REVISION OF A THEORY PRETEST BASED ON OBJECTIVES AND GENERALIZATIONS FOR A BEGINNING COLLEGE FOOD PREPARATION COURSE

By

CHARLINE HYER WHITE

Bachelor of Science

Oklahoma State University

Stillwater, Oklahoma

1942

Submitted to the faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the degree of
MASTER OF SCIENCE
May, 1968

REVISION OF A THEORY PRETEST BASED ON OBJECTIVES AND GENERALIZATIONS FOR A BEGINNING COLLEGE FOOD PREPARATION COURSE

Thesis Approved:

Helen F. Barbour
Thesis Adviser

Hazel J. Baker

N Durham
Dean of the Graduate College

ACKNOWLEDGEMENTS

The author wishes to express her sincere gratitude to her major adviser, Dr. Helen F. Barbour, for her untiring efforts in counseling and guidance throughout this study. A special note of appreciation is extended also to Miss Hazel J. Baker, Assistant Professor of Food, Nutrition and Institution Administration, and to Dr. James B. Mickle, Professor of Dairy Science, who served on the graduate committee of the author.

Dr. Harry Brobst is gratefully acknowledged for consultation regarding the sample audience and item-analysis procedure. To Miss Elizabeth Sanders the author is greatly indebted for a complete proof reading of the entire manuscript.

Loving appreciation goes to the author's son who set up the data for use on the computer which saved many hours of hand calculation, and to her husband, Ralph W. White, and daughter Lucy Mae White, for their constant encouragement and assistance during the entire time this study was being done.

Finally the author is grateful to General Mills, Incorporated, for granting permission to reproduce questions from the Betty Crocker Search for the Homemaker of Tomorrow Tests, 1959-1967.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

Current trends in home economics indicated a need for a re-evaluation of teaching practices in both high school and college home economics classes. Stovall (26, p. 537) stated that ". . . instruction must be that which will help pupils 'to learn how to learn' and thus become independent learners."

Because of the increase in population of college students and the desire to enroll the student in accordance with his level of proficiency, there was an increased need to section students according to their level of achievement, to determine their ability to think critically, to plan course content directed toward student needs, to permit acceleration among the more capable students, to offer individual guidance, and to motivate students to increase present knowledge.

The need was urgent for the development of methods by which students could be placed in classes where they could add to their present knowledge and skills in homemaking. In the present day situation there could be a great amount of unnecessary repetition of previous learning at the college level. In like manner there were many students who suffered from having to compete with others who had had previous experience, thus making it difficult for them to achieve to the degree of which they were capable. Therefore, it was feasible to prepare a food pretest for college entrance that could assist the faculty in evaluating the

aptitudes of students as they were about to enter the university. Such a pretest could also serve in planning a curriculum which would satisfy the needs of the majority.

Early in the development of a National Testing Program by members of the American Home Economics Association, Bonde (5) reported the following factors for consideration: (1) determine students' background, potentiality, and level of comprehension; (2) distinguish between effective and ineffective teaching through evaluation of methods, materials, and personality of teacher; (3) attract high quality students; and (4) aid in placing transfer students. Tests were developed, published and made available for purchase by the American Home Economics Association. However, due to lack of funds, limited use was made of these tests which resulted in discontinuation of them.

Attempts at pretesting, before enrollment in college clothing courses, were made by Hoskins (21) at New Mexico State University and Collins (7) at Southern Illinois University in 1956. Witt (30) also developed a clothing pretest at Oklahoma State University in 1961. In 1962, Steelman (25) constructed a pretest in food preparation at Oklahoma State University and in 1963 another food pretest was developed by Lee (23) at Southern Illinois University. Revision of part of Witt's clothing pretest, at Oklahoma State University, was carried out by Gould (17) in 1963. Cooksey (9) perfected a pretest in introductory nutrition in 1964 at Oklahoma State University.

In general the purpose of these pretests was for exempting, sectioning, motivating, guiding, and developing curriculum. A pretest that had been completely validated and checked for reliability might permit accelerated students to omit the basic course and proceed to a higher

level of study. On the other hand a worthy pretest could assist in placing students in homogeneous groups where instruction could be geared to a level at which they could be challenged and succeed. A pretest could offer an opportunity to preview a course and thus stimulate interest in that direction or tend to alert students to important areas before study was begun. Individual scores on a pretest could serve as a guide to strengths and weaknesses on which the student might profit by counseling. Finally, from an over-all view of the results of a pretest, the instructor would be able to plan his curriculum to adequately meet the needs of the majority of his students.

The Fifth Mental Measurements Yearbook, 1959 (6) contained a review of standardized tests in home economics and other areas. Johnson's Home Economics Interest Inventory was among them. It was designed to help women in home economics refine their choice within the area. Criticism of Johnson's Home Economics Interest Inventory included: complication of scoring; need for cross validation; and lack of sufficient number in the sample. This inventory was considered good. However, recommendations were made for further research involving it.

The pretests of Steelman (25) and Cooksey (9) also needed further refinement before they could be used effectively. The author proposed to revise and further validate the theory pretest developed by Steelman for a beginning college food preparation course. The need for such a program of pretesting existed at Oklahoma State University. Since the writer was most interested in principles of food preparation, ability to recognize standard products and equipment needed in their preparation, only that phase of the program was considered in the present study.

The author accepts the following assumptions: (1) students can perform more efficiently if allowed to proceed on the basis of their

previous knowledge, (2) an evaluation instrument can be devised which will measure the degree of performance in a course, (3) students will be motivated if introduced to course content preceding the first course in food preparation, (4) a pretest can be used as a basis for guidance and counseling of students, and (5) course content can be adjusted to the needs of the students based on findings from a pretest.

It is hypothesized that Steelman's (25) pretest can be improved as follows:

1. Reword questions not having desirable difficulty and discrimination indices as shown by item analysis.

2. Add 50 multiple-choice type questions to the test in an effort to increase validity and reliability.

3. Administer pretest to selected group of college undergraduates to determine, by item-analysis, which questions should be retained for use in the refined pretest.

Clarification of Terms

Basic concepts of nutrition developed by a Cornell University workshop group were accepted by the author as fundamental to food preparation and service. These basic ideas are as follows (19, p. 20):

1. Nutrition is the food you eat and how the body uses it.

2. Food is made up of different nutrients needed for growth and health.

3. All persons, throughout life, have need for the same nutrients, but in varying amounts.

4. The way food is handled influences the amount of nutrients in foods, its safety, appearance, and taste.

Since validity and reliability are crucial in the development of an evaluation instrument, a definition of terms is pertinent. Validity is defined by Ahmann and Glock (1, p. 292) as: " . . . the degree to which an evaluation instrument actually serves the purposes for which it is intended." Arny (2, p. 97) stated: "The reliability of a test indicates its accuracy of measurement and self consistency." Reliability and validity are two very different concepts which are frequently confused.

CHAPTER II

REVIEW OF LITERATURE

One of the major goals of the educational program is to foster a change in the behavioral pattern of the learner. Such a change necessitates that specific objectives be identified. Since objectives are based upon the philosophy of the school, pertinent information and knowledge are necessary for objectives to be meaningful.

Guidelines for Statement of Objectives

Hall and Paolucci (18) have identified three basic consecutive steps which they state are essential to the teaching of home economics. These steps included: (1) determining and stating objectives, (2) providing experiences that make it possible for attaining these objectives, and (3) seeking reliable evidence in relation to the achievement of the objectives.

Change in behavioral patterns can be represented by educational objectives. Tyler (29) believed education to be a process by which the behavioral patterns of people are changed. Hall and Paolucci stated that (18, p. 151):

> If learning is changed behavior, then the primary purpose of teaching becomes motivating and guiding learners to change their behavior. Home economics education objectives, then, in order to give a sense of direction to teaching, need to be defined so that they identify precisely the kind of change that is being developed.

The gap ascertained from a comparison of the present status of the learner and the acceptable norm is generally referred to as need.

The efforts of the school should be concentrated upon the more serious gaps in the present development of students. Needs are established in many ways both within and without the school. Community surveys, questionnaires, home visits, parent days at school, and others can be used. By comparing the gaps located in the development of the student with some set of desirable norms, suggestions can then be made for educational objectives.

Not all needs of students can be met by educational means. Careful consideration should be given when formulating educational objectives. Arny indicated that students profit by discovering for themselves, in advance, their weaknesses. In this manner they are more likely to try to correct them. Arny (2, p. 13) stated further that teachers, in the past, tested only for knowledge or facts, however the more recent concept is that: ". . . . instruction and evaluation should be focused upon goals, rather than upon concepts; upon the use people will make of subject matter rather than upon the subject matter itself."

The number of objectives should be kept to a minimum such that they may actually be accomplished in the time available. They should be carefully coordinated to avoid confusion in the mind and behavior of the student. According to Tyler (29, p. 24), "Educational objectives are educational ends, they are results to be achieved from learning." Consideration must be given the relationship between a learning experience and the age level at which it is contemplated. Psychology of learning assists in determining the length of time required to learn certain things, hence, the time required to attain an objective.

Opportunities should be provided for the student to actively participate in the learning process and not just hear about it. Since education is an active process, Tyler (29) indicated that students learn only when the learner puts forth some effort himself. Provision for an opportunity to use knowledge reduces the forgetting of that knowledge. Tyler (29, p. 26) intimated that: " . . . objectives concentrating on specific knowledge are more attainable and the results more permanent when there are opportunities for this knowledge to be used in the daily lives of the student."

Since most learning experiences produce concomitant outcomes, it is important to select objectives that reinforce each other. Many definitions of objectives are available. Tyler (29, p. 38) used the following: "One can define an objective with sufficient clarity if he can describe or illustrate the kind of behavior the student is expected to acquire so that one could recognize such behavior if he saw it." This author indicated that a useful form for stating objectives was to express them in terms which included both the kind of behavior to be developed in the student and the area of life in which the behavior was to operate.

A two dimensional chart was often used to graphically express the objectives concisely and clearly. Such a chart tended to reveal kinds of learning experiences clearly. According to Tyler (29, p. 40):

> By defining those desired educational results as clearly as possible the curriculum-maker has the most useful set of criteria for selecting content, for suggesting learning activities, for deciding on the kind of teaching procedures to follow . . . setting up and formulation of objectives.

Formulation of objectives is greatly affected by the educator's theory of learning.

Judd and Freeman at the University of Chicago formulated a theory of generalizations. This theory advocated that learning can be explained in terms of the learner's development of principles which he might use in solving a problem or in meeting a new situation (29). Other theories of learning are held by different educators. Arny (2) suggested the necessity for knowing the status of the student in order to make worthwhile changes. Pretesting was submitted as a means of improving the situation.

Use of Pretests

Increased emphasis on liberal arts in education has brought about great concern in the field of home economics for courses such as elementary foods in college. In such courses it is mandatory that the teacher plan for student experiences which will illustrate the basic principles to be understood. It is believed that the understanding of basic principles can be detected through the use of pretests.

Hoskins reported studies which indicated pretests were useful in conditioning the students of lower mental ability for learning. Pretests, then, were considered to be a motivation factor. Hoskins (21, p. 2) proposed to construct a pretest which could:

1. Permit those students with experience who rated high on the test to enroll in a more advanced course.
2. Allow more favorable placement of transfer students.
3. Enable students to better realize their strengths and weaknesses and thus increase interest in the course.
4. Assist the instructor in grouping students according to their experience and needs.
5. Aid in individual guidance.

Unsuccessful attempts at sectioning students of varying degrees of skills were made at Southern Illinois University through the use of pretests. According to Collins (7), there was a need for determining: (1) those who had previous experience and were skilled in clothing construction; (2) those who had previous experience and were not skilled; and (3) those who had no previous experience at all.

The purposes of Collins' (7) study were: to ascertain authoritative opinions of pretests and their use; to formulate a clothing pretest which would be used to precede the first course in clothing at Southern Illinois University; and to construct a test which could be readily scored for purposes of sectioning and development of the curriculum. Homogeneous grouping was believed by some to facilitate teaching. Some studies indicated that the amount of previous experience in home economics had little correlation with the performance exhibited in their college courses. Authorities, however, tend to disagree on this matter.

Collins (7) stated that one of the most valid uses of placement tests was for guidance purposes. She concluded that the amount of previous experience a student had affected her performance in clothing construction classes in college. The student's attitude toward the subject of clothing was determined to be a factor in her success. Placement of students through the use of pretests was considered desirable in order that they could achieve to the fullest extent. Students in the higher scholastic ranks in high school tended to perform well also in college home economics classes.

Witt (30), studying at Oklahoma State University, reviewed many types of evaluation of home economics programs in use since 1923. The Murdock Sewing Scale was published at this time. It was an objective-type evaluation device. Other tests were developed to measure skills in

sewing which seemed the easiest to measure. Score cards for standards for finished garments, speed charts that rated progress made on a garment in a given length of time, and informal type testing were among the earliest devices used.

Standardized tests, according to Witt (30), were needed which would determine the extent of the students' knowledge in subject matter, measure their aptitude with manual and intellectual skills, test their ability to reason, and define aesthetic competencies. Only one such standardized test was reported in the Mental Measurements Yearbook published in 1938.

In 1948, interest centered around developing a testing program for home economics at the college level through the American Home Economics Association. This group considered the desirability of developing a test, or series of tests, that would measure student achievement. Such tests were attempted at West Virginia University by Davis, and four attempts were made at Iowa State University before satisfactory results were obtained (30). All were interested in the matter of placement devices in clothing.

According to Steelman (25, p. 1), little effort had been made to ". . . accelerate qualified students in home economics." Placement of students in advanced standing required a systematic evaluation of the competence of all beginning students in food and nutrition. Because they came to college with such varied backgrounds of experience, and because their native abilities were not equal, some system of evaluation seemed necessary.

Steelman (25) considered the construction of a valid, reliable, discriminating, and objective pretest possible for students taking their

first course in food preparation. The nature of the subject matter, however, did not seem to lend itself to ease of construction. This author's test (25, p. 3) was devised to determine the student's knowledge and aptitudes regarding:

1. Ability to apply principles of food preparation.
2. Recognition of accepted procedures in preparation and service of food.
3. Ability to do critical thinking and use judgement in relation to:
   a. Selection of food.
   b. Combinations of food.
   c. Selection of equipment.
   d. Recognition of standard products.
   e. Use of time, energy, and resources.
4. Attitudes toward food.

The purposes for which Steelman's (25) pretests were constructed were: to accelerate superior students by exempting them from the first course in food preparation; to separate the average to slow students in sections where course content could be geared to their level of aptitude; to aid the faculty in curriculum planning by identifying strengths and weaknesses of the students; to assist in guidance of students; and motivate students to increase their knowledge of food preparation.

Another attempt at pretesting in the area of food was carried out by Lee (23). She proposed to develop a means for placing students in elementary food courses according to their ability; to relate grades on placement tests with course grades; to compare the Ohio State Psychological Examination percentile and course grades; to compare non-majors with majors in home economics; and to observe the effect of high school

science grades, high school home economics and food preparation experiences upon the course grade. The purpose of evaluation, according to Lee, was to locate an individual's needs, interests, and abilities. She considered it an appraisal of progress toward the goals of a course.

Gould (17) devised a pretest intended to determine the difference between students with high and low degrees of skill in clothing construction. She reported that interests and attitudes developed regarding an area of study were as important to the success of a course as the academic achievement of the students. Gould cited the need for discovering the students' background and native ability before they enrolled in a course. By dividing students into homogeneous groups the training process could be facilitated in crowded college classes. She expressed the belief that the use of a pretest could encourage the development of the students' potential ability.

Cooksey constructed a pretest in introductory nutrition which was administered at Oklahoma State University. She pointed out that a testing program at the college entrance level increased the functional value of the subject matter. Cooksey (9, p. 4) stated:

> An important characteristic of human beings is their ability to draw conclusions as a result of experiences and to use these to advantage in meeting related situations. While one must admit that the more intelligent the individual the greater the probability that he will generalize and in turn apply his generalization in new situations; nevertheless, it seems apparent that this method of teaching can markedly increase the ability of the student to draw sound conclusions.

Her purposes, like others, were for exemption, sectioning, placement of transfer students, motivation, curriculum planning, and guidance of students.

Methods Used in Developing Pretests

Dissatisfaction because of duplication of previous knowledge in clothing was expressed by college students according to Hoskins (21). Teachers also had difficulty teaching subject matter which would challenge the experienced student. The use of pretests was recommended to the Home Economics Department at New Mexico State University by an evaluation committee. Hoskins proposed to construct a pretest in the clothing area that could be administered with a minimum expenditure of time, energy, and money. She had the good fortune to be able to secure the suggestions of Beulah I. Coon, Research Specialist, Home Economics Education, Department of Health, Education and Welfare, Office of Education, Washington, D. C., through personal interview. A questionnaire and letter were sent to each of the five colleges and universities in New Mexico requesting their assistance in determining emphases given each area of clothing construction at their particular institution and their cooperation in reviewing generalizations to be formulated. Information regarding previous use of clothing construction pretests was also requested.

Generalizations were formulated from an analysis of the questionnaires returned from the five institutions, textbooks used, and experiences of Hoskins (21). They were then classified into specific areas and submitted to a panel of experts for evaluation. Revisions were made according to suggestions of the panel.

Hoskins (21) divided her test into three areas. Each area was allotted an equal number of multiple-choice and true-false items, however, greater emphasis was given one area than another by varying the number of test items. The final test included 100 questions. Twenty

extra questions were included in the trial test to assist in replacing items found to be invalid or unreliable.

The test items were evaluated by a panel of critics at New Mexico State University. A trial test was administered to 12 Homemaking III students at Las Cruces High School. Based on the results of the trial sampling, the final test was reduced to 100 questions which would permit students to finish in a 55-minute class period (21).

The revised test was administered to students in Vocational Home Economics Departments in cooperating schools with departments of equivalent size. Such schools were thought to have comparable programs in the clothing area. The tests were scored three ways: (1) total number of right answers, (2) number of right answers in each of the three areas, and (3) number of right answers for each area as to type of question answered (21).

Collins (7) constructed a pretest based on a questionnaire that had been sent to selected colleges and universities to determine the success of the use of pretests. On the basis of 60 returns the questionnaire revealed that: 29 per cent of the institutions responding used a pretest for placement of students in beginning clothing courses while 39 per cent used a pretest for exemption purposes. Less than half those using pretests felt they served the purpose for which they were intended. The written objective-type test was most frequently used.

Subject matter for Collins' (7) pretest was based on the outline for the first clothing course offered at Southern Illinois University. Multiple-choice test items were used that permitted machine scoring. Items in the test were grouped according to subject matter for the purpose of locating deficient areas in the student's learning experiences.

In addition to the pretest, students were also required to fill out a questionnaire regarding previous experience, and complete a practical examination before full consideration for sectioning was given. This assisted in the determination of course content in each section. Collins (7) suggested an alternate method of sectioning by combining the objective and practical examination scores.

Witt (30, p. 37), at Oklahoma State University, determined that early attempts at evaluation in home economics followed two types of division: that which tested the value of the total program in home economics, and that which developed "techniques and devices to assess the effectiveness of particular subject matter areas, individual courses or specific aspects of any course."

Witt (30) proposed to revise the existing tests and develop techniques for appraising clothing competencies found among college freshman. She developed a questionnaire for the purpose of obtaining information regarding a student's previous training in clothing selection, construction, and care. This was prepared after reviewing similar questionnaires used at the University of Georgia, University of Southern Illinois, and Oklahoma State University. It was duly tested, revised, and retested. She found that students came to college with a wide range of experiences and that there was indication that some sort of placement device would be useful.

Objectives for clothing courses at Oklahoma State University and Mississippi State College for Women were reviewed by Witt (30) and found to be similar in many respects. Objectives common to tests of both institutions were identified and submitted for consideration to the home economics supervisors in both states. Secondary teachers of home

economics, college staff, and supervisory staff in each state were asked to evaluate their common objectives in clothing courses on the basis of importance. Thus her objectives for the basis of development of individual items in the placement test were determined.

Witt (30) reviewed the pretest at Oklahoma State University to determine how it should be improved and developed other devices for checking the student's aptitude for problem solving. Her pretest consisted of matching, multiple-choice (most popular type objective test), and true-false questions. She concluded that the three major steps in the construction of a pretest were: (1) choosing principles and generalizations to be used in this portion of the test, (2) the construction of problems, and (3) selection of possible conclusions of the problems.

Gould (17) revised three manipulative problems from Witt's (30) station-to-station type test. Witt's item analysis was used as a guide in the revision of these three problems. Six additional problems designed as a station-to-station test were devised by Gould. This laboratory test was administered as a pilot study to 24 students enrolled in one section of Home Economics 114 at Oklahoma State University in 1962. Gould found that running a pilot study was desirable for the purpose of revealing deficiencies in such a test.

Steelman's (25) pretest, developed at Oklahoma State University, consisted of two parts: (1) a theory test and (2) a laboratory test. Each test was an objective test which could be completed in a 50-minute class period. Expenses for the construction and administration of the tests were kept to a minimum.

Steelman used the criteria set up by Barbour (3, p. 50) to formulate generalizations from which her food pretest was developed. These criteria are as follows:

1. Generalizations in nutrition must be based upon reliable experimental evidence.

2. Evidence is acceptable only when the experiment is carefully controlled and subsequently reaffirmed by other investigators.

3. If the results of experiments indicate, but do not clearly establish, the validity of a finding the state of the generalization should be qualified by such terms as "may be", "has been shown", and "there is some evidence."

4. Contradictory results in experimental evidence necessitate the exclusion of some generalizations.

5. Information derived from animal experiments cannot be applied to human beings without qualifications.

6. Choice of non-technical words which convey valid nutrition ideas requires caution and discrimination.

7. In nutrition broad statements often exclude details necessary for useful application.

8. Many statements concerning nutrition which appear in textbooks, periodicals, bulletins and other literature are not based on valid experimental evidence.

It is generally accepted that the major purpose of education is to change the behavior patterns of people (2, 16, 25, 29, 30). Steelman stated (25, p. 5):

> Evaluation is a system of measuring the degree that these changes have occurred. Measurement, a means of obtaining quantitative evidence concerning growth or changes in behavior, is an important phase of evaluation, but it is not the entire process. Evaluation should also consider value standards and the goals which a group is trying to obtain.

Lack of motivation is the most difficult obstacle to overcome in trying to change food habits.

A placement test was devised by Lee (23) based on a list of terms used in food preparation and nutrition. This test was composed of two sections. Section I was a vocabulary test which consisted of terms frequently used in food preparation and nutrition. Words used -- both practical and scientific -- were related to the units studied in elementary food. Section II incorporated these words through use in practical application of basic principles.

The first time the test was administered it was given to 131 students at the beginning of their elementary food class. In the last class period before the final examination, the test was readministered to determine its reliability. There were 116 students who completed both tests. The final test grade was the total number of errors in Sections I and II subtracted from 100.

Facts, generalizations, and application of principles in nutrition formed the basis for the pretest developed by Cooksey (9). Objectives for a beginning course in nutrition were developed and evaluated from a behavioral aspects chart developed by Tyler (29). Key concepts were identified from the objectives and generalizations were formulated. The pretest then developed consisted of 125 multiple-choice items. Careful attention was given to questions which would test reasoning ability, understanding, and application of principles.

Establishment of Validity and Reliability

Validity was defined as the degree to which a testing device measures that which it is supposed to measure. It became increasingly important to determine the validity of a pretest before using it as a criterion for sectioning or exempting students from any college course.

On the other hand, reliability was supposed to establish the degree to which the same results could be expected each time the test was repeated. Reliability also became an important criterion in pretesting. Among the pretests developed preceding the present one, validity and reliability were established in numerous ways. Hoskins (21) submitted her pretest to a panel of experts who reviewed her test and verified its content as having covered the subject matter well. Reliability was established by administering the test to girls in high schools in designated areas in New Mexico. The split-half method of determination was used for calculating the coefficient of correlation. Calculations were made on 103 test papers. The mean and standard deviation were computed. The Spearman Rank-Difference method was used for computing the coefficient of correlation. Item analysis was used on the upper and lower 28 scores to determine item difficulty. The Kuder-Richardson formula was then used to determine the coefficient of reliability from the item difficulty scales. Multiple-choice items were found to be more discriminating than true-false ones.

The pretest constructed by Collins (7) was reviewed by a panel of critics chosen on the basis of: (1) experience and knowledge of clothing construction, (2) experience with testing, and (3) recognition in the field of clothing. However, it was not tested for validity and reliability due to lack of time.

Again, a panel of experts was used by Witt (30) to determine the validity of her pretest. Her test was unique in that faculty members from both Mississippi State College for Women and Oklahoma State University checked the accuracy of the test items and answers. A pilot study using 30 students was made at Mississippi State. Revisions were made

and the pretest was then administered at both institutions. A total of 112 freshmen students started the testing program while only 96 completed it.

Steelman (25), like many others, submitted the test items for her pretest in beginning college food preparation and service to a panel of experts for evaluation. She stated that the reliability could be determined by using an approximation formula from Garret (14). Steelman's test was administered to 138 college students enrolled in food courses but not the course for which it was designed. The theory portion of Steelman's pretest contained 100 questions. The students were encouraged to complete all items in order for the validity of the test to be determined. The mean of the scores for the theory test was 44.40 and the standard deviation was 11.90. The reliability was found to be .83. Seventy-six per cent of the items fell between 30-70 which was considered within the "ideal" range of difficulty. Seventy-six per cent of the items on the theory test were found to be discriminating.

According to Steelman (25), item analysis was commonly used to determine the best items for use in the final test. Items included in the final test were assessed for difficulty and discrimination indices to determine which items to re-edit and which items to omit because of lack of discriminating value.

In accordance with suggestions of experts, an analysis was made by comparison of the upper 27 per cent of the testees with the lower 27 per cent. Thirty-seven subjects were included in each group. Scores from Steelman's (25) test were checked and corrected for guessing. The validity index was read from Flanagan's Table (28) and converted into a discrimination index by using a table prepared by Davis (10). Item difficulty was also determined from a table prepared by Davis.

Cooksey's (9) test in introductory college nutrition was given to 137 students enrolled in beginning nutrition classes who had no previous training in nutrition courses. The same approximation formula used by Steelman (25) for calculating the coefficient of reliability was used. Item analysis was done on responses from the upper and lower 27 per cent of the total group. A formula from Davis (10) was used for item analysis. Cooksey's pretest was found to have .84 reliability. The mean and standard deviation were 76 and 13.2 respectively. The "ideal" range of difficulty was considered between 25-70.

The validity of Lee's (23) pretest was determined by finding the correlation coefficient between the placement tests administered and the final course grade received by the student taking the elementary food course. Comparison of the two sections of Lee's test was the basis for determining reliability.

Steelman (25), Gould (17), and Cooksey (9) used an item analysis technique to further identify imperfections that were found in the tests. Gould used a formula by Ahmann and Glock (1). Gould's pretest was not checked for validity or reliability. Discriminating items on Gould's (17) test were those items answered correctly by more high scoring students than low scoring students. Non-discriminating items were those test items answered correctly by an approximately equal number of high scoring and low scoring students. Item difficulty referred to the per cent of students who answered each question correctly.

Criteria for Development of Evaluation Instruments

Testing is a vital part of any educational program. Bogniard (4) believed it furnished a student with a means of self evaluation, and the

teacher with a tool for evaluation and improvement of instruction. Testing also served as a valuable resource to aid in the selection and identification of capable students.

Factors that should be considered when constructing a test, according to Bogniard (4) should be objectives for the course and effective marking of examinations. The objectives, stated in terms of change in student behavior, should be revealed as a result of teaching. A good examination will employ items to test knowledge; ability to think; and ability to analyze, interpret, deduct, and generalize. Items are needed to measure the development of attitudes and to serve as a motivational device.

The student must understand the marking system and be able to evaluate his own progress. He should be motivated to strive for deeper understanding and application of knowledge. An objective means of scoring is desirable. Tests that cover all the objectives in a learning unit and are sufficiently comprehensive are more valid than those which do not.

At the college level, testing becomes an important function in placement of students, and appraisal of individual and group achievement. The increased interest in testing is due to many factors such as: increased population of college-age students, higher socio-economic status of the families, demands from the professional world for more education, and trends in the economic world which encourage youth to stay in school (8). Testing then becomes more important on an individual basis when it is of great importance to determine the individual's aptitude for skills and technology, strengths and weaknesses, and to define his educational objectives.

The Committee on Measurement and Evaluation (8, p. 37) reported that: "Tests serve in instruction to clarify goals, to determine the initial status of students, to appraise student growth throughout the course, to appraise instructional materials and methods, and to stimulate learning." No testing program can function efficiently without precise definition of the institutional aims and educational objectives that are dictated by the philosophy of the institution.

Gorow (16) indicated that tests serve a variety of purposes. There are those that test for acquired knowledge as a basis for course grades at the end of the term, those that test for mastery of the material, those that are used for diagnostic purposes, and those to be used for practice or as a guide for further planning. He stated that objectives for tests should be limited to behavior that are amenable to classroom paper-and-pencil test. He suggested that (16, p. 7):

> Statements of objectives are most useful when they meet the following criteria:
>
> 1. Clear statement that indicated the "real goals" of instruction -- those which will be tested.
> 2. Attainability in the course or unit.
> 3. Measurability.
> 4. Inclusion of the various kinds of learning which are essential (stated separately and explicitly).
> 5. Specification of the end results of learning -- the abilities which students will have acquired when they have achieved the objectives.
> 6. Expression in behavioral terms -- indicating what the learner will be able to do at the end of the instruction.

Since the fulfillment of the objectives for a course is exemplified in the students' achievement, the kinds of achievement desired are included in the statement of objectives.

Authorities agreed that homogeneous grouping simplified problems of instruction, facilitated a fairer grading system, and stimulated a sense of security among students who might otherwise be frustrated from lack of ability to achieve or lack of motivation from oversimplification (8).

College testing programs for purposes of placement were widespread. Such placement enabled a student to be associated with others of comparable aptitude (8, p. 26) ". . . so that instruction may proceed at a level appropriate to the capacity or readiness of the group." If a student showed unusual aptitude in a specific area, consideration might be given to waive the required course and accelerate the study procedure. Great care must be exercised at this point, however, that the student had sufficient background for the courses to follow.

Courses were not designed in an orderly sequence of events as one might suppose but rather in an atmosphere of constant interaction between instructor and student. Because of this constant interaction and revision, achievement testing became more important in regular instruction than for purposes of course grades.

The experiential background of the student should determine the development of the curriculum. By some means, such as a pretest, much could be learned about the students' background. Pretests then could be a useful tool in obtaining information about specific areas of content and special abilities of individual students in regard to subject matter. The pretest would aid in the formation of homogeneous groups, and development of a suitable course of study.

If it were desirable to determine the acquisition of new learnings, a test similar to the one given at the end of a course might be administered at the beginning of a course. This type of pretest could serve to stimulate interest in subject matter content, point out areas of importance, assay knowledge already acquired, and offer security to those who feared defeat in the face of unfamiliar material.

It became of great importance then to create a pretest that would fulfill the needs for which it was intended. Since no student would perform exactly the same way each time a test was administered, it was highly desirable to formulate test questions that were relevant to the goals of the course; would indicate the desired changes in behavior; and would encourage independent, critical thinking on the part of the student.

It was necessary also that the test be organized in such a way that proper perspective was given to each phase of the program in relation to importance placed on the particular aspect in question. Many authorities (1, 2, 10, 16) and others agreed that items should proceed from the least difficult to the most difficult in a test. Construction of a test such that test items would differentiate between those who knew and those who did not know was obviously an important factor. Items in the 50 per cent range of difficulty discriminated more than items that were too difficult. For this reason item analysis of each item was a useful tool in test construction.

The Committee On Measurement and Evaluation (8, p. 57) reported that:

> Item analysis, if properly interpreted by the instructor, can indicate the quality of his examination effort, suggest areas where learning has been poor, and help determine whether there are individuals of the class who require remedial work. Item analyses, however, are neither infallible nor

> completely informative: they never disclose directly how the individual student reacts to the particular question.

In the selection of items for a test the difficulty of each item was a prime consideration.

Item difficulty was defined as the proportion of students taking the test who actually knew the answer. Allowance for guessing on multiple-choice items could be made by subtracting from the number of correct answers the number of incorrect answers divided by one less than the number of choices in the item.

Attempts had been made to determine the difficulty level of each test item and the extent to which each item discriminated between students being tested (10). Positive discriminating power was attributed to test items answered correctly by all students who scored high on the test and incorrectly by all students who scored low on the test. There was no discriminating power if correct and incorrect answers were divided equally between high and low scoring students. Negative discrimination was indicated when correct answers from low scoring students exceeded those from high scoring students.

Davis (10) recommended that corrections for chance be used in determining the difficulty index of test items because answers to items did not reveal whether a student knew the answer or guessed the correct answer when all items were answered. Item-analysis charts were highly satisfactory when based on sound statistical procedures, on a minimum requirement of labor for use, and on values that were easy to interpret and use. It was assumed that incorrect choices were equally attractive.

Regardless of kind or form, all evaluation instruments should possess a high degree of validity. That is to say they should measure

what they were supposed to measure whether it be course content, application of knowledge, comprehension of complex ideas, or attitude and value development (16). The usefulness of each type test was determined by what and how well it measured the purpose for which it was intended.

Ahmann and Glock (1, p. 292) stated: "Validity is clearly the most important characteristic of an evaluation instrument." Both Gorow (16) and Tyler (29) suggested the use of a two way chart for content analysis. A balance within the coverage of the various areas of content was an important factor in establishing validity in a test. A test which was designed to adequately cover the subject matter, to measure the kinds of learning indicated by the objectives, and to fit a particular group of students took validity into account (16).

Types of validity varied in relation to the goals of an evaluation instrument. Four goals for methods of evaluation were: (1) to determine the degree of performance on sampling of subject matter evaluation, (2) to predict the student's academic success in the future, (3) to estimate the students ability to make practical application of knowledge, and (4) to indicate the degree to which results of an evaluation instrument reflect traits for which students were being tested. Similar types of validity existed which corresponded to these four goals. Ahmann and Glock (1) indicated that some types of validity tended to be more important than others. Validity could be either content, predictive, concurrent, or construct. Content validity, as the name implied, determined the degree to which an evaluation instrument actually covered the subject matter. This type of instrument should represent a complete cross-section of a course. If this were true the instrument would have tested the degree to which the educational objectives had been taught

and understood. Contend validity was established by comparison with textbooks, information from experts, bulletins, brochures, and objectives of the course. The validity of a test depended upon how well the test sample covered the content taught. Therefore, if a sample failed to test all areas taught it lacked content validity.

Many instruments might include more than one type of validity. While content validity was of primary importance in achievement testing, predictive validity was of greater importance in aptitude testing. Predictive validity was the degree to which an instrument was able to predict the future performance or behavior of a student. Correlation between actual achievement and predicted achievement could yield reliable information on predictive validity (1).

By computing a Pearson product-moment coefficient of correlation, represented by the letter r, the degree of correlation or predictive validity could be obtained. The coefficient of correlation represented the degree of tendency for comparison of scores from two variables to fall in a straight line. If the relationship were perfectly linear the r-value would be one, but if not, the value would be less than one. The greater the deviation from a straight line the less the degree of correlation and the closer the r-value approached zero. When high values of one variable were associated with high values of the second variable, or low values of one variable were associated with low values of a second variable, the relationship was positive and the r-value was a positive number. However, if high values of one variable and low values of the other variable, and conversely, were associated, a negative relationship existed and a negative r-value was obtained. Therefore, r-values varied from -1.00 to +1.00. Ahmann and Glock (1, p. 306) stated:

"The size of the number is indicative of the degree to which the points form a straight line, and the sign of the number reveals the direction of the relationship." Positive r-values were much more common than negative values in educational evaluation. Computation of an r-value provided a concise representation of the predictive validity. Many formulas existed for computation of r-values; however, all represented the degree to which the scores fell into a straight line.

Since the decrease in prediction accuracy was not proportional to the decrease in the size of the r-value, tables such as Flanagan's (12) were used for determining the predictive validity. The efficiency of such a system, however, was greatest only in terms of predicting scores within broad categories: that is, "below average", "average", and "above average". Although such predictions were valuable in terms of a group as a whole, they were of no value in predicting performance on an individual basis (1). Predictive validity was necessary to be able to predict how well a student could perform without being exposed to a course at all. In other words it served as a basis for exemption and allowed a student to take an advanced course.

Concurrent validity had to do with determining, by an indirect means, the status of the student at the time the instrument was used. Establishing this type of validity was dependent on the correlation between what a student did in a theoretical situation and in actual practice. Using test scores as an indication of what a student could do in actual practice was an indirect means for establishing concurrent validity. The only difference between predictive and concurrent validity was the time of performance. "Predictive" was related to a future performance and "concurrent" was related to a present performance. The

emphasis was upon a behavioral aspect of the student that was independent of the testing device. The degree of correlation between the performance, present or future, and the evaluative instrument yielded the concurrent or predictive validity as the case might be. Concurrent validity was necessary for purposes of sectioning students and determining the degree of performance a student was able to maintain at the time he was being tested. It was established by correlation between written test and practical application or attitude. This was the same kind of validity one would have if two parts of a test were developed.

Ahmann and Glock (1, p. 297) defined construct as:

> . . . a human characteristic assumed to exist in order to account for some aspect of human behavior. . . .Whenever an evaluation instrument is believed to reflect a particular construct, its construct validity must be investigated. In order for a suitable investigation to be possible, the construct should be sufficiently well defined so that verifiable inferences can be drawn from it. . . .If the instrument possesses construct validity, the results it yields will vary from one kind of individual to another, or from one situation to another, as the theory underlying the construct would predict.

The degree of construct validity was an important factor in defending a stand taken in achievement, aptitude, or personal-social adjustment evaluation.

Validity was of intense concern to those who built evaluation instruments. As previously mentioned, it was the most critical quality an instrument could have. Careful examination of each function of the instrument was an obligation of the author. Evidence of the valid performance of these functions was established as essential (1).

Reliability was another factor to consider in test construction. Only when there was a spread in test scores could there be an accurate

measurement of achievement. This could be directly related to test difficulty which in turn was related to reliability. Since reliability was defined as the degree to which consistent results were obtained, a reliable test would tend to separate those who knew from those who did not know. Students should understand the meaning of the question asked but no clues or suggested answers should be incorporated into the question. Few if any instruments were considered completely reliable due to the chance of error. However, one error had a tendency to offset another with repeated performances. An instrument might be highly reliable without being completely valid. Instrument-centered reliability was increased by an increased number of test items and by more objective means of scoring (2).

There were no reference standards, as physical measurements, with which to compare test scores. Therefore the value of scores could be determined only as they were compared with scores of other students taking the same test. The length of a test effected reliability because of the opportunity for a larger sampling of a student's knowledge. To insure maximum reliability, Gorow (16) suggested that: (1) chance factors be reduced to a minimum; such as eliminate questions with only two choices for answers, (2) write clear instructions so that measurement of performance rather than "what the teacher wants" was possible, and (3) insure consistency in scoring by using a key which was prepared in advance.

The use of a key for scoring increased the objectivity of a test, since this eliminated the mood, judgment, or personality of the person scoring the test. Tests should be constructed to increase their objectivity. It was generally agreed among authorities that the multiple-

choice type test was the most desirable type objective test. Gorrow (16, p. 18) stated:

> To provide a good sample of extensive knowledge, one must use the most "efficient" items -- which require the least reading and responding time; in short, objective items.

Reliability is determined by calculating for: (1) the coefficient of stability, (2) the coefficient of equivalent forms, (3) the coefficient of stability and equivalence, or (4) the coefficient of internal consistency.

The coefficient of stability was found by administering the same test after a limited lapse of time. By giving two tests of equal difficulty, one immediately following the other, the coefficient of equivalent forms was determined. If equivalent forms of a test were administered with a lapse of time between them, a coefficient of stability and equivalence might be found at the same time. By comparing the scores on all even numbered questions with scores on all odd numbered questions, a coefficient of internal consistency could be obtained. Another means of determining the coefficient of internal consistency would be to compare the number of pupils answering the question correctly with those answering the question incorrectly. This was known as an analysis of variance procedure. Coefficients of reliability of .80 and above were generally acceptable for many standardized achievement and aptitude tests. This factor might vary according to the characteristic being tested (1).

Inadequate sampling of content tended to produce low scores in reliability. Therefore the reliability might be improved by increasing the number of test items providing the characteristics of the additional items compare with those of the original ones. Clear and adequate

instructions increased the reliability of a test according to Arny (2). Reliability was also increased when the scores spread over a wide range. The discrimination indices, however, influenced the reliability to a greater extent than the length of the test. By eliminating non-discriminating items the reliability of the test was greatly improved.

If students did not consider a test fair, reliability and validity tended to be influenced. Reliability was also affected by too many different forms being used in a test. Multiple-choice items were considered the most objective and acceptable form to use.

Steelman (25) stated that an objective test that could be scored with an inflexible key or by a rating device which, when used by a different group of people, would yield similar results was most desirable. Accuracy and speed with which test papers could be scored were important.

The review of literature indicated that limited pretesting had been done in the field of food and nutrition. Because of recent trends in this field, it was imperative to develop efficient means for curricular development. Provision for the needs of the learner was one characteristic of an effective curriculum. These needs could be determined by the use of a pretest.

Many factors are involved in the development of an objective pretest. This aggregate of factors determines the reliability and validity of any given test. Thus, the review of literature suggested the factors necessary for the revision of Steelman's pretest.

CHAPTER III

METHOD OF PROCEDURE

Introduction

Several pretests in the area of clothing and textiles have been developed and are in use. Only a few pretests have been reported in the area of food and nutrition. Cooksey (9) and Steelman (25) each constructed tests for use at Oklahoma State University but neither were usable without further validation. Thus, the current writer has attempted a revision of Steelman's theory test for food preparation.

Revision of the existing theory pretest in food preparation allowed questions to be more clearly stated, to be brought up to date with current knowledge and information, and permitted the inclusion of a larger number of valid questions. This test was to consist of multiple-choice questions which Steelman (25) chose to use since they were identified as an effective means of evaluation. One of the weaknesses of her theory test was the failure to establish a high degree of validity. Therefore, in an attempt to overcome this, it was planned to increase the number of test items from 100 to 150.

Objectives

No major curriculum changes in the beginning food preparation course at Oklahoma State University had been made since Steelman's (25) pretest was developed. The objectives and generalizations Steelman formulated

appeared to be adequate, hence, they were accepted by the author without revision. The student objectives developed by Steelman (25, p. 35) for the beginning food preparation course are as follows:

Objectives for FNIA 2113 Family Food

I. Does critical thinking in relation to:
   1. Food purchasing, storage, preparation, and service.
   2. Desirable standards of prepared food.
   3. Management of time, energy, and other resources in food purchasing, storage, preparation, and service.
   4. Evaluation and acceptance of new foods.
   5. Sanitary practices in handling and preparing food.
   6. Food selection, preparation, and service as a means of creative expression.
   7. Desirable combinations of foods.
   8. Equipment needed in food preparation and service.
   9. Wide use of available foods.

II. Directs self in relation to:
   1. Food selection, preparation, and service as a means of creative expression.
   2. Use of terms in handling and preparing food.
   3. Sanitary practices in handling and preparing food.
   4. Management of time, energy, and other resources in food purchasing, storage, preparation, and service.
   5. Wide use of available foods.
   6. Selection of meals for self and others.
   7. Evaluation and acceptance of new foods.

III. Develops desirable attitudes in relation to:
   1. Contribution of processed foods to present-day living.
   2. Combinations of foods.
   3. Psychological, economic, and social effects of food advertising in our society.

IV. Develops an appreciation of:
   1. Contribution of processed foods to present-day living.
   2. Variety of methods used in preparation and service of food.
   3. Contribution of food and its service in fostering family relationships.
   4. Information which appears on packaged products.

5. Contribution of current scientific knowledge of food processing and preparation.
6. Benefits which agencies provide the consumer.
7. Pleasures which may be found in eating.

V. Develops values in relation to:
1. Sanitary practices in handling and preparing food.
2. Food selection, preparation, and service as a means of creative expression.
3. Wide use of available foods.
4. Desirable standards of prepared foods.
5. Management of time, energy, and other resources in food purchasing, storage, preparation, and service.

Generalizations

Steelman (25, p. 36) formulated the following generalizations that formed the basis for the development of the pretest.

Generalizations for FNIA 2113 Family Food

Meats

1. If meat has been inspected and graded by a government agent, it can be considered safe to eat at the time of inspection.

2. When comparing meat prices, the cost of the edible lean should be used rather than the cost per pound.

3. Because neither palatability nor food value correspond to market price, the cheaper cuts of meat may be a more economical source of lean than are expensive cuts.

4. If beef is of good quality, it is bright red in color, it is fine grained and smooth, and the fat is creamy white, firm, and brittle.

5. If veal is of good quality it is grayish pink, the texture is fine grained and smooth and the interior fat is firm and brittle and is grayish or pinkish-white in color.

6. If lamb is of good quality, it is pinkish-red in color, fine grained and smooth, the fat is firm, brittle, and white or pinkish in color

and the bones are soft, red, and spongy, and show cartilage.

7. If pork is of good quality, the flesh is grayish-pink color, fine grained, and the fat is firm and white but not brittle.

8. If meat from older animals is selected, it will usually be less tender than meat from younger animals, since the muscle fibers are more developed in an older animal.

9. If a meat cut is from a portion of the animal that has been exercised a great deal, it will tend to be less tender than cuts from the little-used parts of the animal.

10. Because moisture and heat are conducive to bacterial action, meats should be covered lightly to allow drying and should be stored in the coldest part of the refrigerator.

11. Because ground meats are easily penetrated by bacteria, they should be used as soon as possible.

12. Because flesh foods are highly perishable, they are readily subject to new contamination after cooking even though the original pathogenic organisms were killed when the meat was cooked.

13. Because particles of the bone and foreign particles may be on the surface of the meat, it should be wiped with a damp cloth before cooking.

14. If cut meats are dipped in water, there will probably be a loss of some of the soluble constituents of the meat.

15. If meat is seared before roasting, a greater loss of juices than in unseared meat may result due to the high temperature used.

16. It is advisable to apply salt to steaks and chops after they are browned, because salt retards browning of meat.

17. There is no advantage in salting roasts before serving, because salt only penetrates a short distance in meat.

18. A meat thermometer is the best known way to determine the stage of doneness of the meat, because the internal temperature is directly related to the stage of doneness.

19. Because pieces of meat of the same weight may vary enough in shape, thickness, and in proportion of meat to bone as to cause a difference in roasting time, cooking time charts should be used only as guides to estimate the total cooking time.

20. When meat is removed from the oven, the internal temperature will rise a few degrees, therefore the meat should be removed before the desired stage of doneness is completely reached.

21. When cooking frozen cuts of meat, allow about one and one-half times as long to cook as is necessary for similar cuts of unfrozen meat.

22. If drip and cooking waters are utilized, the food values of meat are well retained.

23. If a cut of meat contains a minimum of connective tissue, it is considered a tender cut of meat and may be cooked by dry heat methods.

24. When cooking meat in the oven by dry heat, a constant low temperature should be maintained to prevent excessive shrinkage and loss of juice and flavor.

25. If a cut of meat is boneless, the cooking time will be longer than for meat with bone, because bone aids in heat penetration.

26. If a good layer of fat covers a roast, some evaporation will be prevented since water does not tend to pass through the layer of fat.

27. When the fat is found well distributed throughout the muscle fibers, the cut of meat will usually be tender.

28. When meat is cooked at low temperatures, the muscle fibers hold together better and there is less crumbling and falling apart of the meat than when it is cooked at high temperatures.

29. When moist heat methods are used to cook meat, the temperature of the connective tissue rises to the point where gelatin is formed much more rapidly than in dry heat, because water and steam are much better conductors of heat than air.

30. If moist heat methods are used to cook meat or poultry, the water should not be hotter than simmering to avoid the toughening effect of high temperature on the fibers.

31. If beef is cooked only to the rare stage, it will be more juicy and flavorful and there will be a larger number of servings than in meat cooked to the well-done stage.

32. Because pork that has not been specially treated may contain trichinae, it is wise to cook it to the well-done stage.

33. Because poultry flesh is often too shallow to insert a regular meat thermometer, it is often difficult to determine the stage of doneness.

34. If no thermometer is used for testing the doneness of poultry, the bird will be done when the joints can be moved easily, especially the thigh joint.

35. Because the legs and wings of poultry are small in relation to the body of the bird, they cook more rapidly.

36. If the legs and wings of poultry are tied close to the body, over-drying can be minimized.

37. If the breast is placed down when roasting poultry, the fat on the back will aid in basting.

38. If the natural form of fish is to be retained, it must be carefully handled and not overcooked.

Eggs

1. If an egg is fresh, the yolk will hold its shape in the center, the white will be more or less firm, the odor will be fresh, and the color will be "good".

2. If the protective, dull coat of the egg shell is washed off before storage, the porous shell may then permit odors, flavors, bacteria, and molds to enter the egg and may permit greater evaporation from the egg.

3. Because the egg can be contaminated from substances on the shell, it is wise to wash

all eggs just before breaking the shell.

4. When eggs are chilled for two or three weeks in the refrigerator, no noticeable changes in the flavor of the egg can be detected even though the egg will change in form to some extent.

5. Because the protein in egg yolk coagulates in air, unbroken egg yolks should be covered with water in a tightly covered container in the refrigerator.

6. When storing egg whites, cover tightly and place in the refrigerator to prevent drying and spoilage.

7. Because eggs have a high protein content, they should always be cooked at low to moderate temperatures to prevent toughening of both yolk and white.

8. If an egg is taken directly from the refrigerator, it will hold its shape better for frying than eggs held at room temperature.

9. If hard cooked eggs are immersed in cold water immediately after cooking, they tend to peel easier and there is less discoloration due to overcooking.

10. If prepared dishes containing a high percentage of eggs are over cooked or cooked at too high a temperature, syneresis or weeping of the curd may occur.

11. If water is used as a heating medium instead of air, baked custard products will be more evenly cooked.

12. When sugar is combined with eggs, the coagulation temperature is raised and more time is required for cooking.

13. When salt and acids are combined with eggs, the coagulation temperature is lowered and less time is required for cooking.

14. When poaching an egg, hot water should be used to coagulate the surface of the egg and protect the interior from the solvent action of the water.

15. Since the fat in yolks greatly interferes with the whipping quality of the whites, great care should be taken to keep all of the yolk out of the whites.

16. If egg whites are at room temperature, they are less viscous and tend to whip better and give greater volume.

17. When egg whites are held after beating, the foam tends to separate and cannot be beaten so lightly again.

18. If an egg white foam is over beaten, the air cells may collapse and give a heavy compact end product.

19. When eggs are used as a leavening agent, their effectiveness will depend upon the amount of air beaten in and retained during the process of food preparation.

20. When egg whites are frozen or dried, their foaming properties are not destroyed.

21. Because eggs will form thin, stable films around tiny globules of oil, they are valuable as emulsifying agents.

22. Because food poisoning develops easily in moist, nonacid foods, cooked egg dishes should be kept stored in the refrigerator.

Milk and Cheese

1. Because the pasteurizing process destroys disease causing micro-organisms, pasteurized milk is considered safe to drink if it has been handled properly in the home.

2. Because milk is an excellent medium for bacterial growth, it should be stored in the coldest part of the refrigerator.

3. If milk is boiled, it has a "flat" flavor due to the loss of the dissolved gases in it.

4. Because milk contains lactose, a sugar, it scorches easily and should be cooked slowly.

5. Because non-fat dry milk solids are inexpensive, wholesome, low in calories, and high in nutritive value, it is wise to consider

them for drinking purposes as well as for cooking.

6. When cooking cheese, a protein food, a low temperature should be used to prevent toughening and stringiness of the protein, and separation of the fat.

Sauces

1. When a starchy material is used for thickening a liquid, the starch granules need to be separated before the hot liquid is added to prevent lumps in the mixture.

2. If each particle of starch is not surrounded by liquid, it cannot swell to its maximum capacity.

3. If starch is overcooked, it converts to dextrin which has less thickening power than the starch.

4. Because acid converts starch to dextrin, it is advisable to cook the starch in the sauce first and add the acid last.

Fruits and Vegetables

1. If vegetables and fruits are properly prepared, cooked, and served, they will help contribute valuable nutrients to the diet.

2. When fruits and vegetables are fresh and attractive, their vitamin content will usually be higher than those that are wilted and off-color.

3. If vegetables are over-mature, there is usually more waste and a longer cooking time is needed than is required for young, slightly immature products.

4. When vegetables are harvested, enzymes are set free which hasten the destruction of certain of the vitamins and cause other forms of deterioration.

5. Because the maturing action continues after harvesting, fresh vegetables, except tomatoes, are considered to be best when harvested at a slightly immature stage.

6. Because deterioration of fruits and vegetables is due primarily to oxidation, generally the lower the temperature, the shorter the time, and the less exposure to oxygen during storage results in a decreased deterioration rate.

7. When held in common storage and in refrigeration above freezing, natural plant products continue to live and undergo some of the changes associated with increasing maturity.

8. If vegetables are stored at low temperatures, the development of pathogenic bacteria and those which make food inedible is retarded.

9. When root vegetables are stored in a cool, dry, well-ventilated place, they will tend to retain their normal characteristics.

10. If leafy and other fresh vegetables are trimmed to remove inedible parts, washed, and stored in a covered container in the refrigerator, dehydration and loss of nutrients will be retarded.

11. If canned goods are to undergo minimum deterioration, they should be stored in a cool, dry place.

12. If a can has been opened, it is safe to refrigerate unused portions of the food in the original can as it has an enamel finish on the inner surface.

13. If reconstituted frozen fruit juices are covered and kept in the refrigerator, they will retain most of their ascorbic acid value for three or four days.

14. If citrus fruits are covered to provide a moist atmosphere, they will not dry out so readily.

15. When held at refrigerator temperatures, the avocado and banana discolor and lose the power of ripening.

16. Because root and tuber vegetables often contain embedded dirt, they should be scrubbed with a vegetable brush before being cooked.

17. When washing leafy vegetables, care should be taken not to bruise the leaves to avoid a loss of nutrients and a less attractive appearance.

18. If warm or hot water is used for washing green leafy vegetables, wilting will occur.

19. When removing the skin from fruits and vegetables, care should be taken to remove as little as possible because many of the food nutrients are located just underneath the outer skin.

20. If the skin of a fruit or vegetable is undesirable, a very thin layer can be slipped off after cooking or the raw food can be scraped or thinly pared to avoid excessive loss of nutrients.

21. Because of the oxidative changes that occur at the surface of cut fruits and vegetables, slicing or other partitioning of plant materials should be done just before they are to be used.

22. If uncooked fruit and vegetable products are ground or mashed, a great loss of flavor, color, and ascorbic acid usually occurs.

23. If cut fruits and vegetables are exposed to the air, ascorbic acid may be lost by oxidation.

24. If pared or cut fruits and vegetables are allowed to soak in water before cooking, water soluble vitamins and certain of the minerals may be lost by dissolving in the liquid.

25. If berries are washed before the stems and hulls are removed, there will be a smaller loss of nutrients than if they are washed after hulling.

26. If berries are sliced rather than crushed, there will be less ascorbic acid lost.

27. If fruits and vegetables are marinated or treated with an acid such as vinegar or lemon juice, the ascorbic acid will be protected and browning will be prevented.

28. If a plant material is heated in a moist atmosphere the tissues will be softened and if heated for a long enough time, the vegetables or fruit will fall apart.

29. Because the softening of the plant structure during cooking allows many of the nutrients and flavoring substances to dissolve in the cooking water, only the amount of water that can be consumed with the food should remain at the end of the cooking period.

30. If excessive cooking water is not evaporated or recovered and used in soups, gravies, or other foods, much of the vitamin and mineral value contained in the vegetables will be wasted.

31. When vegetables are cooked properly, their original shape, flavor, and color will be retained.

32. If heat is reduced to maintain a slow steady boil, there will be less disintegration of the vegetable or fruit and less vitamin destruction than when they are cooked at a rapid boil.

33. If a large amount of surface area is exposed, there will be a greater loss of nutrients than if a smaller surface is exposed.

34. If salt is added at the beginning or the midpoint of the cooking period, the texture and flavor of most vegetables will be improved.

35. If a covered kettle or steamer is used to cook fruits and vegetables, the intensity of the flavor tends to increase, while an open kettle allows some volatile flavoring substances to escape.

36. If vegetables are boiled in their skins only a slight loss of soluble material occurs.

37. If vegetables are cooked in a pressure saucepan, time, fuel, flavor, color, and usually nutrients will be saved due to the short cooking time required.

38. Because canned vegetables are already thoroughly cooked, they need only to be heated or chilled to improve palatability.

39. If the liquid from canned vegetables is evaporated to the desired quantity by rapid boiling in an open container before the vegetables are added, the nutrient and flavor losses from the vegetable to the liquid will be offset.

40. When green vegetables are cooked in the presence of an acid, they tend to turn olive-drab.

41. When the cooking solution is alkaline, green vegetables tend to be bright green.

42. If green vegetables are cooked uncovered during the first few minutes, volatile acids that dull their color will be allowed to escape.

43. When red, purple, blue or violet vegetables are in an acid solution, the red ones retain their original color and the others tend to turn red.

44. When red, purple, blue or violet plant materials are in an alkaline solution, they will retain their blue color or tend to turn blue.

45. If red vegetables are cooked covered and in their skins, there will be a greater preservation of the volatile acids and the red color.

46. When white vegetables are cooked in acid solutions they tend to retain their white color.

47. If white vegetables are cooked in alkaline solutions, they tend to turn yellow.

48. If the yellow vegetables are exposed to small amounts of acid or alkali, there will be little if any color change.

49. If yellow vegetables are overcooked and the sugary juice of the vegetable is scorched, darkening will occur even though the coloring matter in yellow vegetables is very stable under ordinary conditions.

50. If vegetables are preserved by freezing, there is usually a greater retention of fresh flavor, texture, color, and nutritive value than if they are preserved by any other known method.

51. If frozen foods are to retain a high percentage of their nutrients and their good

eating quality, they must be stored near 0°F.

52. When most vegetables are cooked without thawing, they retain their best volume, shape, texture, color, flavor, and food value.

53. If vegetables are thawed prior to cooking, the package should be unopened for best retention of flavor and nutritive value.

54. Frozen greens are best partially defrosted before cooking to avoid over-cooking the outer leaves before the inside of the block is defrosted.

55. If corn on the cob is not defrosted completely before cooking, the cob will not thaw during the cooking process.

56. If fruits are cooked in a sugar sirup, they tend to hold their shape better than when cooked in water.

57. When fruits are cooked in a sirup, the cells take up sugar by osmosis; the fruit becomes more transparent and may tend to shrink slightly.

58. If excess sugar is used with cooked fruits, the delicate flavors will be masked.

59. Because dried fruits have a high sugar content, little sugar needs to be added for sweetening purposes.

60. If a vegetable contains a high water content and a small surface area, it will lend itself to baking.

61. If skins of vegetables are oiled and wrapped in foil before baking, the skins will tend to be soft due to the entrapped steam.

62. When fruits and vegetables are baked in their skins, the steam which is trapped under the skins cooks the interior.

63. If a fruit or vegetable requires very little cooking, broiling is a suitable method to use.

64. When vegetables are fried, the fat should be hot so that a minimum amount of it is absorbed.

65. If vegetables are low in starch and attractive in appearance, they are often good when eaten raw.

Salads and Salad Dressings

1. Because many salad materials are eaten raw, it is essential that they are thoroughly washed.

2. If a watery salad is to be prevented, all materials need to be carefully dried before combining them.

3. If the pieces of a salad are large enough to keep their identity, yet not so large that they are difficult to eat, they will tend to be more attractive and appetizing.

4. Because simple, natural salad arrangements are difficult to improve upon, it is wise to avoid grotesque and fixed arrangements.

5. If salad greens are torn rather than cut, there will be less bruising and discoloring of the vegetables.

6. Because foods with a strong flavor mask more delicate flavors, they should be used sparingly.

7. When selecting a dressing for a salad, consideration should be given to its flavor and consistency in relation to the other parts of the salad.

8. Because the liquid, acid, and fat in a salad dressing will not mix, but one is dispersed in the other, the dressing is an emulsion.

9. Oil at room temperature is desirable when making a dressing, because cold oil is more difficult to break up into small globules than is warm, less viscous oil.

10. After a permanent emulsion is formed, it should be stored in the refrigerator because chilling thickens and stabilizes the product.

11. If salads are stored covered in the refrigerator before serving, there will be less drying, absorption of odors and giving off of odors than if stored uncovered.

Cereals

1. When cereal is cooked, the rate of digestion and palatability of the cereal are improved.

2. Because quick-cooking cereals have been partially cooked before packing, they require a shorter cooking period than untreated cereals.

3. If cereals are cooked slowly, they tend to develop an improved flavor and a softer fiber than if they are cooked rapidly.

4. When all particles of the cereal are equally exposed to the heat and liquid, a desirable uniform gelatinous mass will be formed.

5. If cereals are cooked without stirring, the original form of the cereal can be maintained during the softening process.

6. Because starch of cereal reaches its capacity for maximum absorption of water several degrees below the boiling point of water, a double boiler may be used successfully in cooking cereals.

7. If salt is added to the water at the beginning of the cooking period, the flavor will be improved and the original shape of the cereal will be maintained better than if the salt is added after the cereal is cooked.

8. Because most brands of cereal vary in the amount of water required and the time necessary for cooking, it is recommended that this information be obtained from the package.

9. When ready-to-eat cereal cartons are opened, the remaining cereal should be closed as tightly as possible and stored in a warm, dry place to prevent the absorption of moisture.

Batters and Doughs

1. When the ratio of flour to water is such that mixtures will pour or drop easily from a spoon, they are called batters.

2. If a mixture is thick and can be rolled or kneaded, it is a dough.

3. When cakes contain fat and are leavened with baking powder, steam, and air, they are generally called butter cakes.

4. If a cake does not contain fat and is leavened only with air and steam, it is classified as a sponge cake.

5. When breads are leavened with steam, air, or baking powder, they are usually called "quick" breads.

6. When mixtures have a high proportion of liquid and are baked at a high oven temperature, steam is rapidly formed and acts as a leavening agent.

7. When baking powder is used as a leavening agent it must be dissolved in a liquid before it can release carbon dioxide which aids in leavening the mixture.

8. When yeast is in the presence of moisture, food, and favorable temperatures, it multiplies rapidly and releases carbon dioxide which aids in leavening.

9. If the dough becomes too hot, the yeast cells will be killed, but if the dough is too cool yeast growth will be retarded.

10. Because shortening increases tenderness and aids in leavening, it is used in most batters and doughs.

11. When breads are made with milk they stale less readily than do breads made with no milk.

12. Because wheat flour contains two proteins which form gluten, it is considered superior to all other flours for bread making.

13. If all-purpose flour is substituted for cake flour, the amount called for should be decreased.

14. If the gluten is made elastic by thorough kneading and mixing, it can expand and hold within it the gas bubbles formed by the leavening agents.

15. Lightness, the distinctive characteristic of batters and doughs, is produced when heat expands the air or other gases trapped in the elastic gluten.

16. If a batter or dough is overmixed, the resulting product may have tunnels, peaks, a smooth crust, and an increased tendency to stale readily.

17. When a mixture is understirred, sugary crusts, concave tops and a coarse texture may result.

18. If baked products are stored in cool dry places, staling and molding will be retarded.

19. If muffins, biscuits, waffles, and griddle cakes are stored, they will probably need to be refreshed by heating before being eaten, since they stale so readily.

20. When filling a cake or muffin pan, about one-half to one-third of the depth is generally allowed for expansion.

21. When baking pans are staggered in the oven there can be free circulation of heated air around the pans.

22. If a short baking time is required, pans made of materials that conduct heat rapidly will generally give better browning.

23. When heavy pans made from a material such as glass, iron, and enamel ware are used, the products tend to have a thick heavy crust and in some cases are misshaped.

24. Because shiny pans reflect the radiant heat, a longer baking time may be required than with dull pans.

25. If shiny pans are used, a product with a greater volume may result because it has a longer time to expand before the crust is formed.

26. Because of increased depth, products baked in deep pans require lower temperatures and longer baking times than do products baked in shallow pans.

Pastry

1. When the flour particles are separated from each other by means of fat, the development of gluten will be prevented and the pie crust will be more tender.

2. If large fat particles are rolled between the layers of dough, rather than mixed finely in the dough, the crust will tend to be more flakey.

3. If too much water is used in relation to the amount of fat in pie dough, the flakiness of the pastry may be decreased.

4. If pastry is handled too much, it tends to become tough.

5. If the dough is allowed to stand a few minutes before rolling, the extensibility and elasticity will be increased, making it easier to handle.

6. If pie crusts are pricked with a fork before baking, blistering of the crust will be prevented.

Beverages

1. If water for drip coffee and tea is boiled excessively, it will be flat and insipid and the flavor of the tea and coffee will be affected.

2. When coffee or tea is boiled, there is a greater extraction of tannin which produces a bitter flavor and also causes excessive loss of flavor.

3. Because some aluminum and pewter pots impart a metallic taste to coffee; glass, pottery, enameled ware or stainless steel is recommended for making these beverages.

4. Because unbrewed tea and coffee readily lose their aromatic flavors and fragrances, they should be stored in tightly covered containers.

General

1. When buying food, consideration should be given to unit cost, amount of waste, nutrients supplied by the food, and time, energy, and further expense required for its preparation.

2. When carefully prepared, a food budget should help people to eat economically, to buy intelligently, and to live comfortably.

3. When prepared or partially-prepared foods are used, the cost in dollars is usually increased.

4. When time and money are both considered in the cost of food, convenience foods may be less expensive than home-prepared foods.

5. When equipment is carefully selected and conveniently arranged, much time and energy may be saved in food preparation.

6. When planning menus, consideration should be given to the food already on hand, food available to be purchased, the time needed to prepare the food and the cost of the food.

7. When preparing a meal, a work schedule, either mental or written, will aid in timing so that all foods will be ready for use as desired.

8. If food served is nourishing but unappetizing, nutritional deficiencies may occur because the food is not eaten.

Criteria for Constructing Multiple-Choice Questions

Authorities generally agreed that multiple-choice items were a desirable form of objective question. Steelman (25, p. 25) devised a set of criteria that were used in the revision of this pretest. These criteria are as follows:

1. The lead or stem may be in the form of a direct question or an incomplete statement.
   a. If an incomplete statement is used, it should be meaningful in itself and imply a direct question.
   b. In general, the incomplete sentence form seems to provide greater economy of language.

2. As much of the item as possible should be included in the lead.

3. The lead should generally be stated in positive terms. If negative items are used they should be underlined.

4. A single definite problem should be presented in the lead.

5. Dangling participles or gerunds should not be used in the lead.

6. Each item in a test should be independent of other items.

7. All of the alternatives should follow both plausibly and grammatically from the statement of the problem.

8. The answers should be parallel in grammatical form.

9. The alternatives should be made as brief as possible.

10. The alternatives should be similar.

11. The distractors should be plausible.
    a. Make distractors as familiar as the correct answer.
    b. Relate to the same concept as the correct answer.
    c. Make as reasonable and natural as the correct answer.

12. The arrangement of alternatives should be uniform throughout the test. One under the other is the best arrangement.

13. The length of the alternatives should not vary systematically with their correctness.

14. The position of the correct answer should not vary systematically.

15. "None of the above" as an alternative should be used only with definite facts.

16. "All of the above" gives answer if student knows that two answers are right.

17. Compound responses should be avoided.

18. If alternatives contain two pairs of opposites, the members of each pair should appear together to avoid confusing the subject unnecessarily.

19. Irrelevant inaccuracies should be avoided.

20. Unusual vocabulary should be avoided.

Since information regarding individual items in Steelman's pretest was not available, the instrument was re-administered to 36 twelfth-grade girls at Guymon, Oklahoma and 15 twelfth-grade girls at Perkins, Oklahoma, in April, 1967. From an analysis of these data, test items were revised. The plan was to eliminate, or rework, items that had no discriminating value and to increase the number in order to increase the reliability.

By careful analysis of the number of students using each distractor on the test items, the discrimination index and difficulty index, 69 items from Steelman's (25) test were retained. Questions that were out of date were deleted. Revision of questions included improvement of the distractors to make all four distractors equally attractive to those who did not know the correct answer, or rewording the stem for clarity. The questions were then arranged in order from the least difficult to the most difficult as indicated from the difficulty index of each test item. Eighty-nine additional questions were added from the General Mills Betty Crocker Search for the Homemaker of Tomorrow Tests (15). See appendix for letter confirming permission obtained by telephone conversation to use these questions. It was felt that the Steelman test lacked practical situation examples among the test items, therefore, many of the questions added were of this type.

It was desired to have 125 questions with a high degree of validity and reliability remaining in the final test following administration of the test to the experimental group.

Attention was given to the content areas included in the textbook for the course. By checking a content area chart, no areas of food preparation taught in the course were omitted. Tyler's behavioral aspect chart developed by Steelman for the family food course, and included in the appendix, was then checked. This chart included areas for the development of critical thinking, self direction, attitude, appreciation, and values.

Following the item analysis the questions were arranged in order from the least difficult to the most difficult as suggested by many authorities. The final test included 158 items which were typed on multi-lith mats and printed in booklet form to be given to a select group of college undergraduates.

Administration of the Test

The test booklet consisted of 13 pages of questions, an answer sheet stapled to the back and a cover page giving complete directions for administering the test. The directions on the cover page were as follows:

> TO THE EXAMINER: This test is designed to be taken using a separate answer sheet on which the student records his responses. All answers are to be marked on the answer sheet, not written on the test question sheets.
>
> GENERAL DIRECTIONS: Do not turn this page until the examiner tells you to do so. No QUESTIONS may be asked after the examination has begun. You are required to answer ALL QUESTIONS even when you are not perfectly sure that your answers are correct, but you should avoid wild guessing. DO NOT spend too much time on any one item, but DO ANSWER ALL QUESTIONS.

Inside the test booklet further directions included:

> DIRECTIONS: Each of the following incomplete statements or questions is followed by four possible answers. Select the answer that BEST completes the statement or answers the question. On your answer sheet blacken with pencil the space having the same number as your choice.

Students were asked to sign their name opposite the number on their answer sheet. This procedure was used in order that any incomplete test items could be resubmitted to the testee. The student could be informed of his test score. By using numbers, however, the answer sheets remained anonymous.

Prerequisites for the pretest were that: (1) the student be a high school senior or be enrolled as an undergraduate in college, and (2) the student had not taken FNIA 2113 Family Food, and (3) only girls were permitted to take the test since no male students were enrolled in the family food course.

Dr. Harry Brobst, Head, The Bureau of Testing and Measurements, Oklahoma State University, was consulted as to the method to be used for analysis. He suggested references from which techniques for item-analysis were found. These item analyses yielded discrimination and difficulty indices, and a coefficient of correlation for reliability.

Conditions under which this test was taken were less than ideal. Few students were available during the summer term. More reliable information could be gleaned from administering the pretest at the beginning of the fall semester, 1967, to students enrolled in FNIA 2113 Family Food.

Procedure for Test Analysis

Following the actual administration of the pretest, the papers were scored by hand using a duplicate of the answer sheet from which the correct answers had been punched. The scores were then ranked from the least number of questions missed to the greatest number of questions missed. The answer sheets were grouped into three sections: twenty-seven per cent of the scores from the upper section of answer sheets and 27 per cent from the lower section of answer sheets leaving the remaining 46 per cent in the mid-range of scores. Although the number of testees in this case did not present a impractical situation for counting the entire population, authorities agreed that by using the upper and lower 27 per cent of the socres, a reasonable estimate of the validity of the instrument could be obtained.

At this point a program was prepared for using an IBM computer 7040 at the Computer Center, Oklahoma State University. A sample of the form in which the raw data was tabulated for use on the computer is presented here:

Item No. 23 Answer A

	A	B	C	D
H	8	0	0	0
L	5	0	1	2

From this information the computer was programmed to correct each item for guessing and convert the results into a proportion through the use of the following formula from Davis (10, p. 6):

$$P = \frac{R - \frac{W}{K - 1}}{N}$$

where

P = Proportion
R = The number of testees that answer the item correctly.
W = The number of testees that answer the item incorrectly.
K = The number of choices in the item.
N = The number of testees answering the question.

The proportion figure calculated was then used to find the coefficient of correlation (r) in Flanagan's Table (28, p. 348) which had been stored in memory in the computer. Interpolation by the computer for the appropriate correlation coefficient was required if an odd rather than an even proportion was obtained. Thus the validity index was determined.

After extrapolation of an r-value, a table from Davis (10, p. 13) was referred to by the computer for the discrimination index. This discrimination index was an indication of the value of each question as to whether or not the upper level students could be separated from the lower level students. Those items having a discriminating value below 15 were considered unacceptable and were either discarded or reworked. Consideration was given to the way in which the question was stated, the kind of distractor used, and the timeliness of the content.

The high and low proportions for each item from the first calculations were added and divided by two to obtain the percentage of successes for each test question. This figure was then used to find the difficulty index in a second table by Davis (10, p. 38) which had also been stored in memory on the computer. A difficulty index ranging from 25-70 was considered within the "ideal range", however, an index of 50 was considered the most discriminating. These calculations were made on eight test scores at each tail of the distribution in evaluating the experimental sample.

The mean and standard deviation were calculated using all the sample scores. From these figures the reliability of the test was determined using a formula from Garret (14, p. 341):

$$r_{1I} = \frac{n\,t^2 - m(n - m)}{t^2\,(n - 1)}$$

where

r_{1I} = reliability of the whole test.

n = number of items in test.
t = standard deviation of test scores.
m = mean of test scores.

An example of the final analysis from the computer was as follows:

Item No. 24 Answer C

	A	B	C	D	Total	Proportion
H	0	0	8	0	8	1.00
L	1	0	7	0	8	0.83

R 0.50 Dis. Index 32 Dif. Index 79

Item No. 25 Answer B

	A	B	C	D	Total	Proportion
H	0	7	1	0	8	0.83
L	1	5	1	1	8	0.50

R 0.38 Dis. Index 24 Dif. Index 59

Information from the item analysis was condensed into tables by the computer to facilitate use of the above information. Items analyzed were grouped and listed in numerical order according to: (1) item number, (2) coefficient of reliability, (3) discrimination index, and (4) difficulty index.

CHAPTER IV

RESULTS AND DISCUSSION

Analysis of Pretest Scores

The distribution of scores for the 28 students taking the Food Preparation Theory Pretest was arranged in intervals of 5. In Table 1 the frequency distribution is given. The mean of the scores was 95.14 and the standard deviation was 14.01. The reliability was established as 0.83.

TABLE 1 Distribution of scores* from 28 students taking food preparation theory pretest

SCORE RANGE	FREQUENCY
122-126	1
117-121	0
112-116	0
107-111	3
102-106	6
97-101	5
92- 96	5
87- 91	5
82- 86	1
77- 81	0
72- 76	0
67- 71	0
62- 66	1
57- 61	0
52- 56	0
47- 51	1

* Possible score of 158.

The bulk of the distribution fell between 87-111; well above the middle of the scores for the entire group.

Analysis of Test Items

The distribution of discrimination indices was divided into groups of five ranging from 0-100 per cent discrimination. The larger the discrimination index the greater the discrimination value. Indices below 15 were considered non-discriminating. Presented in Table 2 is evidence that 41 per cent of the items were non-discriminating while 59 per cent of the items were discriminating.

TABLE 2 Distribution of discrimination indices based on upper and lower 27 per cent of subjects taking food preparation theory pretest

DISTRIBUTION RANGE	FREQUENCY
100-	4
95-99	0
90-94	0
85-89	0
80-84	0
75-79	4
70-74	3
65-69	4
60-64	0
55-59	0
50-54	10
45-49	13
40-44	0
35-39	12
30-34	13
25-29	0
20-24	23
15-19	7
10-14	31
5- 9	7
0- 4	27

An indication of the difficulty of an item is given in percentage of individuals who can answer the question correctly. The lower the percentage the more difficult the item is and conversely the higher the percentage the less difficult the item. A distribution of the difficulty indices is shown in Table 3. Again the scores were divided into groups of five. Eighty-eight per cent of the items were found to fall between 25-70. Items outside the ideal range of difficulty need not be eliminated provided they are within the limits of acceptable discrimination (10).

TABLE 3 Distribution of difficulty indices based on upper and lower 27 per cent of subjects taking food preparation theory pretest

DISTRIBUTION RANGE	FREQUENCY
96-100	3
91- 95	0
86- 90	0
81- 85	0
76- 80	10
71- 75	0
66- 70	11
61- 65	14
56- 60	19
51- 55	14
46- 50	34
41- 45	14
36- 40	0
31- 35	20
26- 30	13
21- 25	0
16- 20	6
11- 15	0
6- 10	0
0- 5	0

Item analysis of questions from Steelman's (25) pretest permitted questions such as the following to be revised.

> Egg yolk should not be in egg whites to be whipped because the
> 1. Iron interferes with whipping
> 2. Vitamin A interferes with whipping
> 3. Sugar interferes with whipping
> 4. Fat interferes with whipping

This question had a discrimination index of zero and a difficulty index of 44 following analysis of the sample data of the upper and lower 27 per cent of students in a group of 35 high school students. Two students in the upper 27 per cent chose distractor (1), two chose distractor (2), none chose distractor (3), and five chose distractor (4) which was the correct answer. A proportion of .41 was calculated using Davis' (10) correction-for-guessing formula. Students in the lower 27 per cent answered the question accordingly: two chose distractor (4), two chose distractor (2), none chose distractor (3), and five chose distractor (4). A proportion of .41 was again obtained.

The question was revised in the present test to read:

> Egg whites to be whipped will not reach maximum volume if any egg yolk is allowed to enter because (1) iron interferes with whipping. (2) vitamin A interferes with whipping. (3) sugar interferes with whipping. (4) fat interferes with whipping.

Analysis of the item following the second administration revealed that none of the students in the upper group answered distractor (1), two answered distractor (2), none answered distractor (3), and six answered distractor (4) giving a proportion of .67. Students in the lower group answered in the following manner: three chose distractor (1), two chose distractor (2), none chose distractor (3), and three chose distractor (4). The proportion according to Davis' formula was .17. The

discrimination index this time was 35 and the difficulty index was 46. Although one would expect variations in indices because of the difference in sample populations, the difficulty of the item was not significantly altered. The discrimination index, however, was improved considerably.

CHAPTER V

SUMMARY AND CONCLUSION

A revision of Steelman's (25) theory pretest for a beginning course in food preparation was completed at Oklahoma State University during the Summer Session, 1967. Objectives and generalizations devised by Steelman were utilized. The major emphasis of this study was placed on revision of and addition to items in the pretest, and upon incorporation of computerized analysis of the results.

The final test included 158 test items carefully selected from a combination of Steelman's test (25) and the Betty Crocker Search for the Home Maker of Tomorrow tests (15). These test items had been analyzed from sample data and changes were made when analyses indicated the need for revision.

The theory pretest was administered to 28 selected undergraduate students at Oklahoma State University during the Summer Session, 1967. The reliability of the test was 0.83. Fifty-nine per cent of the items were discriminating and 88 per cent of the items were within the ideal range of difficulty.

One of the purposes for which the revision was planned was to increase the validity of the instrument. As indicated by the lower discrimination index, this was not accomplished. The reliability remained the same. However, the ideal range in difficulty was markedly increased. One factor responsible for the difference in difficulty range

was that the lower limit of difficulty was decreased such that 25 per cent instead of 30 per cent of the range was accepted as too difficult. Other factors which undoubtedly influenced the results were: the composition of the test group; the inability to standardize the administration of the test due to the time of year; the increased length of the test which made the time required for its completion beyond the limits of productivity; and variation in the type of question included.

While considerable time was spent counting the number of answers for each distractor, less time was spent in making calculations. Greater speed and accuracy were achieved through use of the computer than from hand calculations. It was found to be most advantageous to list the computerized test results in numerical tables. Two further calculations, those of tabulating the frequency distribution within each group of scores and the percentage of scores falling within the specified range, could have been included in the program.

The revised evaluation instrument in its current form required approximately two hours to complete. Since no time limit for finishing was set, all students were encouraged to complete the pretest. By careful selection, it is planned to reduce the number of questions to a fit a 50-minute class period. In the beginning 125 questions were thought to be an ideal number of items to include in such an evaluation instrument. However, indications are that only 100 items of the type included in this test could be completed in one class period.

It is recognized that the number of subjects used in administering this test was very small. Any deductions from the results should be considered in light of the size of the group. It is recommended that the pretest in its present form be administered to all students enrolled

in FNIA 2113 Family Food in the fall of 1967. By item-analysis of the data collected at that time, questions that failed to meet the requirements of discrimination and difficulty could be deleted. A refined test of desirable length should remain.

BIBLIOGRAPHY

(1) Ahmann, J. S., and Glock, M. D.: Evaluating Pupil Growth. Boston: Allyn and Bacon, 1963.

(2) Arny, C. B.: Evaluation in Home Economics. New York: Appleton Century-Crofts, Inc., 1953.

(3) Barbour, H. F.: Relationships of values and process concepts of selected students to generalizations in nutrition. Unpublished Ph. D. Dissertation. Ames: Iowa State University Library, 1953.

(4) Bogniard, J. N.: Examinations that educate. Journal of Home Economics. 51: 771, 1959.

(5) Bonde, R. L.: A national testing program. Journal of Home Economics. 41: 582, 1949.

(6) Buros, O. K., ed.: The Fifth Mental Measurements Yearbook. Highland Park, N. J.: Gryphon Press, 1958, p. 568.

(7) Collins, M. H.: A pretest for placement in beginning clothing courses at Southern Illinois University. Unpublished Research Report. Carbondale: Southern Illinois University Library, 1956.

(8) Committee on Measurement and Evaluation: College Testing. Washington, D. C.: American Council on Education, 1959.

(9) Cooksey, D. C.: Pretest in beginning college nutrition based on objectives, concepts, and generalizations. Unpublished M. S. Thesis. Stillwater: Oklahoma State University Library, 1964.

(10) Davis, F. B.: Item-analysis Data. Cambridge, Mass.; Graduate School of Education, Harvard University, 1949.

(11) Eppright, E., Pattison, M., and Barbour, H.: Teaching Nutrition. 2nd ed. Ames: Iowa State University Press, 1963.

(12) Flanagan, J. C.: General considerations in the selection of test items and a short method of estimating the product-moment coefficient from the data at the tails of the distribution. Journal of Educational Psychology. 30: 674, 1939.

(13) Food, Nutrition and Institutional Administration Faculty: Food Preparation Manual. Stillwater: Oklahoma State University, 1962.

(14) Garret, H. E.: Statistics in Psychology and Education. 5th ed. New York: Longmans, Green and Company, 1958.

(15) General Mills: Betty Crocker Search for the American Homemaker of Tomorrow. Knowledge and Attitude Test. Chicago: Science Research Associates, 1959-67.

(16) Gorow,, F. F.: Better Classroom Testing. San Francisco: Chandler Publishing Company, 1966.

(17) Gould, G. F.: A performance pretest for placement of college students in beginning clothing courses. Unpublished M. S. Thesis. Stillwater: Oklahoma State University Library, 1963.

(18) Hall, O. A., and Paolucci, B.: Teaching Home Economics. New York: John Wiley and Sons, 1962.

(19) Hill, Mary M.: A conceptual approach to nutrition education. Journal of the American Dietetics Association. 49: 20, 1966.

(20) Hoover, H. M.: Levels of conceptual understanding. Journal of Home Economics. 59: 89, 1967.

(21) Hoskins, M. N.: Construction of a basic clothing pretest for use in the colleges and universities in New Mexico. Unpublished M. S. Thesis. University Park: New Mexico State University Library, 1959.

(22) Hughes, O.: Introductory Foods. 4th ed. New York: The Macmillan Company, 1962.

(23) Lee, V. T.: The establishment of criteria for placement in college foods courses. Unpublished M. S. Thesis. Carbondale: University of Southern Illinois Library, 1963.

(24) Kelly, T. L.: The selection of upper and lower groups for the validation of test items. Journal of Educational Psychology. 30: 17, 1939.

(25) Steelman, V. P.: Development of an objective written and laboratory pretest based on aims and generalizations for a beginning college food preparation course. Unpublished M. S. Thesis. Stillwater: Oklahoma State University Library, 1961.

(26) Stovall, R.: Education in a world of change. Journal of Home Economics. 54: 537, 1962.

(27) Tinsley, W. V., and Sitton, M.: Teaching intellectual aspects of home economics through the identification of basic concepts. Journal of Home Economics. 59: 85, 1967.

(28) Thorndike, R. L.: Personal Selection, Test and Measurement Techniques. New York: John Wiley and Sons, 1949.

(29) Tyler, R. W.: Basic Principles of Curriculum and Instruction. Chicago: The University of Chicago Press, 1950.

(30) Witt, M. R.: The revision and development of selected evaluation devices for appraising certain clothing competencies of college freshman. Unpublished Ed.D. Dissertation. Stillwater: Oklahoma State University Library, 1961.

APPENDIX

Food, Nutrition and
Institution Administration
Oklahoma State University
Stillwater, Oklahoma
July 28, 1967

Mr. A. L. Champlin
Assistant Public Relations Director
General Mills, Inc.
Minneapolis, Minnesota

Dear Mr. Champlin:

This letter is to confirm the telephone conversation you held with Dr. Helen F. Barbour on July 21, in which permission was granted to reproduce questions from the Betty Crocker Search for the Homemaker of Tomorrow Tests from 1959-1967.

This information is being used in the development of a pretest for the FNIA 2113 Family Food course at Oklahoma State University.

Thank you for your kind consideration.

Very truly yours,

Mrs. Charline H. White
Graduate Student

Helen F. Barbour
Head, FNIA Department

CHW/pc

BEHAVIORAL-CONTENT ASPECT CHART OF OBJECTIVES

Content Aspect of the Objectives	Behavioral Aspects of the Objectives						
	Understands	Applies	Thinks Critically In Relation To	Directs Self In	Develops Attitudes	Develops Values	Appreciates
1. Food in relation to:							
A. Health for self and others	X		X	X	X		X
B. Diverse patterns through which adequate nutrition may be met	X	X	X		X	X	X
C. Social values	X	X	X	X	X		X
2. Contribution of processed foods to present day living	X		X	X	X		X
3. Terms used in handling and preparing food	X	X	X				
4. Sanitary practices in handling and preparing food	X	X	X	X		X	X
5. Psychological, economic, and social effects of food advertising in our society	X		X		X		X
6. Benefits which agencies provide the consumer	X		X	X			X
7. Responsibilities consumers have for supporting and informing agencies	X		X	X	X		X
8. Information which appears on packaged products	X	X	X				X
9. Body current scientific knowledge of food processing and preparation	X	X	X	X			X
10. Food selection, preparation, and service as a means of creative expression	X	X	X	X	X		X

Content Aspect of the Objectives	Behavioral Aspects of the Objectives						
	Understands	Applies	Thinks Critically In Relation To	Directs Self In	Develops Attitudes	Develops Values	Appreciates
11. Wide use of food materials available	X	X	X	X			X
12. Facts and principles of food selection, storage, preparation and service	X	X	X				
13. Selection of meals for self and others	X	X	X	X	X	X	X
14. Management of time, energy, and other resources in food selection, storage, preparation and service	X	X	X	X	X	X	X
15. Equipment needed in food preparation and service	X	X	X				
16. Ways in which foods are prepared	X	X	X	X			X
17. Desirable combinations of foods	X	X	X	X	X		X
18. Desirable standards in food preparation	X	X	X			X	X
19. Worth of one's own food practices	X	X	X		X		X
20. Contribution of food and its service in fostering family relationships	X	X	X	X	X	X	X
21. Pleasures which may be found in eating food	X				X	X	X
22. Evaluation and acceptance of new foods		X	X	X	X		
23. Knowledge food research provides and its contribution to the continued improvement of health	X					X	X

VITA

Charline Hyer White

Candidate for the Degree of

Master of Science

Thesis: REVISION OF A THEORY PRETEST BASED ON OBJECTIVES AND GENERALIZATIONS FOR A BEGINNING COLLEGE FOOD PREPARATION COURSE

Major Field: Food, Nutrition and Institution Administration

Biographical:

Personal Data: Born in Strong City, Oklahoma, January 30, 1920, the daughter of Herbert C. and Winnie D. Hyer.

Education: Graduated from Guymon High School, Guymon, Oklahoma in 1938; received a Bachelor of Science degree in Home Economics Education at Oklahoma State University in 1942; attended Panhandle State College, Goodwell, Oklahoma, during the summer terms, 1957-59; attended Colorado State University, during the summer term, 1960; completed requirements for the Master of Science Degree in Food, Nutrition and Institution Administration at Oklahoma State University in August, 1967.

Personal experience: Instructor of Vocational Home Economics, Sayre, Oklahoma, 1942-45; County Home Demonstration Agent, Collingsworth County, Wellington, Texas, 1945-46; instructor in elementary education, Gruver, Texas, 1956-58; instructor in elementary education, Guymon, Oklahoma, 1958-59; and instructor of Vocational Home Economics, Guymon, Oklahoma, 1959-66.

Member of Omicron Nu, Oklahoma Education Association, National Education Association, Oklahoma Home Economics Association, and the American Home Economics Association.

Typist: Mrs. Phyllis Carruth