R E P O R T   R E S U M E S

ED 016 850

VT 004 295

DEVELOPMENT AND UTILIZATION OF A NATIONAL VOCATIONAL-TECHNICAL SCHOOL ACHIEVEMENT TESTING PROGRAM USING THE PRINTING TRADES AS A PILOT AREA. FINAL REPORT.

BY- MCFADDEN, DENNIS
OHIO TRADE AND INDUSTRIAL EDUCATION SER., COLUMBUS

PUB DATE AUG 67
CONTRACT OEC-5-85-011
EDRS PRICE MF-$0.75 HC-$5.36 132P.

DESCRIPTORS- *PRINTING, *TRADE AND INDUSTRIAL EDUCATION, *STUDENT EVALUATION, *PERFORMANCE TESTS, *ACHIEVEMENT TESTS, TEST RELIABILITY, TEST VALIDITY, STUDENT TESTING, NATIONAL COMPETENCY TESTS, GRADE 12, TASK PERFORMANCE, ACHIEVEMENT, TEST CONSTRUCTION, PERSONNEL EVALUATION, OHIO PRINTING ACHIEVEMENT TEST, OHIO PRINTING PERFORMANCE TEST,

THE PURPOSE OF THIS PILOT PROJECT WAS TO DEVELOP AND VALIDATE TESTS TO ASSESS STUDENT ACHIEVEMENT IN TWELFTH GRADE VOCATIONAL PRINTING PROGRAMS. THE OHIO PRINTING ACHIEVEMENT TEST AND A NEW TEST DEVELOPED BY THE PROJECT, THE OHIO PRINTING PERFORMANCE TEST, WERE ADMINISTERED TO A SAMPLE OF STUDENTS WHO WERE NEARING COMPLETION OF THE 4-YEAR VOCATIONAL PRINTING COURSE IN 77 HIGH SCHOOLS IN 29 STATES AND THE DISTRICT OF COLUMBIA, AND NATIONAL NORMS WERE ESTABLISHED. A JOB PERFORMANCE CRITERION MEASURE WAS DEVELOPED WITH WHICH TO VALIDATE THE TWO ACHIEVEMENT MEASURES, AND MEASURES OF JOB PERFORMANCE WERE OBTAINED FROM THE JOB SUPERVISORS OF GRADUATES WHO HAD PARTICIPATED IN THE NORMING ADMINISTRATION. IT WAS CONCLUDED THAT THE THREE TESTS WERE VALID AND RELIABLE AND THAT VOCATIONAL GRADUATES WHO ENTERED PRINTING OCCUPATIONS HAD SIGNIFICANTLY GREATER MEAN LEVEL OF ACHIEVEMENT THAN THE TOTAL NORM GROUP. A LIST OF REFERENCES, A SUGGESTED PRINTING COURSE OUTLINE, THE JOB PERFORMANCE RATING SCALE, A LIST OF PARTICIPATING SCHOOLS, AND MULTIPLE REGRESSION ANALYSIS TABLES ARE INCLUDED. (EM)

FINAL REPORT

Contract NO. OE-5-85-011

DEVELOPMENT AND UTILIZATION
OF A NATIONAL VOCATIONAL-TECHNICAL SCHOOL ACHIEVEMENT
TESTING PROGRAM USING THE PRINTING TRADES AS A PILOT AREA

August, 1967

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

DEVELOPMENT AND UTILIZATION OF A NATIONAL VOCATIONAL-TECHNICAL SCHOOL ACHIEVEMENT TESTING PROGRAM USING THE PRINTING TRADES AS A PILOT AREA

Contract Number .OE-5-85-011

Dennis Mc Fadden

August, 1967

The research reported herein was performed pursuant to a contract with the Office of Education, U. S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

Ohio Trade and Industrial Education Service
Division of Vocational Education

Columbus, Ohio

TABLE OF CONTENTS

TABLE OF CONTENTS (Contd.)

TABLE OF CONTENTS (Contd.)

LIST OF TABLES

LIST OF TABLES (Contd.)

ACKNOWLEGMENTS

This research report is the result of eighteen months of study. The purpose of the project was to develop valid and reliable measures of achievement in vocational trade and industrial education, using the printing trades as a pilot area of study. It was prompted by the interest of Mr. Sam Burt, Executive Secretary of the International Graphic Arts Education Association.

Mr. Howard Massman, Printing Trades Coordinator, Dayton Public Schools served as Committee Chairman for the study and whose dynamic efforts made this project successful. Members serving on the test development committee were:

Mr. Bill Flack, International Graphic Arts Education Association, Washington, D.C.

Dr. Jerry Olson, Curriculum Coordinator for Occupational, Vocational, and Technical Education, Pittsburgh, Pennsylvania

Dr. Gene Strandberg, Kansas State College, Pittsburg, Kansas

Mr. Arch Macready, Executive Secretary, International Association of Printing House Craftsmen, Inc., Cincinnati, Ohio

Mr. Thomas Spitzig, Manager, Organization Development, McCall Corporation, Dayton, Ohio

Mr. Don Hayes, Printing Instructor, Anderson, Indiana

Job Performance Criterion Measures were developed by:

Mr. Wilbert Rosenthal, President of the International Association of Printing House Craftsmen, Inc., Cincinnati, Ohio

Mr. William Smith, Education Director,
Graphic Arts Technical Foundation, Inc.
Pittsburgh, Pennsylvania

Mr. R. Reid Vance, Treasurer and Past Executive
Secretary of the Printing Industry of Ohio,
Columbus, Ohio

The statistical research was directed by Mr. Dennis McFadden.

Recognition is extended to the following personnel for their guidance throughout the study: Mr. Harry F. Davis, State Supervisor, Trade and Industrial Education Service, Mr. Frank A. Oliverio, Assistant State Supervisor and Dr. Robert M. Reese, Department Head at The Ohio State University. Mr. W. F. Stover, Consultant, Instructional Materials Development Trade and Industrial Education handled the administration of the project and Mr. George Kosbab assisted in the data collection and coordination of the study through the Instructional Materials Laboratory, and served as editor of this report.

The Ohio Department of Finance, Data Processing Center provided the necessary data processing facilities for this study and their cooperation is commended.

The cooperation of the various industrial concerns and their management, state supervisors of trade and industrial education, school superintendents, principals, supervisors, instructors, and students has made this study possible. We are indebted to these people along with the counselors and other school personnel for their time and effort especially during the testing phase of the study.

Byrl R. Shoemaker, Director
Division of Vocational Education

CHAPTER I

OVERVIEW OF THE STUDY

Statement of the Problem

Trade and Industrial Education programs in the school systems of the United States today generally lack valid and reliable instruments for either assessing achievement or predicting vocational success of their students. Specifically, there is lack of valid instruments for measuring the level of occupational preparation, particularly at the end of grade 12, which could be used to predict job performance, lack of generally accepted criteria for evaluating effectiveness of the education and training programs offered; and lack of participation in the school program by local business and industry employers. In sum, there are no adequate means of measuring the effectiveness of the various training programs.

This problem is as broad as it is complex with many areas needing investigation. Identification of differential criteria of success in various occupations, measurement of individual characteristics which contribute to or detract from success in various vocations, experimental evaluations of new methods of teaching vocational education, etc. are but a few of the problems which need investigation. Obviously, no single study can encompass all the facets of a problem so broad as the differential assessment of the effectiveness levels of various trade and industrial education programs. A comprehensive program of research is needed: one which can study student and instructor characteristics, pedagogy, measurement of differential aptitudes or achievement, intermediate and ultimate success criteria, and development of prediction schemes - all as they relate to long-term job success for today's youth.

Not only is there a lack of information about what makes for effective trade and industrial instruction, but we do not even know how to develop instruments for predicting job performance for students in trade and industrial education programs or for generally assessing the effectiveness of the various programs. What is desperately needed is the identification and validation of methodologies for developing evaluation and prediction instruments and also development of methods for identifying relevant success criteria.

Specifically, the purpose of this pilot project is to conduct test development and validation activities designed to assess the achievement of students in printing training programs.

Because of the high cost factor involved in providing trade and industrial courses at the secondary level, a system of evaluation which did judge the merits of instruction on the basis of a job performance criteria was needed. The cost factor for adequate training was much too critical not to provide for the continuous feedback of information directed at keeping pace with the technological changes and the expectations of industry for training in the trade and industrial occupations.

Purpose of the Study

The criterion problem for judging program effectiveness was of interest to the Trade and Industrial Education Service in Ohio. Therefore, they contracted with the United States Office of Education to determine the feasibility of developing valid and reliable measures of student achievement and job performance for evaluating the effectiveness of training. The pilot subject area selected for the study was printing at the twelfth-grade level. Printing was selected because of its representativeness to the kinds of technical knowledges and skills generally emphasized in trade and industrial courses, and, in particular, to the kinds of problems posed for measuring student achievement and subsequent occupational performance.

Design of the Study

The design adopted in support of the purpose of this study was a longitudinal correlational analysis in which a dual set of criteria was specified. The first step was the development and standardization of an achievement test series designed to measure the technical knowledge and skills relevant to the aims of a vocational course in printing at the twelfth-grade level of instruction. The second step was the development and standardization of a job performance measure against which to validate the newly developed achievement test battery. For it was the intent of this study to determine a system of measures for relating the end of training achievement to relative success in the occupation for which that training had been conceived.

A longitudinal design, providing for the attainment of two sets of measures on the same group of students nearing completion of a secondary vocational printing course, was to be tested with the achievement measures. Later, after eight to ten months, estimates of job performance were then to be obtained for the students who became employed in the printing trades. The two sets of measures on the same group of students were then to be judged as an index for evaluating the training.

Objectives of the Study

The specific objectives in support of the general purpose of this study were as follows:

A. Measuring Student Achievement

1. To define student achievement representative of the educational objectives for a vocational printing course at the twelfth-grade level.

2. To select existing valid and reliable measures of student achievement and to provide for the development of such measures.

3. To provide for the national standardization of the conceived achievement test battery on a representative student population nearing the completion of a vocational printing course at the twelfth-grade level.

B. Measuring Employee Job Performance

1. To define employee job performance representative of newly employed graduates of vocational secondary printing courses.

2. To provide for the development of valid and reliable measures of job performance.

3. To obtain from the students originally tested and subsequently employed in the printing trades measures of job performance.

C. Relating Measures of Achievement to Job Performance

1. To provide an index for relating on-the-job performance measures to the original student achievement obtained from the standardization.

Organization of the Report

Chapter I describes the problem which led to the present research together with the rationale underlying the purpose and the objectives of the study.

Chapter II gives the assumptions and hypotheses upon which the achievement tests, job performance measure, and the index for relating the two were found. This chapter also describes the procedures and methods used to accomplish the purpose of the study.

Chapter III presents the results obtained in support of the proposed objectives.

Chapter IV analyzes and interprets the results of the study.

Chapter V discusses the conclusions and implications of the results.

Chapter VI provides a review of the study's findings and conclusions.

CHAPTER II

METHODS

Measuring Student Achievement

Several conferences were arranged with representatives of the International Graphic Arts Education Association for the purpose of defining student achievement in terms of the technical knowledges and skills generally desired for vocational secondary printing. An analysis of the curriculum had been completed in the form of a course outline describing the content areas normally covered in such a course of training (Appendix A). An examination of this outline suggested the measurement of two general categories of achievement. One was concerned with the knowledge and understanding of fundamental operations involved in printing. The other general category was more concerned with the ability to apply such knowledge and understanding to problem-solving situations normally encountered in the printing industry. After having identified these two categories of achievement, the attention was directed toward examining what methodological problems might emerge in designing and structuring testing situations for the demonstration of such achievement.

Concerning the knowledge and understanding category, the committee adopted the Ohio Printing Achievement Test. The existing form of this test was a multiple-choice-type achievement measure whose item construction was based on the suggested course outline. Its technical design included a 0.97 reliability coefficient and a 8.71 standard error at measurement. Other item analysis data were also available. There were 16 sub-tests in the Ohio Printing Achievement Test with a total of 340 multiple-choice items. The sub-tests were as follows: (1) Orientation, (2) Printing Planning, (3) Hand Composition, (4) Machine Composition, (5) Photo Composition, (6) Camera Operation, (7) Film Processing, (8) Letter Press Platemaking, (9) Letterpress Presswork, (10) Applied Science, (11) Lithographic Stripping and Platemaking, (12) Lithographic Presswork, (13) Binding Work, (14) Paper

Technology, (15) Ink Technology, and (16) Applied Mathematics. The test is administered in two parts with Part I covering sub-tests 1-10 and Part II sub-tests 11-16. Administration time for Part I was 2½ hours while 1½ hours were allowed for Part II.

The second general category of achievement emphasized the student's ability to make use of what he knows in problem-solving situations encountered during the printing process. Measuring such achievement suggested a performance testing situation in which achievement is demonstrated under actual or simulated conditions of printing. How to standardize the test administration was the primary methodological problem posed for this kind of testing situation. This problem was especially critical since the quality and quantity of equipment, tools, and materials varied extensively between schools.

Another aspect considered was the time and cost of scoring individually with the help of skilled craftsmen where valid and reliable scoring procedures could be guaranteed. Aside from the practical constraints of time and cost incurred from administering such a test, there remained the very real possibility that sources of scorer error resulting from inadequate standardization would seriously disguise the student's true achievement.

Some aspects of achievement could have been tested with this technique because the conditions for standardization would not be too difficult to establish. However, not all areas of achievement in this second category were amenable to this technique. The adoption of a system of measurement whose design and structure could satisfy the constraints of administration time, adequate standardization, and objective scoring was needed.

<u>System of Measurement Employed</u>

The recognition type of measurement was selected to satisfy the above constraints for the performance measure. This type of measurement provides for such important aspects of achievement as the ability to (1) recognize essential characteristics of performance (product or process), (2) choose the solution for a defined operational problem, (3) judge the accuracy of specimens or products, (4) identify critical operations, functions, or parts of a particular series, and (5) examine, locate, and identify product defects and indicate probable causes. These kinds of testing operations were intrinsically similar to actual performance operations under conditions where the validity and reliability of measurement was controlled.

This type of testing situation was most amenable to other features, for example, the development of (1) test operations intrinsically similar to the problem-solving situations encountered in the printing process, (2) suitable time allowances for administration because of its applicability to group testing, and (3) objective scoring systems employable by personnel not trained in printing. Many of the desirable elements of this measurement system were incorporated when standardized conditions for stimulating the actual printing conditions were not feasible.

The selected measurement objectives in this second general category of achievement were identified as five phases of printing.

1. Planning - The ability to visualize the appearance of a job as a final product requires the achievement of the following:

   A. The ability to determine the optical center for designing a sheet layout

   B. The ability to fit copy to page specifications.

   C. Knowledge of the fundamental principles of design.

   D. Knowledge of the kinds and types of paper normally used in the printing industry.

II. Composition - The ability to perform the fundamental operations involved in composing a job for printing requires the achievement of the following skills:

A. Knowledge of tools and equipment used in the composing room.

B. The ability to compose troublesome letters of type.

C. The ability to read the form in a galley and identify errors of composition.

D. Knowledge of the proper division of words.

E. The ability to read and mark a proof for typographical errors.

F. Knowledge of the California Job Case.

G. Knowledge of machine composition.

III. <u>Imposition and Lockup</u> - The ability to place the pages in a form so that they will be in proper position when the sheet is printed and folded requires the achievement of the following:

A. The ability to read a sheet layout designed to determine the proper composition of a multiple page form.

B. The ability to position a form for lockup.

IV. <u>Presswork</u> - Knowledge of the basic procedures in printing a form requires the achievement of the following:

A. An understanding of the common principles underlying the operation of the basic methods of printing.

B. Knowledge of Platen press operations.

V. Lithography - An understanding of the common operations involved in offset printing requires the achievement of the following:

A. The ability to make a layout and identify proper positioning of an illustration so that the image will be in the correct position on the plate.

B. The ability to judge and recognize essential characteristics of camera operation and film processing.

C. The ability to recognize and judge the essential characteristics in lithographic platemaking.

D. The ability to judge and recognize the essential characteristics in lithographic offset presswork.

An analysis of these objectives generated several assumptions regarding their measurement. For example, the notion was advanced that each of the five measurement objectives was characterized by the property of homogeneity because the instruction provided for an integration of the fundamental knowledges and skills necessary to attain the desired level of proficiency. On the other hand, each objective was also thought to be relatively independent because of the high degree of specificity involved in attaining the desired general outcome of achievement.

Nature of Testing Operations

The next procedure, that of specifying the objectives in terms of acts to be performed on selected materials brought the committee to the rationale for selecting test materials for use in the Ohio Printing Performance Test. The identification of relevant and crucial measurement objectives did not in any way imply that the student's true ability or achievement would be adequately measured by the system of measurement adopted. The path leading from measurement objectives to valid test materials operations was the most crucial of the test development activities. The procedure adopted to assure equivalency between test behavior and true achievement was to define the testing operations first as if equipment, tools, and printing materials were being employed. These performance operations would provide the committee with the substance from which the more indirect recognition-type test materials could be derived. This procedure would help to guarantee that the selection of the more indirect task or problem would be intrinsically similar to the actual task encountered during printing operations. For example, in the printing area of imposition and lockup, an actual task might require the student to demonstrate his ability to select and place furniture reglets and quoins to position a form in the chase so it could be inserted in a printing press. From an analysis of this task, the committee suggested that the measurement of this ability might be specified by presenting the student with a photograph of a form positioned for lockup with directions to identify those parts which have been improperly positioned in the chase. A similar procedure was employed for deriving recognition tasks for measuring each of the five measurement objectives.

In analyzing the actual achievement series in this manner, the committee recommended the following specifications be observed as a basis for selecting the more indirect testing operations.

1. The problems should specify operations to measure the achievement of the formulated measurement objectives.

2. The problem should represent a valid sample of the most critical elements of the actual problems encountered in printing.

3. The problem should be amenable to testing operations which could be objectively scored.

4. The problem should specify the testing operations on materials normally encountered in the school shop or laboratory.

5. The problems should be relatively independent of one another.

6. Each problem should deal with a central theme.

7. Problems should specify the measurement of abilities normally achieved in the school shop and laboratory.

For example, a description of a problem so derived by the analysis is presented to further illustrate the development of the recognition test.

Objective - The ability to judge and recognize the essential characteristics in lithographic presswork.

Problem - The student is presented with four prints of identical copy which represent offset presswork. Each print has a discrete defect resulting from the operation of the offset press. For the first half of this problem, the student is presented with a list of presswork troubles and is directed to select the trouble which corresponds to a defect on the offset prints.

In the second half of the problem, the student is presented with a list of probable causes which may have produced the undesirable prints during the operation of the offset press. The student is then directed to select the probable causes which may have affected the undesirable prints.

Most of the objectives specified were measured by similar kinds of test materials. More than 50 per cent of the problems specified the use of actual materials encountered by the students in the school shop or laboratory. The remaining problems used illustrations and photographs. Matching exercises were also included and some problems specified actual performance as in drawing a layout, composing lines, and editing copy. Equipment and materials were easily standardized for these operations.

Relative Weights

Twenty-three problems were derived through the procedure and incorporated into the Ohio Printing Performance Test consisting of five sub-tests. In establishing the percentage weights, the committee was asked to judge how much weight each type of problem should have in determining the total score, and by so doing affix the relative importance of each problem to that total test score. The obtained rational weights were then judged to be an indication of the contribution that each sub-test should make to the total variance of a printing test. Underlying this notion of variance was the contention that the established weights for a sub-test reflected the importance of the behaviors being measured in relation to the actual process of printing and the difficulty with which they were achieved. The establishment of percentage weights did, then, imply the adequacy of a sub-test to provide for a distribution of scores over an appreciable range. Table 1.1 presents the weights derived for each sub-test of the Ohio Printing Performance Test.

Table 1.1

OHIO PRINTING PERFORMANCE TEST
SUB-TEST WEIGHTS

Contents	No. of Points	Relative Weight, per cent
Printing Planning	75	15.2
Composition	140	28.3
Imposition & Lockup	69	14.0
Presswork	60	12.1
Lithography	150	30.4
Total	494	100

Structure of the Ohio Printing Performance Test

In designating the format for the Ohio Printing Performance Test such factors as motivation problem presentation, ease of self-administration consistency of direction, appropriate language, instruction to test administrators length of test, and scoring key were considered.

Motivation - The test was designed to provide for an organization of content which would be of intrinsic interest to the students. The sub-tests were arranged in a pattern normally followed in actual printing practices beginning with printing planning and ending with offset presswork. The arrangement of the problems was similar to the actual procedures which a student would normally pursue in the planning and completing a printing job.

Problem Presentation - The presentation of the problems enabled the student to concentrate on one aspect of content at a time. Each problem was presented on a separate sheet which was easily removed from a folder designating a sub-test. For example, in Printing Planning, a folder was designed to contain five problems.

Ease of Self-Administration - General directions were provided for completing the entire test. In order to eliminate the difficulty of having the students try to remember the general directions throughout the test, directions for completing each of the separate sub-tests were printed on the folders.

Consistency of Directions - Each section of the test was classified by the nature of the problem and the directions and format were held constant throughout the test. This approach was selected since it would minimize the time spent on learning how to respond to a problem and maximize the time for demonstrating achievement.

Appropriate Language - Care was exercised in describing problems in language appropriate for the students taking the tests. Trade terms like copyfit, proof, and plates were selected if they held a constant meaning for students completing vocational printing programs.

Instructions to Administrators - A set of instructions was provided for the test administrators as part of the standardization procedures used in this development project. No provisions were made for anv variations in the use of the "Instructions to Test Administrators." They were strongly urged to adhere to the directions as stated.

Length of Test - Since the test was being designed to represent a measure of the student's problem-solving ability, test length was determined by the number of problems necessary to adequately sample the more critical aspects. Achievement rather than speed was the primary objective of this test development project. The primary consideration was one of validity and not time involved for completing the test. The committee reasoned that the test could be realistically administered in a three-hour session. Reasonable time limits were established for each sub-test. These time limits were liberal and served only to keep the slowest students from spending an interminable amount of time on separate parts of the test. However, students were permitted and encouraged to work through the test at their own speed.

Scoring Key - The scoring key for Ohio Printing Performance Test was designed to be clerically scored by personnel not acquainted with the area of printing. Directions for scoring specified the number of points to be credited for each correct response. Standard correcting procedures were used when guessing was obvious to the scorer. No penalty points were established for incorrect answers. All directions necessary for scoring the test were contained in the scoring booklet.

Tryout Administration

A tryout administration was conducted during the month of February, 1966. Used as subjects were 18 senior students enrolled in a vocational printing program at Dayton, Ohio, J. H. Patterson Cooperative High School. This preliminary tryout of the Ohio Printing Performance Test revealed that:

1. The directions prepared for administering the test were adequate.

2. The students experienced no problems in responding to the general specific directions for completing the sub-tests.

3. The liberal time limits for the sub-tests were appropriate as all students finished well in advance of the specified time allowances.

4. An internal consistency coefficient as determined by the Spearman Brown Prophecy Formula estimated the reliability of the instrument as 0.89 for this group of students.

An additional study was then conducted using the 18 students who had participated in the administration of the Ohio Printing Performance Test. The next week the Ohio Printing Achievement Test was administered to these students. The purpose of the study was to examine how well the two measures correlated with shop grades. It was hypothesized that both measures would be significantly correlated to a criterion of shop grades. This criterion was particularly relevant since the observations were based on activities equivalent to a criterion of job performance (Item 1).

Estimates of Reliability

The estimates of reliability selected for the two achievement tests were needed to determine how much of the variation in the set of scores might be the result of systematic differences among the individuals of the specified population and how much to inaccuracies in measurement. Two possible sources of unreliability were identified: inconsistencies associated with the measuring instrument itself and inconsistencies associated with the scoring system.

The selected technique for estimating the reliability of the Ohio Printing Performance Test was the application of the split-half method. This technique required two assumptions: (1) that the level of student achievement rather than speed is considered, and (2) that the two split halves of the test are comparable in content. Regarding the first assumption, it is to be recalled that the tasks selected for inclusion in the test were designed to measure the student's level of achievement only. Liberal time limits were established only to prevent the very slowest of students from spending an interminable amount of time on separate sub-tests. This procedure was adopted to provide for an orderly administration of the test on the basis that at least 90 per cent of the students would finish in advance of the time limits. Speed, therefore, was not a consideration in the assessment of student achievement.

Regarding the second assumption, the test was constructed to permit a split in the basis of odd-even number of items. This procedure would minimize the variability of content between the halves because of the homogeneous aspects of the five sub-tests. The split half was used because the technique provides for an estimate of reliability which can be extracted from a single administration. Regarding the reliability of the scoring system, interscorer reliability was estimated from an analysis of the scores obtained from four independent scorers.

Several reliability estimates were computed from the norming administration of the Ohio Printing Achievement Test.

Norming Administration

Arrangements were made to administer the two achievement measures to a representative sample of students nearing completion of a vocational printing course at the secondary level. These students were in 8? schools and 29 states plus the District of Columbia.

Nature of Population

Norms were established for each measure by using as the population, students in the twelfth grade throughout the United States who were nearing completion of a four-year vocational printing course. It was assumed that this population was normally distributed with regard to the levels of achievement. The method for scaling these achievement tests on a national level was to normalize a single distribution of student scores by sampling a large percentage of the schools that offer a vocational printing course.

Sample Representation

The sample selection procedure was to select schools which were homogeneous with respect to providing similar learning opportunites to their students. It was assumed that most of the Federal-reimbursed vocational training programs throughout the United States were relatively homogeneous with respect to such factors as the course of study and length of school year. It was hypothesized that any marked deviation from these conditions was not related to any particular school enrollment or geographic location. The rationale for this assumption was based upon two factors in the training of students for entry into the printing industry. The first is that most states have attempted to establish similar objectives for training which are acceptable to the printing industry. The second is that standards are applied to schools seeking to initiate, develop, or maintain vocational printing courses, regardless of geographic region or district.

Admittedly, part of the variability of individual differences was associated with the scope of the particular course. It was assumed, however, that such variability or error in the achievement of students as a result of participation in courses which offer varied learning opportunities would occur independently and be distributed randomly throughout the population.

Sample Selection

The adopted procedure for the sample selection was to invite all schools which offered a state-approved vocational printing course to participate. All schools choosing to participate were to be included in the norming administration. This selection was based on the assumption that any variations in learning opportunities between schools was not restricted to a particular geographic region and that such variations would be randomly distributed among the total number of schools in the United States.

On December 10, 1965, letters were sent to Trade and Industrial Education State Supervisors requesting a list of schools offering secondary vocational printing programs. Replies were received from 48 states. Seven states indicated that they had no secondary vocational printing programs.

Letters and project materials were sent to all schools. As the lists were received from the state supervisors, 238 schools representing 41 states were contacted and asked if they would be willing to participate in the norming administration of these measures. All schools that indicated a willingness to participate were included in this administration. The administration of the two achievement measures was conducted during a three-week period beginning March 14 and ending April 1, 1966.

Dissemination of Normative Data

All project participants received a listing of the students' raw scores for each of the two measures:

(1) Ohio Printing Achievement Test - A total score and 16 sub-test scores.

(2) Ohio Printing Performance Test - A total score and five sub-test scores.

Norm tables illustrating percentile rankings for all the raw scores were sent to the participating schools. The norm tables permitted school personnel to compare their students' performance with the performances of the total national sample.

Normative Data

Normative data indicating student level of achievement with respect to the two measures specified for this study was obtained through the use of the Ohio State Department of Finance Test Score Analysis Program. This program yeilded the following totals:

1. Raw score analysis which included:

   (a) Raw score list and student identification

   (b) Frequency distribution

   (c) Cumulative frequency

   (d) Percentiles.

2. Summary Statistics included·

   (a) Total number of students

   (b) Mean

   (c) Median

   (d) Standard deviations

   (e) Sum of X

   (f) Sum of X squared

   (g) Sub-test correlations.

The total score analyses detailed above were applied to the sub-test of the Ohio Printing Achievement Tests and the Ohio Printing Performance Test.

Measuring Job Performance

The intent of this study was to determine the feasibility of developing valid and reliable criterion for evaluating the effectiveness of vocational printing. The development and standardization of the achievement measures and the dissemination of the results marked only the first phase of this research study. In order to determine the relevance of achievement to subsequent performance on the job, measures of job performance were needed for those students who found employment in the printing industry. For those students estimates of job performance were obtained to test the adequacy of the training received. The development of such measures or estimates was the purpose of the second phase of this study.

Two conferences were held with representatives of the graphic arts industry for the purpose of defining a strategy for assessing job performance in the printing trades. During the conferences emphasis was placed on the need to include behaviors which were observable, e.g., descriptive samples of job skills from which a rater could logically infer a judgment about the employee's proficiency. Within this general context, the procedures adopted for identifying and selecting behavior traits of job performance consisted of the following:

1. The trait was to reflect behavior providing for a large spread of individual differences. Of the traits identified, those in which employees differ most from one another were preferred, because it was to be the purpose of this measure to differentiate among the degrees to which an employee was of value to the company. Thus, even for a trait of utmost importance (honesty), if all employees demonstrated it to the same degree, it would be of little use in estimating the effectiveness of a single employee to his company.

2. The language used to describe the traits should provide a known and constant meaning for the raters. Unless this is true, ratings by different supervisors could not be meaningfully compared with one another.

3. Traits should reflect behavior which is pertinent to job performances. If the traits to be evaluated were not specifically tied

down to the work situation, there was the ever present danger that raters might base their judgments on observations made in other than work situations and on characteristics which will have little relevance to the employee's job performance.

4. The trait should reflect behavior relevant to job performance of a newly graduated vocational person at the early stages of his employment. In order to measure the relevance of training received in the vocational course and not training received on the job, the intent of the study was to provide for the measurement of job performance after 8 to 10 months of employment when exposure on the job would not have added greatly to his vocational competence.

The scope of the traits selected, therefore, was limited to these expectations of job behavior within this time period of employment.

On the basis of the above rationale, the following 14 traits were selected as the most critical and the most readily measurable of the elements of job performance at this stage of employment (See Appendix B).

1. Dependability

2. Safety

3. Quantity

4. Care of tools

5. Resourcefulness

6. Neatness

7. Accuracy

8. Industriousness

9. Reaction to criticism

10. Adaptability

11. Communication skills

12. Organization of work

13. Technical knowledge

14. Job skill.

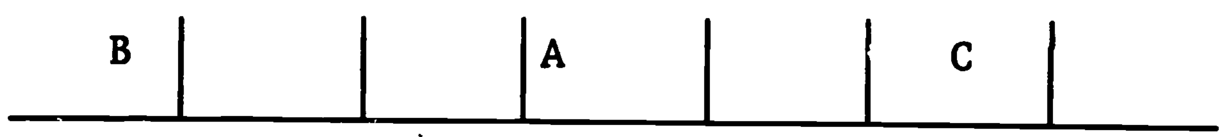
Format of Job Performance Rating Scale

Measures of job performance were to be secured for all newly employed persons in printing who had participated in the norming administration of the two achievement measures. Great care was exercised to assure that different supervisors throughout the United States would be able to use the rating scale with equal efficiency once it was received through the mail. The graphic scale format was adopted because of its simplicity of administration. This type of scaling assured that a careful balance was maintained between validity of the measure and simplicity in its administration. Any attempt to increase the validity through a more complex rating system might have added confusion to the inexperienced rater. This could easily result in introducing error variability in the form of a rating scale expertise factor on the part of the raters.

The adopted graphic scale format provided for a continuous straight line with adjectives placed at the extreme points on the scale to guide the rater. The line indicating a continuum of the trait was divided into seven units of equal length. For example, the trait accuracy appeared in the scale as follows:

Accuracy: Consider the extent to which his work is free of errors and meets plant standards for quality

Often makes mistakes — Rarely makes mistakes



A check in the middle (see A) of the scale indicated average performance whereas a check at the extreme left (see B) indicated the lowest level of performance and at the extreme right (see C) the highest. The length of the line was set at five inches, a standard size frequently employed as the maximum length for rater discremination. Guilford (4) suggests that longer lines have actually decreased the spread of scores of ratings because of the tendency to cluster ratings centrally on a long line. No attempt was made to control for either the halo effect, the leniency factor by alternating randomly the desirable point on the scale, or by presenting each rating item on a separate page. It was felt that either of these control factors tended to make the job of rating somewhat cumbersome for the rater particularly since the optimal effects of such control had yet to be demonstrated. Remmers (7) in a study of teacher characteristics found no systematic differences between one arrangement of desirable points on the scale for obtaining student ratings of instructors. Guilford (4) also suggests that for the untrained rater the desirable end of the scale should come first, because in rating people the typical rater likes to think of the good qualities first. He goes on to suggest that to capitalize on this tendency might be something of value. However, this appears to favor the introduction of the leniency factor. Remmers (8) also found that there was little difference in the results when the arrangements of the 10 personality traits of the Purdue Rating Scale for instruction were changed. As a result of these studies, it was decided to present the undesirable end of the scale first to inhibit the leniency factor.

Scoring System

The scoring system was based on the assumption that no finer discrimination would be employed for scoring than that employed by the units marked off on the line. No numbers were included on the scale itself in order not to unnecessarily introduce distracting elements on the line. A number system of 1-7 was employed for purposes of scoring. However, scoring across items was done by summing the numbers on the assumption that the traits were equal in importance. This assumption was based on Richardson's (9) analysis that nominal weights of one would weight themselves as a result of a different variance obtained for each item.

The format of the job-performance criterion measure consists of the following:

1. Personal information - To obtain the department, the length of time employee had worked for rater, and how often the rater saw the employee in the working situation.

2. Introductory Statement - Explaining both the purpose and use of the rating scale.

3. Considerations in the use of the Scale - Attention was drawn to the kinds of errors often committed when appraising performance through the use of a human instrument, the rater. Particular attention was focused on the population of employees being considered. The rater was repeatedly reminded that this study was concerned with the employee's

performances during the early stages of employment. As such, the employee being rated must be carefully compared to other or previous employees with whom the rater had observed at similar stages of employment.

4. Directions - Illustrative examples of how to use the rating scale were presented to minimize error resulting from chance misinterpretations.

5. Rating Items - Fourteen rating items were included in the final form.

Tryout Administration

To test the effectiveness of this newly developed job performance criterion measure, two qualitatively different tryout administrations were conducted.

The first concerned interviews with seven printing firms in the central Ohio area to discuss the adequacy of the scale in lieu of responses to the following questions:

1. Are the directions for using the scale sufficiently succinct and exact?

2. Are the behavior traits relevant to job performance of newly employed personnel?

3. Are the behavior traits sufficiently representative of the crucial elements of job performance at this level?

Positive responses were obtained for the three categories of questions asked throughout the interview underscoring the willingness on the part of the industry to accept this measure as being valid.

The second tryout administration concerned a research study conducted at the Samuel P. Rosenthal

Company, Cincinnati, Ohio. The purpose of this study was to obtain measures of reliability and validity in a situation similar to one in which the measure would be used by supervisors in different cities.

The Rosenthal Printing Company had used a rating procedure for appraising employee performance similar in content to the one adopted for this study. Its structure was adjectival as opposed to graphic. A comparative study was then conducted between the Rosenthal measure and the graphic scale using 25 randomly selected employees. The purpose of the study was to:

1. Obtain an estimate of reliability based on the technique of equivalent forms

2. Test the null hypothesis of no difference between an adjectival rating format and the graphic scale

3. Test the validity of the graphic scale against a supervisor's ranking of the employees.

Follow-up Study

In order to obtain measures of job performance on the students originally tested, a follow-up study was conducted to identify the students who entered the printing industry. In order to accomplish this phase, a graduate questionnaire (Appendix C) was mailed to the students. This mailing occurred during the fall of 1966. The questionnaire was designed to obtain the following information:

1. Employment or non-employment in the printing trades.

2. Place of employment.

3. Date of employment.

4. Supervisor's name.

Two follow-up letters were used where necessary as part of this procedure.

Supervisory Ratings

The next phase of this study required the direct contact of supervisors to obtain measures of job performance for vocational graduates employed in the printing industry. In support of this phase, articles concerning the nature of this study appeared in the following national trade journals just prior to the distribution of the job performance measures.

1. Printing Production: October, 1966, p. 14.

2. Inland Printer: October, 1966, p. 159.

3. Printing Views for the Midwest Printer: November, 1966, p. 47

4. Printing Impressions: November, 1966, p. 87.

Since practically every employer in the printing industry subscribed to one or more of these trade journals, the strategy of the advanced promotion was to encourage participation of supervisors to respond positively to the request materials. Following the release of the articles, the job performance criterion measures were sent to the respective supervisors. Two follow-up letters were included as part of the procedures for maximizing the returns of this data.

Analysis of Data

In order to provide an index for evaluating the relevance of vocational training in printing to job performance, a statistical analysis was planned to validate the two achievement measures against measures of job performance. This analysis, a correlation technique, was conducted by The Ohio State University Numerical Computational Laboratory. The scoring of the measure itself was conducted by the Trade and Industrial Education Instructional Materials Laboratory.

The representativeness of the newly obtained population of employees to the original norming population of students was examined by an analysis of variance technique recommended by Meyers. (6) This technique was used to test the null hypothesis that no differences existed between the student achievement levels of the norming sample and the follow-up sample who obtained employment in the printing industry. This procedure was adopted to minimize generalizations of the obtained correlations between student achievement and job performance. This was necessary if differences between the two categories of population were found.

Reliability estimates were obtained from a single administration of the job performance criterion measure by employing internal consistency techniques.

CHAPTER III

RESULTS

Measuring Student Achievement

This chapter deals with the results of the norming administration, estimates of reliability for both the achievement measures, and evidence which bears upon the assumptions underlying the procedures adopted for developing the Ohio Printing Performance Test and the Job Performance Criterion measure. In addition, evidence bearing directly upon the validity of both achievement measures is presented.

Tryout Administration

The Ohio Printing Performance Test and an existing form of the Ohio Printing Achievement Test were each validated against a criterion of shop grades using as subjects 18 senior students from Dayton Patterson High School, Dayton, Ohio. The shop grades employed in this study were the ones normally provided at the end of the first semester. They are based upon a five point scale with an A=5 points, B=4 points, C=3 points, etc. See Table D-1 in Appendix D.

A correlation coefficient of 0.45 was obtained for the Ohio Printing Achievement Test but it was found not to be significant. A correlation coefficient of +0.62 for the newly developed Ohio Printing Performance Test was found to be significant at the 0.01 level.

Means and standard deviations for the variables measured appear in Table 2.1.

Table 2.1

MEASURES OF CENTRAL TENDENCY AND VARIABILITY
DAYTON PATTERSON HIGH SCHOOL

Measures	N	Mean	Standard Deviation
Ohio Printing Achievement Test	18	177.94	17.51
Ohio Printing Performance Test	18	269.00	38.20
Shop Grades	18	3.98	0.71

Norming Administration

Seven-hundred ninety-five students participated in the norming administration of the Ohio Printing Achievement Test and 745 for the Ohio Printing Performance Test. The totals represented 77 high schools offering certified vocational printing courses including 29 states and the District of Columbia (See Appendix E). A summary of the level of achievement for the students is presented in Table 3.1.

Table 3.1

SUMMARY OF AVERAGE PERFORMANCE

Test	Mean	Median	S.D.	Sum of Scores	Sum of Squares
OPPT	238.57	238.75	57.40	177,732	44,855,344
OPAT	153.76	151.05	38.33	122,241	19,963,757

Reliability - Ohio Printing Performance Test Scoring System

An analysis of variance technique by Winer (12) was used to estimate the interrater reliability of the Ohio Printing Performance Test scoring system. Four independent scorers were asked to score each test of a group of 74 randomly selected tests. This random selection was made on the basis of 10 per cent of all completed tests resulting from the norming administration. An estimate of the reliability of the mean scores obtained by each of the four independent scorers was 0.9899 (See Appendix D, Table D-2).

Since an identical scoring system was used by each of the scorers with explicit directions for scoring on the test, it was assumed that any deviations between the mean scores obtained by the four independent scorers represented a source of variation caused by errors of measurement alone. Mean scores and standard deviations obtained by each of the scorers appear in Table 4.1.

Table 4.1
GROUP MEANS AND STANDARD DEVIATIONS OBTAINED BY FOUR INDEPENDENT SCORERS ON THE SAME GROUP OF OHIO PRINTING PERFORMANCE TESTS

Statistic				
Mean	224.46	223.77	223.59	223.44
S.D.	55.06	55.72	55.71	54.89
Sum of Scores	16,610	16,549	16,546	16,535
Sum of Squares	3,956,582	3,930,756	3,929,848	3,922,687

Ohio Printing Performance Test Reliability

Since there was but a single administration of the Ohio Printing Performance Test, it was necessary to obtain estimates of reliability for both measures by an internal consistency method. The technique selected for the Ohio Printing Performance Test was the split half and the test was artifically divided into two halves of equal length on the basis of an odd-even distribution of item scores.

In order to examine the effectiveness of the procedures adopted for spliting this test on the basis of an odd-even distribution of item scores, empirical evidence for analyzing the comparability of the two split halves was obtained and appears in Table 5.1.

Table 5.1

COMPARABILITY OF TWO SPLIT HALVES BASED ON AN ODD-EVEN DISTRIBUTION OF SCORES

Statistic	Test X	Test Y
Mean	112.22	111.98
S.D.	25.26	23.48
N	74.	74.

The same 74 subjects selected for obtaining an estimate of reliability for the scoring system were used for estimating the reliability of the test. The raw scores obtained for both halves of the test were correlated with each other using the Pearson Product Moment Correlation Formula (See Appendix D, Table D-3 for the Summary data used in deriving this coefficient).

Since the coefficient derived from this data gives the reliability of a test only half as long, the reliability of the full test was estimated by use of the Spearman-Brown Prophecy Formula. Estimates of reliability for the Ohio Printing Performance Test appear in Table 5.2.

Table 5.2

OHIO PRINTING PERFORMANCE TEST RELIABILITY ESTIMATES

Method	N	Reliability
Pearson-Product Moment Correlation Coefficient	74	0.84
Spearman Brown Prophecy Formula	74	0.90
Standard Error	74	18.42

Ohio Printing Achievement Test Reliability

Concerning the Ohio Printing Achievement Test, estimates of reliability were obtained for the total norming population using such techniques as Kuder Richardson 20 and 21, Pearson-Product Moment Correlation Coefficient, and the Spearman Brown Prophecy Formula. These estimates appear in Table 5.3.

Table 5.3
OHIO PRINTING ACHIEVEMENT TEST RELIABILITY ESTIMATES

Method	N	Reliability	Standard Error
Kuder Richardson #20	795	0.95	8.43
Kuder Richardson #21	795	0.94	9.20
Spearman Brown Prophecy Formula	795	0.96	7.67
Pearson-Product Moment Correlation Coefficient	795	0.92	10.74

Normality of Obtained Set of Frequencies

A comparison was made between the obtained frequencies of the Ohio Printing Performance Test scores and the theoretical frequencies expected in a normal distribution having the same mean and standard deviation as the obtained distribution. The scores obtained from the norming administration of these measures were grouped in intervals of 15 raw score points and the expected frequencies for the same class intervals of a normal distribution were then derived. See Appendix D, Table D-4 for the derived expected frequencies and the observed frequencies.

A chi square goodness of fit test recommended by Guilford (3) was used to test the null hypothesis that the distribution of observed scores do not depart from normality. The tails of this frequency distribution were collasped leaving a total number of intervals at 22 for the test of significance. Data for testing goodness of fit are presented in Table 6.1.

Table 6.1

TESTING GOODNESS OF FIT

Expected Frequency	Observed Frequency	Cell Discrepancies	Squared Cell Discrepancies	Cell Square Contingencies
fe	fo	(fo-fe)	$(fo\text{-}fe)^2$	(fo-fe) /fe
Sums				
747.56	745	-2.56	453.41	*17.979

*Chi Square

Sub-test Correlations

Sub-test correlations were obtained for the Ohio Printing Performance Test using as a sample the 74 students randomly selected in order to test the assumption that the printing process is characterized by stages of operations which are relatively discrete and independent. Table 7.1 presents the obtained correlations using the Pearson-Product Moment Formula and the respective levels of significance.

Table 7.1

INTER CORRELATIONS OF SUB-TESTS AND LEVELS OF SIGNIFICANCE

Sub-tests	Pearson-Product Moment Correlation Coefficient
Printing Planning and Composition	0.44**
Printing Planning and Imposition-Lockup	0.16
Printing Planning and Presswork	0.42**
Printing Planning and Lithography	0.34**
Composition and Imposition-Lockup	0.28
Composition and Presswork	0.42**
Composition and Lithography	0.36*
Imposition-Lockup and Presswork	0.36*
Imposition-Lockup and Lithography	0.51**
Presswork and Lithography	0.37*

* 0.01 Level of Significance   ** 0.001 Level of Significance

Reliability estimates for these sub-tests were also obtained on the basis of an odd-even distribution of item score values. Table 7.2 presents the estimated reliabilities.

Table 7.2

RELIABILITY ESTIMATES FOR THE FIVE SUB-TESTS OF THE OHIO PRINTING PERFORMANCE TEST

Sub-test	Reliability
Printing Planning	0.52
Composition	0.94
Imposition-Lockup	0.75
Presswork	0.95
Lithography	0.90

The sub-tests correlation appearing in Table 7.1 were somewhat restricted because of the existence of measurement errors reflected by the respective reliabilities presented in Table 7.2. In order to determine the true variance for the obtained correlations, a correction was made for attenuation with Guilford (3). The corrected correlation is presented in Table 7.3.

Table 7.3

INTERPART CORRELATIONS CORRECTED FOR ATTENUATION

Interpart Correlation	Correction for Attenuation
Printing Planning vs. Composition	0.63
Printing Planning vs. Imposition-Lockup	0.26
Printing Planning vs. Presswork	0.60
Printing Planning vs. Lithography	0.50
Composition vs. Imposition-Lockup	0.34
Composition vs. Presswork	0.45
Composition vs. Lithography	0.40
Imposition-Lockup vs. Presswork	0.43
Imposition-Lockup vs. Lithography	0.62
Presswork vs. Lithography	0.40

A multiple regression analysis was conducted to determine each sub-test's contribution to the total test variance of the Ohio Printing Performance Test. This procedure is outlined in Appendix F.

Measuring Job Performance

To test the new rating scale, a tryout administration was held at the Rosenthal Printing Company in Cincinnati, Ohio. The new scale was used in conjunction with an adjectival scale (developed by Rosenthal) and a ranking of employees for determining reliability and validity. The tryout administration of the newly developed rating scale at the Rosenthal Printing Company yielded an estimate of reliability at 0.79 through an analysis of variance technique recommended by Winer (12). This design averaged the intercorrelations of the 14 rating items of the graphic scale over the same group of subjects. Results of this analysis are summarized in Appendix D, Table D-5.

A statistical test was made on the null hypothesis of no difference between the adjectival scale (Rosenthal) and the newly developed Job Performance Criterion Measure. Using the Pearson-Product Moment Formula, a correlation coefficient of 0.85 was found to be significant at the 0.001 level of significance. Summary data from which this correlation was derived appears in Appendix D, Table D-6.

Kendall's Tau (5) was employed to validate the Job Performance Criterion Measure against a ranking of the employees by their supervisors. A correlation of 0.92 was obtained by the following formula.

$$R_{xy} = \frac{1 - 6X\ (\text{Summation } D^2)}{N\ (N2 - 1)}$$

$$(\text{Summation } D^2 = 185,\ N = 25)$$

Follow-Up Study

Seven-hundred questionnaires were mailed to graduates who had participated in the norming administration of the two achievement measures. More than 85 per cent was returned for a total of 597. Table 8.1 categorizes the distribution of these returns.

Table 8.1
GRADUATE QUESTIONNAIRE DISTRIBUTIONS

Class of Post-Graduate Activity	No. of Returns	Per Cent
Working in Printing	250	42.1
Work other than Printing	130	21.8
In School	103	17.3
Armed Services	74	12.4
Not Working	38	6.7
Total	595	100.0

Eighteen of the questionnaires for those employed in printing were disqualified either because of insufficient information on the questionnaire or entry into the Armed Services. This left a total of 232.

Requests for ratings of job performance for these students were mailed to their respective supervisors. More than 85 per cent of 198 rating scales was returned for data analysis. Twenty-three were disqualified because of the rater's refusals to consider all of the items. The remaining 175 ratings were employed as measures against which to validate the two achievement measures.

Table 8.2 gives the mean and standard deviation of the three variables.

Table 8.2

MEANS AND STANDARD DEVIATIONS

	N	Mean	S.D.
Ohio Printing Performance Test	175	267.4	39.37
Ohio Printing Achievement Test	175	170.9	51.86
Job Performance Criterion Measure	175	64.1	14.80

Reliability

Since there was a single administration of this measure, estimates of reliability were obtained by two internal consistency methods, an analysis of variance recommended by Winer (12) and an odd-even split half technique. An estimate of the average intercorrelations between rating items over the same sample of 175 subjects was found to be 0.9997 with a standard error of measurement less than 1.0. Summary data for this estimate appears in Table D-7 of Appendix D.

Raw scores obtained for both halves of the Job Performance Criterion Measure were correlated through the use of the Pearson-Product Moment Correlation Formula. Summary data used in deriving the coefficient appears in Table D-8 of Appendix D.

The reliability for the full rating scale consisting of 98 raw score values for the 14 items was obtained through the use of the Spearman Brown Prophecy Formula. Estimates of reliability based upon this split-half technique appear in Table 9.1.

Table 9.1
JOB PERFORMANCE CRITERION RELIABILITY ESTIMATE

Method	
Pearson R^{xy}	0.91
Spearman Brown Prophecy Formula	0.95
Standard Error of Measurement	4.74

Comparability indexes for the split halves appear in Table 9.2.

Table 9.2
COMPARABILITY OF TWO SPLIT HALVES BASED ON AN ODD-EVEN DISTRIBUTION OF SCORES

Statistic	Test X	Test Y
Mean	31.46	32.69
S.D.	7.45	7.86
N	175	175

Validation of the Achievement Measures

Multiple Prediction

The final data analysis of this study concerned a multiple prediction and regression analysis between the two achievement measures and the job performance criterion measures. The multiple correlation prediction estimate ($R_{1.23}$) was found to be 0.46 and significant at the 0.001 level. See Appendix D, Table D-9 for the data from which the multiple correlation estimate was derived.

Summary data for deriving Product Moment Correlation Coefficients among the three measures appears in Table D-10 of Appendix D.

The multiple regression analysis concerned the prediction problem of estimating Job Performance Criterion Measure (X1) values from values obtained for both the Ohio Printing Achievement Test (X2) and the Ohio Printing Performance Test (X3). This called for a multiple regression analysis from which a multiple prediction regression equation could be derived. This equation for the three variable problem has the general form,

$$X_1 = a + b_{12.3}X_2 + b_{13.2}X_3.$$

Summary data for solving the regression coefficients for the multiple regression equation appear in Table 10.1.

Table 10.1

SOLUTION OF THE REGRESSION COEFFICIENTS
FOR THE MULTIPLE-REGRESSION EQUATION

1	2	3	4	5	6	7	8
OPAT	-0.31	0.15	-0.05	0.38	-0.12	170.9	20.1
OPPT	0.63	0.40	0.25	0.29	0.18	267.4	-48.1
			0.20 = R^2				-28.0
							M_1 = 64.1
							a = 36.1

1. Measure designation
2. Beta coefficients
3. Correlation of achievement measures to Job Performance Criterion Measure
4. Variance contributed by each achievement measure and multiple R^2
5. Ratio of Job Performance Criterion Standard Deviation to Achievement Measure Standard Deviation
6. Derived weights for each achievement measure (i.e., $b_{12.3}$ and $b_{13.2}$)

7. Achievement measure means
8. Value of a.

The mean of the weighted composite for the two achievement measures was found to be 64.0 with a standard deviation of 6.5.

Homogeneity of Variance

A test of homogeneity for variance was made on the null hypothesis of no difference between the variances obtained for the two achievement measures in the follow-up study and the variances obtained in the norming administration. Table 11.1 presents the sample stati tics for this analysis.

Table 11.1
TEST FOR HOMOGENEITY OF POPULATION VARIANCE

Test		Variance	df	F observed
OPAT:	Norming	1,469.2	784	0.95
	Follow-up	1,550.0	174	
OPPT:	Norming	3,294.7	744	1.14
	Follow-up	2,883.2	174	

The level of significance for this test was chosen at 0.05. Since the null hypothesis for the Ohio Printing Achievement Test variances was clearly not rejected, only one test of significance was conducted. The critical region for rejecting the null hypothesis for the Ohio Printing Performance Test variances was found to be 1.17 with F_{95} (744,174) df. Since F observed was less than this value, the null hypothesis was not rejected.

Prediction Statistics

A more thorough analysis was made by comparing the predictor measures separately and in combination with the Job Performance Criterion Measure, using indexes of prediction recommended by Guilford (4). The index of prediction for r_{21}, r_{31}, r_{23}, and $r_{1.23}$ were compared in terms of their respective standard error of estimates, coefficient of alienation, and its complement, the index of forecasting efficiency. Table 12.1 presents the statistics obtained for this analysis.

Table 12.1

INTERPRETATIVE STATISTICS FOR THREE VARIABLES OF A MULTIPLE REGRESSION ANALYSIS

	OPAT vs. Criterion X 21	OPPT vs. Criterion X 31	OPAT vs. OPPT X 23	OPAT-OPPT vs. Criterion X 1.23
Pearson R. Correlation Coefficient	0.15	0.40	0.73	0.46
S. E. of Estimate	14.50	13.62	10.21	13.0
Coefficient of Alienation	0.98	0.92	0.69	0.88
Index of Forecasting Efficiency	2.00	8.00	31.00	12.00

Levels of Performance

The final step in this analysis was to test the null hypothesis of no difference between the means obtained for the two achievement measures in the norming population and the means obtained in the follow-up study. Table 13.1 presents the sample statistics for this analysis.

Table 13.1
TESTING HYPOTHESIS ABOUT THE DIFFERENCE BETWEEN TWO MEANS

		Mean	Standard Error of a Difference	Standard Score
OPAT:	Norming	153.76	3.26	5.26
	Follow-up	170.90		
OPPT:	Norming	267.40	4.43	6.51
	Follow-up	228.57		

Since a standard score of 2.48 is required for significance at the 0.01 level, the null hypothesis was rejected for both achievement measures.

CHAPTER IV

DISCUSSION

This chapter examines the evidence that the measures of student achievement provide for a valid and reliable appraisal of training objectives for vocational printing at the twelfth-grade level.

Measuring Student Achievement

Evidence obtained from the tryout administration supported the assumption that both measures appraised achievement relevant to actual production activities. When the Ohio Printing Achievement Test and the Ohio Printing Performance Test were each validated against a criterion of shop grades, both measures were moderately correlated to the criterion. Further analysis revealed that the Ohio Printing Performance Test was significantly related while the Ohio Printing Achievement Test was not. No interpretation was made concerning these results. Cronback (1) states, one should not rush to interpret small sample statistics because of the considerable variations which can occur from sample to sample. This is particularly relevant since the distribution of t values at the 0.05 and 0.01 levels are quite sensitive to any change of the validity coefficient for small sample sizes. The evidence obtained was based on a consideration of student achievement observed under conditions similar to actual printing production. Since the end goal of this study was to validate the achievement measures against such criteria, it was particularly important to obtain such results prior to the norming administration.

Following the norming administration, a study was conducted to determine whether the sample was sufficiently representative of the specified student population. It was assumed that this population was normally distributed with respect to the achievement level being appraised. Using a chi square test of significance, it was found that the obtained distribution of scores for the Ohio Printing Performance Test did not represent a significant departure from the normal curve. In addition, it was observed that the mean score value and median value for the distribution were equivalent suggesting a condition commonly associated with a student population thought to be normally distributed. This evidence was thought to be in support of the notion that the sample of students used in the national norming of both measures were sufficiently representative of the specified population.

The acceptance of the normative sample as an adequate representation of the student population, led to a consideration as to what degree the obtained distribution of score values for both the Ohio Printing Performance Test and Ohio Printing Achievement Test represented systematic individual differences with respect to "true" achievement. In order to answer this question, it was necessary to examine the reliabilities of both measures.

Since the reliability of the Ohio Printing Performance Test could not be effectively interpreted without examining to what degree inaccuracies in measurement might result from the scoring system employed, an estimate of interscorer reliability was obtained. This reliability was estimated at 98 per cent. Concerning the interpretation of this estimate, Winer (12) suggests that if such an experiment were repeated with another random sample of four scorers but with the same subjects, the correlation between the mean scores obtained from the two sets of data would be approximately the same. The obtained estimate, therefore, underscored the adequacy with which untrained personnel can effectively use such a scoring system to measure consistently and accurately the students' responses to the test problems.

Reliability estimates obtained for both measures were 0.90 and above, under conditions where speed was not a factor and when the scoring systems employed for the two were nearly perfect. Thorndyke (10) suggests that one method of interpreting reliability estimates is to examine what proportion of 'true" achievement can be accounted for by the tests themselves. Using this analysis, it was found that better than 81 per cent of the obtained distribution of test score variance for both measures could be attributed to "true" differences in levels of achievement. Applying the standard error of measurements obtained for the two measures, Ferguson (2) suggests that the chances would be one in three that the obtained score would differ as much as the value of the standard error of measurement from the "true" score. Considering the probability factor underlying this estimate as well as the total number of points possible for each measure, one can readily observe that such deviations from "ture" score values are indeed small.

The next consideration was to examine the effectiveness of the procedures adopted for developing the Ohio Printing Performance Test in terms of the assumptions underlying student achievement. The committee hypothesized that because of the nature of the instruction given at each stage of the printing process, the obtained interpart correlations should reflect relatively discrete and independent scores. Because of the low reliabilities obtained for both the Printing Planning and Imposition-Lockup sub-tests, interpart correlations were corrected for attenuation. When corrected, moderate positive correlations were found for Printing Planning versus Composition, Printing Planning versus Presswork, and Imposition-Lockup versus Lithography, while the remaining intercorrelations reflected a low but definite relationship. The obtained correlations were interpreted to mean that although some instructional overlap was to be expected, the evidence was in support of the hypothesis of independent and discrete sub-tests.

Evidence from the multiple regression analysis was employed to examine the relationship of the percentage of total test score variance predicted on a <u>rational basis</u> for the sub-tests to the statistically derived variances for each sub-test measure resulting from this analysis. It was found that significant contributions to total test score variance were made by each of the sub-test measures. There was, however, some departure from the original rational weights which required further analysis. Table 14.1 presents the comparison of the variances obtained for the sub-test measures to those predicted on a rational basis.

Table 14.1

COMPARISON OF RATIONAL AND STATISTICAL VARIANCES FOR THE SUB-TEST MEASURES OF THE OHIO PRINTING PERFORMANCE TEST

Sub-test	Predictor or Rational, per cent	Derived or Statistical, per cent	Average Difference, per cent
Printing Planning	15.2	13	38.9
Composition	28.4	19	60.9
Imposition & Lockup	14.0	10	37.1
Presswork	12.2	20	52.7
Lithography	30.2	38	44.0

The percentage of total test variance was somewhat lower than predicted for the measures of Printing Planning, Composition and Imposition-Lockup, while higher than predicted for Presswork and Lithography. An examination of the average difficulties revealed that the variance of each was restricted compared to the range of possible score values. Whereas, the variance of the Composition measure was somewhat restricted because of the ease of the tasks. On the other hand, it appears that Presswork and Lithography approached the ultimate level of discrimination and subsequently a larger spread of score values. Thus, the correlation of sub-test scores to the total test score would statistically favor those measures in which average difficulties reflected a maximum spread of score values. In addition, had the average difficulties reflected a maximum spread of scoring for the measures of Printing Planning, Composition and Imposition-Lockup, the contribution to total test variance may have been as predicted since such contribution for a sub-test measure is influenced by both the restriction of the frequency distribution and its correlation with all other sub-tests and the total test score. Both achievement measures were conceived by considering classes of qualitatively different technical skills and knowledges attained in vocational printing and judged relevant to early success on the job. The next phase of this discussion follows such an analysis by examining the relevance of student achievement to a job performance criterion.

Measuring Job Performance

A job performance criterion measure was developed with which to validate the two achievement measures. Evidence employed to examine the validity and reliability of this measure was obtained from a tryout administration among 25 employees at the Rosenthal Printing Company. The reliability estimates obtained by an analysis of variance of the average intercorrelation between items and an equivalent form technique were 0.79 and above. At a minimum, 64 percent of the obtained distribution for this measure could be attributed to "true" differences in levels of job performance. The validity of the job performance criterion measure was obtained against a ranking of the employees by supervisory personnel. The obtained coefficient (0.85) suggested that better than 75 percent of the obtained distribution of rankings could have been predicted by this measure alone. Considering that respectable rating scale measures (Cronback (1) normally obtain reliability estimates in the 0.80's and above, the obtained estimates were judged to be a sound measure of job performance. However, when one considers that validity coefficients are rarely achieved above 0.60 (Cronback (1) or 36 percent of the predictable variance, the obtained coefficient was judged to be adequate.

Consideration was given to what balance should exist between the validity of the job performance measure and its reliability. Guilford (4) suggests that when one obtains a significantly high validity coefficient, estimates of reliability are limited to the 70's and 80's. The rationale for such an interpretation is based on the assumption that a high reliability generally reflects high average intercorrelations between items on a general performance ability, whereas a high validity reflects lower average intercorrelations between distinct items comprising a composite of job performance abilities. The higher the validity, the less chance one has in obtaining a significantly high reliability. Since we were interested in 14 traits in which the behaviors to be measured were judged by the committee as constituting the most critical of job performance, the balance obtained between reliability and validity was judged adequate and acceptable.

A sample size of 25 is not necessarily presumed to be as variable as the small samples of 10 to 20 subjects. At this distribution level, t-tests are generally not as sensitive to small changes in validity coefficients, making it possible to interpret the adequacy of this newly developed job performance measure on the basis of reasonable stable statistics.

Following the tryout administration at the Rosenthal Printing Company, measures of job performance were obtained from the job supervisors of those graduates who had participated in the norming administration. The analysis of variance technique for estimating reliability revealed that each item discriminated about the same since an extremely high reliability here reflects the condition of no significant differences between the obtained variances for each item. Such a condition appears to suggest that an equal weight assumption underlying the scoring system for this test appears to have been appropriate. Such a condition also suggests a high intercorrelation among items as well as items of equal difficulty. All three of these incepretations imply the maximum conditions for estimating reliability. To internally maximize reliability, one risks reducing the validity, for as it was suggested, the two are often incompatible. Maximum validity requires low intercorrelations and items differing in difficulty (Tucker (11). The attainment of a nearly perfect reliability estimate raises some concern as to what implications this holds for validity. A high average intercorrelation between items implies the presence of a "halo" effect and reduces the discrimination that comes from traits of job performance assumed to be independent.

When one deals with average intercorrelations, one removes much of the variability between items caused by the prescence of "means" in the data analysis underlying this technique. Such averages tend, therefore, to inflate somewhat unrealistically, the reliability estimate. In order to extract a more reasonable estimate of inter item variability, a split-half technique was employed. The estimate here was found to be 0.95 with a standard error of 4.74 units. Because no time existed between the ratings on the basis of comparable halves, the notion that a "halo" effect is inflating this estimate must continue to remain a serious consideration in the interpretation of these estimates.

"Halo" effects normally occur in all ratings but are more prone to occur when a supervisor cannot compare an individual to a comparative group of individuals. The supervisor was asked to rate the newly employed graduate by comparing him to others he had known at similar stages of employment and not to compare him with other employees who were on the job significantly longer. Whereas, in the tryout administration, it was possible to compare each individual with his peers. In the absence of observations within groups of employees at the same level of employment, the halo effect becomes difficult to control. The question still remains as to whether consistent ratings across all items provide for total scores which discriminated among individuals and what if any comparison can then be made between the variance of score values for the achievement measures and the variance of score values from the job performance criterion measure.

In order to examine these questions, a multiple prediction analysis was conducted to observe the relationship between the achievement measures and the job performance measure. The multiple prediction correlation was found to be moderately correlated (0.46) and significant at the 0.001 level. A further analysis here revealed that the relationship found between the Ohio Printing Achievement Test and the criterion was slight (0.15) but positive. This correlation was not found to be significant at my reported level of alpha. Concerning the Ohio Printing Performance Test, the relationship reflected a moderate correlation (0.40) which was found to be significant at the 0.001 level. The index of forecasting efficiency further reflected the predictive efficiency of the three correlation conditions. For it was found that with the Ohio Printing Achievement Test, errors of prediction were reduced only by two per cent, whereas the Ohio Printing Performance Test reduced the error by eight per cent. When the two are jointly correlated with the criterion, errors of prediction are reduced by as much as 12 per cent. Although these efficiencies may seem quite small, they must be treated in a relative and not an absolute sense. For as Thorndyke (10) has pointed out, it is probable that predictions based on interviews alone is less than five per cent. With this as our base, the picture of efficiency of the joint effect of our two measures appears much better.

Another method employed for interpreting the predictive efficiency of the joint contribution of the two achievement measures was to examine the standard error for predicting scores on the Job Performance Criterion Measure from the scores obtained for both achievement measures. The complete regression equation obtained from the multiple regression analysis was found to read:

$$X_1' = 36.1 - 0.12X_2 + 0.18X_3,$$

where X_1' = predicted score on the Job Performance Criterion Measure and X_2 and X_3 are the symbols representing the score values obtained for the Ohio Printing Achievement Test and the Ohio Printing Performance Test measures. The values of (-0.12) and (+0.18) represent the beta weights for each measure. To interpret this equation, one might say that for every unit increase in X_2, X_1 is increasing -0.12 unit and that for every increase in X_3, X_1 is increasing 0.18 unit. Table 15.1 provides a comparison between the predicted score and the obtained score on the Job Performance Criterion Measure for students A, B, and C.

Table 15.1

SOME PREDICTIONS OF JOB SUCCESS CRITERION SCORE VALUES FROM MEASURES IN TWO VARIABLES

	Students A	B	C
X_2 Ohio Printing Achievement Test	188	136	155
X_3 Ohio Printing Performance Test	300	257	254
$b_{12.3}X_2$	-22.15	-16.02	-18.26
$b_{13.2}X_3$	+53.97	+46.27	+45.69
X_1 Predicted Mark	67	66	63
X_1 Obtained Mark	54	61	75

The standard error of estimate for this multiple regression equation was found to be approximately 13.0. The interpretation here is that the probability is 1 in 3 that the predicted value will differ more than ± 13 from the "true" value.

The final consideration was to examine in what way the sample of students employed in the follow-up study were representative of the students participating in the norming administration. The first analysis concerned the variances between the two samples for the two measures. The fact that no significant differences were found was interpreted to mean that the sample of students for the follow-up study came from the same population of students drawn upon to form the normative sample. Such an interpretation suggested a normative distribution for the follow-up study. Knowing the form of the distribution enables us to know whether a low validity coefficient could possibly have resulted in a curtailment of the new follow-up distribution of the scores for both the Ohio Printing Achievement Test and the Ohio Printing Performance Test. From this analysis, it was concluded that no such curtailment occurred.

Another analysis considered what differences may have occurred between the mean levels of achievement of the two groups for both measures. Significantly higher mean levels of achievement were found for the sample of students in the follow-up study. This suggests that the mean aggregate of students who did significantly better in the achievement tests when compared to the norm were also able to obtain jobs in printing. Also, when one attempts to predict job performance, he does so for these students who find employment in the printing industry. Predicting job performance, however, was not the ultimate objective of this validation study. For it was the intent of this phase to simply provide evidence with which to examine whether the two achievement measures could predict job performance at a level better than chance. Evidences based upon the idea of forecasting efficiency clearly show the achievement measures to have been adequately validated within this context.

Overview Multiple Prediction Measures

The index of forecasting efficiency and standard error of estimate reflecting the joint contribution of the achievement measures in predicting job performance must, also, be integrated in light of the possible sources of error variance concerning the validity of the job performance measure. Consideration has already been given to the "halo" effect resulting from the lack of systematic comparisons within groups of individuals. Although the "halo" effect does tend to reduce validity because of less than efficient discrimination between items, the total rating scores may still discriminate differences between subjects involved in this follow-up study. The necessary condition for this occurrence is the assumption that all supervisors adopt similar standards for appraising job performance. In other words, variability between raters concerning the same subjects should be minimized. Since there were no plans for comparing different supervisor ratings on the same subjects, the error associated with varying standards of appraisal was considered to be randomly distributed and independent.

The results (i.e., $R_{12.3} = 0.46$) suggest that this assumption for random error was not supported. Since the reliability and validity were optimum for the tryout administration, one must interpret the obtained multiple R correlation coefficient within the context of possible errors of validity for the Job Performance Criterion Measure. The obtained mean and standard deviation for the Job Performance Criterion Measure had a positively skewed distribution. This curtailment of the obtained distribution, arising undoubtedly from the sources of error variance previously noted, tended to reduce the validity coefficient. Therefore, the efficiency of the two achievement measures to predict a criterion measure at a level significantly better than chance is impressive since rigorous control of validity was not totally possible.

CHAPTER V

CONCLUSIONS AND IMPLICATIONS

The conclusions reached in this study are reported by those concerning the study design and those directly arrived fro .he test results.

Conclusions of Study Design

(1) The evidence presented in this study suggests that the procedures, methods, and assumptions employed were adequate in providing for the development of valid and reliable achievement measures of vocational printing at the twelfth-grade level of instruction.

(2) Concerning the achievement measures, errors associated with inaccuracies of measurement were rigorously controlled in terms of the standardization procedures adopted for administering and scoring the tests.

(3) In view of the possible sources of error variance associated with the Job Performance Criterion Measure and in view of the fact that the achievement measures did predict the criterion at a level better than chance, it was concluded that measures of student achievement could be validated against measures of job performance. This conclusion is particularly sound when one considers that the magnitude of the multiple R correlation coefficient was equivalent to a value normally obtained under conditions where the valid'ty is usually more rigorously controlled.

Conclusions of Tests

(4) The Ohio Printing Performance Test drew its content and testing operations almost exclusively from considerations of job production activities, a criterion situation selected because of its relevance for evaluating training in terms of its effect on early job performance. This test proved to be highly valid as supported by the following evidence.

a. A 0.66 validity coefficient obtained with a criterion of shop grades often considered as being notoriously unreliable.

b. Internal consistency evaluation in which evidence obtained supported the basic assumption of discrete and independent sub-test measures.

c. Multiple regression analysis in which rational weights did not depart significantly from those derived statistically.

d. An optimum balance between reliability (.90) and validity (i.e., optimal discrimination between part measures or low interpart correlations and high part correlations to total test score).

e. A significant and substantial validity correlation (0.40) against a criterion of job performance in which a rigorous control of validity was judged not to have been possible.

(5) Within the context of the strong evidence supporting the validity of the Ohio Printing Performance Test, the Ohio Printing Achievement Test is also considered to be valid. This

conclusion is supported with the following evidence.

a. The index of forecasting efficiency was increased from 8 per cent for the Ohio Printing Performance Test alone to 12 per cent when the Ohio Printing Achievement Test was added.

b. A concurrent validity correlation of 0.73 for the two was found to be significant at the 0.001 level.

(6) For those graduates who entered printing occupation, their mean level of achievement for both measures were significantly greater than the norm.

(7) Evidence from the tryout administration of the Job Performance Criterion Measure clearly suggested that an optimum balance was obtained between validity and reliability.

Some evidence was obtained to suggest that extremely high reliabilities resulting from the "halo" effect reduced the validity of the measure. This was due to a restricted condition of comparing a single employee to employees the supervisor had known at similar stages of employment. Defining the standard group in terms of individuals separated by points in time made it extremely difficult to make discriminating judgments for appraising an employee's job performance. The assumption of equivalent standards for all supervisors appraising employees was not supported and further reduced the validity by curtailing the form of the obtained distribution.

Implications

The conclusions in this study suggest the following implications:

(1) The significance of this study is that

the procedures, assumptions, and techniques underlying the development of a recognition-type testing situation may offer a solution for appraising student performance. The technique of criterion task analysis can demonstrate the skills relevant to early job performance and provide for assumptions underlying the nature of this kind of valid achievement testing. Such assumptions can be fruitfully employed to examine statistically the internal validity of the measure. This is particularly important since shop grades as criterion sources are usually unreliable and not generalizable beyond the individual school. Therefore, because of the ease with which this kind of testing can be validated, other achievement, ability, and aptitude measures can be examined concurrently with it.

(2) The fact that the administration of such a test was amenable to the simultaneous testing of large groups of students within short periods of time implies that schools may become more encouraged to participate in research activities of an evaluative type, particularly when the time allowances for testing are compatible to what schools can realistically allow.

(3) Concerning future prediction validation strategies, efforts must be made to realistically control the sample in order to reduce the errors resulting from variable standards of appraisal and the errors resulting from the "halo" effect.

(4) Future validation strategies relating achievement measures to job performance measures can provide an index reflecting the effectiveness of the instruction. However, such research should definitely be a phase of the overall validation strategies. The conditions which favor the "halo" effect and varying standards of appraisal need to be controlled possibly by working with small regional samples.

(5) Finally, it is feasible to construct valid and reliable measures of instruction for assessing both the end of training achievement and the relationship of that training to subsequent job performance. Therefore, it is feasible that achievement testing can detect general trends concerning the effectiveness of the instruction, thereby, allowing for the improvement and maintenance of quality education in sound vocational programs.

CHAPTER VI

SUMMARY

Measuring Student Achievement

The primary objective of this pilot study was to determine the feasibility of developing valid and reliable measures of student achievement in vocational printing at the twelfth-grade level. In support of this objective, conferences were held with representatives of the International Graphic Arts Education Association for the purpose of identifying educational objectives relevant to student achievement. Two categories were identified. One was concerned with the measurement of the student's knowledge and understanding of fundamental operations in printing, while the other concerned the student's ability to apply this knowledge and understanding to problem-solving situations normally encountered in printing operations. The Ohio Printing Achievement Test, a multiple-choice-type achievement test, was selected to measure the first category. Whereas measurement of the second cateogry required the development of a recognition-type achievement test now designated as the Ohio Printing Performance Test.

The Ohio Printing Performance Test provided for the indirect measurement of such important aspects of achievement as the ability to recognize essential characteristics of performance, choose the solution for a defined operational problem, judge the accuracy of specimens or products, locate and identify product defects, and indicate probable causes. The essential feature of this testing situation was the possibility of developing testing operations intrinsically similar to the actual problems commonly encountered in printing under conditions where group testing and standardized scoring systems were feasible.

Five sequential stages of the printing process were identified as Printing Planning, Composition, Imposition-Lockup, Presswork, and Lithography. An analysis into the nature of the stages generated several assumptions concerning their measurement. For example, each stage was characterized by a general performance ability because the nature of the instruction generally provided for an integration of fundamental problem-solving skills in order to attain the level of proficiency desired. On the other hand, the measurement of each stage was also thought to reflect relatively independent achievement because of the high degree of specificity involved in each. In other words, achievement was predicted to be characterized by a composite grouping of problem-solving abilities reflecting the proficiency desired within each of these five stages of the printing process.

A procedure was adopted for identifying and developing testing operations intrinsically similar to actual printing operations. This procedure derived the indirect recognition-type testing situation on the basis of analyzing a "true" production activity. Twenty-three tasks were derived through this analysis procedure. An achievement test with five sub-tests, each representing a stage of the printing process, was developed. Weights were established for each of the sub-tests by asking the committee to judge how much weight each should have in determining the total score. Underlying this notion was the assumption that weights for a part reflected not only the importance of the behaviors being measured in relation to the printing instruction and to the difficulty with which they are achieved.

Following the development of the Ohio Printing Performance Test, a tryout administration of both achievement measures was conducted using senior students enrolled in a vocational printing program at Dayton, Ohio's J. H. Patterson Cooperative High School. The purpose of the study was to examine the validity of the two against a criterion of shop grades derived from the students' performances in actual production activities.

However, because of the small sample size, no interpretations concerning these results were made other than noting the obtained correlations were in a positive direction.

The sample selection procedure for establishing national norms was to normalize a single distribution of student scores by sampling a large percentage of the schools which offer vocational printing programs. Letters were sent to state supervisors requesting a list of schools offering such training. Replies were received from 48 states of which 7 indicated they had no vocational printing programs. Subsequent letters were then sent to 238 schools requesting their participation in the norming administration. All schools who replied in the affirmative were accepted.

Seven-hundred and ninety-five students participated in the norming administration of the Ohio Printing Achievement Test and 745 in the Ohio Printing Performance Test. The totals represented 77 vocational printing programs which included 29 states and the District of Columbia.

Norms were established for both measures using the national population students nearing completion of a four-year vocational printing course. It was assumed that this population was normally distributed with regard to levels of achievement under the conditions of equivalent opportunities for learning. It was further assumed any variability in achievement would occur independently and be distributed normally throughout the population, that is, any variations in learning opportunities between schools were not systematically restricted to any geographic region or school enrollment.

In order to examine the degree to which the distribution of obtained scores for both measures were "true" measures of achievement, reliabilities were estimated for each. Since the Ohio Printing Performance Test was clerically scored, its reliability could not be effectively interpreted without examining to what degree inaccuracies of measurement might be the results of human errors in scoring. A comparison was, therefore, made between the means scores obtained by four independent scorers on the same group of subjects. This comparison revealed that the scoring system was nearly perfect (0.98) in terms of its reliability and objectivity.

Since the "true" achievement was unaffected by the adopted scoring procedures, the next step was to examine the reliabilities of the measures themselves. The estimates obtained were 0.90 and above suggesting that better than 80 per cent of the "true" scores for the obtained distribution of achievement scores was accounted for by the tests themselves. The remaining 20 per cent or less was attributed to errors which remained unidentified.

The next consideration was to determine whether the obtained distribution of test scores whose true distribution was, in fact, accountable in terms of these achievement tests was normally distributed. It was found that by examining the Ohio Printing Performance Test frequency distribution, no significant differences of departure from normality were found. Concerning the relationship between the distributions of the two measures, a high correlation (0.73) was found reflecting a very dependable relationship between the two measures. This also suggested the presence of a normal distribution for the Ohio Printing Achievement Test.

Since the study involved the development of a new measure of achievement, the next consideration was to analyze the data to test the assumptions and effectiveness of the procedures underlying the development of the Ohio Printing Performance Test. Intercorrelations between the sub-test scores did support the assumption that performance within a stage of printing was relatively independent of the other four stages. In addition, a multiple regression analysis suggested that the importance assigned by the committee to each stage of the printing operations was supported by the obtained achievement test scores. This evidence was interpreted to mean that the behaviors tested were valid in terms of the assumptions underlying their measurement.

Measuring Job Performance

The question of validity concerned how well the achievement measures predicted job performance. In other words, the intent was to define a validation strategy which took as its substance the evaluation of the achievement measures on the basis of job performance.

To determine this relevance, measures of job performance were needed for those students originally tested who subsequently became employed in the printing industry. Conferences were held with representatives from the printing industry for the purpose of defining job performance. Fourteen traits were selected and identified as the most critical, the most crucial and the most readily measurable of the job performance elements at this stage of employment. The traits were then adapted to a graphic scale format for distribution to the employee's supervisor.

Evidence with which to examine the validity and reliability of this newly developed measure was obtained from a tryout administration among 25 subjects as the Rosenthal Printing Company. The validity of the measure was obtained against a supervisory ranking of employees from highest to lowest. It was found that the sensitivity of the newly developed rating scale to discriminate differences in employee performance was substantially similar (0.85) to the rank ordering of the employees by the supervisors. Estimates of reliability or estimates of the "true" differences in performance as measured by this scale was 0.79. Therefore, the new job performance criteria measure was accepted as both valid and reliable.

To obtain from among the students tested measures of job performance, a follow-up study was conducted to determine what proportion of the students entered the printing industry. A questionnaire was mailed to the students for the purpose of determining their place of employment and the name of their immediate supervisor. Seven-hundred questionnaires were mailed to these students for whom measures for both achievement tests were obtained. More than 85 per cent of the questionnaires was returned for a total of 597. Of this total, 232 were employed full time in the printing industry.

Requests for ratings of job performance for these new employees were mailed to the respective supervisors. More than 85 per cent was returned of which 175 were returned for the remaining analysis of predicting job performance.

The condition underlying the validation strategy was that the two achievements measures must predict scores on the newly developed job performance rating scale at a level better than chance. The evidence obtained was clearly in support of this predetermined criterion level. For example, the joint effect of the two measures on predicting job performance was 12 per cent better than chance. This level of prediction was based on a substantial correlation of 0.46 in the same group of subjects between the joint effect of the two measures and the scores obtained for the job-performance criterion measure.

A source of error is "halo" effect or consistently rating an individual the same across all items. This normally occurs when a supervisor cannot compare an individual to a comparative group. In this study, the supervisor was asked to rate the newly employed graduate by comparing him with others he had known at similar stages of development and not with other employees who were on the job significantly longer. This condition inherent in the study made it impossible to control for the presence of the "halo" effect simply because of the absence of observations within groups of employees. Therefore, it was concluded that different supervisors employed different standards of appraisal which limited the measure's sensitivity for detecting "true" differences in performance. However, within the context of a 0.46 correlation and a prediction of 12 per cent better than chance it becomes quite impressive that the adequacy of the two tests does provide for valid and reliable measures of student achievement.

Basic Conclusions and Implications

Several substantial conclusions were reached as a result of this study. They are:

1. The Ohio Printing Achievement Test is a valid and reliable instrument.

2. The Ohio Printing Performance Test is a valid and reliable instrument.

3. The Job Performance Criterion Measure used in this study is a valid and reliable instrument.

4. The graduates who entered printing occupations had a significantly greater mean level of achievement than the norm group.

As a result of the above conclusions, two strong implications arise. One is that valid and reliable measures can be constructed for assessing student achievement and job performance. The second implication is that achievement testing can detect general trends concerning the effectiveness of instruction which may be used in the improvement of instruction.

LIST OF REFERENCES

1. Cronback, Lee J. Essentials of Psychological Testing. New York: Harper and Bros.; 1960

2. Ferguson, George A. Statistical Analysis in Psychology and Education. New York: McGraw-Hill Book Co.; 1959

3. Guilford, J. P. Fundamental Statistics in Psychology and Education. New York: McGraw-Hill; 1965

4. Guilford, J. P. Psychometric Methods. New York: McGraw Hill; 1936

5. Kendall, M. G. Rank Correlation Methods. London: Griffin; 1948

6. Meyers, Jerome L. Fundamentals of Experimental Design. Boston: Allyn and Bacon; 1966

7. Remmers, H. H. "Reliability and halo effect of high school and college students judgments of their teachers." Journal of Applied Psychology., 1936, 18, 619-630.

8. Remmers, H. H. & Elliott, D. Manual, The Purdue Rating Scale for Instruction. West Lafayette, Ind.: University Book Store; 1960

9. Richardson, M. W. The Combination of Measure. Social Science Research Council Bulletin 48 New York: The Council; 1941., 377-401.

10. Thorndyke, Robert L. & Elizabeth Hagan. Measurement and Education In Psychology and Education. New York: John Wiley and Sons; 1961

11. Tucker, L. R. "Maximum validity of a test with equivalent items." Psychometrika., 1946, 11, 1-13.

12. Winer, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill Book Co.; 1962.

APPENDIX A

SUGGESTED PRINTING COURSE OUTLINE

APPENDIX A

PRINTING

Suggested

COURSE OUTLINE

PUBLISHED BY
INSTRUCTIONAL MATERIALS LABORATORY
TRADE AND INDUSTRIAL EDUCATION
THE OHIO STATE UNIVERSITY
1885 NEIL AVENUE - ROOM 112
COLUMBUS, OHIO 43210

Trade and Industrial Education
Division of Vocational Education
State Department of Education
Columbus, Ohio

PRINTING

COURSE OUTLINE

I. ORIENTATION

   A. History of Graphic Communication

      1. Beginning of Graphic Communication

      2. Development of Roman Alphabet

      3. Invention of Moveable Type

      4. Historical Development of Modern Printing

   B. Printing Process

      1. Letterpress

      2. Lithography

      3. Intaglio

      4. Screen Process

      5. Special Processes

   C. Scope of Graphic Arts Industry

      1. Socio-Economic Importance

      2. Occupational Opportunities
         a. Number of Jobs
         b. Wages
         c. Types of Printing Establishments
         d. Allied Establishments
            (1) Supplies
            (2) Equipment
         e. Training Required

II. PRINTING PLANNING

   A. Manuscript Preparation

   B. Illustration Preparation

C. Layout and Design

1. Principles

2. Use of Color

3. Hand Lettering

4. Rough Sketch

5. Finished Layout

D. Production Specifications

1. Selection of Process

2. Paper and Ink Selection

3. Type Selection

4. Scaling

E. Camera Copy

1. Reproduction Proofs

2. Paste-up Copy

3. Overlays

III. COMPOSITION

A. Safety

B. Hand Composition

1. Printers' Point System

2. Parts of a Piece of Type

3. Lay of California Job Case

4. Spacing Material

a. Spaces and Quads
b. Leads and Slugs

5. Composing Stick

6. Typesetting
   a. Justification
   b. Quadding Out
   c. Centering
   d. Word Spacing
   e. Letter Spacing
   f. Word Division
   g. Indentions
   h. Troublesome Letters

7. Placing Type on Galley

8. Proofing
   a. Galley
   b. Stone
   c. Reproduction Proofs

9. Proofreading

10. Correcting

11. Make-up

12. Tabular Composition

13. Imposition and Lock-up
   a. Tools and Materials
   b. Open Platen Press Lock-up
   c. Automatic Platen Press Lock-up
   d. Cylinder Press Lock-up

14. Distribution

C. Machine Composition and Material Casting

1. Linotype and Intertype
   a. Nomenclature
   b. Operating Principles
      (1) Keyboard
      (2) Assembling
      (3) Casting
      (4) Distributing
      (5) Auxiliary Attachments

(a) Saw
(b) Quadder
(c) Blower
c. Tape Operation
(1) Tape Punch
(2) Keyboard Unit
d. Type Metal

2. Ludlow Typograph
a. Nomenclature
b. Operating Principles
(1) Matrix Cases
(2) Sticks
(3) Casting
(4) Distribution of Matrices
c. Type Metal

3. Monotype
a. Keyboard
(1) Nomenclature
(2) Operating Principles
b. Composition Casters
(1) Nomenclature
(2) Operating Principles
c. Type Metal

4. Material Makers
a. Type and Sorts
(1) Monotype
(2) Thompson
b. Strip Material
(1) Elrod
(2) Monotype
(3) Linotype and Intertype

D. Photo Composition

1. Intertype Fotosetter
a. Nomenclature
b. Operating Principles

2. Linofilm
a. Nomenclature
b. Operating Principles

3. ATF Typesetter
    a. Nomenclature
    b. Operating Principles

4. Photon
    a. Nomenclature
    b. Operating Principles

5. Monophoto
    a. Nomenclature
    b. Operating Principles

6. Display
    a. ATF Hadego
    b. Headliner
    c. Filmotype
    d. Protype
    e. Others

E. Copy Producing Typewriters

1. Varityper
    a. Nomenclature
    b. Operating Principles

2. Justowriter
    a. Nomenclature
    b. Operating Principles

3. Special Typewriters

IV. COPY CAMERA WORK

A. Safety

B. Theory

1. Wave Length Concept of Light

2. Photographic Film
    a. Emulsions
        (1) Photomechanical
        (2) Commerical
        (3) Color Sensitivity
            (a) Regular
            (b) Orthochromatic
            (c) Panchromatic

b. Bases
c. Antihalation Coating

C. Copy Camera

1. Nomenclature

2. Operating Principles

D. Film Processing

1. Chemistry of Photography

2. Developing

3. Short Stop

4. Fixing

5. Rinsing

6. Intensifying

7. Reducing

E. Exposure

1. Line Copy

2. Halftone
    a. Glass Screen
    b. Contact Screen
    c. Autoscreen

3. Positives
    a. Contact
    b. Autopositives

4. Filters

5. Color
    a. Line
    b. Process

V. PLATEMAKING

A. Safety

B. Letterpress Plates

1. Photoengraving
   a. Film Imposition
   b. Line
   c. Halftone
   d. Combination
   e. Plate Preparation
   f. Exposure
   g. Etching
   h. Routing and Finishing

2. Polyester Plates
   a. Film Imposition
   b. Exposure
   c. Wash out
   d. Finishing

3. Electronically Scanned Plates

4. Duplicate Plates
   a. Sterotypes
   b. Electrotypes
   c. Rubber
   d. Plastic

C. Lithographing Stripping

1. Preparation of Negatives
   a. Ruling
   b. Line-up or Center Marks
   c. Inserts or Strip-ins
   d. Opaquing

2. Making the Layout
   a. Press Gripper
   b. Plate Gripper
   c. Line-up or Center Marks

3. Marking the Flat
   a. Imposition of Negatives
   b. Screen Tints

4. Photocomposing Machines
    a. Lanston
    b. Rutherford
    c. Other Step and Repeat Methods

D. Lithographic Plates

1. Albumir
    a. Film Imposition
    b. Plate Preparation
    c. Exposure
    d. Development

2. Deep Etch
    a. Film Imposition
    b. Plate Preparation
    c. Exposure
    d. Development

3. Presensitized
    a. Film Imposition
    b. Exposure
    c. Development

4. Xerox
    a. Plate Preparation
    b. Exposure
    c. Transfer
    d. Fusion

5. Multi-Metal

6. Direct Image

VI. PRESSWORK

A. Safety

B. Letterpress

1. Kinds of Presses
    a. Platen
        (1) Open
        (2) Automatic

b. Cylinder
   (1) Horizontal
      (a) Open
      (b) Automatic
   (2) Vertical
   (3) Single Color
   (4) Multicolor
   (5) Perfector
c. Rotary
   (1) Sheet Fed
   (2) Web Fed
   (3) Wrap Around
   (4) Tandem
   (5) Multicolor
   (6) Perfector
d. Indirect Letterpress (Dry Offset)

2. Press Preparation

3. Make-ready
   a. Overlay
   b. Underlay
   c. Interlay

4. Press Operation
   a. Feeder
   b. Printing Unit
   c. Inking
   d. Delivery

5. Special Operations
   a. Die Cutting
   b. Scoring
   c. Embossing
   d. Perforating

C. Lithographic

1. Kinds of Presses
   a. Direct
   b. Offset
      (1) Sheet Fed
      (2) Web Fed
      (3) Tandem
      (4) Multicolor
      (5) Perfector

2. Press Preparation

3. Press Operation
    a. Feeder
    b. Dampening
    c. Inking
    d. Printing Unit
    e. Delivery

VII. BINDERY WORK

A. Safety

B. Cutting

C. Folding

D. Scoring

E. Collating

F. Stitching

G. Round Cornering

H. Drilling

I. Perforating

J. Padding

K. Binding

L. Trimming

M. Packaging

VIII. PAPER TECHNOLOGY

A. Manufacture of Paper

B. Kinds of Paper

1. Fine
    a. Newsprint
    b. Book

(1) Letterpress
(2) Offset

c. Writing
d. Cover
e. Cardboard

2. Coarse
   a. Craft
   b. Manila

C. Finishes

1. Coated

2. Uncoated

3. Embossed

D. Weights

1. Basis (Substance)

2. "M" Weight

3. Equivalent

E. Characteristics

1. Grain

2. Felt Side

3. Water Mar

4. Rag Content

IX. INK TECHN LOGY

A. Composition of Inks

1. Pigment

2. Vehicle

3. Modifiers (including driers)

B. Manufacture of Ink

1. Grinding

2. Mixing

3. Milling

C. Mixing and Matching Color

D. Handling and Care of Inks

E. Ink Characteristics

1. Viscosity

2. Tack

3. Length

4. Drying

5. Opacity

6. Transparency

X. PRINCIPLES OF SCIENCE AND MATHEMATICS

A. Applied Science

1. Effect of heating and cooling on expansion of materials. (Change of dimensions.)

2. Effect of heating and cooling on state of matter. (Change of matter from one form to another.)

3. Differences in absorption and radiation of energy between dark rough surfaces and light, smooth, polished surfaces.

4. Compostion of matter, including protons, neutrons, electrons, atoms, molecules, elements.

5. Centrifugal forces developed by bodies in rotation. (Example: Force tending to discharge material from a rotating body.)

6. Forces acting on a body immersed or floating in a liquid.

7. Transfer of heat from one body to another.

8. Inertia and momentum. (Body at rest and body in motion.)

9. Effects of friction on work processes and product quality.

10. Relationship of force to distortion in an elastic body.

11. Resistance of materials to change in shape. (Examples: Bending, twisting, stretching.)

12. Inter-relationship of acids, bases and salts.

13. Effect of humidity on certain forms of matter.

14. Collodial Chemistry.

15. Affinity.

16. Compounds and Mixtures.

17. Emulsification.

18. Oxidation.

19. Electrostatics.

B. Applied Mathematics

1. Addition and subtraction of whole numbers.

2. Multiplication and division with whole numbers.

3. Reduction of fractions. (Example: 12/16 = 3/4)

4. Addition and subtraction of proper (Example: 3/4) and improper (Example: 11/8) fractions.

5. Multiplication and division of proper and improper fractions.

6. Changing mixed numbers to improper fractions. (Example: 4-3/4 = 19/4)

7. Addition and subtraction of decimal fractions.

8. Multiplication and division of decimal fractions.

9. Rounding off decimals and whole numbers (Example: .4877 = .488 when rounding to three decimal places.)

10. Changing percents to fractions and fractions to percents.

11. Finding a percent of a number and what percent one number is of another.

12. Measures of length. (Example: Inches, feet, etc.)

13. Measure of time and speed. (Example: time-seconds, minutes, etc.;speed - feet per minute, R.P.M., etc.)

14. Measures of weight.

15. Measures of temperature.

16. Liquid and dry measures.

17. Solution of problems involving numerical algebraic expressions. (Example: -12 + 7 -6(6-1) = -40.)

18. Use of formula to determine the area of a circle.

19. Determination of area, perimeter and diagonals of quadrilaterals (4 sided figures).

20. Determination of area and circumferance of circles.

21. Ratio and proportion.

PRINTING

Achievement Test Areas

1. ORIENTATION
2. PRINTING PLANNING
3. HAND COMPOSITION
4. MACHINE COMPOSITION
5. PHOTO COMPOSITION
6. CAMERA OPERATION
7. FILM PROCESSING
8. LETTERPRESS PLATEMAKING
9. LITHOGRAPHIC PLATEMAKING and STRIPPING
10. LETTERPRESS PRESS WORK
11. LITHOGRAPHIC PRESS WORK
12. BINDERY WORK
13. PAPER TECHNOLOGY
14. INK TECHNOLOGY
15. APPLIED SCIENCE
16. APPLIED MATHEMATICS

PRINTING

BIBLIOGRAPHY

Auble, J. Woodard, Arithmetic For Printers, 237 N. Monroe Street, Peoria, Illinois: Charles A. Bennett Co., Inc., 1954. (Out of print)

Carlsen, Darvey E., Graphic Arts, 237 N. Monroe Street, Peoria, Illinois: Charles A. Bennett Co., Inc. 1965.

California State Department of Education, Industrial Arts and Science, 721 Capital Mall, Sacramento, California, 95814: California State Department of Education, 1965.

Cleeton, Glen U., Pitkin, Charles W. and Cornwell, Raymond L., General Printing, Bloomington, Illinois: McKnight & McKnight Publishing Co., 1963.

Cogoli, John E., Photo-Offset Fundamentals, Bloomington, Illinois: McKnight & McKnight Publishing Co., 1960. (To be revised 1966-67)

Delmar Publishers, Inc., Printing Layout and Design, Mountainview Avenue, Albany 5, New York: Delmar Publishers, Inc., 1955.

Eisenberg, James and Kafka, Francis J., Silk Screen Printing, Bloomington, Illinois: McKnight & McKnight Publishing Co., 1957.

Graphic Arts Technical Foundation, Offset Platemaking (Surface), 4615 Forbes Avenue, Pittsburg, Pa. 15213: Graphic Arts Technical Foundation

Graphic Arts Technical Foundation, Offset Platemaking (Deep Etch), 4615 Forbes Ave., Pittsburgh, Pa. 15213: Graphic Arts Technical Foundation, 1963.

Graphic Arts Technical Foundation, Offset Photography (Line), 4615 Forbes Avenue., Pittsburg, Pa. 15213: Graphic Arts Technical Foundation.

Graphic Arts Technical Foundation, Offset Stripping (Black and White), 4615 Forbes Ave., Pittsburg, Pa. 15213: Graphic Arts Technical Foundation, 1964.

Graphic Arts Technical Foundation, Offset Photography (Halftone), 4615 Forbes Ave., Pittsburg, Pa. 15213: Graphic Arts Technical Foundation, 1964.

Graphic Arts Technical Foundation, Color Stripping, 4615 Forbes Ave., Pittsburg, Pa. 15213: Graphic Arts Technical Foundation, 1964.

Government Printing Office, U.S. Navy Handbooks, Lithographer 3 and 2. Catalog No. D208.11-L 71/963, Washington, D.C.: U.S. Government Printing Office

Glafkides, Pierre, Photographic Chemistry, Vol. 1 & Vol. 2, 915 Broadway, New York, New York: American Photographic Book Publishing Co.

Greenleaf, Allen R., Chemistry For Photographers, 33 W. 60th Street, New York 23, New York: American Photographic Publishing Co. (Out of print)

Kagy, Frederick D., Graphic Arts, 1825 Harwood, Homewood, Illinois 60430: The Goodheart-Wilcox Co., Inc. 1965.

Karch, R. Randolph, Graphic Arts Procedures, 848 E. 58th Street, Chicago, Illinois 60637: American Technical Society, 1965.

Karch, R. Randolph, Printing and the Allied Trades, 20 E. 46th Street, New York, New York 10017: Pitman Publishing Corp. 1962.

Mergenthaler Linotype Co., Linotype Keyboard Operation, 29 Ryerson, Brooklyn, New York: Mergenthaler Linotype Co.

Mergenthaler Linotype Co., Linotype Machine Principles, 29 Ryerson, Brooklyn, New York: Mergenthaler Linotype Co.

Mertle, J. S. & Monsen, Gordon L., Photomechanics and Printing, 1308 Menominee Drive, Oshkosh, Wisconsin: Mertle Publishing Co., 1957.

Mills, George J., Platen Press Operation, Available from author, 4136 Bigelow Blvd., Pittsburg, Pa. 15213, 1959.

Polk, W. Ralph and Polk, Edwin W., The Practice of Printing, 237 N. Monroe Street, Peoria, Illinois: Charles A. Bennett Co., Inc., 1964.

Sayre, I. H., Photography and Platemaking for Photolithography, 30 W. Washington Street, Chicago, Illinois 60602: Lithographic Textbook Publishing Co., 1965.

APPENDIX B

JOB PERFORMANCE RATING

Appendix B

I.D. Number

JOB PERFORMANCE RATING

NAME OF EMPLOYEE________________________________

NAME OF COMPANY_________________________________

DEPARTMENT______________________DATE____________

RATED BY____

JOB TITLE OF RATER______________________________

How long has this employee worked for you?

Please check (✓)

_______ under one month
_______ one to two months
_______ three to five months
_______ six months to a year

How often do you see him in a work situation?

_______ See him at work all the time
_______ See him at work several times a day
_______ See him at work several times a week
_______ Seldom see him in a work situation

General Directions

USE OF RATING SCALE

We are asking you to rate the job performance of a newly employed person working for you. In making this rating, do not let general impressions or one outstanding trait affect your judgment. Try to forget your personal feelings about the worker and rate him only on the way he performs his work. Below are some additional points which may prove helpful in rating this person.

1. Read the directions for scoring the performance thoroughly before rating.

2. Be particularly careful that each item is read carefully before the rating is begun. Serious errors of judgment result when definitions and terms are misinterpreted or misread.

3. Rate only the items listed in the rating sheet. Do not attempt to "read in" anything more than what is stated in the definition of the items. We are interested in obtaining a rating of job performance only as they are defined and presented on the rating sheet.

4. Practice and experience usually improves a worker's skill. Remember that we want a rating of how well this employee is performing during the early months of his initial employment experience. Do not rate him as poorer than another because he has not been on the job as long.

5. FOR EACH ITEM, THEREFORE, COMPARE THE EMPLOYEE TO OTHER WORKERS YOU HAVE KNOWN AT A SIMILAR STAGE OF WORK EXPERIENCE.

6. Rate the employee according to the work he has done over a period of several weeks. Do not rate on the basis of one "good" day, or one "bad" day, or some single incident. Think in terms of the employee's usual or typical performance.

Specific Directions

On the following page begins a list of fourteen characteristics that, when taken together, offer an estimate of the employee's proficiency on the job. No employee is, of course, perfect in all of the characteristics listed for your consideration, and his level of performance will vary with respect to any one of them. The rating scale does enable you to be discriminating in your judgment of his performance by providing a scale for each characteristic ranging from the lowest level of performance to the highest. A rating on each characteristic is then obtained simply by placing a check mark on the scale at a point which most nearly describes the employee. For example:

Accuracy: Consider the extent to which his work is free of errors and meets plant standards for quality.

Often makes mistakes | | | | | | Rarely makes mistakes

1. A check in the middle of the scale indicates average performance. A check at the extreme left indicates the lowest level of performance and at the extreme right, the highest.
2. A check may therefore be placed at any point along the line according to your judgment of his performance.

Job Performance Rating Scale

(Answers will be kept confidential)

1. Dependability: Consider the employee's reliability in carrying out instructions.

Unreliable — Exceptionally Dependable

2. Safety: Consider the extent to which the employee practices safe work habits.

Takes Unnecessary Risks Careless Worker — Minimizes Chances for Accidents Extremely Safe Worker

3. Quantity: Consider the amount of work performed in relation to desired expectations.

Considerably Below Expectations — Considerably Above Expectations

4. Tools and Equipment: Consider how well the employee maintains tools and equipment.

Unsatisfactory Care of Tools and Equipment — Takes Exceptional Care of Tools and Equipment

5. Resourcefulness: Consider the extent to which the employee handles work problems without assistance.

Experiences Considerable Difficulty — Highly Effective

6. <u>Neatness</u>: Consider the extent to which the employee is neat and orderly in his work.

Unsatisfactory — Satisfactory

7. <u>Accuracy</u>: Consider the extent to which his work is free of errors.

Frequently Makes Mistakes — Rarely Makes Mistakes

8. <u>Industriousness</u>: Consider the employee's willingness to move ahead on job assignments.

Needs Constant Prodding — Moves Ahead on His Own

9. <u>Reaction to Criticism</u>: Consider the employee's attitude with respect to criticism.

Resents Criticism — Accepts Criticism Willingly

10. <u>Adaptability</u>: Consider the employee's ability to adjust to changes in job situations.

Fails to Adjust — Adjust Rapidly

11. <u>Communication</u>: Consider the effectiveness of the employee in expressing himself.

Ineffective — Highly Effective

12. Organization of Work: Consider the extent to which the employee follows a logical work pattern in completing his assignment.

Haphazard Does not plan ahead	Efficient Worker Plans Effectively

13. Technical Knowledge: Consider the extent of the employee's knowledge about his job. (That is, his understanding of the technical knowledges, equipment, materials and methods that have to do directly or indirectly with his job.)

Has little knowledge Knows enough to get by	Has broad knowledge Knows enough to do his job thoroughly

14. Job Skill: Consider the extent of the employee's aptitude or facility for this kind of job. (That is, the employee's "knack" for performing the job easily and well.)

Limited aptitude for this job	Exceptionally well suited for this kind of job

GENERAL COMMENTS:

APPENDIX C

GRADUATE QUESTIONNAIRE

Appendix C

GRADUATE QUESTIONNAIRE

(ANSWERS WILL BE KEPT CONFIDENTIAL)

Printing Achievement Test Participants
1966 Graduates

The Ohio Trade and Industrial Education Service

Name (Print)______________________________Date__________

(Last) (First) Middle)

Present Address____________________________Phone__________

(Street)

____

(City) (State)

1. What are you now doing? (Check one or more)

______A. Working for pay, full time
______B. Working for pay, part time
______C. In business for self
______D. Not working, but looking
______E. In armed services
*______F. In school, full time
*______G. In school, part time
**______H. Other, please describe below

*Type of School (Technical Institute, College, etc.)____

____

**____

2. Are you employed in the printing trades?

______Yes ______No

If yes, check below the department in which you are now working.

______A. Composition
______B. Press room
______C. Lithography
______D. Bindery
______E. Photo engraving
______F. Other, please describe below

____

____

3. Have you been accepted into an apprenticeship training program? ______Yes ______No

4. At what place do you now work?

Employer or Firm______________________________

Address______________________________
(Street) (City) (State)

Supervisor's Name______________________________

(Last) (First) (Middle)

Date you started______________________________

APPENDIX D

SUMMARY TABLES

Table D-1
SUMMARY DATA FOR OBTAINING PEARSON PRODUCT MOMENT CORRELATION COEFFICIENTS BETWEEN EACH OF THE PROFICIENCY MEASURES AND THE CRITERION OF SHOP GRADES (Y VARIABLE)

X Variables	ΣX	ΣY	ΣX^2	ΣY^2	ΣXY
Ohio Printing Achievement Test	3203	68	575,479	266	12,120
Ohio Printing Performance Test	4842	68	1,328,820	266	18,579

Table D-2
ANALYSIS OF VARIANCE OF OHIO PRINTING PERFORMANCE TEST SCORING SYSTEM

Source of Variation	Sum of Squares	df	Mean Square
Between Students	876,843	73	12.012
Within Students	27,649	222	124.54
Total	904,492	295	

Table D-3
SUMMARY OF DATA USED FOR ESTIMATING THE RELIABILITY OF THE OHIO PRINTING PERFORMANCE TEST

	X	Y	X^2	Y^2	XY
Sum	8305	8287	981,847	968,920	966,652

Table D-4

OBTAINING THE EXPECTED FREQUENCIES IN THE CLASS INTERVALS OF THE OHIO PRINTING PERFORMANCE TEST ON THE ASSUMPTION THAT THE TRUE DISTRIBUTION IS NORMAL

Scores	fe Expected Frequencies	fo Observed Frequencies
406-420	0.78	0
391-405	1.6	2
376-390	3.3	1
361-375	6.2	4
346-360	10.9	11
331-345	17.4	21
316-330	26.4	35
301-315	37.3	36
286-300	49.5	46
271-285	62.5	64
256-270	71.5	74
241-255	76.8	71
226-240	77.3	76
211-225	72.9	68
196-210	64.5	58
181-195	52.0	45
166-180	41.5	52
151-165	29.1	34
136-150	19.6	19
121-135	12.5	19
106-120	7.4	5
91-105	3.8	2
76- 90	1.9	2
61- 75	0.08	0
	Sum 747.56	Sum 745

Table D-5
ANALYSIS OF VARIANCE OF THE JOB PERFORMANCE CRITERION MEASURES

Source of Variation	Sum of Squares	df	Mean Squares
Between Employees	261	13	20.08
Within Employees	126	336	.38

Table D-6
SUMMARY DATA FOR OBTAINING PEARSON PRODUCT MOMENT CORRELATION COEFFICIENTS BETWEEN THE ROSENTHAL ADJECTIVAL MEASURE AND THE JOB PERFORMANCE CRITERION MEASURE (Y VARIABLE)

X Variable	ΣX	ΣY	ΣX^2	ΣY^2	ΣXY
Rosenthal Adjectival Measure	1,421	1,746	81,969	125,676	101,005

Table D-7

ANALYSIS OF VARIANCE OF THE JOB PERFORMANCE CRITERION MEASURE

Source of Variation	Sum of Squares	df	Mean Squares
Between Employees	54,018	174	3104
Within Employees	1,950	2,262	0.8621
Total	55,968	2,436	

Table D-8

SUMMARY OF DATA USED FOR ESTIMATING THE RELIABILITY OF THE JOB PERFORMANCE CRITERION MEASURE

	X	Y	X^2	Y^2	XY
Sum	5,505	5,721	182,841	197,791	189,291

Table D-9

INTERCORRELATIONS AMONG THREE VARIABLES, INCLUDING THE JOB PERFORMANCE CRITERION MEASURE AND THE TWO ACHIEVEMENT MEASURES

Variable	X_2	X_3	X_1
X_2	-	0.73**	0.15
X_3	0.73**	-	0.40**
X_1	0.15	0.40**	-

** 0.001 Level of Significance

X_2 Ohio Printing Achievement Test

X_3 Ohio Printing Performance Test

X_1 Job Performance Criterion Measure

Table D-10

SUMMARY DATA FOR OBTAINING INTERCORRELATIONS BETWEEN THREE VARIABLES

Variables	ΣX	ΣY	ΣX^2	ΣY^2	ΣXY
(r_{21}) O.P.A.T. vs. Criterion	29,904	11,211	5,380,016	756,327	1,930,439
(r_{31}) O.P.P.T. vs. Criterion	46,796	11,211	12,982,986	756,327	3,051,620
(r_{23}) O.P.A.T. vs. O.P.P.T.	29,904	46,796	5,380,016	12,982,986	8,257,031

APPENDIX E

PARTICIPATING SCHOOLS AND STATES

Appendix E

National Printing Study

Schools Which Participated in National Printing Study - 1966

SCHOOL	CITY
ARIZONA	
Tucson High School	Tucson
ARKANSAS	
Metropolitan High School	Little Rock
CALIFORNIA	
Santa Cruz High School	Santa Cruz
COLORADO	
Colorado Springs, School #11	Colorado Springs
CONNECTICUT	
A. I. Prince Vocational-Technical School	Hartford
Bul' rd-Havens Technical School	Bridgeport
H. C. Wilcox Technical School	Meriden
J. M. Wright Technical School	Stamford
DISTRICT OF COLUMBIA	
Bell High School	Washington
Chamberlain High School	Washington
Phelps High School	Washington

NATIONAL PRINTING STUDY (Cont.)

SCHOOL	CITY
FLORIDA	
Dixie Hollins High School	St. Petersburg
Lindsey Hopkins Education Center	Miami
Mid-Florida Technical School	Orlando
IDAHO	
Boise High School	Boise
ILLINOIS	
East Alton-Wood River High School	Wood River
J. S. Morton High School	Cicero
Proviso East High School	Maywood
INDIANA	
Bartholomew High School	Columbus
Bloomington High School	Blocmington
Elkhart High School	Elkhart
Gerstmeyer High School	Gerstmeyer
Hammond High School	Hammond
Marion High School	Marion
New Albany High School	New Albany
IOWA	
Fort Dodge Senior High School	Fort Dodge
KANSAS	
Labette County Community High School	Alamont
Pittsburg High School	Pittsburg
S. E. Kansas Area Vocational Technical School	Columbus

NATIONAL PRINTING STUDY (Cont.)

SCHOOL	CITY
KENTUCKY	
Northern Kentucky State Vocational School	Covington
Owensboro Area Vocational School	Owensboro
Somerset Area Vocational-Technical School	Somerset
MARYLAND	
Bladensburg Senior High School	Bladensburg
Fairmont Heights High School	Washington, D. C.
MASSACHUSETTS	
Attleboro Trade High School	Attleboro
Newton Technical High School	Newtonville
Pittsfield Vocational High School	Pittsfield
Springfield Trade High School	Springfield
Weldon Vocational High School	Medford
Weymouth Vocational Technical High School	Weymouth
MICHIGAN	
Bay City Central High School	Bay City
Hamtramck High School	Hamtramck
Lansing Eastern High School	Lansing
MINNESOTA	
Duluth Area Vocational-Technical School	Duluth
MISSOURI	
Kansas City Public Schools	Kansas City
O'Fallon Technical High School	St. Louis

NATIONAL PRINTING STUDY (Cont.)

SCHOOL	CITY
NEBRASKA	
Lincoln High School	Lincoln
NEW JERSEY	
Essex County Vocational School	Newark
NEW YORK	
McKee Vocational Technical High School	Staten Island
Sewanhaka High School	Floral Park, L.I.
Thomas A. Edison Vocational Technical High School	Jamaica
OHIO	
Akron Hower Vocational High School	Akron
Ashland High School	Ashland
Ashtabula High School	Ashtabula
Canton Timken Vocational High School	Canton
Courter High School	Cincinnati
Dayton Patterson Cooperative High School	Dayton
Lima Senior High School	Lima
Parma Senior High School	Parma
Parma Valley Forge	Parma Heights
Toledo Macomber Vocational High School	Toledo
Wadsworth Senior High School	Wadsworth
OKLAHOMA	
Midwest City High School	Midwest City
Muskogee Central High School	Muskogee
Oklahoma City Central High School	Oklahoma City
Oklahoma City Douglass High School	Oklahoma City
Will Rogers High School	Tulsa

NATIONAL PRINTING STUDY (Cont.)

SCHOOL	CITY
PENNSYLVANIA	
Bucks County Technical School	Fairless Hills
Eastern Montgomery County Vocational Technical School	Willow Grove
Easton Area Vocational Technical School	Easton
Pittsburgh Connelley Vocational Technical High School	Pittsburgh
TENNESSEE	
Booker T. Washington High School	Memphis
TEXAS	
Brackenridge High School	San Antonio
Fort Worth Technical High School	Fort Worth
L. W. Fox Vocational Technical High School	San Antonio
Sidney Lanier High School	San Antonio
VERMONT	
Brattleboro Union High School	Brattleboro
Burlington High School	Burlington
WEST VIRGINIA	
Huntington East Vocational Technical School	Huntington
McKinley Vocational High School	Wheeling
WISCONSIN	
Memorial High School	Eau Claire

APPENDIX F

MULTIPLE REGRESSION ANALYSIS

Appendix F

MULTIPLE REGRESSION ANALYSIS

The same random selection of students (N=74) was employed for the multiple regression analysis as was used in the sub-test correlations. The purpose of this study was to obtain evidence with which to compare the assigned rational weights to the statistical weights derived by this analysis. A linear relationship was assumed to exist among the five parts and the total test and the Beta coefficients were obtained by solving normal equations through the use of the Doolittle Method (Guilford (3). Table F-1 presents the intercorrelations among the six variables, using the total test score as the criterion and the five test parts as the independent variables. Means and standard deviations used to determine the contribution of each part to total test variance are also presented. The variances obtained for each part are presented in Table F-2 under Column 4.

Table F-1

INTERCORRELATIONS AMONG SIX VARIABLES, INCLUDING THE TOTAL TEST AND THE FIVE RESPECTIVE SUB-TESTS

Variable	X_2	X_3	X_4	X_5	X_6	X_1
X_2	-	0.4442	0.1572	0.4194	0.3404	0.6251
X_3	0.4442	-	0.2789	0.4199	0.3632	0.7030
X_4	0.1572	0.2789	-	0.3589	0.5120	0.6088
X_5	0.4194	0.4199	0.3589	-	0.3653	0.7132
X_6	0.3404	0.3632	0.5120	0.3653	-	0.8172
X_1	0.6251	0.7030	0.6088	0.7132	0.8172	-
M_x	29.10	76.24	21.81	15.84	64.43	223.55
S.D.	11.23	15.77	8.94	31.96	25.59	55.66

X_2 = Printing Planning
X_3 = Composition
X_4 = Imposition-Lockup
X_5 = Presswork
X_6 = Lithography
X_1 = Ohio Printing Performance Test

Table F-2

SOLUTION OF THE REGRESSION COEFFICIENTS FOR THE MULTIPLE REGRESSION EQUATION

1	2	3	4
X_2	0.2011	0.6251	0.1257
X_3	0.2752	0.7029	0.1934
X_4	0.1604	0.6088	0.0976
X_5	0.2857	0.7132	0.2037
X_6	0.4606	0.8172	0.3764

$R^2 = 0.9962$

2 = Beta Coefficients

3 = Correlation of Sub-tests tc Total

4 = Variance Contributed by Each Sub-test