

The Woman's College of  
The University of North Carolina  
LIBRARY



CQ  
no-302

COLLEGE COLLECTION

Gift of  
Marjorie Ann Cooper

DEVELOPMENT OF A TEST TO MEASURE THE ABILITY  
OF HOME ECONOMICS TEACHERS TO APPLY  
CERTAIN ACCEPTED EDUCATIONAL CONCEPTS

by

Marjorie Ann Cooper

A Thesis Submitted to  
the faculty of the Graduate School at  
The Woman's College of the University of North Carolina  
in Partial Fulfillment  
of the Requirements for the Degree  
Master of Science in Home Economics

Greensboro

May, 1962

Approved by

*Hildegarde Johnson*  
Hildegarde Johnson, Director

APPROVAL SHEET

This thesis has been approved by the following committee of the Faculty of the Graduate School at the Woman's College of the University of North Carolina, Greensboro, North Carolina.

Hildegard Johnson  
Thesis Director

Oral Examination  
Committee Members

J. H. Good Jr.  
D. Perry

Harold K. Staley

Madeleine B. Street

May 5, 1962  
Date of Examination

## ACKNOWLEDGEMENTS

The researcher wishes to express her indebtedness and appreciation to the countless number of persons who were directly and indirectly involved in each phase of this study, particularly the staff and college students in home economics education at The Woman's College, University of North Carolina; to her committee for their patience and guidance; to her family for their understanding and reassurance in moments of doubt, and especially to Dr. Hildegard Johnson, whose inspiration, tedious efforts and constant encouragement have made this moment of satisfaction possible.

## TABLE OF CONTENTS

CHAPTER	PAGE
I. INTRODUCTION . . . . .	1
Purpose. . . . .	2
Philosophy Guiding the Development of the Test .	2
Relation of Present Study to Larger Project. . .	4
Limitations. . . . .	8
Special Use of Terms in the Present Study. . . .	8
II. REVIEW OF THE LITERATURE . . . . .	10
Progress in the Area of Achievement Testing. . .	11
Measuring Characteristics of a Test. . . . .	14
Item construction. . . . .	14
Item selection . . . . .	17
Reliability. . . . .	20
Research Studies Having Implications for the Present Study. . . . .	23
III. PROCEDURE FOR DEVELOPING THE TEST . . . . .	29
Item Construction. . . . .	29
Determining test content . . . . .	29
Stating questions in subjective form . . . . .	32
Development of the Objective Form of Test Items.	36
Choice of the objective form of test items . .	36
Responses to the trial form of the test. . . .	38

CHAPTER	PAGE
Developing a Scoring Key . . . . .	39
Securing Responses . . . . .	41
Analysis of the Test Data. . . . .	42
IV. ANALYSIS OF THE TEST . . . . .	44
Reliability. . . . .	45
Procedure for Item Analysis. . . . .	46
Methods of Reporting the Item Analysis . . . . .	47
Results of the Item Analysis . . . . .	48
Revision of the Test . . . . .	67
Procedure for revision . . . . .	67
Results of the test revision . . . . .	70
V. SUMMARY. . . . .	80
BIBLIOGRAPHY . . . . .	87
APPENDIX A . . . . .	90
APPENDIX B . . . . .	105

## LIST OF TABLES

TABLE	PAGE
1. Analysis of Responses of Upper and of Lower 27 Per Cent of the Teachers: Situation I . . . .	49
2. Analysis of Responses of Upper and of Lower 27 Per Cent of the Teachers: Situation II. . . .	50
3. Analysis of Responses of Upper and of Lower 27 Per Cent of the Teachers: Situation III . . . .	51
4. Analysis of Responses of Upper and of Lower 27 Per Cent of the Teachers: Situation IV. . . .	52
5. Analysis of Responses of Upper and of Lower 27 Per Cent of the Teachers: Situation V . . . .	53
6. Analysis of Responses of Upper and of Lower 27 Per cent of the Teachers: Situation VI. . . .	54
7. Analysis of Responses of Upper and of Lower 27 Per Cent of the Teachers: Situation VII . . . .	57
8. Analysis of Responses of Upper and of Lower 27 Per Cent of the Teachers: Situation VIII. . . .	58
9. Analysis of Responses of Upper and of Lower 27 Per Cent of the Teachers: Situation IX. . . .	59
10. Analysis of Responses of Upper and of Lower 27 Per Cent of the Teachers: Situation X . . . .	60
11. Analysis of Responses of Upper and of Lower 27 Per Cent of the Teachers: Situation XI. . . .	61
12. Analysis of Responses of Upper and of Lower 27 Per Cent of the Teachers: Situation XII . . . .	62

TABLE	PAGE
13. Analysis of Responses of Upper and of Lower 27 Per Cent of the Teachers: Situation XIII. . .	63
14. Analysis of Responses of Upper and of Lower 27 Per Cent of the Teachers: Situation XIV . . .	64
15. Analysis of Responses of Upper and of Lower 27 Per Cent of the Teachers: Situation XV. . . .	65
16. Number of Items Having r Values Above and Below .15 . . . . .	66
17. Summary of Decisions Concerning Items in Revised Form A. . . . .	72
18. Summary of Decisions Concerning Items in Revised Form B. . . . .	78

## CHAPTER I

### INTRODUCTION

A beginning teacher is confronted daily with new and challenging classroom situations. Both consciously and unconsciously she develops an individual pattern of behavior in dealing with these situations. She has been exposed to a large volume of educational theory and has had limited opportunities to apply this theory in classroom situations. The question may be raised of her awareness of the relationship between this educational theory and the classroom situations which now confront her. Before the pattern of behavior becomes too rigid, could she be guided to increase her ability to apply important educational concepts? The researcher believed that an evaluation device which would help the beginning teacher to see her progress in developing the ability to apply educational concepts in classroom situations would be helpful to her professional growth. No such device known to the researcher was available in the area of home economics education.

Teacher educators could make considerable use of an achievement test of this nature in their programs of instruction. College students enrolled in courses termed

Methods of Teaching Home Economics are beginning to develop their ability to apply educational concepts. They have additional opportunities to develop this ability during their student teaching experiences. At this time they are guided to apply their knowledge and understanding of the educational concepts in real situations. Evaluation of the students' progress in developing this ability is an integral part of a sound program of instruction.

Purpose. The purpose of the present study was to develop an achievement test which would measure a teacher's ability to apply in hypothetical classroom situations ten basic concepts accepted as important by authorities in home economics education. If a teacher understands the implications of the concepts and has some skill in using them, there should be evidence that she possesses ability to apply this knowledge and understanding in dealing with the new problems presented in hypothetical classroom situations on a paper-and-pencil achievement test. The title of the test, "You Are There," suggests the projective approach to be used in eliciting the respondent's honest reactions to the test situations.

Philosophy guiding the development of the test. Both an intellectual grasp of the concepts and an understanding of their implications for teaching are needed before teachers are ready to apply them in their own classroom

situations. A committee of college and university examiners developed a taxonomy of educational objectives. They placed the class of objectives dealing with application one step higher in the hierarchical order than those objectives dealing with comprehension.<sup>1</sup> The class of objectives dealing with comprehension are, in turn, higher than those objectives dealing with recognition or recall of knowledge. A list of the classes of educational objectives placed in hierarchical order by this Committee appears as follows:

Knowledge  
Comprehension  
Application  
Analysis  
Synthesis  
Evaluation 2

Each class of objectives makes use of and is built on the behaviors found in the preceding classes. Thus, in the Handbook edited by Bloom we read,

...problems requiring a knowledge of principles and concepts are correctly answered more frequently than problems requiring both knowledge of the principle and some ability to apply it in new situations.<sup>3</sup>

Definition of educational beliefs in terms of behaviors is an important step in the process of evaluation. A major objective in education is to change behavior in a desirable direction. If the educational concepts are

---

<sup>1</sup>Benjamin S. Bloom, et al., Taxonomy of Educational Objectives (New York: David McKay Company, Inc., 1956), p. 18.

<sup>2</sup>Ibid.

<sup>3</sup>Ibid., p. 19.

effectively taught in college methods courses, the student teacher or beginning teacher should be able to apply them in her own classroom situations. Tyler pointed out that evaluation involves getting evidence about behavior changes in the students, and that more than one appraisal is desirable to identify changes that may be occurring and which indicate whether or not desirable objectives are realized.<sup>4</sup> If the students in home economics courses in methods of teaching have changed their behavior in a desirable direction, they should exhibit some evidence of this achievement through effective classroom teaching as they apply the important educational concepts. An evaluation device which would give an accurate appraisal of this achievement, therefore, would take into account the behavior to be expected if the concepts are being applied effectively.

Relation of present study to larger project. A series of research projects having the ultimate purpose of improving classroom teaching have been conducted recently at The Woman's College, University of North Carolina. Current research work by members of the School of Home Economics, Education Area, at this institution had direct

---

<sup>4</sup>Ralph W. Tyler, Basic Principles of Curriculum and Instruction (Chicago: University of Chicago Press, 1950), pp. 69-78.

influence on the present study. During the school year 1959-1960 a research project was conducted to determine which concepts in home economics education are most widely accepted by leaders in the field.<sup>5</sup> The researchers formulated thirty-seven concepts which were potentially important in home economics education. These concepts were organized into a questionnaire to which reactions were secured from a sample of home economics teacher educators and supervisors throughout the United States. The ten concepts ranked as most important by these authorities in home economics education were arbitrarily selected for use in the present study. These ten concepts are listed in descending order of importance as follows:

1. Teaching methods should be selected on the basis of their effectiveness in bringing about desired pupil development.
2. Classroom activities should be provided which can carry over into the pupil's personal and home experiences.
3. Pupils should be guided to evaluate their own progress toward personal goals.
4. The homemaking program should deal with attitudes, interests, habits, understandings and appreciations as well as abilities.
5. A class should evaluate from time to time its progress toward group goals.
6. A teacher should continuously evaluate her own progress as a person and as a teacher.

---

<sup>5</sup>Hildegarde Johnson, et al., "Our Educational Beliefs", Journal of Home Economics, LIII (March, 1960) pp. 175-178.

7. Discussions and other learning experiences in homemaking classes should help pupils to better understand their own behavior and the behavior of other people.
8. The homemaking teacher should attempt to guide each pupil in such a way that his progress is satisfactory in relation to his initial skill and ability.
9. Democratic procedures should be used in the homemaking classroom.
10. In all areas of the homemaking program pupils should be stimulated to independent thinking.<sup>6</sup>

Studies begun in the school year 1959-1960 were continued in an attempt to determine how well these concepts were being applied in home economics classrooms and to determine where teacher educators should concentrate their efforts toward improving present programs of instruction. The purpose of the first of these studies, by Clawson, was to appraise the status of the application of three of these concepts in classroom teaching in schools in a small area around Greensboro, North Carolina.<sup>7</sup> This evaluation was done through observation in the classroom, analyzing pupil questionnaires, and tape recording an interview with the teacher. The concepts studied were numbers one, two and seven.

A second study, by Godwin, dealt with discovering

---

<sup>6</sup>Ibid., pp. 176-178.

<sup>7</sup>Barbara Nelle Clawson, "Appraisal of Home Economics Teachers' Ability to Apply Three Basic Concepts of Home Economics Education" (Master's thesis in progress, Consolidated University of North Carolina, Greensboro, 1962).

some strengths and needs which are recognized by home economics teachers in North Carolina in their efforts to apply these concepts.<sup>8</sup> The teachers were asked to react directly to statements of the concepts which had been arranged in questionnaire form and mailed to them. Their responses indicated their recognized needs and strengths in relation to each concept.

A status study of high school home economics programs in North Carolina is being conducted by Curry.<sup>9</sup> The project grew out of a national study in which high school home economics programs were appraised, with the researcher analyzing the North Carolina data. These findings are being evaluated by comparing them with recommendations of home economics educators. The findings will indicate such things about high school programs of home economics in North Carolina as (1) the extent of home economics offerings, (2) the division of time among the various phases of home economics, (3) enrollments at each grade level in vocational and non-vocational programs and

---

<sup>8</sup>Helen Miller Godwin, "Educational Concepts Which Home Economics Teachers in North Carolina Accept and Believe They Apply" (unpublished Master's thesis, Consolidated University of North Carolina, Greensboro, North Carolina, 1961).

<sup>9</sup>Helen Lanning Curry, "Appraisal of Home Economics Programs in Secondary Schools in North Carolina" (Master's thesis in progress, Consolidated University of North Carolina, Greensboro, 1962).

(4) strength of the adult programs, the FHA club programs, and the home experience programs.

Limitations. There were certain limitations which affected the techniques and procedures used in the present study. Due to the parallel studies being made in the same institution, the sample of teachers in North Carolina who were asked to cooperate in the present study was not as random as those used in the other two studies. The purpose of the present study, however, did not demand as representative sample as the parallel studies.

The researcher made an effort to construct test items in such a way that content validity would be satisfactory. The emphasis of the present research was on the reliability of the test and recommendations for changes based on an item analysis of teachers' responses. Several revisions of the test will be recommended for further test analysis.

Special uses of terms in the present study.

**Concept:** One of the ten concepts selected from the 1959-1960 research project conducted at the Woman's College, University of North Carolina.

**Test Situation:** A hypothetical classroom situation followed by five or more statements of teacher or pupil behaviors.

**Test Item:** One of the numbered statements following each test situation.

**Students:** College students majoring in home economics education who were asked to participate in earlier stages of the test development.

**Supervising Teachers:** Home economics teachers who were supervising student teachers and who were asked to react to the trial form of the objective test.

**Teachers:** The sample group of 150 high school home economics teachers in North Carolina who were asked to react to the items in "You Are There".

## CHAPTER II

### REVIEW OF THE LITERATURE

Literature concerning progress in the area of testing for educational achievement and a few studies in which suitable instruments were developed for testing achievement of teachers-in-training were reviewed for the present study. The lack of adequate instruments of this nature became apparent as the researcher reviewed recent literature in the area of home economics education.

A test is effective or ineffective only in relation to its particular purpose. While tests are often used for more than one purpose, such tests do not offer maximum effectiveness for any purpose except that for which the test was constructed. Once the test purpose is decided, the desirable and appropriate test characteristics can be determined and considered in the construction of the test. The general characteristics of a good test are validity, reliability, discriminative power and appropriate difficulty level.

## I. PROGRESS IN THE AREA OF ACHIEVEMENT TESTING

An educational achievement test has been described by Lindquist as a

...device or procedure for assigning numerals (measures) to the individuals in a given group indicative of the various degrees to which an educational objective or set of objectives has been realized by those individuals....A test of any objective may be regarded as consisting in part of a situation or series of situations designed to elicit the desired behavior, or some other behavior which is presumably related to and will, therefore, predict the desired behavior, and in part of a procedure for assigning numerals to the properties of the behavior thus elicited, or to the product of that behavior<sup>1</sup>.

The researcher viewed the purpose of the present test as an attempt to create a series of situations in a paper-and-pencil test which would require a teacher to apply her knowledge and comprehension of basic concepts, therefore providing evidence of her probable success in handling the various problems in a classroom situation.

Lindquist defined two kinds of measurement, direct and indirect. A type of direct measurement of an educational objective would be to observe teachers during actual periods of teaching in order to evaluate their progress toward consistent application of educational concepts. What the teacher does in the classroom is considered criterion behavior. An example of indirect measurement of

---

<sup>1</sup>Everet F. Lindquist, Educational Measurement (Washington: American Council on Education, 1951), p. 142.

an educational objective would be the use of a paper-and-pencil test in which descriptions of hypothetical classroom situations substitute for the criterion behavior. The present study deals with indirect measurement of progress toward a major objective in home economics education.

The most suitable type of achievement test for the present study was described by Lindquist as a "verbalized behavior" type of test.<sup>2</sup> Whenever the criterion behavior series, such as actual classroom teaching situations, is impractical to reproduce or to simulate, the "verbalized behavior" type of test may serve as an acceptable and practicable substitute, and might do so even though the criterion series could be employed. Lindquist mentioned the possibility of educators using this type of test with teacher candidates to collect evidence of behavioral patterns they believe they would follow in handling descriptive situations.<sup>3</sup> Test constructors have only begun to exploit the potentialities of this very promising type of test.

A taxonomy of educational objectives has been developed by a committee of college and university examiners.<sup>4</sup> The taxonomy can serve many purposes in improving educational achievement tests. The taxonomy

---

<sup>2</sup>Ibid., p. 141.

<sup>3</sup>Ibid., p. 150.

<sup>4</sup>Benjamin S. Bloom, et al., Taxonomy of Educational Objectives (New York: David McKay Company, Inc., 1956), pp. 1-207.

consists of the definition of major types of educational objectives and their organization into a hierarchy which is descriptive of the complexity of the objectives and relationships among them.<sup>5</sup>

Educational objectives were classified into three major categories: the cognitive, the affective, and the psychomotor. The committee defined the cognitive domain as including "...those objectives which deal with the recall or recognition of knowledge and the development of intellectual abilities and skills."<sup>6</sup> The affective domain includes objectives which describe emotional rather than purely intellectual behaviors. This behavior denotes changes in interest and attitudes and values, the development of appreciations, and adequate adjustment. The psychomotor domain is described as the manipulative or motor-skill area. Many objectives which are held as important in home economics education would be classified in the affective domain. It has been difficult to describe the behaviors related to these objectives, since the covert feelings are as significant for this domain as are the overt manifestations of behavior. Thus it stands to reason that testing for progress toward achievement of these objectives has been, and still is, a complex problem.

---

<sup>5</sup>Cf. ante, p. 3.

<sup>6</sup>Bloom, op. cit., p. 28.

A handbook for the cognitive domain has been prepared. This handbook includes constructive suggestions for measuring each class of objectives and offers specific examples of the different item types which have been used by examiners. Once the objectives for all domains have been classified, teachers and test authors working on evaluation problems can refer in the handbook to the discussions of the problems of measuring such objectives. The availability of such a handbook for reference should encourage further progress in the area of achievement testing.

There are many standardized achievement tests for measuring progress toward cognitive objectives in the various subject matter areas. Efforts of test builders have been directed toward the development of such tests. Lindquist emphasized progress that has been made in the technical improvements and refinements of those tests measuring achievement of cognitive objectives. He deplored the serious dearth of tests which measure attainment of affective objectives.<sup>7</sup>

## II. MEASURING CHARACTERISTICS OF A TEST

Item construction. The major problem in developing a test is the construction of appropriate items. There has been a tendency to minimize the importance of item

---

<sup>7</sup>Lindquist, op. cit., p. 124.

construction and to rely heavily upon item selection techniques in the development of a good test. This tendency may be due to the difficulty involved in item writing and editing and the relative ease with which an item analysis may be completed. Ingenious statistical techniques are valuable guides in developing a final test form, but they can never replace the scholarship, ingenuity, and painstaking efforts of the item writer.<sup>8</sup> When a "verbalized behavior" type of test is constructed the researcher desires to structure the situation so that the respondent reproduces the thought processes involved in behaving in the real life situation. The best possible substitute for observing the behavior in a real life situation is desired.

Ebel indicated that the problem of item writing, which is essentially a creative process, has been neglected in the literature on testing.<sup>9</sup> He considered item writing an art requiring an uncommon combination of skills for which no set of rules can apply. Ebel expressed the kind of creativity required of test constructors in this way:

...excellence in item writing demands...imagination and ingenuity in the invention of situations that require exactly the desired knowledge and abilities. It demands ability to identify the crucial element in each problem situation so that the corresponding item will be as direct and concise as possible. Most of all, it demands skill and judgment....

---

<sup>8</sup>Ibid., p. 266.

<sup>9</sup>Ibid., p. 188.

The test constructor must have critically supervised practice to develop these abilities and to apply sound principles and suggestions that determine whether good items or mediocre ones will be produced. The specific abilities required for the successful writing of educational achievement tests can be summarized in general terms. They include thorough mastery of subject matter, a well-developed set of educational values which are operative, an understanding of the individuals for whom the test is intended, mastery of verbal communication, and familiarity with the special techniques of item writing. Ebel said that in the past these high standards for item writing have not generally been met.

One of the most difficult problems facing the item writer, Ebel pointed out, is the production and selection of ideas upon which test items may be based. Chance ideas and inspirations, as well as appropriate materials, stimulate this process of producing item ideas. The writer must use discriminating judgment in the selection of item ideas considering their appropriateness, importance, and probably discriminating ability.<sup>10</sup>

Some kinds of items are more difficult to construct than others. It is relatively easy to develop items measuring recall of information and successively more

---

<sup>10</sup>Ibid., p. 191.

difficult to develop items measuring comprehension, the application of knowledge, and the ability to make sound judgments based on a body of knowledge.

The desired level and range of difficulty of test items should receive careful attention and technical consideration before attempts are made to construct items. Vaughn pointed out that "primary factors affecting the difficulty of an item are the nature of its content and the type of behavior it requires of the examinee".<sup>11</sup> Vaughn reminded those who construct tests to consider the influence of the item form and the directions to the examinees on item difficulty. Subjective judgment in estimating item difficulty at the stage of item construction is to be encouraged, Vaughn pointed out.<sup>12</sup>

Item selection. The try-out form of the test usually contains more items than will be needed in the final form. Items are selected on the basis of their difficulty and their discriminating power. Item difficulty will be discussed first followed by techniques used to determine discriminating power.

The most obvious expression of item difficulty and the only one used by test constructors is the per cent of the tryout group that marks the item correctly. Davis

---

<sup>11</sup>Ibid., p. 174.

<sup>12</sup>Ibid., p. 174.

recommended three formulas for computing per cent of correct responses.<sup>13</sup> He stated that difficulty indices are extremely reliable when based on samples as large as 400. For convenience, however, estimates of the per cents are frequently based on data obtained from only the highest and lowest 27 per cent of the sample. Davis cautioned the test constructor that it is important to consider the problem of correcting measures of item difficulty for failure of some examinees to read and respond to some items. He discussed at length techniques to use in correction for guessing and adjustment of the formulas for omissions of items.

Test constructors have recognized for many years the importance of determining the contribution that each item makes to the discriminating power of a test. If the relative value of each item were known, the test constructor could select only the best for inclusion in the final form of the test. A wide variety of ingenious statistical procedures have been devised to provide discrimination indices. These procedures differ greatly in the amount of labor they require. Among these are graphic methods, critical ratios, chi square, point biserial  $r$ , the biserial  $r$ , the tetrachoric  $r$ , and the phi coefficient. Davis said, "The choice among these depends partly on the purpose for which the test and the item analysis data are to be used,

---

<sup>13</sup>Ibid., p. 280.

partly on the convenience with which each statistic serves that purpose, and partly on the ease and economy of computation required by the practical circumstances.<sup>14</sup>

Total scores on the test are used as the criterion measures with which the individual items in the test are correlated. The relationships between the total scores and item scores are referred to as internal-consistency item discrimination indices. Davis said, "The fact that some items prove to have more discriminating ability than others means that for the group tested they are better measures of whatever the whole test actually measures."<sup>15</sup>

For all procedures of item analysis, an upper and lower group must be identified. The upper group is composed of persons who score high on the test, and lower group of those who score low. Kelley demonstrated mathematically that the best proportions to use of the tails of the criterion distribution are 27 per cent in the upper group and 27 per cent in the lower group.<sup>16</sup> Davis said that omission of the middle 46 per cent of the distribution is justified by the fact that

"since the magnitude of a correlation coefficient is determined by extreme cases to a much greater extent than by cases near the middle of the bivariate surface, an estimate of the coefficient may be obtained...by utilizing only the data in the tails

---

<sup>14</sup>Ibid., p. 289.

<sup>15</sup>Ibid., p. 286.

<sup>16</sup>T. L. Kelley, "The Selection of Upper and Lower Groups for the Validation of Test Items," (Journal of Educational Psychology, Vol. 30, 1939), pp. 17-24.

of the two distributions."<sup>17</sup>

The computation of biserial correlation coefficients is laborious and time consuming. Flanagan and Kelley developed a simpler formula for approximating the biserial coefficient. Later Flanagan published an abbreviated table of the biserial correlation coefficients.<sup>18</sup>

Reliability. How accurately and consistently a test measures what it is supposed to measure may be designated as the test reliability. A joint Committee on Test Standards of three professional research organizations described three kinds of measures of reliability: coefficient of internal consistency, coefficient of equivalence, and coefficient of stability.<sup>19</sup>

The coefficient of internal consistency, which is the most commonly used measure of reliability, may be computed by using the split-half method or the analysis of variance method. The split-half method requires the subdivision of a single test into two presumably equivalent groups of items, scoring each separately, and correlating

---

<sup>17</sup>Lindquist, op. cit., p. 297.

<sup>18</sup>Helen M. Walker and Joseph Lev, Statistical Inference (New York: Henry Holt and Company, 1953), pp. 472-475.

<sup>19</sup>Joint Committee of AERA, APA, and NCMUE, "Technical Recommendations for Psychological Tests and Diagnostic Techniques" (Washington: American Psychological Association, Inc., 1954), p. 28.

the resulting two scores. The amount of agreement estimates the reliability of the half test which can be converted to an estimate of the reliability of the total test by application of the Spearman-Brown formula.<sup>20</sup> This means of determining test reliability was considered most suitable for use in the present study. The analysis of variance method may be described as an analysis of the variance among individual items, and the determination of the error variance therefrom. This general approach is similar to that of subdividing a test, but analysis of variance is not applicable to a test which involves the element of speed and which is administered with a single time limit. It is implicitly assumed that the individual has attempted each item because consistency of performance cannot be evaluated unless the subject had an opportunity to read and attempt each item.<sup>21</sup>

The coefficient of equivalence is computed when two equivalent forms of the test are available for administration. Both forms of the test are administered to the same people at the same time. This type of reliability measure lends itself to direct and simple interpretation, attributing the proportion of variance of any test score

---

<sup>20</sup>Lindquist, op. cit., pp. 579-80.

<sup>21</sup>Ibid., pp. 586-87.

distribution to systematic differences between individuals and not to chance errors.<sup>22</sup>

The coefficient of stability is computed when two forms of the test are given with an intervening period of time, or the same test is submitted to the same person after an intervening period of time. An individual should receive the same score every time on a perfectly reliable test unless there has been progress in learning during the test-retest interval.

Certain factors affect test reliability. The first of these is the length of the test. In general, the longer a test is, the more reliable it is. Increasing the length of a test, however, does not increase the reliability if the additional material has lower item reliability.

Another factor affecting test reliability is the condition under which the items are administered. Thorndike states that, "any variation in testing conditions from one test administration to another may be expected to be a source of variance in test performance. This variance must be considered error variance and will have the effect of reducing the reliability of the test."<sup>23</sup> This reduction of reliability due to variations in testing conditions will take place only if a test-retest coefficient of reliability

---

<sup>22</sup>Ibid., p. 561.

<sup>23</sup>Ibid., p. 603.

is computed. Thorndike said that variation in testing conditions may operate to raise split-test coefficients of reliability.

The discussion of reliability cannot be concluded without mentioning the nature of reliability in connection with chance error. There is always some degree of unreliability in any set of measurements if the unit of measurement is fine enough. As reliability is increased, unreliability is decreased, and the set of measurements can be used as a more reliable and sound basis for judgments.<sup>24</sup>

### III. RESEARCH STUDIES HAVING IMPLICATIONS FOR THE PRESENT STUDY

The purpose of a study by Roberts was to develop an evaluation device to discover if students enrolled in Methods of Teaching Home Economics understood basic principles to the extent that they could apply them in written problem situations.<sup>25</sup> The Application of Principles Test presented hypothetical teaching problems, with suggested solutions, and a list of reasons to support the solutions. Roberts supported the purpose of her study by stating that a teacher's judgment and ability to make decisions are tested each time she is confronted with a

---

<sup>24</sup>Ibid., pp. 560-63.

<sup>25</sup>Ada Frances Roberts, "The Application of Educational Principles by Home Economics Education Majors" (Unpublished Master's thesis, University of Georgia, Athens, 1951).

problem situation in teaching. To meet these problems successfully, the teacher needs an understanding of the basic educational principles involved.

Before developing test items Roberts reviewed principles of education listed in home economics literature, and developed problem situations related to these principles. When constructing problem situations, Roberts relied on her own teaching experience and the observations of other teachers as well as suggestions and ideas from home economics professors and student teachers.

A tentative form of the test which included three kinds of items was constructed. The first section consisted of only one item, an example of the item form desired for the final test. This sample item included a description of a problem situation, three tentative solutions to the problem with space allowed for a written alternative solution, and a group of seven reasons with spaces for additional reasons. Section two contained eight problem situations and directions for the respondent to (1) choose five situations and present an appropriate solution for each, and (2) write a reason or reasons to support each solution. Reasons were to include the statement of one or more educational principles. The third section consisted of three problem situations with the following directions: (1) study two of the situations, (2) state one good feature about each situation, (3) state one poor feature about each

situation and (4) give a reason or reasons, including one or more educational principles for each statement.

This tentative test form was given to two home economics education classes at the University of Georgia. Their responses were used to construct test items similar to the one described in the first section of the tentative test. Roberts added seven problem situations and prepared a scoring key which indicated correct solutions and correct reasons supporting the solutions.

This form of the test was submitted to thirty-two home economics education majors. Their comments for improvements were requested after the test had been administered. Several kinds of test data were compiled and summarized.

Roberts found that correct solutions were chosen frequently while appropriate reasons, including the principles, were chosen less frequently. The appropriate principles were also chosen quite frequently. Of the twenty-four principles used, only four were selected correctly by all of the persons at each test period. This suggests that basic principles had not been adequately emphasized to the point of application in previous home economics education courses. The test was suggested for use with undergraduate home economics education majors, including transfers, with student teachers, and with graduate home economics majors before planning their courses of study.

Brown was also concerned with evaluating ability of home economics teachers to apply certain principles, those related to the developmental needs of adolescent girls.<sup>26</sup> The purpose of the study was to develop paper-and-pencil tests which would measure the ability of teachers to recognize and apply principles concerned with human growth and development of adolescents.

The first of two tests developed was called a Recognition of Principles Test. A group of 123 principles which were related to the physiological, social, and integrative needs of adolescents were selected from the literature. Four judges unanimously agreed on the principles selected.

These principles were then used to develop a short-answer type of test, with a three-level choice of response: (a) true for almost every adolescent, (b) true for many adolescents and (c) true for very few adolescents. The jury of judges responded to each item to determine the correct level of response. Agreement was 100 per cent on those items which were retained. The test was scored by assigning a value of one to each right answer.

The second test, called "The Case of Jacqueline Croner," was designed to measure the ability of teachers to

---

<sup>26</sup>Sara Ann Brown, "Techniques for Evaluating the Ability of Teachers to Apply Principles Concerned with the Developmental Needs of Adolescent Girls" (Unpublished Doctor's dissertation, University of Iowa, Ames, 1949).

apply principles (1) when interpreting data descriptive of adolescents who had problems and (2) when selecting educational means for helping pupils with these problems. One comprehensive case study compiled from data in a cumulative record of a high school girl was chosen from a group of 165 case studies. The case study of Jacqueline Croner, followed by questions, was used as an essay form of the test with home economics teachers and college seniors majoring in home economics education. Their responses were used to make statements which explained Jacqueline's behavior and which proposed educational means for helping her with her problems. Their responses also were used to discover misconceptions which teachers and future teachers have about causes of behavior and ways to help adolescents solve their problems. These statements became the 144 items in the short-answer form of Test B, "The Case of Jacqueline Croner." Brown selected a three-point scale for recording group reactions: agree, uncertain, and disagree.

These tests were first administered at West Virginia University to several trial groups of home economics teachers and senior home economics education majors to discover difficulties in administration, vocabulary, time and overall clarity. Eighty-eight West Virginia home economics teachers were then selected to participate in the study. These were teachers whom the writer could visit

and observe, and with whom she could have conferences.

Materials were sent to the teachers in the following sequence:

1. The essay form of Test B, "The Case of Jacqueline Croner".
2. The short-answer form of Test B, "The Case of Jacqueline Croner".
3. Test A, the Recognition of Principles Test.

Reliability coefficients were computed for Tests A and B, using odd and even items in half-tests for correlation between scores. The coefficients of reliability were satisfactory in each test.

Measures of validity and item analyses were computed for each test. Validity was found to be satisfactory for each test and few revisions were necessary for items in either test. Brown concluded that one case study was satisfactory as a sample of a teacher's ability to apply principles in this area.

## CHAPTER III

### PROCEDURE FOR DEVELOPING THE TEST

An attempt was made in this study to develop an achievement test which would measure a home economics teacher's ability to apply certain accepted educational concepts. An evaluation device of this type could serve many purposes for teacher educators and for researchers in home economics.

The researcher investigated a variety of procedures for developing test items. Throughout the development of the test there was concern for selecting procedures which measured the teacher's ability to apply concepts, rather than merely to recall or comprehend the concepts. Comprehension may be assumed when there is evidence that a teacher can apply the concept successfully. It is the teacher's ability to apply concepts which is important if the quality of teaching is to be improved.

#### I. ITEM CONSTRUCTION

Determining test content. The first step in developing the test was to select the educational concepts which would determine the test content. The researcher realized

that the finished test would be more useful in college classes if responses from the students could be completed in one class period. In addition, better cooperation would be secured from teachers who were asked to respond to the test as a basis for securing norms and measures of reliability if the test were not too long. For these reasons it was necessary to limit the area of test content. The ten educational beliefs<sup>1</sup> which were ranked as most important by the authorities in home economics education who cooperated in an earlier study<sup>2</sup> were selected for use in the present study.

As a second step the ten concepts were defined in terms of pupil and teacher behaviors. This step is in line with Tyler's emphasis on defining educational objectives in terms of behavior changes to be expected if the desirable pupil development is achieved.<sup>3</sup> To illustrate more fully the implications of these beliefs, it was necessary to define them with positive and negative behavioral implications, that is, by describing what the teacher would do if she believed in the concept and what she would avoid doing

---

<sup>1</sup>Cf. ante, p. 5.

<sup>2</sup>Hildegarde Johnson, et al., "Our Educational Beliefs", Journal of Home Economics, LIII (March, 1960) pp. 175-178.

<sup>3</sup>Ralph W. Tyler, Basic Principles of Curriculum and Instruction, (Chicago: University of Chicago Press, 1950), pp. 69-78.

if she effectively applied the concepts. These lists of behavioral implications for each of the ten concepts were compiled by recording incidents from the researcher's previous teaching experience and interpreting illustrations of classroom situations from numerous references. Some of the original statements were combined after considering the reactions of college students and staff members in home economics education. Some statements were revised or restated, and new implications of the concepts for pupil or teacher behavior were added.

Another important step in the development of the test was the identification of situations in which the above behaviors for the selected concepts might be observed. Obviously these situations would be largely classroom situations.

A third step in developing the test was the selection of the form of the test items to be used in the completed test. The decision was made to use an objective item form involving a hypothetical classroom situation followed by statements of alternative teacher or pupil behaviors. Some of the statements of the teacher or pupil behaviors should indicate mistakes that are typical of classroom situations.<sup>4</sup>

---

<sup>4</sup>Hester Chadderdon, *et al.*, "Development of Paper and Pencil Tests to Evaluate the Ability to Apply Generalizations in Home Economics" (Des Moines: State of Iowa, 1947), p. 10.

Stating questions in subjective form. An essay or open-end question approach was selected as an appropriate procedure for obtaining written descriptions of behavioral implication for the ten concepts from college students. This was an important step, since it is very difficult to foresee all the kinds of mistakes which beginning teachers might make, and to state them in their language as plausible reactions to a test situation.

Several forms of essay questions were used to secure student expressions of behaviors. Responses to items of each form were secured from home economics students in Methods of Teaching Home Economics at The Woman's College, University of North Carolina. The reactions from the college students were anonymous in each trial approach to encourage fuller cooperation and more accurate responses. This procedure seemed especially helpful with the open-end questions. Wherever feasible, the students were asked to respond to only the number of questions they could comfortably answer within an hour.

The first essay questions were direct statements of the educational concepts. Each concept was followed by a question in which the student was asked whether or not she accepted it. If her response was negative, she was asked to support her decision with reasons. If her response was affirmative, she was asked to list examples of how a home economics teacher might put that particular concept into

practice in classroom situations. An example of this approach is as follows:

Do you believe that teaching methods should be selected on the basis of their effectiveness in bringing about desired pupil development? If no, why? If yes, what are some examples of how a teacher could put this belief into practice?

Interpretations by students of the concepts in practical situations provided many useful ideas for developing hypothetical classroom situations and alternative reaction statements. This approach, however, was not as helpful as had been anticipated in obtaining descriptions of the behavioral implications.

The second form of essay question was identical with the first except that the respondent was asked to describe the pupils' behavior as well as the teacher's behavior in cases where the concept was accepted. A question was stated in the following manner for each of the ten concepts:

Do you believe that democratic procedures should be used in the homemaking classroom, yes or no? If no, why? If yes, what are some examples of how a teacher or the pupils would behave if the teacher believed this to be important?

The example cited was designed to solicit their opinions of behavioral implications for the application of the concept, "Democratic procedures should be used in the homemaking classroom." These responses, stated in the language of the students, were used to supplement the growing lists of behavioral implications for each of the ten selected educational concepts.

A third form of question was used to obtain the reactions of students on how a teacher would behave in classroom situations if she did not apply an appropriate educational concept. Such student responses would be helpful in designing statements which would be plausible yet undesirable or irrelevant reactions to follow the situations. The students were presented direct statements of the selected concepts, prefaced by the following remarks:

Below are listed four educational beliefs held as important by the leaders in the area of home economics education. Under each theory (or belief) please describe some of the negative behaviors of a teacher who does not apply this belief in her home economics classes.

Reactions were recorded for all ten of the concepts, and the responses were incorporated into the test situations and growing list of behavioral descriptions.

A fourth essay-type question was used experimentally with the students in the college methods classes, an approach more nearly resembling the form desired for items in the completed test. The difference in this approach, as compared to those used earlier, was to withhold any direct statement of the concept being tested, but to use a hypothetical classroom situation based on an opportunity to apply one or more of the ten concepts. One of the hypothetical situations used to obtain student responses was:

Miss Curtis has in her 10th grade class a student

named Sue who has an I. Q. of 95, does D work in general in her high school classes and averaged D work in her first year home economics class. Sue comes from a home where the family doesn't own a sewing machine, so she has done no sewing since her 9th grade home economics class. Miss Curtis also has in the same class a 10th grade student named Carol who has an I. Q. of 118, has maintained a B+ in her high school work, and had a B average in her 9th grade home economics class. Carol's family owns a sewing machine and Carol has made a few simple garments during her summer vacation. The class is approaching a unit on sewing. Should Miss Curtis expect the same quality in the finished garments of the girls? Why?

This situation was designed to secure student interpretations of the appropriate behavior for the teacher who applied concept number eight, "The homemaking teacher should attempt to guide each pupil in such a way that his progress is satisfactory in relation to his initial skill and ability." An entire class of twenty-four students responded to this situation. Many of the responses were lengthy and results were awkward to tabulate, yet this type of question seemed to be most helpful in stating items.

In summary, all of the open-end statements and essay questions were valuable as steps in gathering information for construction of objective test items. Each hypothetical classroom situation was to be followed by alternative teacher or pupil behaviors. Statements written by students were helpful in designing the alternative behavior statements to follow each situation. The misconceptions they expressed were especially helpful in designing plausible yet undesirable behavioral statements to various situations

in testing the respondents' ability to discriminate between correct and incorrect responses. The subjective approaches were helpful, also, to the author in designing hypothetical classroom situations that teachers currently encounter.

## II. DEVELOPMENT OF THE OBJECTIVE FORM OF TEST ITEMS

An objective form of test items had been proposed early in the study in which a teacher would be asked to project herself into a realistic situation and decide what she would do.

Choice of the objective form for test items. Several objective forms for test items were considered in preliminary forms of the test, usually of the combined-response variety. One form of test item consisted of a hypothetical classroom situation followed by three possible solutions and a list of reasons for the selection of a specific solution. Plausible reasons were included for each potential solution. First, the respondent would choose one solution. As a second step, the respondent would select one or more reasons for the solution from the list of six or more reasons. The basic concept being applied was described in one of the reasons. Requiring the person reacting to the test to choose the reasons for her decision was expected to provide added evidence of her ability to apply the concept.

The objective form which was considered most satisfactory was designed in a similar way. The item was composed of a hypothetical classroom situation followed by statements of possible teacher behaviors. An example of this form is as follows:

I. In your ninth-grade homemaking class, you have two pupils from the total of eighteen who are considered slow learners. Determine the advisability of each of the following ways of handling the homework assignments throughout the year under these circumstances.

1. Give the whole class the same assignments, but ignore the grades on written assignments for the slow learners.
2. Don't give the slow learners written assignments; let their classwork determine their grades.
3. Give the slow learners performance activities to carry out at home while the other pupils have written assignments.
4. Give all pupils performance activities to do at home, as well as written homework assignments.
5. Allow some choice in the amount of homework to be turned in by each pupil, while encouraging them to do whatever they are capable of doing; allow oral as well as written reports of these assignments.
6. When giving written assignments, lower the amount required only for the slow learners.

The person reacting to the test item was asked to imagine that she was the teacher and to respond to each statement by selecting one of five levels of response, the one nearest to the way she would behave if she were the teacher. The five levels of response ranged from a position of strongly agreeing with the statement through a neutral position to that of strongly disagreeing with the behavioral

statement. The code to responses was as follows:

Strongly Agree  
Agree in General  
Uncertain  
Disagree in General  
Strongly Disagree

The difference in this form of test item and the first one described was the importance of having the respondent react to every statement of behavior following a test situation on the above five-point scale. Items were developed by the researcher and examined by specialists in home economics education to determine if the hypothetical situations and the statements of behavior expected from the teacher involved the use of the ten educational concepts.

Responses to the trial form of the test. A trial form of the test containing fifteen situations and 110 statements of possible teacher or pupil behaviors was prepared. The title of the test, "You Are There," was selected at this time. Eight home economics supervising teachers were asked to participate in the study by responding to the trial form of the test. Each teacher was asked to react critically to the situations described and encouraged to write in any difficulty in responding to the test. These comments were helpful in arranging the sequence of items, and clarifying directions as well as wording of the items.

### III. DEVELOPING A SCORING KEY

Three specialists in home economics education at The Woman's College, University of North Carolina, were asked to serve as judges for the purpose of developing a scoring key. They were given independent copies of "You Are There" and an IBM answer sheet on which to record their responses. Their reactions would determine what the correct response was for each item for the scoring key and which items should be deleted. Responses of the three judges were summarized and analyzed for evidence that they agreed or disagreed concerning the correct level at which each item should be checked. Agreement was defined as:

1. Unanimous agreement of the three judges at any of the five levels of response, for example:

A	B	C	D	E	or	A	B	C	D	E
(0)	(0)	(0)	(0)	(3)		(0)	(3)	(0)	(0)	(0)

2. Agreement on whether the correct response should be at the agree or disagree end of the scale, for example:

A	B	C	D	E	or	A	B	C	D	E
(1)	(2)	(0)	(0)	(0)		(0)	(0)	(0)	(1)	(2)

3. Agreement similar to 2 except that one judge was neutral in response, for example:

A	B	C	D	E	or	A	B	C	D	E
(0)	(0)	(1)	(2)	(0)		(1)	(1)	(1)	(0)	(0)

Items to which the judges responded in any of the above ways were accepted for the test.

The test author met with the committee of judges to review their reactions to the situations and behavioral

statements. Items on which they disagreed were discussed. Examples of items in which judges disagreed are as follows:

A B C D E    or    A B C D E    or    A B C D E  
 (1)(0)(1)(1)(0)    (0)(1)(1)(1)(0)    (1)(0)(1)(0)(1)

In several cases a judge changed her reaction to an item, or the item was reworded in such a way that the judges could agree on a correct level of response. These items were retained. In cases where the judges failed to agree on responses after discussing the item, that item was rejected. Nine of the 110 items were discarded, leaving 101 usable items. Following revisions, the test was multilithed for mailing.

An IBM answer sheet was selected for use in recording reactions to the test. Standardized answer sheets facilitated scoring of large numbers of tests and promoted greater accuracy in recording responses. A copy of the scoring key is included in the Appendix, page 107.

A weight of two points was arbitrarily assigned the level of response to fifty items upon which the judges unanimously agreed concerning the correct level of response.<sup>4</sup> For the remaining fifty-one items, the levels of response selected by any of the judges were given a weight of one point.<sup>5</sup> A response was considered incorrect and received a weight of zero when no judge had responded at that particular level of response. A copy of the 101 weighted

<sup>4</sup>Cf. ante, p. 39.

<sup>5</sup>Ibid.

correct responses is included in the Appendix, page 107. The total possible score for the test was 151. Scoring stencils were prepared separately for odd and even numbered items to aid in the item analysis.

#### IV. SECURING RESPONSES

A sample of 150 home economics teachers in North Carolina, who were not already cooperating in another phase of the larger research project which included the present study, was selected using a random method. The cooperation of both Negro and white vocational home economics teachers was solicited.

A procedure was used for securing responses which would keep the respondent anonymous, yet would facilitate follow-up measures. Asking the teachers to remain anonymous as they responded was an approach intended to promote frank and honest reactions to the items. To each teacher was mailed a large envelope containing a copy of "You Are There,"<sup>6</sup> an IBM answer sheet, a return envelope, a postcard,<sup>7</sup> and a letter<sup>8</sup> explaining the purpose of the study. The postcard was addressed to the institution sponsoring the research, and the teacher was asked to mail the signed postcard at the same time that she mailed the unsigned IBM

---

<sup>6</sup>See Appendix, p. 91.

<sup>7</sup>See Appendix, p. 104.

<sup>8</sup>See Appendix, p. 103.

answer sheet. When the postcard was returned, the teacher would not receive a follow-up reminder to return the answer sheet, thus reducing the cost involved in the follow-up effort. Since the test was mailed to the teachers late in the school year, only one reminder was sent out requesting them to react to the test.

The percentage of response was not expected to be high for a study of this nature. A total of 95 answer sheets was returned. This response represented 63 per cent of the sample.

#### V. ANALYSIS OF TEST DATA

The test was divided into two subtests of equal length using alternate responses. Four scoring stencils, two for each subtest, were punched and each test was hand scored. The total score was recorded on each answer sheet as the sum of the weighted correct responses on the two subtests.

The coefficient of reliability was determined by correlating the scores on the two subtests and applying the Spearman-Brown modified formula. The range, the mean, and the median were computed and the frequency distribution was determined.

Flanagan's method of estimating  $\rho$  in a bivariate normal population was used to determine the ability of each

item to differentiate between the upper 27 per cent and the lower 27 per cent of the teachers. There were twenty six teachers in each of the two groups. This method is discussed by Walker and Lev.<sup>9</sup> The score on the test is a continuous variable, and whether or not the teachers respond correctly is a dichotomous variable. Values of  $r$  can be read from a table prepared by Flanagan and recorded in Walker and Lev.<sup>10</sup> The proportion of the upper and of the lower groups who reacted correctly to each item was computed before referring to the table. The values of  $r$  vary from  $-1$  to  $+1$ . Negative values indicate that more teachers in the low than in the high group responded correctly. Items receiving negative values are termed reversals. Responses were accepted as having differentiated between the groups if the coefficient of correlation was  $.15$  or above.

---

<sup>9</sup>Helen M. Walker and Joseph Lev, Statistical Inference (New York: Henry Holt and Company, 1953), pp. 275-281.

<sup>10</sup>Ibid., pp. 472-475.

## CHAPTER IV

### ANALYSIS OF THE TEST

An achievement test was developed in the area of home economics education to measure the ability of home economics teachers and college students to apply ten important educational concepts. Data collected by administering the items to a sample of ninety-five home economics teachers in North Carolina were analyzed as a basis for suggested improvements in the test. The test consisted of fifteen situations, each followed by an average of seven possible reactions stated in terms of teacher or pupil behavior. Each of the 101 numbered statements following the hypothetical classroom situations was called an item. The teachers were asked to react to each item using a five-level code to responses, selecting one level of response for each item.

The maximum score on the test was 151. Scores ranged from 36 to 111, with a mean of 72, and a median of 83. The frequency distribution was positively skewed, indicating that more teachers in the group received scores above the mean than below it.

## I. RELIABILITY

The coefficient of reliability was determined by correlating scores on the odd and even halves of the test. The test was divided into two subtests of equal length by placing alternate items in the subtests. The resulting coefficient of reliability was .65 for a test one-half the length of the test. The Spearman-Brown modified formula was applied to estimate the reliability of the entire test. Reliability of the test was .78.

The coefficient of reliability of the present test, .78, is relatively high for a test of this nature. It is possible, however, that it can be further increased. One way to raise the reliability is to lengthen the test. Forty-five minutes to over 90 minutes were required for teachers to complete the present test. Most college class periods in North Carolina are no longer than 50 minutes; therefore, it does not seem advisable to lengthen it. However, an alternate way to increase reliability is to substitute new items for those omitted because they are not sufficiently discriminating. This may be done without increasing the length of the test. Suggestions of this nature were included in the recommendations for revision of the test.

Another factor which affects the reliability of a test is clarity of directions. The only indication that the directions of this test were not clear was that five

teachers omitted a response. This was a small number of omissions.

## II. PROCEDURE FOR ITEM ANALYSIS

An item analysis was made using a method developed by Flanagan and described by Walker and Lev.<sup>1</sup> This is a method for computing  $r$ , which is an estimate of  $\rho$  in a bivariate normal population. The estimated  $\rho$  is an expression of the ability of each item to differentiate between the 27 per cent who obtained the highest scores and the 27 per cent who obtained the lowest scores on the test. Items were arbitrarily accepted as having satisfactorily differentiated between these two groups if values of  $r$  were .15 or above.

After scoring, the 95 tests were arranged in numerical order from the lowest to the highest. The 26 tests having the highest scores were grouped as the upper 27 per cent of the 95 sample returns and the 26 tests having the lowest scores were grouped as the lower 27 per cent. Responses of the two groups were tabulated separately as a basis for the item analysis. For example, responses at the various

---

<sup>1</sup>Helen M. Walker and Joseph Lev, Statistical Inference (New York: Henry Holt and Company, 1953), pp. 275-281.

levels to Item 1 are as follows:

	<u>A</u>	<u>B</u>	<u>C</u>	<u>D</u>	<u>E</u>
Upper group	3	5	1	7	10
Lower group	2	8	4	10	1

Correct responses to this item are D and E.

Two-way tables were then constructed to show the number and percentage of persons in the upper and in the lower group who responded correctly. The entries in the table for Item 1 are:

	<u>Number</u>	<u>Per cent</u>
Upper group	17	65
Lower group	11	44

These proportions were used to secure values of  $r$  from a table prepared by Flanagan and reported in Walker and Lev.<sup>2</sup> The  $r$  values were used to determine how well the item discriminated between the upper and lower groups.

Data for the remaining 100 items were similarly recorded.

### III. METHODS OF REPORTING THE ITEM ANALYSIS

Results of the item analysis were reported in Tables 1 through 15. Each table contains information for one situation and the items related to it. Each situation was discussed separately. From each table the following

---

<sup>2</sup>Ibid., pp. 472-475.

information can be obtained: the number of teachers in the high and in the low groups reacting correctly to each item and the  $r$  value, which is an expression of the degree to which each item differentiates between the two groups. Whenever an  $r$  value was below .14, indicating lack of sufficient discrimination, this fact was indicated in the table with an asterisk (\*). Items with negative  $r$  values are termed reversals. More teachers in the low group than in the high group reacted correctly to these items.

An indication of item difficulty may be obtained by considering the total number in the two groups who reacted correctly. There were 52 teachers in the two groups, upper and lower. Whenever the number of correct responses was low the item was considered to be difficult. The few omissions of responses were not indicated in the tables, but were accounted for in the method used for estimating  $\rho$ .

#### IV. RESULTS OF THE ITEM ANALYSIS

The data in Table 1 indicate that five of the seven items in Test Situation I differentiate adequately between the two groups. The  $r$  values of these five items ranged from .16 to .46. Item 2 was a reversal. Most of the items appeared to be average in difficulty, Item 3 being the easiest in this test situation. Item 6 was the most effective item in discriminating between the two groups.

Test Situation I was designed to involve three closely

TABLE 1  
ANALYSIS OF RESPONSES OF UPPER AND OF LOWER  
27 PER CENT OF THE TEACHERS: SITUATION I

Item	Number responding correctly		r
	High Group	Low Group	
1	17	11	.23
2	11	15	-.18*
3	24	21	.10*
4	19	14	.17
5	21	17	.16
6	23	13	.46
7	15	8	.26

\*Indicates lack of sufficient discrimination.

related concepts, numbered three, five and eight.<sup>3</sup> They were stated as follows:

Pupils should be guided to evaluate their own progress toward personal goals.

A class should evaluate from time to time its progress toward group goals.

The homemaking teacher should attempt to guide each pupil in such a way that his progress is satisfactory in relation to his initial skill and ability.

<sup>3</sup>Cf. ante, pp. 5,6.

All of the items in Test Situation II, Table 2 differentiated adequately between the two groups except

TABLE 2  
ANALYSIS OF RESPONSES OF UPPER AND OF LOWER  
27 PER CENT OF THE TEACHERS: SITUATION II

Item	Number Responding Correctly		r
	High Group	Low Group	
8	15	9	.20
9	15	6	.37
10	15	2	.58
11	10	8	.05*
12	21	7	.52
13	11	1	.56
14	18	9	.36

\*Indicates lack of sufficient discrimination.

Item 11. Items 10, 12, and 13 received r values above .50 indicating that they do an especially effective job of discriminating between the two groups. The particular concepts involved in Test Situation II were numbers six and nine:

A teacher should continuously evaluate her own progress as a person and as a teacher.

Democratic procedures should be used in the homemaking classroom.

Five of the eleven items in Test Situation III, Table 3, discriminated effectively between the two groups,

TABLE 3  
ANALYSIS OF RESPONSES OF UPPER AND LOWER  
27 PER CENT OF THE TEACHERS: SITUATION III

Item	Number Responding Correctly		r
	High Group	Low Group	
15	25	23	.22
16	10	12	-.10*
17	21	19	.07*
18	23	8	.60
19	8	11	-.14*
20	8	9	-.06*
21	22	24	-.29*
22	20	15	.20
23	4	2	.16
24	16	4	.50
25	24	24	-.13*

\*Indicates lack of sufficient discrimination.

especially items 18 and 24. Five of the six items which did not discriminate were reversals. Item 17 received an r value below the level of acceptability. The test situation was designed to involve concepts six and ten:

A teacher should continuously evaluate her own progress as a person and as a teacher.

In all areas of the homemaking program pupils should be stimulated to independent thinking.

Items 27, 30, and 31 differentiated adequately between the upper and lower groups, as shown by the data in Table 4 for Test Situation IV. Item 26 received an r value

TABLE 4

ANALYSIS OF RESPONSES OF UPPER AND OF LOWER  
27 PER CENT OF THE TEACHERS: SITUATION IV

Item	Number Responding Correctly		r
	High Group	Low Group	
26	21	18	.12*
27	24	17	.39
28	16	16	.03*
29	13	16	-.12*
30	22	14	.36
31	20	13	.28

\*Indicates lack of sufficient discrimination.

of .12 which approaches the level of acceptability. Item 29 was a reversal. The particular concept involved in this test was:

In all areas of the homemaking program pupils should be stimulated to independent thinking.

Only Item 37 was given an  $r$  value indicating sufficient ability to differentiate between the two groups in Test Situation V, Table 5. The fact that a high number in

TABLE 5  
ANALYSIS OF RESPONSES OF UPPER AND OF LOWER  
27 PER CENT OF THE TEACHERS: SITUATION V

Item	Number Responding Correctly		$r$
	High Group	Low Group	
32	21	21	0*
33	8	7	.05
34	23	22	.08*
35	19	19	0
36	16	21	-.23*
37	11	4	.31
38	22	20	.08*

\*Indicates lack of sufficient discrimination.

each group responded correctly to Items 32, 34, and 38 and the low  $r$  values suggested that the items were easy and lacking in ability to discriminate between the groups. Item 33 was difficult for persons in both groups and also was not discriminating. The statement suggested the use of grades as a motivation for pupil development in this situation, which is a common misconception among classroom

teachers. The fact that Item 34 was more lengthy than other statements in the test may have caused many of the persons who reacted to the test to guess that the reaction was desirable behavior for the teacher. Item 36 was a reversal and may have been an ambiguous statement since there was a considerable difference of opinion among the judges concerning their reaction to the situation. Test Situation V was designed to involve two particular concepts, one and four:

Teaching methods should be selected on the basis of their effectiveness in bringing about desired pupil development.

The homemaking program should deal with attitudes, interests, habits, understandings and appreciations as well as abilities.

All of the items in Test Situation VI, Table 6,

TABLE 6

ANALYSIS OF RESPONSES OF UPPER AND OF LOWER  
27 PER CENT OF THE TEACHERS: SITUATION VI

Item	Number Responding Correctly		r
	High Group	Low Group	
39	8	3	.27
40	20	8	.46
41	1	1	0*
42	26	20	.55
43	25	17	.50
44	26	23	.43

\*Indicates lack of sufficient discrimination.

proved effective in differentiating between the two groups, except Item 41. Only one person in each group responded correctly to this item, a most unusual result. The statement was lengthy and therefore may have falsely suggested that the behavior described was desirable, or it may be that telling pupils how to feel is a commonly accepted practice. The researcher believes there is no question about the correctness of the judges' decision that E is the proper level of response. Test Situation VI was designed to involve two particular concepts which read:

Teaching methods should be selected on the basis of their effectiveness in bringing about desired pupil development.

Discussions and other learning experiences in homemaking classes should help pupils to better understand their own behavior and the behavior of other people.

All of the items in Test Situation VII, Table 7, differentiated adequately between the two groups, according to the data collected. Most of the persons in the high scoring group reacted correctly to all the items, except Item 49, which seemed to be a more difficult statement for both groups. Item 51 was especially effective as a discriminator between the two groups with an  $r$  value of .60. The test situation was designed to appraise a teacher's ability to apply concepts one, two, and five:

Teaching methods should be selected on the basis of their effectiveness in bringing about desired pupil

development.

Classroom activities should be provided which can carry over into the pupil's personal and home experiences.

A class should evaluate from time to time its progress toward personal goals.

TABLE 7

ANALYSIS OF RESPONSES OF UPPER AND OF LOWER  
27 PER CENT OF THE TEACHERS: SITUATION VII

Item	Number Responding Correctly		r
	High Group	Low Group	
45	19	16	.17
46	22	15	.32
47	26	24	.35
48	25	17	.49
-----			
49	12	2	.49
50	23	19	.16
51	24	10	.60

In Table 8, Test Situation VIII, the only item which did not differentiate adequately between the two groups was Item 53. An equal number of persons in each group responded correctly. All persons in the high scoring group responded correctly to Item 54. Concept one which was

involved in this test situation was stated:

Teaching methods should be selected on the basis of their effectiveness in bringing about desired pupil development.

TABLE 8

ANALYSIS OF RESPONSES OF UPPER AND OF LOWER  
27 PER CENT OF THE TEACHERS: SITUATION VIII

Item	Number Responding Correctly		r
	High Group	Low Group	
52	24	21	.35
53	15	15	.04*
54	26	22	.44
55	16	11	.20

\*Indicates lack of sufficient discrimination.

The data collected for Test Situation IX, Table 9, revealed a wide range of r values, from  $-.27$  to  $.69$ . The items which discriminated adequately between the two groups were 57, 58, and 62. Recorded r values for Items 56, 60, and 61 indicated that these items were reversals. Though Item 59 was not a reversal, the assigned r value was below the level of acceptability. The most effective item following this test situation was Item 57 with an r value of  $.69$ . This test situation involved concepts two and

eight:

Classroom activities should be provided which can carry over into the pupil's personal and home experiences.

The homemaking teacher should attempt to guide each pupil in such a way that his progress is satisfactory in relation to his initial skill and ability.

TABLE 9

ANALYSIS OF RESPONSES OF UPPER AND OF LOWER  
27 PER CENT OF THE TEACHERS: SITUATION IX

Item	Number Responding Correctly		r
	High Group	Low Group	
56	9	13	-.20*
57	24	7	.69
58	24	12	.55
59	23	22	.07*
60	9	10	-.04*
61	11	18	-.27*
62	25	20	.39

\*Indicates lack of sufficient discrimination.

Except for Item 66, which was a reversal, all the items in Test Situation X, Table 10, differentiated adequately between the high and low scoring groups. Item 63 was the most effective item and yet one of the least difficult items

for both groups of teachers. The number of correct

TABLE 10

ANALYSIS OF RESPONSES OF UPPER AND OF LOWER  
27 PER CENT OF THE TEACHERS: SITUATION X

Item	Number Responding Correctly		r
	High Group	Low Group	
63	26	23	.42
64	10	5	.25
65	10	6	.19
66	20	24	-.27*
67	24	22	.21

\*Indicates lack of sufficient discrimination.

responses in each group indicated that Items 64 and 65 were the most difficult items. The concept which was involved in this test situation reads:

Classroom activities should be provided which can carry over into the pupil's personal and home experiences.

Each item following Test Situation XI, Table 11, differentiated adequately between the two groups with the exception of Item 69. Item 69 was easy for persons in both groups and was not discriminating. Items 71 and 74 were the most difficult items according to the total number of correct responses. The most discriminating item was Item 73

with an  $r$  value of .44. The concepts which the test

TABLE 11

ANALYSIS OF RESPONSES OF UPPER AND OF LOWER  
27 PER CENT OF THE TEACHERS: SITUATION XI

Item	Number Responding Correctly		$r$
	High Group	Low Group	
68	7	2	.31
69	26	25	0*
70	20	10	.39
71	16	6	.39
72	24	20	.23
73	21	10	.44
74	11	6	.23
75	22	15	.31

\*Indicates lack of sufficient discrimination.

situation was designed to involve, numbers nine and ten, are stated:

Democratic procedures should be used in the homemaking classroom.

In all areas of the homemaking program pupils should be stimulated to independent thinking.

All the items in Test Situation XII, Table 12, differentiated adequately between the two groups, except items 76 and 77. The judges agreed that both items were

expressions of undesirable behavior for the teacher in this situation. An especially effective item related to this test situation was Item 79. The concept involved in Test Situation XII was stated as follows:

The homemaking program should deal with attitudes, interests, habits, understandings and appreciations as well as abilities.

TABLE 12

ANALYSIS OF RESPONSES OF UPPER AND OF LOWER  
27 PER CENT OF THE TEACHERS: SITUATION XII

Item	Number Responding Correctly		r
	High Group	Low Group	
76	18	18	.02*
77	6	8	-.11*
78	19	8	.41
79	25	9	.70
80	25	22	.29
81	24	15	.45

\*Indicates lack of sufficient discrimination.

The data recorded in Table 13 for Test Situation XIII indicate that all of the statements of reactions, 82 through 88, differentiated effectively between the high and low scoring groups. Two of the items, numbered 86 and 87, proved to be especially effective in discriminating between

the two groups and did, in fact, receive the highest  $r$  values recorded for any item in the test. The particular

TABLE 13

ANALYSIS OF RESPONSES OF UPPER AND OF LOWER  
27 PER CENT OF THE TEACHERS: SITUATION XIII

Item	Number Responding Correctly		$r$
	High Group	Low Group	
82	4	2	.18
83	22	12	.41
84	4	2	.18
85	23	5	.69
86	22	2	.75
87	23	2	.78
88	17	2	.63

concepts which were involved in the test situation were numbers one, four and seven. They are stated as follows:

Teaching methods should be selected on the basis of their effectiveness in bringing about desired pupil development.

The homemaking program should deal with attitudes, interests, habits, understandings and appreciations as well as abilities.

Discussions and other learning experiences in homemaking classes should help pupils to better understand their own behavior and the behavior of other people.

Another unusually effective group of items was the

group following Test Situation XIV, Table 14. The  $r$  values

TABLE 14

ANALYSIS OF RESPONSES OF UPPER AND OF LOWER  
27 PER CENT OF THE TEACHERS: SITUATION XIV

Item	Number Responding Correctly		$r$
	High Group	Low Group	
89	16	3	.55
90	9	4	.27
91	19	2	.69
92	23	17	.37
93	17	5	.50
94	10	1	.54
95	24	14	.58
96	11	3	.41

ranged from .27 to .69 indicating that all of the statements discriminated adequately between the two groups. The test situation was followed by several items which were apparently difficult for members of both groups, but especially for the persons in the low scoring group. The test situation was designed to involve concept six which was stated:

A teacher should continuously evaluate her own progress as a person and as a teacher.

The data collected for the last test situation,

number XV, Table 15, indicate that two of the five items

TABLE 15

ANALYSIS OF RESPONSES OF UPPER AND OF LOWER  
27 PER CENT OF THE TEACHERS: SITUATION XV

Item	Number Responding Correctly		r
	High Group	Low Group	
97	9	19	-.38*
98	17	18	-.01*
99	16	6	.40
100	11	13	-.08*
101	13	5	.35

\*Indicates lack of sufficient discrimination.

discriminated adequately between the two groups. The other three items, numbered 97, 98, and 100 were reversals. The concept which this situation was designed to involve was stated:

The homemaking teacher should attempt to guide each pupil in such a way that his progress is satisfactory in relation to his initial skill and ability.

Thirty-one items, out of the total 101, did not discriminate between the high and low scoring groups at the .15 level. Three of the fifteen test situations contained items all of which discriminated at the desired level. In five of the test situations only one item did not

discriminate. In two of the situations, two items failed to discriminate adequately, and in the two remaining situations six items did not differentiate between the two groups.

TABLE 16

NUMBER OF ITEMS HAVING  $r$  VALUES  
ABOVE AND BELOW .15

Test Situation	$r = .15$ or Above	$r = .14$ or Below
I	5	2
II	6	1
III	5	6
IV	3	3
V	1	6
VI	5	1
VII	7	0
VIII	3	1
IX	3	4
X	4	1
XI	7	1
XII	4	2
XIII	7	0
XIV	8	0
XV	2	3
Total	70	31

## REVISION OF THE TEST

Procedure for revision. The researcher prepared two revisions of the test for further test analysis. The first revision, designated as Form A, contained all the items which were not reversals. Items with  $r$  values between .00 and .15 seem to add little to the discriminating power of the test, but they do help the respondent to consider a number of possible reactions to the test situations. Whenever fewer than four items could be retained for a test situation, the entire test situation was omitted. A maximum score for each test situation in the first revision of the test, Form A, and a new total score were recorded.

The second revision of the test, Form B, was an effort to retain some items which the researcher believed to have merit if the items were reworded or the scoring of the items modified. The test is in the earliest stages of development and further group reactions are recommended before certain items are discarded as being ineffective. Whether an item with an  $r$  value of .14 or below was retained in this revision was determined by examining the empirical data secured from the sample of 95 home economic teachers. The  $r$  values recorded in the tables presented with the item analysis are based solely on the scoring key developed by the three judges. Frequently all of the judges agreed with an item, but two of them marked A, strongly agree, and one

of them marked B, agree in general. In this case a respondent received one point for marking either A or B and was credited with responding correctly in the item analysis. In several items better discrimination between the upper and lower group could be achieved by considering A as the only correct response. This does not change the decision of the three judges with respect to the desirability or undesirability of the stated behavior, but improves the effectiveness of the test item by narrowing or broadening the number of acceptable levels of response for any given item. An illustration of a typical item which was treated in this manner was Item 25. Both level A and level B were accepted as correct in the first scoring key and the recommendation was to accept level A only in a modified scoring key.

	<u>Correct Responses Original Key</u>	<u>Correct Responses Modified Key</u>
Upper 27 per cent	24	17
Lower 27 per cent	24	8

The  $r$  value for Item 25 would be .34 if the modified scoring key were accepted, indicating satisfactory discrimination between the two groups.

Proposed changes in the scoring key might involve inclusion of the C (middle or neutral) level of response as a correct answer, but could not result in accepting levels

of response on both sides of the neutral position on the scale. To illustrate a typical item revision, the following steps might be helpful.

Step 1. Select an item assigned an  $r$  value below the level of acceptability (noted by an asterisk in tables).

Examples:

	<u>Present <math>r</math> Value</u>
Item 21	-.29
Item 38	.08

Step 2. Examine the empirical data for the item for suggested trends in both the high and low scoring groups toward broadening or narrowing the correct levels of response.

Example:

<u>Item 21</u>	<u>A</u>	<u>B</u>	<u>C</u>	<u>D</u>	<u>E</u>
Upper group	9	13	0	3	1
Lower group	3	13	8	1	0

In the original scoring key, levels A, B, and C were the correct levels of response. It may be observed that the high scoring group chose levels A and B, omitting C entirely whereas one third of the low scoring group chose level C.

Step 3. Select for the revised key the level(s) of response favored by the upper group and record the new total number of correct responses for each group according to the accepted change.

Example:

<u>Item 21</u>	<u>A</u>	<u>B</u>	<u>(C Omitted)</u>	<u>Total Correct</u>
Upper group	9	13	X	22
Lower group	3	13	X	16

Step 4. Calculate the proportion of the group who responded correctly and record these.

Item 21	<u>Total Correct</u>	<u>Total Number of Responses</u>	<u>Proportion</u>
Upper group	22	26	84.6
Lower group	16	25	64.0

Step 5. Use these proportions to calculate a new  $r$  value for the item based on the modified scoring key. In this example the new  $r$  value is .27, a value which is above .15, the level of acceptability.

Step 6. Record the suggested changes in the scoring key.

This procedure was followed to recommend revisions in the scoring key for other items. In cases where following the procedure did not result in a new  $r$  value above .15, the item was further examined to see whether a minor change in wording might improve the discriminating power of the item. Discussion of these reworded statements will be presented. The researcher suggests that at least five items be used with each test situation in future trials, and recommendations for improvements included this consideration. Items which were assigned  $r$  values lower than the criterion of acceptability, .15, and which could not be improved in effectiveness through the above methods were discarded.

Results of the test revision. The first revised form of the test contained thirteen test situations followed by a range of four to eight items. Test Situations

V and XV were omitted. Within other test situations the following fourteen items were discarded as being ineffective in discriminating between the two groups: 2, 16, 19, 20, 21, 25, 29, 41, 56, 60, 61, 66, 69, and 77. The total possible score for each test situation for this revision of the test, Form A is recorded in Table 17. A maximum score of 119 resulted when the 75 items were retained.

Form B, the second revised form of the test, contained reworded items and a modified scoring key for certain other items. All items included in Form A were included in Form B. Items which were retained, items discarded in each test situation, and the modification of the scoring key for specific items will be discussed. Whenever substitutions were recommended for ineffective items, these were included in the discussion of the test situation.

The reversal following Test Situation I, Item 2, could not be improved by either of the methods described earlier and the item was discarded (Table 1, page 49).

Six of the eleven items in Test Situation III received  $r$  values below .15 (Table 3, page 51). Five of these six items were discarded in the first revision of the test, Form A. No changes were recommended for Items 19 and 20, and they were also omitted in the second revision, Form B. Item 16 was retained because it approached an effective level of discrimination, with an  $r$  value of

.12, if response levels A and B were accepted as correct

TABLE 17  
SUMMARY OF DECISIONS CONCERNING  
ITEMS IN REVISED FORM A

Test Situation *	Items Retained	Items Discarded	Possible Score for Items Retained
I	6	1	9
II	7	0	13
III	6	5	10
IV	5	1	8
VI	5	1	6
VII	7	0	11
VIII	4	0	5
IX	4	3	6
X	4	1	5
XI	7	1	11
XII	5	1	8
XIII	7	0	12
XIV	8	0	15
Totals	75	14	119

\*Test Situations V and XV were discarded.

responses. Level A is in the same direction on the scale as

B, the choice of the judges, and would increase the number of correct responses in the upper scoring group to 21 as compared to 18 in the lower group. Item 21 had an acceptable  $r$  value of .27 if the C level, the neutral level of response, was not given credit as being correct. This change reduced the number of correct responses in the lower group to 16, while the 22 in the upper group remained constant. In Item 25, the discriminating power was much improved with an  $r$  value of .34 by considering response level A the only correct response.

The items in Test Situation IV (Table 4, page 52), which were examined for possible changes, were Items 26 and 29. Item 26, which received an  $r$  value of .12, received an improved  $r$  value of .15 when C was omitted as a correct response. This change in the scoring key was recommended. Item 29 was a reversal, but approached the level of acceptability with a change in the scoring key. The  $r$  value became .13 when the C level of response was omitted, and the A level of response was included. The modification was recommended for the scoring key to retain five items in this test situation for future group reactions.

Only Item 37 was given an  $r$  value indicating sufficient ability to differentiate between the two groups in Test Situation V (Table 5, page 53). The researcher believed that the hypothetical situation had potential value and suggested that it be retained in Form B. Item 32

was omitted since it was non-discriminating and no improvements resulted from a change in scoring. Item 33 was difficult for persons in both groups and also lacked sufficient ability to differentiate between the two groups. The statement suggested the use of grades as a motivation for pupil development in this situation, which is a common misconception among classroom teachers. The researcher suggested that this item be retained to get another group reaction before it is discarded as being ineffective, since this is an especially important application of concepts numbered one and four. If the judges had limited the correct response to Item 34 to A, the  $r$  value would become .42. This change was recommended. For Item 35, the  $r$  value would be changed to .27 if the correct response were limited to A. Item 38 was accepted, also, with a new  $r$  value of .46, by narrowing the correct response to the A level. On the basis of the empirical data the researcher suggested that Items 33, 34, 35, 37, and 38 in Test Situation V be retained for future group reactions.

Only one item, number 41, was examined for possible changes in Test Situation VI. The reactions to this item (Table 6, page 54) indicated that it was the most difficult item in the test. The researcher recommended that the following item be substituted for 41: "Discuss the desirable behaviors they observed; then role play the situation again to see if pupils have changed in the

direction of more mature behavior, especially Phyllis." With the suggested revision of the item statement the scoring key was changed to indicate that A was now the correct level of response.

All of the items in Test Situation VII (Table 7, page 57) differentiated adequately between the two groups, and no revisions were necessary.

The only item in Test Situation VIII which did not differentiate adequately between the two groups was Item 53 (Table 8, page 58). An equal number of persons in each group responded correctly. The researcher suggested that the item might be improved by rewording the statement to read, "Allow the slow learners to substitute other activities for written homework assignments." This statement was substituted for Item 53, with D and E as the correct levels of response. Item 55 which received an r value of .20 might be improved by making two statements from the original one, since the length of the statement may have helped the respondent guess that the behavior was desirable. The researcher's suggestion for these two statements was: "Allow some choice in the amount of homework to be turned in by each pupil, while encouraging them to do as much as they are capable of doing," and, "Allow oral as well as written reports of the assignments." The A level of response would be considered correct for each of these statements.

The recorded  $r$  values for Items 56 and 60 in Test Situation IX (Table 9, page 59) indicate that these items are reversals. The researcher would not recommend improvements for either of them. The discriminating power of Item 61, which was also a reversal, was improved with a new  $r$  value of .24 omitting response level C, and accepting response levels A and B as the correct ones. Item 59 received a new  $r$  value of .40 by limiting the correct level of response to A. With the recommended changes, there were five items in the test situation.

Item 66 in Test Situation X was a reversal with an  $r$  value of  $-.27$  (Table 10, page 60). The  $r$  value became .29, however, when A was considered the only correct level of response. Two of the three judges selected A.

Modification of the scoring key was recommended for only one item in Test Situation XI, (Table 11, page 61). Narrowing the scoring key to accept A as the only correct level of response for Item 69 resulted in an  $r$  value of .67 indicating a high degree of discrimination between the two groups.

Test Situation XII contained two items, 76 and 77, which did not differentiate adequately between the two groups, (Table 12, page 62). There was no apparent reason for these reactions and no suggestions for improvement were made at this time. Item 77 was also discarded in revised Form B. The researcher suggested that Item 76 be

retained for one more group reaction before it is discarded as being ineffective, giving a total of five items for this test situation.

All the items in Test Situations XIII (Table 13, page 63) and XIV (Table 14, page 64) were effective in discriminating between the two groups. These were included in this revision of the test without changes.

Three items, numbered 97, 98, and 100, were all reversals in Test Situation XV (Table 15, page 65). This test situation was dropped in the first revision of the test, Form A. For the second revision, the scoring key was modified for Items 97, 98, and 100, based on the empirical data. The  $r$  value for Item 97 became .39 when A was accepted as the only correct level of response. Item 98 received an  $r$  value of .15 when E was the only acceptable level of response. Item 100 was also improved in discriminating power by narrowing the correct levels of response to A and B, omitting C. With this suggested change in the scoring key, the item received an  $r$  value of .30.

The researcher recommended revisions in the scoring key based on the empirical data for fourteen items. Four statements were reworded and recommended on the basis of the improvements to be used with this second revision of the test in future trials. At least five items for each

TABLE 18  
 SUMMARY OF DECISIONS CONCERNING  
 ITEMS IN REVISED FORM B

Test Situation	Items Retained	Items Discarded	Possible Score for Items Retained (Modified Key)
I	6	1	9
II	7	0	13
III	9	2	14
IV	6	0	9
V	5	2	9
VI	6	0	8
VII	7	0	11
VIII	5 *	0	8
IX	5	2	8
X	5	0	7
XI	8	0	13
XII	5	1	8
XIII	7	0	12
XIV	8	0	15
XV	5	0	9
Totals	94	8	154

\*Item 55 was divided into two statements.

test situation were included in Form B. Several items retained would not add to the discriminating ability of the test, but they might increase the effectiveness of the other items in the test situations. For eight items (Numbers 2, 19, 20, 32, 36, 56, 60, and 77) the researcher made no suggested improvements and these items were omitted from both revisions of the test. The scoring key for Form B resulted in a maximum score of 154. The total number of items in this revised form of the test was ninety-two.

## CHAPTER V

### SUMMARY

The purpose of this study was to develop an achievement test to measure a home economics teacher's ability to apply ten basic concepts in home economics education. Each concept was defined in terms of behavioral implications for the teacher and pupils as a preliminary step in item construction. The "verbalized behavior" type of achievement test was selected as most appropriate for securing evidence of a teacher's ability to apply her knowledge and comprehension of the basic concepts.

The researcher designed hypothetical classroom situations involving an opportunity to apply one or more of the ten selected concepts. These situations followed by open-end questions were given to college students majoring in home economics education. Subjective responses of students were used to develop a list of ways that the teacher or the pupils might behave in each situation. Some of the behaviors were desirable and some were undesirable from the standpoint of applying the concepts. These hypothetical situations were examined by specialists in home economics education to insure appropriate content

as well as effective wording and desired level of difficulty. The resulting number of hypothetical classroom situations and behavioral statements were further refined as test situations and items for the objective form of the test, "You Are There."

The tryout form of the test contained fifteen situations and 110 behavioral statements, called items. A five-level code to responses was used to guide the respondent in reacting to each item, and responses were recorded on IBM answer sheets. The tryout form of the test was submitted to eight home economics supervising teachers in an area surrounding Greensboro, North Carolina to see if the test was appropriate in length, if the directions were clear, and if the wording of the test situations was clear.

The scoring key was developed by summarizing the responses of three judges who were specialists in home economics education at the Woman's College, University of North Carolina. The judges met as a committee with the researcher to discuss the items on which they had disagreed and revisions were suggested. The five levels of response were weighted for each item according to the agreement of the judges. As a result, the test had a total possible score of 151. There were fifteen test situations and 101 items in the copy of "You Are There" which was multilithed for mailing.

A sample was selected of 150 home economics teachers

in North Carolina who had not already been asked to cooperate in other research projects during the 1960-61 school year. The sample included both Negro and white vocational home economics teachers. A large envelope was mailed to each teacher containing a copy of "You Are There", an IBM answer sheet, a return envelope, a postcard and a letter explaining the purpose of the test. A total of 95 unsigned answer sheets was returned, giving a 63 per cent response, with only one follow-up effort.

The test was divided into two subtests of equal length using alternate responses, and each test was hand scored. The total score was recorded on each answer sheet as the sum of the weighted correct responses on the two subtests.

The coefficient of reliability was determined by correlating the scores on the two subtests and applying the Spearman-Brown modified formula. The coefficient of reliability was .78, which is relatively high for a test of this nature. Scores ranged from 36 to 111, with a mean of 72, and a median of 83. The frequency distribution was positively skewed, indicating that more teachers in the group received scores above the mean than below it.

An item analysis was made of the responses of the twenty-six teachers in each of the upper and lower 27 per cent of the group. Items were considered as differentiating adequately between these two groups if the value of  $r$  was

found to be .15 or above.

Thirty-one items, out of the total 101, did not discriminate between the high and low scoring groups at the .15 level. From the total fifteen test situations, three situations contained items all of which discriminated at the desired level. In five of the test situations only one item did not discriminate effectively. In two situations, four items failed to discriminate adequately, and in the two remaining situations, six items did not differentiate between the two groups.

Two revisions of the test were recommended for further test analysis. Revised Form A was identical with the form used in the present study except that two test situations and all items which were reversals were omitted. Items which were not reversals and yet had  $r$  values which were below .15 were retained. These items have some value when used with a sufficient number of other items to appraise the concept being tested in the hypothetical classroom situation. This form of the test contained 75 items, and had a maximum score of 119.

The second revision of the test, Form B, was an effort to retain some items which were discarded on Form A and which the researcher believed to have merit if the items were reworded or the scoring of items modified. Four statements were reworded and recommended for inclusion in Form B. On the basis of the empirical data, secured from

the sample of 95 home economics teachers, modifications in the scoring key were recommended for fourteen items. Eight items were discarded because it was not possible to improve the discriminating power of the item by either of the methods described. A total of fifteen situations and ninety-four items were included in Form B and the maximum score was 154.

The test is in the earliest stages of development and the following steps are recommended for improvement:

1. Develop several new test situations and items to replace those which were omitted in Form A. Test reliability is related to length of the test, and a longer test than the revised form may be completed in the same time allotted for the test in the present study.

2. Develop a scoring key in which there are three rather than five levels of response to each item.

3. Submit Form A and Form B to a new sample of teachers. Score Form B twice, once on the Form B key developed by the researcher and again on the key with three levels of response.

4. Compute a coefficient of reliability for Form A, Form B, and for Form B scored as though there were three levels of response.

5. Analyze the items in each form, including the new items.

6. Select from the revised forms of the test the one having the highest reliability and discriminating power.

7. Submit the test selected above to a new and larger sample of home economics teachers. Prepare frequency distributions and percentile ranks of their scores to be used as norms in interpreting future scores.

The researcher suggests that the final test could serve many useful purposes for home economics teacher educators. The test is recommended for use as an evaluation device to measure progress of the student in methods courses and student teaching. If the students have had adequate opportunities to develop the ability to apply each of the ten concepts, the test should secure evidence of this learning.

The final test is recommended for use as a teaching aid to emphasize the behavioral implications of the basic concepts throughout a course in Methods of Teaching. Single test situations or a combination of two or more situations may be used at various times in a college class.

The final test would be useful, also, in the work of area supervisors. The test could be administered to individual teachers or groups of teachers to help them recognize their strengths and needs as a basis for improving

their teaching. A further potential value of the test would be to help supervisors in identifying some teachers who might become effective supervising teachers.

An evaluation device of this nature could serve a further purpose in home economics education. Responses of teachers-in-service would be indicative of some of the kinds of help teachers need in graduate courses and would aid in planning these programs of instruction.

Journal of Educational Research  
and Psychological Technology  
Journal, XIX (1926), 1-10

Clark, William B. 1923. *Teacher's  
Education*. Boston:  
Little, Brown, 1923.

Clark, W. B. 1923. "The  
Ability of Teachers to  
with the Development of  
Specialized Instructional  
Work, 1923, 1924.

Clark, W. B. 1923. "The  
Faculty Work in the  
Generalization in  
Need for Specialized  
Instruction."

#### BIBLIOGRAPHY

Clark, W. B. 1923. "The  
Teacher's Ability to  
with the Development of  
Specialized Instructional  
Work, 1923, 1924.

Clark, W. B. 1923. "The  
Faculty Work in the  
Generalization in  
Need for Specialized  
Instruction."

Clark, W. B. 1923. "The  
Faculty Work in the  
Generalization in  
Need for Specialized  
Instruction."

Clark, W. B. 1923. "The  
Faculty Work in the  
Generalization in  
Need for Specialized  
Instruction."

Clark, W. B. 1923. "The  
Faculty Work in the  
Generalization in  
Need for Specialized  
Instruction."

## BIBLIOGRAPHY

- American Educational Research Association. "Educational and Psychological Testing." Review of Educational Research, XXIX (February, 1959).
- Bloom, Benjamin S. (ed.). Taxonomy of Educational Objectives. Handbook I: Cognitive Domain. New York: David McKay Company, Inc., 1956.
- Brown, Sara Ann. "Techniques for Evaluating the Ability of Teachers to Apply Principles Concerned with the Developmental Needs of Adolescent Girls." Unpublished Doctor's dissertation, University of Iowa, Ames, 1949.
- Chadderdon, Hester, et al. "Development of Paper and Pencil Tests to Evaluate the Ability to Apply Generalizations in Home Economics." Des Moines: Board for Vocational Education, State of Iowa, 1947.
- Clawson, Barbara Nelle. "Appraisal of Home Economics Teachers' Ability to Apply Three Basic Concepts of Home Economics Education." Master's thesis in progress, Consolidated University of North Carolina, Greensboro, 1962.
- Committee on Test Standards, AERA and NCMUE. Technical Recommendations for Achievement Tests. Washington: National Education Association (January, 1955).
- Cronbach, Lee J. Essentials of Psychological Testing. New York: Harper and Brothers, 1949.
- Curry, Helen Lanning. "Appraisal of Home Economics Programs in Secondary Schools in North Carolina." Master's thesis in progress, Consolidated University of North Carolina, Greensboro, 1962.
- Godwin, Helen Miller. "Educational Concepts Which Home Economics Teachers in North Carolina Accept and Believe They Apply." Unpublished Master's thesis, Consolidated University of North Carolina, Greensboro, 1961.

- Goodenough, Florence L. Mental Testing; Its History, Principles and Applications. New York: Rinehart and Company, 1949.
- Johnson, Hildegard, et al. "Our Educational Beliefs," Journal of Home Economics, LIII (March, 1961), pp. 175-178.
- Joint Committee of AERA, APA, and NCMUE. "Technical Recommendations for Psychological Tests and Diagnostic Techniques." Washington: American Psychological Association, Inc., 1954.
- Lindquist, Everet F. (ed.). Educational Measurement. Washington: American Council on Education, 1951.
- Roberts, Ada Frances. "The Application of Educational Principles by Home Economics Education Majors." Unpublished Master's thesis, University of Georgia, Athens, 1951.
- Smith, Eugene R. and Ralph W. Tyler. Appraising and Recording Student Progress. New York: Harper and Brothers, 1942.
- Stecklein, John E. "What Is A Good Test?" Bureau of Institutional Research, University of Minnesota. Bulletin on Classroom Testing, No. 3, 1954.
- Tyler, Ralph W. Basic Principles of Curriculum and Instruction. Syllabus for Education 305. Chicago: University of Chicago Press, 1950.
- United States Department of Health, Education, and Welfare, Office of Education. Understanding Testing Purposes and Interpretations for Pupil Development. Washington: U. S. Government Printing Office, 1960.
- Walker, Helen M. and Joseph Lev. Statistical Inference. New York: Henry Holt and Company, 1953.

APPENDIX A

## YOU ARE THERE

The situations described in the test are like those which some teachers have met. After each situation, there is a list of statements describing the choices which might be available to the teacher in meeting the situation. You are asked to assume the role of the teacher and make decisions on the basis of your own beliefs as if "You Are There." The way you really would react is the "right" answer for you.

## Instructions:

1. There are 15 situations and 101 items in this test. Respond to every item on the printed answer sheet.
2. Be frank and honest. Avoid a reaction that you might make to an item just because you think you ought to. This would defeat the purpose of the test.
3. Consider each item by itself, forgetting about your response to other items. You do not have to be consistent.
4. In situations where you cannot base your response on experience, answer the way you think you would react.
5. Blacken only one answer space for each item, the one closest to the way you feel.

## CODE TO RESPONSES

	A	B	C	D	E
If you <u>Strongly Agree</u> , blacken space A.....	■				
If you <u>Agree in General</u> , blacken space B.....		■			
If you are <u>Uncertain</u> , blacken space C.....			■		
If you <u>Disagree in General</u> , blacken space D..				■	
If you <u>Strongly Disagree</u> , blacken space E....					■

11. The pupils seem to depend on the opinions of several leaders in the class.
12. "I didn't know Suzi had so many good ideas till she worked in a group with me."
13. "Nancy seems to be awfully prejudiced, but I can understand why after what she said today in class."
14. "I'm glad we don't have to listen to lectures in this class very often. I'm more interested in working out our own solutions to the problems we discuss."

III. One day as your Homemaking II class is dismissed after taking a written test covering the entire foods unit, you over-hear a pupil say, "Why does Mrs. \_\_\_\_\_ always give true-false tests? Doesn't she know how to do anything else?" Rate each of the following ways a teacher might react in such a situation.

15. Notice which pupil made the comment and lower her grade on the test.
16. Plan a way to secure further anonymous suggestions from pupils about methods of evaluating their progress.
17. Check with all other teachers to see if the critical pupil makes such comments in their classes.
18. Read for more information on how to measure pupil progress and revise your tests applying new methods and techniques.
19. Make your next test of completion items.
20. Omit true-false items on future tests.
21. Ask students to design likely test items and draw ideas from these as you design a test.
22. Ignore such remarks made by pupils.
23. Compare your testing techniques with those of other teachers to determine the effectiveness of your tests.
24. Consider taking a course in evaluation when you next return to college for a summer session.
25. Consider having pupils think about things that they do in Homemaking II which are helpful to them and give suggestions for improvements of their work in class.

CODE TO RESPONSES

- A--Strongly Agree      B--Agree in General
- C--Uncertain        D--Disagree in General
- E--Strongly Disagree

I. In a second year homemaking class, you are preparing for a unit in clothing construction. Each of your pupils has made a cotton skirt and blouse, and each one is now planning to construct a simple dress. As their teacher you are responsible for seeing that each pupil makes reasonable progress during this unit. How would you rate each of the following ways of measuring this progress?

1. Give written quizzes and tests throughout the unit, one as each step in construction is accomplished.
2. Give some paper and pencil tests with the items constructed to measure their understanding of the techniques and principles being taught.
3. Guide the pupils in keeping personal progress records throughout the unit, to be used with the final evaluation of the garment.
4. Pause for discussions of progress in relation to the original goals and objectives during the unit of work, discussions both with individuals and with the class as a group.
5. Set up a class progress chart listing each pupil's name to be displayed and used throughout the unit.
6. Allow pupils to share in planning the evaluation devices to be used to measure their progress, including the final score card.
7. Guide pupils in outlining their plans of work and in summarizing their progress regularly. Read their reports carefully.

II. You have been trying to develop a democratic atmosphere in your classroom during the first semester of the school year in your Homemaking I class. As you evaluate the success of your efforts, rate the following observations and comments according to the extent to which you agree or disagree with their desirability.

8. The pupils sometimes disagree with you and with each other, and offer suggestions freely.
9. Leadership in the class and in small groups often changes from one pupil to another.
10. "We seem to be working each other harder than if Miss \_\_\_\_\_ were making all the assignments."

11. The pu  
the cl
12. "I did  
group
13. "Nancy  
after
14. "I'm g  
often  
the p

III. One day as you  
covering the er  
al  
thing else?"  
a situation.

15. Notic  
test.
16. Plan  
about
17. Check  
makes
18. Read  
revis
19. Make
20. Omit
21. Ask s  
these
22. Ignor
23. Comp  
to d
24. Cons  
to c
25. Cons  
Home  
for

11. The pupils seem to depend on the opinions of several leaders in the class.
12. "I didn't know Suzi had so many good ideas till she worked in a group with me."
13. "Nancy seems to be awfully prejudiced, but I can understand why after what she said today in class."
14. "I'm glad we don't have to listen to lectures in this class very often. I'm more interested in working out our own solutions to the problems we discuss."

III. One day as your Homemaking II class is dismissed after taking a written test covering the entire foods unit, you over-hear a pupil say, "Why does Mrs. \_\_\_\_\_ always give true-false tests? Doesn't she know how to do anything else?" Rate each of the following ways a teacher might react in such a situation.

15. Notice which pupil made the comment and lower her grade on the test.
16. Plan a way to secure further anonymous suggestions from pupils about methods of evaluating their progress.
17. Check with all other teachers to see if the critical pupil makes such comments in their classes.
18. Read for more information on how to measure pupil progress and revise your tests applying new methods and techniques.
19. Make your next test of completion items.
20. Omit true-false items on future tests.
21. Ask students to design likely test items and draw ideas from these as you design a test.
22. Ignore such remarks made by pupils.
23. Compare your testing techniques with those of other teachers to determine the effectiveness of your tests.
24. Consider taking a course in evaluation when you next return to college for a summer session.
25. Consider having pupils think about things that they do in Homemaking II which are helpful to them and give suggestions for improvements of their work in class.

CODE TO RESPONSES

- A--Strongly Agree      B--Agree in General
- C--Uncertain          D--Disagree in General
- E--Strongly Disagree

IV. During a class discussion in your Homemaking II class, several pupils disagreed with a statement in one of their textbooks which one member of the class quoted as a correct answer. As the teacher, rate the following examples of possible reactions.

- 26. Encourage the pupils to accept the statement since it is stated by an authority in print.
- 27. Encourage the pupils to see if further information is available before deciding to accept or reject the statement.
- 28. Let the matter go, since there will always be differences of opinion anyway.
- 29. Compliment the pupils who disagreed for being so observant.
- 30. Reprimand the pupils who disagree to discourage any further class interruptions of this nature.
- 31. Tell the class you will check for further available information so they won't have to bother wasting class time.

V. One member of your Homemaking I class opposes every suggestion made by other members of the class, especially on the days when goals are outlined for a unit. You have noticed that she refuses to cooperate in the activities the group undertakes. However, she hands in all assignments on time and does excellent work on an individual basis. How would you rate each of the following ways her teacher could handle the situation?

- 32. Let her alone for fear of antagonizing her further.
- 33. Tell her that an essential part of her grade depends upon cooperation with other members of the class, and she will have to cooperate.
- 34. Have the whole class plan and carry out an activity in small groups, assigning groups at random and using a class-developed set of criteria for evaluating ways of working together.
- 35. In early experiences in small groups allow her to work with the few class members she prefers as companions.
- 36. Give her more individual tasks than the other members of the class.

37. Encourage activities that interest their pupils.

38. Allow pupils to point out their own mistakes.

VI. Your Homemaking I class is so noisy that one pupil, who is a very good student, is leaving class because she can't see anything. How would you handle this situation? Do not use the following approach:

39. Avoid discussing the problem too emotionally.

40. Encourage the pupils to view the situation from their own point of view.

41. Follow up with the pupils who are leaving class, then re-examine the situation.

42. Include the pupils in the interpretation of the situation.

43. Invite the pupils to direct the discussion.

44. Arrange for a class discussion perhaps with a class representative.

VII. During several conferences you have covered that many of the pupils are shocked at things that are only old-fashioned. How would you handle this problem for some of the pupils? Do not use the following goals and objectives in these conferences: objectives that you do not emphasize the importance in checking the work.

- 94
37. Encourage her to work with someone she prefers on an out-of-class activity related to the unit being studied, and jointly present their findings to the class as a whole.
  38. Allow her to be your assistant during a class demonstration, and point out the ways she is helping you, with sincere comments.

pils dis-  
r of the  
ing ex-

stated

ailable

s of

at.

ner

mation

de by other  
ned for a  
ivities the  
and does  
f the

n  
l

ss-

VI. Your Homemaking I class is studying a unit on child care. You have noticed that one pupil, named Phyllis, does not like small children. As your pupils are leaving class one day, you hear Phyllis comment to another pupil, "I don't see anything so wonderful about having a little sister. Mine is a pest, and my folks always take up for her whenever we argue." As you continue to develop your unit of work, consider the worth of each of the following approaches.

39. Avoid classroom discussions to prevent pupils from becoming too emotionally involved with the remarks being made.
40. Encourage pupils to act out their feelings and points of view in skits and role-playing situations.
41. Following the role-playing situations in class, tell the pupils how they should feel about little brothers and sisters; then role-play the situations again to see if pupils practice these feelings, especially Phyllis.
42. Include test items having some problem-solving situations for interpreting the behavior of the persons involved.
43. Invite some parents to bring small children to class for a directed experience with your pupils.
44. Arrange for the class to observe young children at play, perhaps in a nursery or kindergarten school, and plan for a class discussion of their observations the following day.

VII. During several conferences with your first year homemaking pupils you discover that many of them have one or more of their grandparents living with them. Sometimes these pupils complain to you that their grandparents are shocked at things they do and that ideas expressed by the grandparents are only old-fashioned. These differences in points of view have become a real problem for some of their families. As you guide your class in setting the goals and objectives for a unit on family relations, you are reminded of these conferences and pupil comments. Rate each of the following teacher objectives that you considered in planning this unit with respect to their importance in changing pupil attitudes toward older people.

CODE TO RESPONSES

A--Strongly Agree      B--Agree in General  
C--Uncertain          D--Disagree in General  
E--Strongly Disagree

45. Develop desirable attitudes toward elderly members of our family and society in general.
46. Increase our interest and understandings of the needs of older people.
47. Increase our ability to understand and remember what authorities say about problems in three-generation families.
48. Develop a greater willingness to compromise in areas of conflict with grandparents.

Rate each of the following ways to achieve the most important of the above objectives:

49. Invite elderly persons to visit your classroom and contribute their opinions on well-chosen topics.
50. Have class members interview grandparents or elderly people on some well-chosen topic.
51. Provide classroom opportunities for pupils to act out their feelings and points of view in mock-family situations using skits and role-playing techniques.

VIII. In your ninth-grade homemaking class, you have two pupils from the total of eighteen who are considered slow learners. Determine the advisability of each of the following ways of handling the homework assignments throughout the year under these circumstances.

52. Don't give the slow learners written assignments.
53. Give the slow learners realistic activities to carry out at home while the other pupils have written assignments.
54. Give all pupils realistic activities to do at home, as well as written homework assignments.
55. Allow some choice in the amount of homework to be turned in by each pupil, while encouraging them to do whatever they are capable of doing; allow oral as well as written reports of these assignments.

IX. Your Homemaking I...  
You are concerned achieving. She has started this year best friend, Gail is to play dumb so and doesn't seem to might help Linda.

56. Assign I...  
for clas
57. During a...  
reasons
58. Invite e...  
to discu
59. Invite e...  
discussi  
by Today
60. Advise I...  
for good
61. Have a...  
on how t
62. Suggest...  
by the b...  
traits o...  
to posse

X. One of the units...  
housing and you ha...  
might include. Me...  
the following clas

63. Followin...  
this uni
64. Have pup...  
latest s...  
and pres
65. Bring a...  
to compi
66. Use clas...  
arrangen...  
in her o
67. Spend pe...  
accessor...  
lamps, e...  
out the

IX. Your Homemaking III class is in the second week of a unit on family relations. You are concerned about Linda, one of the juniors, because she is under-achieving. She had an A-average during her freshman and sophomore years, and started this year making B's and C's on several quizzes. You overheard her best friend, Gail, telling her that, "The only way you'll ever interest Freddy is to play dumb so he'll think he's smarter." Linda is average in appearance, and doesn't seem to have a lot of poise. React to the following ways you might help Linda.

56. Assign Linda and Gail to separate groups whenever possible for class activities.
57. During a conference with Linda, seek to discover other possible reasons for her lowered grades.
58. Invite a panel of high school pupils, including boys and girls, to discuss the topic, "The Qualities I Admire Most in Others."
59. Invite several successful homemakers to participate in a panel discussion on the topic, "What Characteristics Are Needed Most by Today's Homemakers."
60. Advise Linda to choose her friends with more care and to try for good grades no matter what boys think.
61. Have a "Glamour Clinic" to give the pupils some practical help on how to make themselves more attractive.
62. Suggest that Gail and Linda do a survey of the opinions held by the high school boys in response to the question, "What traits do you think are most important for a high school girl to possess?"

X. One of the units to be included in your second semester of Homemaking II is housing and you have been planning some of the classroom experiences you might include. Most of your pupils come from low-middle class homes. Rate the following classroom experiences which you find are possible.

63. Following group and individual conferences at the beginning of this unit, allow the pupils to help plan the problems to be studied.
64. Have pupils read texts, pamphlets, and other references for the latest styles in design of furnishings and in home decoration, and present information in forms of reports, both oral and written.
65. Bring a stack of your old House Beautifuls which pupils may clip to compile a booklet of ideas for their home of the future.
66. Use classroom kits of scaled models of furnishings to plan room arrangements and have each pupil apply these principles to a room in her own home, sharing results during a later class period.
67. Spend part of the time allowed for the unit helping pupils to plan accessories for their bedroom such as pillows, bulletin boards, lamps, curtains, or picture arrangements. Encourage them to carry out their plans as home experiences.

CODE TO RESPONSES

- A--Strongly Agree      B--Agree in General
- C--Uncertain         D--Disagree in General
- E--Strongly Disagree

XI. As a learning experience for a unit on child development your Homemaking II class would like to have a play school for one or two mornings. This is a new idea to your pupils, and to the parents in the community who would be asked to bring their pre-school aged children. Rate each of the following approaches in planning for this classroom experience.

- 68. Prepare a written outline for the pupils of the responsibilities they would assume and divide the class into groups according to their interests in these responsibilities.
- 69. Help the class decide which committees are needed.
- 70. Allow the class to decide what things are important in committee work and grade girls on the basis of these criteria.
- 71. Prepare a bibliography for the various problems to be solved in relation to the nursery school experience, citing references in pamphlets, magazines, books, and texts. Have as many of these references on hand as possible.
- 72. Work with each committee, helping them only when you see that they need a clue about how to proceed or when they are settling on an inferior plan.
- 73. Plan the program yourself and delegate the responsibilities to your pupils.
- 74. Leave the planning to your pupils, offering suggestions only when asked to do so.
- 75. Ask class members to choose the chairman for each committee---planning, arrangements, evaluating, etc.---considering abilities of class members.

XII. Your Homemaking I class is studying the various customs and beliefs of different cultures in a unit of family relations. Only one girl in your class, named Ericka, is of a different culture. She has not been accepted too well by other class members. As Ericka's teacher, how would you rate each of the following examples of possible reactions to this situation?

- 76. Ignore the fact that Ericka is of a different culture.
- 77. Avoid drawing unnecessary attention to Ericka during this unit.
- 78. Discuss with the class that there are differences in cultures, and that each of us has to learn to accept ways in which we are different.

- 79. Prepa  
activ
- 80. Take  
and h
- 81. Discu  
to se  
butio
- XIII. Yesterday your  
olds playing in  
the second week  
period they saw  
demolish it, an  
for this behavi  
you choose from  
in class?
- 82. Give  
four-
- 83. Begin  
of ho
- 84. Begin  
in th  
allow  
own c
- 85. Have  
situ  
hand  
scho
- 86. Assur  
with
- 87. Assur  
with  
back
- 88. Lead  
beha  
unde

79. Prepare for this unit by letting Ericka assist in classroom activities which bring recognition for her good points.
80. Take advantage of Ericka's difference in cultural background and help her interpret her culture to the class.
81. Discuss with the class differences in cultures and help them to see that members of every culture have important contributions to make to our American way of life.

XIII. Yesterday your Homemaking II class observed a group of fifteen four-year-olds playing in the nursery school over a period of 30 minutes. This is the second week of study in their child care unit. During the observation period they saw Tommy ride his tricycle into Bobby's block-castle and demolish it, and then bump Bobby with the tricycle. Tommy made no apology for this behavior, but rode off immediately on the tricycle. How would you choose from among the following ways of using this observation today in class?

82. Give each pupil a list of basic understandings concerning four-year-olds as a basis for discussing their observations.
83. Begin the class period by commenting on your own evaluation of how the nursery school teacher handled the above situation.
84. Begin the class with a period of supervised study of a chapter in the text related to behavior of pre-school children; then allow part of the period to be used for discussion of their own observations.
85. Have several groups of class members role-play the above situation, each group showing how they think they might handle the situation should it occur in their own play school.
86. Assuming that your class did the role-playing, follow this with a discussion of how small children develop habits.
87. Assuming that your class did the role-playing, follow this with a discussion of how some things in the environment and background of a child may cause him to be aggressive.
88. Lead your class to say what they believe about the "whys" of behavior in small children (to state some principles about understanding children).

CODE TO RESPONSES

A--Strongly Agree      B--Agree in General  
C--Uncertain          D--Disagree in General  
E--Strongly Disagree

XIV. You have observed that several of your classes seem to be noisy and restless, especially your Homemaking II class. Several pupils have neglected their homework assignments and several others are beginning to waste class time with "little" conversations. Rate each of the following ways you could react to this situation.

89. Provide opportunities for pupils to offer suggestions for class activities.
90. Use more discipline during the class and set up higher standards of conduct for the pupils.
91. Have more individual and small group conferences to determine the needs and interests of the pupils.
92. Vary your teaching methods whenever possible.
93. Consider taking an advanced course in the area of the teaching of home economics when you return for summer school work.
94. Ask your area home economics supervisor to arrange a visit to help you analyze the situation and offer suggestions to you.
95. Put up with the situation since all pupils are restless and somewhat lacking in interest in school work.
96. Recognize that since you are conscious of your problem, you are growing as a teacher.

XV. From a diagnostic survey of sewing experience you learn that 18 of the 24 pupils in your Homemaking I class have had almost no experience with machine sewing. There are two pupils who have constructed several garments, including a wool skirt. As you develop tentative classroom experiences for a unit in clothing, rate each of the following ways of applying the findings from your survey.

97. Ask your more experienced pupils to assist you with construction demonstrations and to help other pupils with their problems.
98. Praise the experienced pupils more often as examples for the other members of the class to follow.

99. Schedule purpose each pu eventua
100. Allow pattern sewers
101. Allow range during

99. Schedule individual and small group conferences for the purpose of giving more individual guidance and evaluating each pupil's progress, being sure to include all pupils eventually.
100. Allow the more experienced pupils to select their own patterns, but choose a basic pattern for the beginning sewers.
101. Allow each pupil to choose her pattern from a limited range of patterns which you have previewed with them during class.

T H A N K Y O U !

We appreciate your cooperation in the  
development of this test.

SCHOOL OF HOME ECONOMICS

## YOU ARE THERE

Yes, you are a home economics teacher right now. For this reason you are in the best position to let us know how teachers on the job respond to the enclosed test.

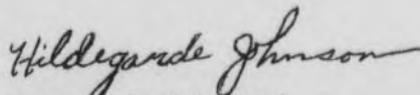
The test has been designed to measure a student's ability to apply the educational principles taught in her methods courses. It is still in the process of development and is the subject of research for my graduate degree in home economics education. Many steps are involved in developing such a test and we can do no more until we have your reactions to what has been done so far.

How can you help? By giving your honest and frank reaction to each item of the enclosed test as you assume the role of the teacher. Your experience in teaching home economics will be valuable in helping us to check the quality of each item and in securing a measure of reliability. We hope that you will send it back without your name on the answer sheet. However, we need to know which teachers respond early so we will not bother them with a reminder. For this reason a post card is enclosed.

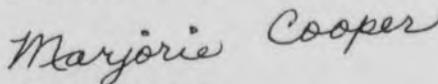
How soon? Probably you do not have the time today, for your plans have been made to include other important activities. We ask that you look over your calendar and select a convenient time within the next ten days when you can give the test your undivided attention.

What do we want back? First, the answer sheet, folded and placed in the enclosed envelope which has been self-addressed and stamped. Second, the enclosed post card letting us know you have mailed the answer sheet.

Last request: May we ask again that you choose your time and place for this professional opportunity within the next ten days and that you keep the date? Thank you. Good-bye until then.



Hildegarde Johnson  
Professor of Home Economics Education



Marjorie Cooper  
Graduate Assistant

## POSTCARD

Dear \_\_\_\_\_,

When you received our letter and an enclosed test a few days ago, you probably had your schedule well filled with end of the year activities. Let us reassure you that your responses to the test are important. Again we urge you to take time to give us the benefit of your experience as you respond to each item. We want only the unsigned answer sheet returned, and the card notifying us of your cooperation. Thank you again for taking time to pretend that "You Are There."

Hildegarde Johnson,  
Professor of Home Economics Education  
Marjorie Cooper,  
Graduate Assistant

5-22-61

ORIGINAL RECORDS ACT FOR

THE COURT

APPENDIX B

Education

Page 200 of 201

ORIGINAL SCORING KEY FOR

"YOU ARE THERE"

Total Score = 151

NAME Scoring Key Used in This Study DATE DATE OF BIRTH AGE SEX

SCHOOL CITY GRADE OR CLASS INSTRUCTOR PART

1 2 3 4 5 6 7 8

1 A B C D E 2 A B C D E 3 A B C D E 4 A B C D E 5 A B C D E 6 A B C D E 7 A B C D E 8 A B C D E 9 A B C D E 10 A B C D E 11 A B C D E 12 A B C D E 13 A B C D E 14 A B C D E 15 A B C D E

31 A B C D E 32 A B C D E 33 A B C D E 34 A B C D E 35 A B C D E 36 A B C D E 37 A B C D E 38 A B C D E 39 A B C D E 40 A B C D E 41 A B C D E 42 A B C D E 43 A B C D E 44 A B C D E 45 A B C D E

61 A B C D E 62 A B C D E 63 A B C D E 64 A B C D E 65 A B C D E 66 A B C D E 67 A B C D E 68 A B C D E 69 A B C D E 70 A B C D E 71 A B C D E 72 A B C D E 73 A B C D E 74 A B C D E 75 A B C D E

91 A B C D E 92 A B C D E 93 A B C D E 94 A B C D E 95 A B C D E 96 A B C D E 97 A B C D E 98 A B C D E 99 A B C D E 100 A B C D E 101 A B C D E 102 A B C D E 103 A B C D E 104 A B C D E 105 A B C D E

121 A B C D E 122 A B C D E 123 A B C D E 124 A B C D E 125 A B C D E 126 A B C D E 127 A B C D E 128 A B C D E 129 A B C D E 130 A B C D E 131 A B C D E 132 A B C D E 133 A B C D E 134 A B C D E 135 A B C D E

BE SURE YOUR MARKS ARE HEAVY AND BLACK. ERASE COMPLETELY ANY ANSWER YOU WISH TO CHANGE.

16 A B C D E 17 A B C D E 18 A B C D E 19 A B C D E 20 A B C D E 21 A B C D E 22 A B C D E 23 A B C D E 24 A B C D E 25 A B C D E 26 A B C D E 27 A B C D E 28 A B C D E 29 A B C D E 30 A B C D E

46 A B C D E 47 A B C D E 48 A B C D E 49 A B C D E 50 A B C D E 51 A B C D E 52 A B C D E 53 A B C D E 54 A B C D E 55 A B C D E 56 A B C D E 57 A B C D E 58 A B C D E 59 A B C D E 60 A B C D E

76 A B C D E 77 A B C D E 78 A B C D E 79 A B C D E 80 A B C D E 81 A B C D E 82 A B C D E 83 A B C D E 84 A B C D E 85 A B C D E 86 A B C D E 87 A B C D E 88 A B C D E 89 A B C D E 90 A B C D E

106 A B C D E 107 A B C D E 108 A B C D E 109 A B C D E 110 A B C D E 111 A B C D E 112 A B C D E 113 A B C D E 114 A B C D E 115 A B C D E 116 A B C D E 117 A B C D E 118 A B C D E 119 A B C D E 120 A B C D E

136 A B C D E 137 A B C D E 138 A B C D E 139 A B C D E 140 A B C D E 141 A B C D E 142 A B C D E 143 A B C D E 144 A B C D E 145 A B C D E 146 A B C D E 147 A B C D E 148 A B C D E 149 A B C D E 150 A B C D E

EXPLANATION

115 Weight response at this level one. 147 Weight response at this level two. 149

SCORING KEY FOR

FORM A

BE SURE YOU HAVE MARKED ALL THE ANSWERS  
ON THE ANSWER SHEET. ONLY ANSWERS FROM THE ANSWER SHEET WILL BE GRADED.

Total Score = 119

NAME Revised Form A LAST FIRST MIDDLE DATE DATE OF BIRTH GRADE OR CLASS INSTRUCTOR AGE SEX M OR F PART

1	A B C D E	1	A B C D E	61	Omit	91	2	A B C D E	121	A B C D E
2	Omit	32	Omit	62		92	2	A B C D E	122	A B C D E
3		33	Omit	63		93	2	A B C D E	123	A B C D E
4	2	34	Omit	64		94	2	A B C D E	124	A B C D E
5		35	Omit	65		95	2	A B C D E	125	A B C D E
6	2	36	Omit	66	Omit	96	2	A B C D E	126	A B C D E
7	2	37	Omit	67		97	Omit	A B C D E	127	A B C D E
8	2	38	Omit	68		98	Omit	A B C D E	128	A B C D E
9	2	39		69	Omit	99	Omit	A B C D E	129	A B C D E
10	2	40	2	70		100	Omit	A B C D E	130	A B C D E
11		41		71	2	101	Omit	A B C D E	131	A B C D E
12	2	42		72		102		A B C D E	132	A B C D E
13	2	43		73		103	2	A B C D E	133	A B C D E
14	2	44		74		104	2	A B C D E	134	A B C D E
15		45	2	75		105		A B C D E	135	A B C D E

BE SURE YOUR MARKS ARE HEAVY AND BLACK. ERASE COMPLETELY ANY ANSWER YOU WISH TO CHANGE.

SCHOOL CITY STATE ZIP CODE

16	Omit	46	2	76		106		A B C D E	136	A B C D E
17		47		77	Omit	107		A B C D E	137	A B C D E
18	2	48		78	2	108		A B C D E	138	A B C D E
19	Omit	49	2	79	2	109		A B C D E	139	A B C D E
20	Omit	50		80		110		A B C D E	140	A B C D E
21	Omit	51	2	81	2	111		A B C D E	141	A B C D E
22		52		82		112	2	A B C D E	142	A B C D E
23		53		83		113		A B C D E	143	A B C D E
24	2	54		84		114		A B C D E	144	A B C D E
25	Omit	55	2	85	2	115		A B C D E	145	A B C D E
26		56	Omit	86	2	116		A B C D E	146	A B C D E
27	2	57	2	87	2	117		A B C D E	147	A B C D E
28		58	2	88	2	118		A B C D E	148	A B C D E
29	Omit	59		89	2	119		A B C D E	149	A B C D E
30		60	Omit	90		120		A B C D E	150	A B C D E

SCORING KEY FOR

FORM B

Total Score = 154

5 DATE OF BIRTH AGE SEX

6 GRADE OR CLASS INSTRUCTOR

7 NAME OF TEST PART

	A	B	C	D	E		A	B	C	D	E		A	B	C	D	E		A	B	C	D	E		
1						31						61							91	2				121	
20	mit					320	mit					62								92	2				122
31						33						63							93	2				123	
42						34	2					64							94	2				124	
51						35	2					65				2			95			2		125	
62						360	mit					66	2							96	2				126
72						37	2					67							97	2				127	
82						38	2					68			2				98			2		128	
92						39						69	2						99	2				129	
10	2					40	2					70							100					130	
11						41	2					71	2						101	2				131	
12	2					42						72							102					132	
13	2					43						73				2			103					133	
14	2					44						74				2			104					134	
15				2		45	2					75							105					135	

BE SURE YOUR MARKS ARE HEAVY AND BLACK. ERASE COMPLETELY ANY ANSWER YOU WISH TO CHANGE.

1 NAME Revised Form B LAST FIRST MIDDLE DATE CITY SCHOOL

2

3

	A	B	C	D	E		A	B	C	D	E		A	B	C	D	E		A	B	C	D	E			
16						46	2					76							106					136		
17						47						77	Omit					107								137
18	2					48						78	2						108					138		
19	Omit					49	2					79	2						109					139		
20	Omit					50						80	2						110					140		
21						51						81	2						111					141		
22						52	2					82			2				112					142		
23				2		53						83							113					143		
24	2					54						84							114					144		
25	2					55	2					85	2						115	EXPLANATION				145		
26						56	Omit					86	2						116		Weight response at this level one.			146	147	
27	2					57	2					87	2						117					148		
28				2		58	2					88	2						118	2	Weight response at this level two.			148	149	
29						59	2					89	2						119					150		
30				2		60	Omit					90								120						