

PENCIL AND PAPER AND OPERATIONAL PERFORMANCE TESTS  
AS MEASURES OF STUDENT MASTERY  
FOR SMALL ENGINE INSTRUCTION

*William Benjamin Symons*      and      *Richard H. Wilson*  
*Agricultural Industries*      *Agricultural Education*  
*Southern Illinois University*      *The Ohio State University*

Many educators, administrators, and advisory persons associated with the vocational education profession are concerned about the validity of using traditional objective-type examinations to determine student ability to perform an array of skill-type assignments.

A rallying point for most areas of vocational education is the "hands on" philosophy of instruction. Many vocational educators would conclude that along with "hands on" training must come "hands on" evaluation in order to verify that individuals were capable of performing tasks for which they were being trained. This study was developed to determine if an operational performance evaluation was valid and feasible for an area of small engine instruction, a subject common to the programs of many teachers of agriculture.

*Hypotheses*

The hypotheses were as follows:

1. There will be a positive relationship between student scores on pencil and paper and operational performance tests used in this study.
2. There will be a positive relationship between student scores on the non-manipulative portion of the operational performance test and student scores on the pencil and paper test.
3. Students completing the operational performance test prior to the pencil and paper test will score higher on the pencil and paper test than students who completed the pencil and paper test prior to the operational performance test.

*Journal of the American Association of  
Teacher Educators in Agriculture  
Volume 20, Number 2, pp.11-17  
DOI: 10.5032/jaatea.1979.01011*

4. Students completing the pencil and paper test prior to the operational performance test will not score significantly higher on the operational performance test than students who completed the operational performance test first.

### *Design*

The design of the study is shown below. Student scores on the two tests served as the statistical unit of analysis.

Replication	Sequence A		Sequence B	
$R_1$	$O_1$	$O_2$	$O_2$	$O_1$
$R_2$	$O_1$	$O_2$	$O_2$	$O_1$
$R_3$	$O_1$	$O_2$	$O_2$	$O_1$
$R_4$	$O_1$	$O_2$	$O_2$	$O_1$

$O_1$  = performance test

$O_2$  = pencil and paper test

$R_1, R_2, R_3, R_4$  = intact classes from four schools randomly selected from a population of 24

### *Evaluative Instruments*

*Pencil and Paper Test.* An instrument was developed which would measure both the learner's capability on a pencil and paper test and an operational performance test. Instrumentation on the pencil and paper test consisted of 50 multiple choice items.

*Operational Performance Test.* In order to create instrumentation for measuring student operational performance, a series of 36 test displays was developed. This series of displays consisted of engine components, instruction and response captions, and related hand tools. One or more operational performance test displays were placed on 2 foot square plywood panels for storage, transportation and test administration. The test scenario called for presenting panels in sequence before individual students who moved in rotation.

### Findings

*Descriptive Statistics on Tests.* A summary of descriptive statistics for operational performance test and pencil and paper test is shown in Table 1.

Table 1  
SUMMARY OF DESCRIPTIVE STATISTICS FOR TESTS

Statistic	Operational Per- formance Test	Pencil and Paper Test
Mean Score	22.16	29.55
Median	22	31
Mode	22	23
Maximum Score	32	44
Minimum Score	11	10
Reliability Estimates -		
Kuder-Richardson 20	0.786	0.838
Kuder-Richardson 21	0.725	0.779
Mean Item Difficulty	0.385	0.409
Mean Item Discrimination	0.370	0.346
Number of Items	36	50
Number of Students	64	64

In testing hypotheses one and two, a correlation coefficient for each was calculated using the following equation:

$$r = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sqrt{(\sum x^2 - \frac{(\sum x)^2}{n}) (\sum y^2 - \frac{(\sum y)^2}{n})}}$$

The correlation coefficient calculated for testing hypothesis one was  $r = 0.69$  which is significant at the 0.01 level as plotted in Figure 1. The number "2" found in the Figure indicates that two students had the same scores on both tests.

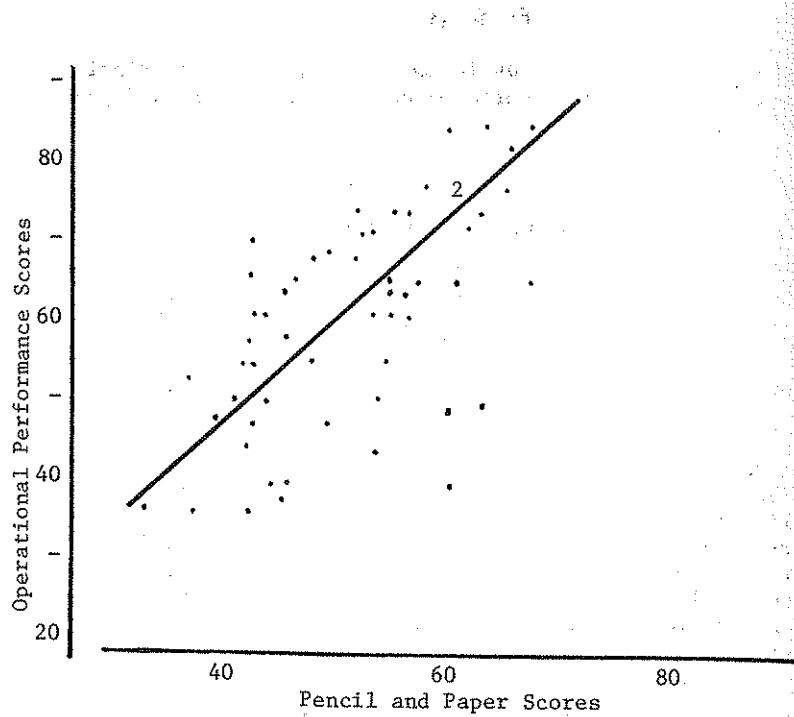


Figure 1

STUDENT SCORES ON PENCIL AND PAPER TEST AND  
OPERATIONAL PERFORMANCE TEST

The correlation coefficient calculated for testing hypothesis two was  $r = 0.23$ , which is not significant at the .05 level. This hypothesis was rejected. The compared student scores on the two tests are shown graphically in Figure II. The number "2" found in the Figure indicates that two students had the same scores on both tests.

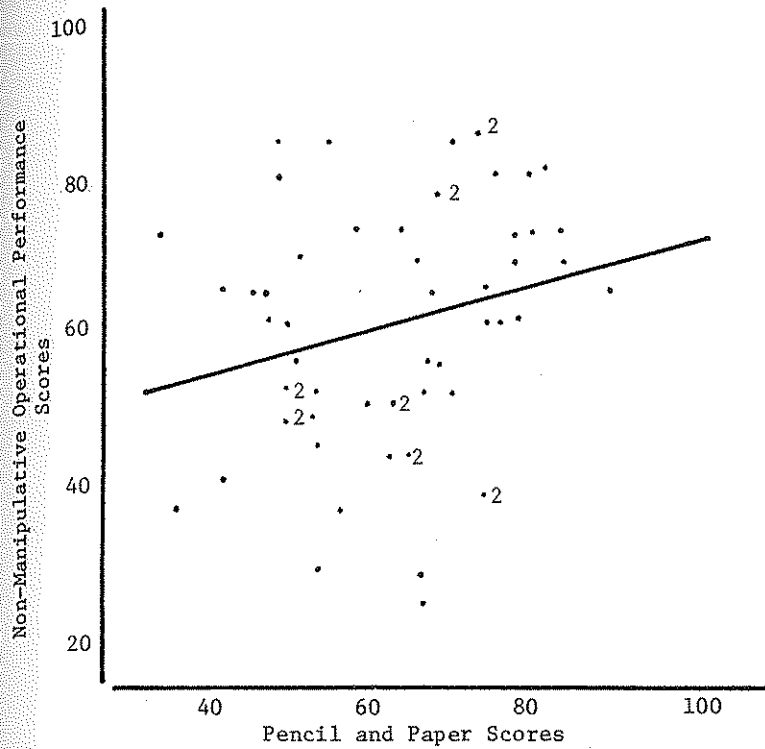


Figure 2

STUDENT SCORES ON PENCIL AND PAPER TEST AND  
NON-MANIPULATIVE PORTION OF OPERATIONAL  
PERFORMANCE TEST

Hypotheses 3 and 4 were tested using a two-way analysis of variance. The results are shown in Tables II and III, respectively. Hypothesis 3 was rejected and Hypothesis 4 was accepted.

Table II  
TWO-FACTOR ANOVA ON STUDENT SCORES  
ON PENCIL AND PAPER TEST

Source of Variation	DF	SS	MS	F	P
School	3	2023.29	674.43	5.4	<.01
Test Sequence	1	72.25	72.25	0.59	NS
School x Test Sequence	3	735.45	242.15	1.97	NS
Within School - Test Sequence Groups	56	6876.8	122.8		
Total	63	9707.79			

Table III  
TWO-FACTOR ANOVA ON STUDENT SCORES  
ON OPERATIONAL PERFORMANCE TEST

Source of Variance	DF	SS	MS	F	P
School	3	5705.04	1901.68	13.94	<.01
Test Sequence	1	56.25	56.25	0.41	NS
School x Test Sequence	3	910.41	303.47	2.22	NS
Within School - Test Sequence Groups	56	7678.30	136.39		
Total	63	14350.00			

In both ANOVA's, R<sub>1</sub> (School A) scores were found to be significantly different from the other three schools sampled. Follow-up information gathered tended to indicate that the scores for school A were higher than others due to the specialization on small engines of one school year which greatly exceeded the length of time instruction was given in the other three schools.

#### References

- Cronbach, Lee J. *Essentials of Psychological Testing*. New York: Harper and Brothers, 1949.
- Glass, Gene V. and Stanley, Julian C. *Statistical Methods in Education and Psychology*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1970.
- Guilford, J. P. *Fundamental Statistics in Psychology and Education*. New York: McGraw-Hill Book Co., 1965.
- Lefkowitz, E. F. "The Effect of Pictorial Stimuli Similarity in Teaching and Testing." Unpublished Ph.D. Dissertation, Pennsylvania State University, 1955.
- Scholl, Paul A. "Experimental Design of an Objective Verbal-Pictorial Test for Audio-Visual Equipment Operations Skills." Paper read at the National Convention of Departments of Audio-Visual Instruction (NEA), Houston, Texas, March 24-29, 1968.
- Thelen, Herbert A. "Testing by Means of Film and Slides with Synchronized Recorded Sound," *Educational and Psychological Measurement*, Spring 1945, 5:33-48.
- Warmbrod, J. Robert, Principal Investigator. "Criterion-Referenced Instruments for the Assessment of Specialized Vocational Agriculture Programs." Final Report, Columbus, Ohio, 1974.
- Williams, Fred E. "An Experimental Comparison of Equivalent Verbal and Verbal Pictorial Tests in a University Audio-Visual Course." Unpublished Ph.D. Dissertation, Indiana University, 1961.