A PERFORMANCE PRETEST FOR PLACEMENT OF COLLEGE STUDENTS IN BEGINNING CLOTHING COURSES

by

GROVALYNN FOREMAN GOULD

Bachelor of Science

Oklahoma State University

Stillwater, Oklahoma

1957

Submitted to the faculty of the Graduate School of
the Oklahoma State University
in partial fulfillment of the requirements
for the degree of
MASTER OF SCIENCE
May, 1963

A PERFORMANCE PRETEST FOR PLACEMENT OF COLLEGE STUDENTS IN BEGINNING CLOTHING COURSES

Thesis Approved:

Juanita M. Nord
Thesis Adviser

Nell P. Logan

Robert MacVicar
Dean of the Graduate School

ACKNOWLEDGEMENT

Indebtedness is acknowledged to Dr. Juanita M. Noel, thesis adviser, for her interest, encouragement and guidance throughout the study; to Dr. Nell P. Logan for her inspiration and assistance; to Dr. Eugene L. Swearingen for his suggestions regarding the statistical treatment of the data; to the Simplicity Pattern Company and the McCall's Pattern Company for the demonstration size patterns used in the test; to the General Foods Corporation for the generous fellowship which made this year of study possible; and to Mary Jane Walters for typing the manuscript.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

Education should stress the development of each student to his maximum. Theoretically, a teacher can teach one student at a time more effectively than thirty students at one time. One reason is that the teacher can stress the parts in which the student is weak, and go over quickly the material which the student understands. In the present educational system it is impossible for each student to have a private tutor. In fact, with the increase in college age population in the United States, it seems inevitable that the load of each college teacher must be increased.

One attempt to expedite the teaching-learning process in crowded college classes is to divide students into homogeneous groups within a specific course. Students with similar strengths and weaknesses can more nearly be treated as one student. This division is particularly urgent in subjects in which some students have had previous instruction while others have had little or none.

A paper-and-pencil pretest is used by the Department of Clothing, Textiles, and Merchandising in the College of Home Economics at Oklahoma State University as an aid in sectioning students who are enrolled in the basic clothing course. The pretest is unsatisfactory as it now exists. It may be erroneous to assume that a student who makes a high score on a paper-and-pencil test in clothing can also sew well.

It has been suggested that a performance test accompany the paper-and-pencil test for a more reliable appraisal of the individuals' understanding in this area.

Performance tests developed previously for use in clothing classes have been expensive and time consuming; therefore, it has not seemed practical to use them for placement purposes. If it could be established statistically that students who are in the upper ranks on a paper-and-pencil clothing pretest will also be in the upper ranks on a performance clothing pretest, the idea that a performance test is necessary for placement purposes could be rejected.

The problem in this study was the development of a clothing construction performance pretest. The problem was divided into two sub-problems:

1. To construct a performance pretest suitable for use in a college freshman clothing construction course.
2. To study the relationship, if any, that exists between ranks of individuals on the performance test and on the written test.

Hypothesis

The major hypothesis underlying the study was that a pretest can be developed which will differentiate between those students with a high degree of clothing construction skill and those students with a low degree of clothing construction skill.

A sub-hypothesis relevant to the study was that there will be no significant relationship between the scores obtained on the performance pretest and the scores on the paper-and-pencil pretest to indicate that success on one test can accurately predict success on the other.

Assumptions

These assumptions were basic to the study:

1. Education is a process of changing the behavior patterns of human beings.
2. Evaluation is a process of ascertaining the extent to which behavior is changing.
3. Evaluation includes the use of a variety of techniques and devices for obtaining evidence of changing behavior.

Definition of Terms

Performance and practical test were used interchangeably in the study to refer to a series of processes actually performed by students involving manipulation of equipment and materials used in clothing construction.

Paper-and-pencil tests are usually objective in nature, and are used primarily to measure knowledge of facts.

Pretests are evaluative devices used at the beginning of a course to ascertain the extent of knowledge of the subject prior to specific instruction.

Correlation may be defined as the degree of relationship between two variables. It refers to the extent that decisions are enhanced by taking into consideration an associated variable. The variables considered in this study are scores on the performance test and scores on the paper-and-pencil test.

Discriminating items are those which sufficiently more high scoring students than low scoring students answer correctly.

Nondiscriminating items are those which approximately equal

proportions of high scoring students and of low scoring students answer correctly.

Item difficulty refers to the per cent of the students attempting the question who answer it correctly.

Limitations of the Study

The study was limited to the development of a performance test to be used in combination with a paper-and-pencil test for placement in a basic clothing course at Oklahoma State University, and to a correlation of scores made on this test with those made on the paper-and-pencil pretest presently being used at Oklahoma State University.

Participants in the study were freshmen students at Oklahoma State University who were enrolled in four sections of Home Economics 114 during the spring semester, 1963.

The test is not intended to predict a student's future success in a clothing construction course. Scores made on the test should aid in grouping students with similar needs.

Organization of the Study

The study was organized into five chapters.

Chapter I has presented a statement of the problem, the hypotheses, assumptions, definition of terms, and limitations of the study.

Chapter II gives a review of the literature relevant to the study.

Chapter III describes the development of the performance test and the treatment of the data.

Chapter IV presents the summary, conclusions, and recommendations for further study.

CHAPTER II

REVIEW OF LITERATURE

Many educators recognize evaluation as an integral part of the teaching-learning process. "Effective evaluation does not materialize through caprice, intuition or accident." (11). College instructors are endeavoring to develop a more effective process of evaluation in clothing laboratories.

A student brings knowledge and skills gained from previous experiences with him into the classroom. The amount and type of previous experience varies with the individual and his environment. The experience in conjunction with native ability produces vast differences in students. Arny (2) points out the seriousness of differences in the following statements:

> The need for an evaluation of what students know and can do at various stages of their education becomes clearer each year. All teachers recognize that there is considerable range in ability and achievement among students; but few have any conception of how great these differences are.

Arny (2) studied hundreds of freshmen in twenty typical schools in one state and found that some students know more as a beginning freshman than some seniors who have had all the home economics offered. This accentuates the need for discovering where a student is when he begins his first course, so that a challenging program may be offered, and he may develop to the greatest extent possible.

Development and Use of Performance Tests

In order to get a clearer picture of the knowledge and experience of the student, many devices have been tried. Emphasis has shifted from objective-type tests to a greater variety of instruments for securing evidence of changes taking place as a result of instruction in home economics.

Brown (9) states:

> Research which relates to evaluation can be divided into two major types: (a) the determination of the effectiveness of the programs in home economics; and (b) the development of techniques and/or devices which will provide reliable evidence of the accomplishments of the program. The use of objective paper-and-pencil tests as the only means of determining the effectiveness of the home economics program disregards many of the most important values in home economics.

Palmer (17) emphasizes the importance of using a variety of methods.

> If we are to find out whether we have attained a certain objective, we must apply tests directly to that objective. We must not depend upon testing attainment in areas in which measurement is easy, such as that of factual information, and then assume that, if the measurable objective has been achieved, it is reasonable to suppose that the immeasurable or intangible ones have also been achieved.

Many educators recognize the need for obtaining data in a wide variety of areas. At the same time, however, there is recognition of the limitations imposed because of the lack of suitable measuring devices.

Although objectivity in evaluation is valued, complete objectivity is neither always possible nor always desirable. Palmer (17) verifies this statement.

> Although evaluation should be as objective as possible, it should not be restricted to objective instruments. If this is done we are in danger of allowing our goals to be controlled by the availability of instruments. Since even

> the results secured by such instruments must have their significance and implications determined by human judgement, it is sensible to recognize that complete objectivity in evaluation is never possible.

It is impossible for a performance test developed for use in a clothing laboratory to be completely objective.

It becomes increasingly evident that teachers of clothing need a wide variety of evaluation devices for use in identifying competencies of students. Possession of needed factual information does not insure presence of well-learned habit patterns. "Without a minimum standard of motor coordination achieved through practice, the factual knowledge of how to do the job may be practically useless." (5). A student who is superior in theory may not be superior in laboratory work and vice versa. Performance tests have been developed in an attempt to determine the quality of laboratory work.

Arny (3) lists these four uses of performance tests in home economics:

(1) to diagnose difficulties
(2) to predict achievement
(3) to motivate learning
(4) to measure achievement

Whenever a performance test is used for one of these purposes it should meet certain criteria. Arny (3) lists characteristics of a good performance test.

(1) Fundamental skills should be checked.
(2) Relatively inexpensive materials should be used.
(3) The task should be completed within the usual class period or less.
(4) The results should be scorable in objective terms.
(5) The testing situation should represent one met in real life.

Tasks which will actually cover that which is to be measured must be selected for the performance test. The tasks should not only differentiate between good and poor workers, but should yield several

different degrees of proficiency. Performance tests are often developed, but the cost of administration is prohibitive. A compromise between high reliability and reasonable cost of administration may be necessary. (5).

Measuring Devices Used in Home Economics

A variety of measuring devices has been developed by home economists in an attempt to evaluate achievement in clothing classes more effectively. As early as 1919, a scale for the measurement of hand sewing was published by the Bureau of Publications of Teachers College. (16).

The scale, developed by Katharine Murdock, consisted of fifteen half-tone reproductions of samples of hand sewing which varied in general merit from extremely poor sewing to that of very excellent quality. It was patterned after the Thorndike Handwriting Scale and was used to furnish a measuring rod for general merit of hand sewing of students in elementary and high school. (16).

Brown (8) studied the validity of the scale, and concluded that the instrument was of value in the placement of transfer students or those who have had a background of experience in clothing construction. Likewise, it was useful in identifying individual differences of students.

The original Murdock scale did not measure each type of stitch separately. In some situations, measuring each type of stitch was necessary; therefore, a new scale, called the Murdock Analytic Sewing Scale, was developed to supplement the original scale. (16). The Murdock Analytic Sewing Scale measured five different types of stitches separately. Both teachers and students could use this evaluative device.

The Murdock Sewing Scales measured only hand sewing. With the use of the sewing machine another device was needed for measuring machine stitching. The Knapp and Williams Scale was developed for measuring ability in machine sewing. These five factors were decided on as the elements which contribute to good machine sewing: spacing, constructive elements, tension, length of stitch, and neatness. The scale consisted of photographic reproductions of samples showing varying degrees of excellence in each of these factors. (19).

Both the Murdock Sewing Scales and the Knapp and Williams Scale were performance tests. The Trilling and Bowman Test was developed during the same era. It was an objective-type test designed to test the acquisition of information and the ability to reason in situations involving use of material presented in a textile and clothing course. (19).

In using these tests it was disclosed that teachers differ from each other and the same teacher will differ on given days in rating students' work. There was a need for an accurate and objective means of measuring abilities. Standardized tests in textiles and clothing seemed to be the answer. Trilling and Williams (19) believed it was necessary to analyze subject matter in terms of mental processes involved before standardized tests in textiles and clothing could be developed. The following mental processes were suggested as being an integral part of clothing classes.

(1) acquisition of skill
(2) exercize of problem-solving abilities
(3) acquisition of information
(4) development of appreciation

In 1921, Trilling and Hess (18) advanced the theory that there could be no effective standardized tests in home economics because

home economics "is still vague in its purposes, and really has no well-defined objectives." Therefore, they proposed that informal tests be worked out by the teacher to measure some definite ability or product. The teacher must clearly define reasons for giving the test, and the outcomes desired. Then intelligent selection of the material will measure these outcomes. (18).

The idea of standardized achievement tests in home economics was revived in 1947. The American Home Economics Association worked with the Cooperative Test Service of the American Council on Education for developing tests to measure student achievement in home economics. Both a Foods and Nutrition Test, and a Textiles and Clothing Test were developed at that time. (2).

In 1949, clothing and textiles tests were sent to all institutions who wished to participate in a testing program for college students. Arny (4) postulates that students should be exempted from home economics courses containing only material with which they are familiar. When that can be done, the home economics program may be strengthened in these ways:

(1) The caliber of students taking more than a minimum amount of home economics in high school is likely to improve so that there will no longer be negative correlation between scholastic ability and number of units of home economics on students' high school records.

(2) Widespread criticism of many elementary courses in college to the effect that they repeat what is being taught in a good school at the senior high school level may no longer be justified.

(3) Really able students, who represent the group from which future leaders in the field may be drawn, can complete undergraduate requirements in less than four years and some may be well on the way to advanced degrees by the time they would normally have a bachelor's degree.

A variety of efficacious evaluation devices is a precedent to a program such as this.

Clothing Placement Tests at Various Institutions

In this space age when mathematics and science are receiving increased emphasis, the value of instruction in home economics is being criticized, especially that instruction at the college level. Much criticism has been directed specifically toward clothing construction. Educators must critically evaluate the curriculum in the area of clothing. Werden (21) summarizes the problem thus:

> In the light of current economic, sociological and psychological trends we must give clothing construction its rightful place along with the other phases of textiles and clothing that are essential to a good, sound academic program.

The obvious necessity and urgency for the development of more effective evaluation devices in the area of clothing and textiles has been felt by many educators. In the area of pretesting alone much research has been pursued in recent years.

Bray (7) developed a paper-and-pencil pretest to be used for the placement of college students in clothing courses at Macdonald Institute in Guelph, Ontario, Canada. A test which had been developed during a home economics workshop at the University of Minnesota in August, 1947, was given to students entering the clothing course at Macdonald Institute in September, 1947. It was also used as a retest at the end of the course. This objective test contained one hundred fifty items, most of which were planned to test the ability to apply knowledge in specific situations.

A statistical analysis of the test was made and the test was re-

vised on the basis of these data. The revised test was given to eighty-five students in September, 1948. An analysis was made of the revised test and the results were compared with results from the previous test. Bray (7) concluded that the test was a valid instrument for segregating those students with ability and previous training in clothing construction. The test was more discriminating when used as a pretest than when used as a retest.

It was recommended in this study that better results could be achieved if some other devices were used in addition to the paper-and-pencil test. Bray (7) believed that the typical practical test was too time consuming and too expensive to be useful for the purpose of placing students, but suggested that "two or three of the more difficult construction processes might be combined to give the teacher some idea of the students' techniques."

It is generally agreed that previous clothing construction experiences influence the performance of a college freshman in a clothing construction class. In 1949, Wright (23) conducted a study at Purdue University to determine the effect of students' previous clothing construction work on achievement in a freshman clothing construction laboratory at college level. One hundred seventy-nine students enrolled in a freshman clothing laboratory participated in the study.

Achievement was based on three phases:

(1) Knowledge as measured by an objective pretest-retest
(2) Skills as measured by actual construction processes
(3) Attitudes as measured by a questionnaire and an attitude scale. (23).

Data were obtained by means of a fifty minute objective test, given as a pretest and as a retest, a fifty minute practical test, and a questionnaire. The practical test included one problem in each area of difficulty:

(1) Simple - dart, button, hook or snap, machine gathers
(2) Intermediate - binding, facing, particular seam
(3) Difficult - bound or worked buttonhole, set-in-sleeve

Wright (23) concluded that previous experience in clothing construction is a factor affecting achievement in a university clothing course, but that the amount, rather than the type of previous experience indicated greater interest and achievement on the part of the students.

Sectioning into homogeneous groupings on the basis of previous clothing construction work has been the custom in some universities. Recent studies indicate that a statement of previous experience alone is unreliable as compared with a measurement of achievement upon completion of the freshman course.

As a part of a continuing curriculum study at Purdue University, the introductory course in clothing and textiles was selected for an experimental study in 1951. The reason for the selection of this course accents the importance of the freshman clothing course.

> The beginning course in any area of study is of great importance. It may serve to introduce a student to a new area of knowledge, but it may also be required to serve as an agent to bring together earlier learnings into a basic pattern. The success of such a course lies not only in the academic achievement of the students, but also in the interests and attitudes that are created toward the new field. (13).

A checklist was used to determine the previous experiences and training of the student. It was found that students enter the classes "with widely differing backgrounds of training and experience, varying from no formal training to a combination of high school courses, 4-H Club training, and home training." (13).

Henkel and Seronsy (13) concluded that achievement as measured by a reliable test is more basic in predicting course grades than is a record of previous learning experiences. "Previous training as

measured by a checklist bears no relation to achievement as measured by course grades."

In 1956, Collins (10) undertook to develop a pretest for use in Home Economics 127 at Southern Illinois University, to make it possible to place students with similar needs in section together.

A questionnaire was sent to seventy-one colleges and universities in an attempt to determine number and type of pretests currently being used for placement or exemption in clothing courses. Results from the thirty-two institutions in which these pretests were used indicated that thirteen pretests served the purposes for which the tests were intended, nine pretests satisfied the purposes in part, and two pretests did not fulfill the purposes. Only seven practical examinations were used as pretests, generally in combination with an objective examination. (10)

A test was developed using pretests from other institutions as a guide. The paper-and-pencil test included sixty objective items, and the practical test was composed of seven problems: a plain seam with pinked and edge-stitched finishes, a fell seam, a bias binding, a bias facing, a hem such as would be used on a dress, and a skirt zipper. A large brown envelope containing supplies for the problems was given each student and two hours were allowed for the practical test. (10).

It was recommended that at some future time the test which was developed be tested for validity and reliability, and needed alterations made. It was suggested that the test be used as an exemption device for students in options not directly related to the field of clothing construction. (10).

The major emphasis in home economics in the past has been upon the construction of pretests. Little attempt has been made to establish validity and reliability, or to correlate the scores obtained from one

variable with the scores obtained from other selected variables. The work has centered primarily around developing pretests, largely paper-and-pencil tests, for use in a particular course in a particular institution.

In 1959, Hoskins (14) developed a pretest for use in all the colleges and universities in New Mexico offering home economics in the curriculum. Questionnaires were sent to home economics departments in New Mexico colleges and universities. Textbooks used in the first courses were examined. Generalizations, formulated from this data, were used as a guide to areas to be tested and to the amount of emphasis to be given each area.

A paper-and-pencil test was developed containing sixty multiple choice and eighty true-false items. The test was given to high school girls who had previously had two years of home economics or four years of 4-H. It was felt that their clothing experiences would be comparable to those of college freshman. Based on scores obtained from these students, the coefficient of correlation was .99 and the coefficient of reliability was .717. (14).

Hoskins (14) suggested that the pretest be used for both placement and exemption. She states:

> While the author feels that the understanding of the principles underlying the skills involved in clothing construction is essential, the level of the skills needs to be evaluated as well if the pretest is to be used successfully as an exemption instrument. It is recommended that a practical test accompany the written pretest. In the practical test the student would do independent work for a specified amount of time and would be graded on ability to follow instructions and on the speed, accuracy, and quality of the work completed.
>
> In addition to the use of the pretest for exemption, results from the practical and written pretest could be an aid in the amount and kind of emphasis in course work, could help increase motivation in the students, and be an aid in placement of transfer students.

Development of Clothing Pretests at Oklahoma State University

The first recorded work toward developing a pretest for use in the basic clothing course at Oklahoma State University was done in 1959, by Walsh. (20). It was verified by use of a questionnaire that wide variations exist in the number of previous clothing experiences of the freshman students. A careful study of the development of tests was made with particular attention being given to pretests. An objective test was then developed to be used as a placement device in the basic clothing course at Oklahoma State University.

As the test items were constructed, each item was analyzed and content validity of the items was established by:

(1) direct comparison with objectives of instruction
(2) comparison with expert opinion
(3) comparison with text books and other source material (20).

Walsh only developed the test; she did not administer it to prelimenary tryout groups. She states:

> The writer does not submit the pretest as a flawless instrument. There is much room for improvement. The most effective way to insure having a better test is to use the one now developed, study the results and offer criticisms and suggestions for improvements and then continue to use their successors.

The Walsh test was used over a period of years and revisions were made. The pretest presently being used at Oklahoma State University is a revision of the Walsh test.

In 1961, Witt (22) revised the written placement test at Oklahoma State University and developed a station-to-station test to evaluate the students' manipulative and judgmental skills pertaining to clothing construction, selection, and care.

Data were gathered by administering the tests to freshmen clothing students at Mississippi State College for Women and Oklahoma State University. A coefficient of reliability of .74 was computed for the written test and .58 for the station-to-station test. (22).

Witt's study points to a need for evaluating different types of clothing construction skills in order that students may be placed more satisfactorily in clothing courses. Students who scored high on one problem in the test did not necessarily receive a similar high score on a second problem. Students comments implied that they arrived at answers to some of the written questions by guessing, but this was not possible in performing the manipulative problems.

The recommendation was made that further studies be conducted to improve the evaluation devices developed in the study. (22). An evaluative device must be revised frequently if it is to continue to be effective.

Summary

Home economists have long recognized the need for more effective evaluation devices. Rating scales, dexterity tests, performance tests, and standardized tests have been used at various times for various purposes.

Administrators and teachers in many colleges and universities are concerned with developing better devices for the placement of students in the hope that achievement will be accelerated. Educators in few institutions are completely satisfied with the tests they are currently using.

The following conclusions were delineated from the review of

literature:

1. Students enter college with varying degrees of previous experience and ability in clothing construction.

2. Skills cannot be measured entirely by paper-and-pencil tests.

3. Students' interest and achievement increases when they are placed in homogeneous groups according to ability.

4. There is a need for evaluation and revision of pretests which are currently being used as placement devices.

5. It is generally recommended that a performance test accompany the written test.

6. Performance tests which have previously been developed for use in clothing laboratories are too expensive and too time consuming to be practical for use as placement devices.

CHAPTER III

DEVELOPMENT OF PERFORMANCE TEST

The problem undertaken in this study was to develop a clothing construction performance test. Chapter III deals with the development of the performance test and the analysis of the data.

Pilot Study

The first step in the development of the performance test was to construct problem situations requiring the use of manipulative skills. Witt (22) used three manipulative problems in a station-to-station type test. These three problems were studied, and the problems constructed for the performance test were patterned after them. The three manipulative problems in the Witt station-to-station test were: (1) Buttonhole Placement and Length, (2) Pattern Alteration, and (3) Cutting and Joining a Bias. These three problems were revised using Witt's item analysis as a guide.

Six additional problems were devised: (1) Pattern Placement, Cutting, and Marking, (2) Applying a Snap, (3) Applying a Hook and Eye, (4) Constructing a Plain and a French Seam, (5) Stitching a Dart, and (6) Hemming and Gathering. The specific purposes of the pilot study were: (1) to estimate an appropriate number of problems to be included in the test, (2) to determine which of the problems were

more discriminating, (3) to see whether the station-to-station method of administering the test was satisfactory, and (4) to determine which questions should compose the check sheet used by the teacher in scoring the test.

The problems were designed as a station-to-station test. Nine stations were designated in a clothing laboratory and each station was supplied with instructions, materials, and equipment necessary for completing one problem. Twenty-four students enrolled in one section of Home Economics 114 at Oklahoma State University participated in the pilot study in December, 1962.

The test was so administered that two or three students could begin the test at each of the nine stations, but not more than three students were allowed at one station at the same time. The students were directed to move from station to station and perform the construction processes. There was no specified order of procedure from one station to another nor was a time limit set for completion of the problems. Students were encouraged to work rapidly, however, and to complete as many of the problems as possible in one hour.

Data from the pilot study indicated that four or five manipulative problems were suitable for a one hour test. The average number of problems completed in one hour was four, and only one student attempted as many as six problems.

Many deficiencies in the test were revealed by the data from the pilot study. For example, confusion was created by the constant moving of students between stations. Some problems required a greater amount of time than others; consequently, traffic became congested at

some stations while other stations were left vacant.

Adequate supplies were on hand at the various stations for all students; however, the first students arriving at a station tended to use more than their share of supplies, discarding those on which mistakes were made. This practice was disadvantageous from two standpoints: (1) Providing more supplies meant extra time and expense, and (2) If a student were allowed a number of trial and error processes, his final product would not be truly representative of his work.

The students working at the same station on the same problem tended to be influenced by each other's work; therefore, the probability that each student's work represented his own thinking was decreased.

Approximately one hour was required to assemble the equipment and supplies and to arrange the room for the station-to-station test. A problem in time management was thus created for the administrator of the test, particularly with regard to "make-up" tests.

Performance Test

The test was revised on the basis of the results from the pilot study. The five problems which portrayed the most discriminating power were selected for inclusion in the test. The wording on some problems was changed to clarify the meaning. Questions on the score sheet were revised where it was deemed expedient. The five problems selected were: (1) Cutting and Joining a Bias, (2) Placement and Making of Buttonholes, (3) Pattern Placement, Cutting, and Marking, (4) Construction of Plain and French Seam, and (5) Hemming and Gathering. A scoring device in the form of a "yes-no" check sheet was developed, listing specific points to be observed. One point was given for each "yes" answer.

The method of administering the test was changed because of the previously stated disadvantages of the station-to-station test. The supplies needed for completing the problems, together with the instructions for each problem, were placed in a large manila envelope. Small equipment such as tape measures, rulers, thimbles, scissors and tracing wheels were placed on a table for students to use when needed. Irons and sewing machines were provided for the use of all students. The time limit for the test was one hour and fifteen minutes, and at the end of that time, all problems were to be placed in the envelopes and handed to the instructor.

The revised performance test was given to seventy-seven students enrolled in four sections of Home Economics 114 at Oklahoma State University during the spring semester, 1963. The results indicated improvement over the test used in the pilot study. Elimination of predetermined areas for specific problems reduced the idle moments spent in waiting for equipment. The tendency toward conformity was reduced because students sharing a table were working on different problems. Enclosure of supplies in an envelope reduced the time needed for setting up the room, thus making the test easier to administer.

One of the criticisms of clothing construction performance tests has been the expense involved in using the test. The cost of the supplies used in this test was approximately ten cents per student.

Item Analysis

Although teacher-made tests represent the expenditure of much time and effort, they no doubt contain numerous unsuspected flaws. One measure of the worth of a test can be estimated by statistical treatment of the data. One purpose of an item analysis is to determine the difficulty

and the discriminating value of each item in a test so that its strengths and flaws can be detected.

An item analysis of the test was performed using data obtained from the seventy-seven participants in the study. The writer used the formula stated by Ahmann and Glock (1) for determining difficulty and discriminating power. Table I shows the difficulty level of each of the forty-six items on the test. Forty-one per cent of the test items have a difficulty level between forty and seventy per cent. Twenty-two per cent of the items have a difficulty level above seventy per cent. Thirty-seven per cent of the items have a difficulty level below forty per cent. Frequently, the recommendation is made that achievement tests include only those test items with mid-range levels of difficulty, between forty and seventy per cent. (1).

TABLE I

DIFFICULTY LEVEL OF ITEMS USING RESPONSES OF SEVENTY-SEVEN STUDENTS

Item No.	Difficulty (Per cent)
Problem 1	
1	66
2	22
3	87
4	31
5	23
6	34
7	40
8	49
9	8
Problem 2	
1	53
2	32
3	22
4	39
5	30
6	31
Problem 3	
1	21
2	60
3	22
4	77
5	66
6	64
7	65
8	61
9	62
10	51
Problem 4	
1	73
2	74
3	86
4	61
5	52
6	23
7	32
8	29
9	16
10	42

TABLE I (continued)

Item No.	Difficulty (per cent)
Problem 5	
1	52
2	73
3	91
4	53
5	60
6	82
7	84
8	62
9	38
10	40
11	94

Item difficulty sometimes influences item discriminating power; therefore, it was necessary to determine the discriminating power of the test items. Ahmann and Glock (1) consider any discriminating values above .40 to be good, any values between .40 and .20 to be satisfactory, and any values between .20 and 0 to be poor. Negative values identify items that differentiate among students in the wrong direction.

The discriminating power of the forty-six items on the test is shown in Table II. Fifty-four per cent of the items have good discriminating power; thirty-five per cent have satisfactory discriminating power, and eleven per cent have poor discriminating power. No items on the test had a negative discriminating value.

Item analysis data cannot be completely accurate and completely meaningful. Limitations are introduced by the use of a small sample of students, and in the special circumstances that surround the administration of every test. Generally, the degree to which these factors influence item analysis data is slight; however, in a given instance it could be prominent.

The difficulty values and discrimination values for each of the forty-six test items are merely representations of the level of difficulty and discriminating power of these items as they appeared under the specified circumstances. The values would likely change if the test or the environment were changed. (1).

TABLE II

DISCRIMINATING POWER OF ITEMS USING RESPONSES OF UPPER AND LOWER THIRTY-THREE PERCENT OF SEVENTY-SEVEN STUDENTS

Item No.	Discrimination (per cent)	Rating
Problem 1		
1	58	Good
2	27	Satisfactory
3	27	Satisfactory
4	42	Good
5	50	Good
6	54	Good
7	73	Good
8	58	Good
9	12	Poor
Problem 2		
1	50	Good
2	69	Good
3	42	Good
4	69	Good
5	58	Good
6	58	Good
Problem 3		
1	19	Poor
2	27	Satisfactory
3	31	Satisfactory
4	50	Good
5	46	Good
6	35	Satisfactory
7	42	Good
8	12	Poor
9	62	Good
10	38	Satisfactory
Problem 4		
1	27	Satisfactory
2	35	Satisfactory
3	31	Satisfactory
4	46	Good
5	46	Good
6	38	Satisfactory
7	69	Good
8	54	Good
9	31	Satisfactory
10	73	Good

TABLE II (continued)

Item No.	Discrimination (per cent)	Rating
Problem 5		
1	35	Satisfactory
2	62	Good
3	19	Poor
4	58	Good
5	62	Good
6	38	Satisfactory
7	23	Satisfactory
8	58	Good
9	35	Satisfactory
10	38	Satisfactory
11	4	Poor

Correlation of Scores on Performance Test With Scores on Paper-and-Pencil Test

A decision concerning the degree of skill attained in clothing construction might be enhanced by an actual demonstration of that skill. A reliable performance test could assist in measuring the degree of skill; however, most performarce tests are expensive and time consuming to use. Many clothing teachers believe that developing a paper-and-pencil test which could predict skill in clothing construction would be desirable.

A correlation was calculated in order to determine the relationship between scores on a paper-and-pencil test and scores on a performance test. The seventy-seven students participating in the study had been assigned to clothing classes on the basis of scores made on a paper-and-pencil test prior to enrollment in the course. These paper-and-pencil scores were available for use in the correlation.

Since scores are only a rank, and insignificant in themselves, Spearman's formula for rank correlation was used. (6). A coefficient of rank correlation of .70 was calculated, using the scores on the paper-and-pencil test as the independent variable and the scores on the performance test as the dependent variable. A correlation coefficient of .70 indicates that forty-nine per cent of the variability on the performance test can be explained by, or is associated with, the paper-and-pencil test.

Tabulation of data (Table III) reveals that fifty-three per cent of the students scored higher on the paper-and-pencil test than on the performance test, while forty-four per cent scored higher on the performance, than on the paper-and-pencil test. Three per cent scored the same

on both tests. Using the paper-and-pencil test as the independent variable, the greatest deviation upward was +43.5 points, and the greatest deviation downward was -34.5 points.

The preceding statistics and the correlation coefficient of .70 indicated that scores on the two tests were related to some degree; however, a high score on one test did not insure a high score on the other test.

TABLE III

SCORES AND RANKS OF STUDENTS ON PAPER-AND-PENCIL TEST AND PERFORMANCE TEST

Score on Written Test	Score on Performance Test	Rank on Written Test	Rank on Performance Test	Difference in Ranks
79	28	1	27.5	-26.5
78	38	2	3.0	- 1.0
77	32	3	15.5	-12.5
76	28	4	27.5	-23.5
74	36	5	6.5	- 1.5
73	28	6	27.5	-21.5
71	31	8	19.0	-11.0
71	30	8	21.5	-13.5
71	32	8	15.5	- 7.5
70	21	12	46.5	-34.5
70	37	12	5.0	+ 7.0
70	34	12	11.0	+ 1.0
70	30	12	21.5	- 9.5
70	38	12	3.0	+ 9.0
68	23	16	40.0	-24.0
68	32	16	15.5	+ .5
68	34	16	11.0	+ 5.0
67	34	19	11.0	+ 8.0
67	28	19	27.5	- 8.5
67	31	19	19.0	0.0
66	36	22	6.5	+15.5
66	27	22	31.5	- 9.5
66	35	22	8.5	+13.5
65	25	26	33.0	- 7.0
65	23	26	40.0	-14.0
65	18	26	52.0	-26.0
65	24	26	36.0	-10.0
65	29	26	23.5	+ 2.5
64	22	30.5	43.5	-13.0
64	19	30.5	49.0	-18.5
64	35	30.5	8.5	+22.0
64	17	30.5	54.5	-24.0
63	39	33.5	1.0	+32.5
63	32	33.5	15.5	+18.0
62	33	35.0	13.0	+22.0
61	22	36.0	43.5	- 7.5
60	38	37.0	3.0	+34.0

TABLE III (continued)

Score on Written Test	Score on Performance Test	Rank on Written Test	Rank on Performance Test	Difference in Ranks
59	24	39.0	36.0	+ 3.0
59	18	39.0	52.0	-13.0
59	21	39.0	46.5	- 7.5
58	14	43.0	63.0	-20.0
58	29	43.0	23.5	+19.5
58	14	43.0	63.0	-20.0
58	18	43.0	52.0	- 9.0
58	15	43.0	58.5	-15.5
57	16	47.0	56.0	- 9.0
57	24	47.0	36.0	+11.0
57	24	47.0	36.0	+11.0
56	15	49.5	58.5	- 9.0
56	31	49.5	19.0	+30.5
54	15	52.0	58.5	- 6.5
54	12	52.0	71.0	-19.0
54	14	52.0	63.0	-11.0
53	27	54.0	31.5	+22.5
51	10	55.0	74.5	-19.5
50	14	57.0	63.0	- 6.0
50	22	57.0	43.5	+13.5
50	11	57.0	73.0	-16.0
49	12	60.5	71.0	-10.5
49	14	60.5	63.0	- 2.5
49	28	60.5	27.5	+33.0
49	12	60.5	71.0	-10.5
48	13	63.0	67.5	- 4.5
47	19	65.5	49.0	+16.5
47	22	65.5	43.5	+22.0
47	17	65.5	54.5	+11.0
47	7	65.5	76.0	-10.5
46	23	68.0	40.0	+28.0
45	13	69.5	67.5	+ 2.0
45	15	69.5	58.5	+11.0
44	28	71.0	27.5	+43.5
43	13	72.0	67.5	+ 4.5
42	19	73.0	49.0	+24.0
41	13	74.0	67.5	+ 6.5
40	24	75.0	36.0	+39.0
36	10	76.0	74.5	+ 1.5
22	4	77.0	77.0	0.0

CHAPTER IV

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

The problem in this study was the development of a clothing construction performance pretest. The problem was divided into two subproblems: (1) to construct a performance pretest suitable for use in a college freshman clothing construction course, and (2) to study the relationship, if any, that existed between ranks of individuals on the performance test and on the written test.

Hypotheses were: (1) A pretest can be developed which will differentiate between those students with a high degree of clothing construction skill and those students with a low degree of clothing construction skill. (2) There will be no significant relationship between the scores obtained on the performance pretest and the scores on the paper-and-pencil pretest to indicate that success on one test can accurately predict success on the other.

It was assumed that: (1) Education is a process of changing the behavior patterns of human beings. (2) Evaluation is a process of ascertaining the extent to which behavior is changing. (3) Evaluation includes the use of a variety of techniques and devices for obtaining evidence of changing behavior.

A review was made of the related literature and the need for the study was substantiated. A study of the pretests developed previously at Oklahoma State University led to concentration on the performance

test, known as a station-to-station test, which had been developed in 1961. The station-to-station test was revised and additional problems were developed. The test was given as a pilot study to one section of Home Economics 114, in December, 1962.

Changes were made in the structure of the station-to-station test on the basis of the results from the pilot study. The major revision was changing from a station-to-station type test to an individual set of problems.

The revised performance test was given to freshmen students enrolled in a basic clothing course on the second day that the classes met during the spring semester, 1963. The results of the test were tabulated, analyzed, and appropriate statistical treatment was used. An item analysis was made of the performance test, and the scores made on the performance test were correlated with the scores made on the paper-and-pencil test. A correlation coefficient of .70 was calculated.

Conclusions were drawn, and recommendations for further study are given.

Conclusions

The following conclusions were delineated from the data:

1. The wide range of scores on both the paper-and-pencil test and the performance test indicated that students enter college with varying degrees of knowledge and skill in clothing construction.
2. The test has merit for aiding in the placement of students in a basic college clothing course.

The hypotheses underlying the study were sustained. The item analysis verified the effectiveness of the performance test in differentiating between students with a high degree and those with a low degree of skill in clothing construction. The correlation coefficient of .70 indicated that although there was some relation between the scores on the tests, it would be erroneous to assume that success on one test could accurately predict success on the other test.

Recommendations

These recommendations appear justified in light of the analysis of the data:

1. Conduct further studies to improve the performance test.
2. Develop additional problems so that problems can be rotated from time to time, thus eliminating specific student preparation.
3. Reduce the number of problems given at one time so that cost and preparation time may be curtailed.
4. Use additional tests to determine the reliability and validity of the performance test.
5. Conduct item analyses using a larger number of scores, and using scores obtained over a longer period of time.
6. Calculate correlation coefficients using other selected variables, such as course grades or grades on completed garments.
7. Conduct studies concerning the amount of time required to complete each problem and the amount

of credit designated to that problem. More credit might be given for those problems requiring a greater amount of time.

8. Conduct further studies to determine the feasibility of using this and other placement tests as exemption devices in clothing classes at Oklahoma State University.

A SELECTED BIBLIOGRAPHY

1. Ahmann, J. Stanley, and Marvin D. Glock. Evaluating Pupil Growth. Boston: Allyn and Bacon, Inc., 1959.

2. Arny, Clara Brown. "Evaluating Our Teaching," Journal of Home Economics, XL (January, 1948), 7-8.

3. Arny, Clara Brown. Evaluation in Home Economics. New York: Appleton-Century-Crofts, Inc., 1953.

4. Arny, Clara Brown. "The 1949 Testing Program for College Students," Journal of Home Economics, XLI (January, 1949), 15-16.

5. Bean, Kenneth L. Construction of Educational and Personnel Tests. New York: McGraw-Hill Book Company, Inc., 1953.

6. Blair, Morris Myers. Elementary Statistics. New York: Henry Holt and Company, 1952.

7. Bray, Edyth Louise. "The Development and Use of a Pencil and Paper Test for Determining Placement of College Students in Clothing Courses and for Measuring Achievement After Instruction." Unpublished Master's thesis, University of Minnesota, 1949.

8. Brown, Clara M. "Investigations Concerning the Murdock Sewing Scale," Teacher's College Record, XXIII (November, 1922), 459-470.

9. Brown, Sara Ann. "Evaluation of Home Economics," Encyclopedia of Educational Research (3rd. ed.) New York: The Macmillan Company, 1960.

10. Collins, Mildred Hart. "A Pretest for Placement in Beginning Clothing Courses at Southern Illinois University." Unpublished Research Report, Southern Illinois University, August, 1956.

11. Fleck, Henrietta. How to Evaluate Students. Bloomington, Illinois: McKnight and McKnight, 1953.

12. Hall, Olive A. Research Handbook for Home Economics Education. Minneapolis: Burgess Publishing Company, 1962.

13. Henkel, Jean, and Louise Baird Seronsy. "First Course in Clothing and Textiles," Journal of Home Economics, XLIII (March, 1951), 195-197.

14. Hoskins, Mercedes Nelson. "Construction of a Basic Clothing Pretest for Use in the Colleges and Universities of New Mexico." Unpublished Master's thesis, New Mexico State University, May, 1959.

15. Lehman, Ruth T. Appraising the College Program in Home Economics. Washington: American Home Economics Association, 1950.

16. Murdock, Katharine. "A New Analytic Sewing Scale," Teacher's College Record, XXIII (November, 1922), 453-458.

17. Palmer, R. Roderick. "Evaluation: A Repraisal," Educational Research Bulletin, XXXVI (April 10, 1957), 137-143.

18. Trilling, Mabel B., and Adah Hess. "Informal Tests in Teaching Textiles and Clothing," Journal of Home Economics, XIII (October, 1921), 483-489.

19. Trilling, Mabel B., and Florence Williams. "Standardized Tests in Textiles and Clothing," Journal of Home Economics, XII (November, 1920), 486-491.

20. Walsh, Grace M. "The Development of a Pencil and Paper Pretest for Placement of College Students in First Course in Clothing, Textiles, and Merchandising at Oklahoma State University." Unpublished Report, Oklahoma State University, May, 1959.

21. Werden, Jane. "The Place of Clothing Construction in the College Program," Journal of Home Economics, LII (May, 1960), 340-341.

22. Witt, Mildred Rea. "The Revision and Development of Selected Evaluation Devices for Appraising Certain Clothing Competencies of College Freshmen." Unpublished Ed. D. dissertation, Oklahoma State University, 1961.

23. Wright, Janet Smith. "The Effect of Students' Previous Experience on Achievement in a University Course in Clothing Construction." Unpublished Master's thesis, Purdue University, June, 1949.

VITA

Grovalynn Foreman Gould

Candidate for the Degree of

Master of Science

Thesis: A PERFORMANCE PRETEST FOR PLACEMENT OF COLLEGE STUDENTS IN BEGINNING CLOTHING COURSES

Major Field: Clothing, Textiles, and Merchandising

Biographical:

Personal Data: Born December 2, 1934, Altus, Oklahoma, the daughter of Pat and Norene Foreman.

Education: Graduated from Altus High School in 1953; received the degree of Bachelor of Science in Home Economics Education from Oklahoma State University in May, 1957; completed the requirements for the Master of Science degree in May, 1963, at Oklahoma State University.

Professional Experience: Taught vocational home economics, Estancia High School, Estancia, New Mexico, 1959-1962.

Professional Organizations: American Home Economics Association, Kappa Delta Pi, Phi Upsilon Omicron.