	3. Low (0.01-0.20)	More than: 15%	1%
	4. Zero or negative	Less than: 5%	0%
	B. Score		
	1. Mean	40	39.91
	2. Standard deviation		15.35
	3. Range of scores		2-79
	4. Reliability coefficient	More than 0.85	0.93
	5. Standard error of measurement		4.06
	6. Standard error of the mean		0.40

Summary

The norms are considered preliminary because of the sampling procedure. Forty-four percent of the schools which were asked to participate returned data. This may not be a truly random sample and may be biased to those schools which felt they had a good horticulture program and were willing to have their students evaluated in comparison to others. Some of the schools

(Continued on page 30)