A Study of Answers to True False Test Questions Marked to Indicate Confidence in Correctness

ALMA LONG, Purdue University

Continuing a three-year study of an orientation test dealing with Home Economics subject matter and principles applicable to general practices in every day home making, a study has been made of values obtained from questions of information in the correctness of which the student has expressed confidence. Some aspects of student behavior other than evidences of knowledge of subject matter may have important relationship with success in a particular field of work. Also, further measures of reliability of objective tests are needed before pupils can be classified accurately in terms of essential ability in the subject by means of test scores.

Previous analyses of results of the test given to students entering Home Economics departments of state schools in Indiana gave some evidence of a high degree of reliability of scores obtained in repetition of the test. There are also reasonably high coefficients of correlation between college aptitude and the home economics scores of groups classified by amount of previous training, although these correlations are not high enough to justify the use of the home economics scores in predictions for individuals.

Weighted average values have been placed on each of the one hundred forty questions dealing with information and the application of principles included in the test. Differences in scores on the various questions earned by various groups indicate some important relationships between the kind of training the pupil has had in home economics and the kind of subject matter most familiar to the classified groups subjected to the test.

The purpose of the most recent phase of this study has been to find more accurate indications of the contribution such a test can make toward the identification of those pupils whose training and experience in home economics vary. Three hundred and two papers collected in September, 1941, were studied with this in mind. At the time the test was given, the pupils were asked to indicate confidence by "double checking" any answers which they felt were correct. The papers were then scored, disregarding all indications of confidence, and again taking account of double checked answers. The original scores on questions were very similar to scores obtained from similar groups in previous years.

The general plan for scoring true-false questions which had previously been used, that is rights minus wrongs, was applied to the additional scores obtained from double checking. By addition of correct double checked answers the range of scores was increased very consid-

erably, although many scores at the lower end of the distribution changed very little. Not only the mean of the double check distribution but also the standard deviations were considerably higher than those of the original scores.

Numerous questions are raised by the amount of change which took place and by the differences in gain by individuals, due to confidence scores:

1. To what extent would increases in score be a function of superior intelligence, or of superior training and experience?
2. Would all pupils in the superior intelligence or training levels tend to respond in the same general manner?
3. To what extent would personality traits such as daring or conservatism promote or hinder the indication of answers thought to be right? In some instances would the device promote recklessness which might be costly to the pupil?
4. Would recency of training or much experience contribute greatly to the accuracy with which correct answers would be indicated with confidence?
5. To what extent would practice in any phase of home economics without formal training tend to give pupils an accurate knowledge of some of the principles and practices which form rather basic parts of a home economics education program at the high school level? Would sureness scores result in penalizing some students?
6. What proportionate parts of the final score would be indicative of superior ability, interest in the subject or long practice?

Complete answers to these questions would extend far beyond the analysis of these test scores. There were, however, some well defined groupings and dispersions of scores which may assist somewhat in our interpretation of a value to be derived from a test situation dealing with information, principles and practices in home economics education.

Differences in total scores occurred which tend to coincide with what might reasonably be expected of groups of individuals whose amounts of training in home economics at the high school level varied. Previous analyses of scores on this same test indicated with a high degree of consistency that pupils who had participated in Four-H Club work scored somewhat better on the test, other factors held constant, than those who had not had such practice.

Indications of confidence in correct answers resulted in much greater average gains for those who had had Four-H Club experience than for other groups. Their average gains were greater than the average gains of other groups, although their average scores on college aptitude tests were somewhat lower. The extent of these gains on the home economics test and the nature of the questions upon which a particular group

distinguished itself in this recent test probably expresses some important and fundamental differences between the responses of a specifically trained and experienced young home economist and the one who gains her proficiency through general practice at home or from general reading of publications which deal with modern home making.

When the scores were arranged in percentile rank for (a) the original score and (b) the new score obtained by double checking, high coefficients of correlation between the distributions of the original and the double checked confidence scores were shown. These were .81 ± .036 for students who had no home economics in high school, .89 ± .018 for those who had some home economics but no Four-H Club work and .82 ± .31 for those who had both Four-H Club work and training in high school. In each of the groups there were individuals whose original and confidence scores were low. There was a noticeable amount of shifting of some individuals whose original scores were in the middle categories, some of which resulted in significant re-location of individuals who had much experience and training in home economics. The results of the changes in score are shown in Tables I and II.

In previous analyses of the test data, the coefficients of correlation between college aptitude rank and score on the home economics test ranged from .12 ± .005 for those who had no home economics in high school to .355 ± .003 for those who had several semesters. From this it appears that success in the home economics test is somewhat less related to general intelligence scores as represented by college aptitude test scores than to some specific training and experience in a technical subject.

The matter of indicating confidence in answers to test questions appears to be a complex of specific information, native intelligence and probably also some basic emotional attitudes toward a probability or chance. In a few cases, individuals in the lowest category of intelligence ranking indicated confidence in all of their answers, right or wrong, resulting in a minimum of gain for themselves. On the other hand a few individuals in the higher intelligence category made no errors in double checking and consequently increased their scores. The largest and most consistent gains were made by individuals in the upper fifty percentiles of college aptitude scores who had home economics in high school plus some Four-H Club work.

To indicate real differences between groups having different backgrounds of training, scores which were more than one sigma from the mean were studied. These superior scores were earned by 9.9% of the entire groups, who had home economics training plus Four-H Club work; 6.6% of the group, who had no Four-H Club work, and 3.3%, who had no previous training in home economics. Those found in the second or third sigma below the mean include 1.3% of the entire group, who had home economics training plus Four-H work, 2.9% of the population, who had no Four-H Club work and 6.9%, who had neither home economics training nor Four-H Club work in high school.

TABLE I.—Differences in distribution of scores of groups having different amounts of training in home economics, and whose scores were more than one sigma from the mean. Original scores and those corrected for "Double Checking." Distributions exclude those who chose not to double check any answers.

Total		Original Scores				Corrected Scores			
		Above		Below		Above		Below	
N		N	%	N	%	N	%	N	%
117	No Home Ec. in H.S.	17	14.4	20	16.9	10	8.5	21	17.9
84	H. Ec. in H.S., No 4-H	14	16.4	21	24.6	20	23.8	9	10.7
101	H.Ec. in H.S. plus 4-H	8	8.0	20	20.0	30	29.7	4	3.9

The above changes in scores shown in Table I resulted in significant differences in the proportions of individuals in the higher and lower sigmas of the distribution. Table II shows the changes arranged according to gain or loss in proportions found in the higher and lower sigmas.

TABLE II.—Changes in size of groups whose confidence scores were one sigma or more from the mean.

Above mean	Direction	Crit. Ratio
No Home Economics in High School	Loss	1.1
Home Economics in H.S.—No 4-H Club	Gain	2.3
Home Economics in H.S. plus 4-H Club	Gain	3.2
Below mean		Crit. Ratio
No Home Economics in High School	Gain	0.0
Home Economics in H.S.—No 4-H Club	Loss	2.4
Home Economics in H.S. plus 4-H Club	Loss	3.6

In summarizing the results of the investigation, it may be said that an opportunity for scores to be increased by the indication of confidence in the answers given to true-false questions, seems to tend quite definitely toward indicating some relationships between higher scores and greater amounts of training. Those individuals who obviously have had little specific training in the field tended to receive the lower test scores regardless of any advantages of superior general intelligence. Some students having low general ability tended to remain in the lower categories of score regardless of their many indications of confidence in correctness. Others in the lower intelligence categories obviously expressed themselves as being very unsure and would probably remain in those ranks excepting as they applied themselves with continued effort to gain competency in the work.

In the middle groups, many individuals had learned much and were sure of many items. These people represent a group whose efforts and interest would seem to combine toward real achievement. In the upper intelligence levels, there were some who had benefited by superior training and who consequently achieved a great deal. They seemed to tend to check relatively few items for confidence. They also made relatively few errors in their confidence scores.

It is doubtless important to recognize the need for accurate knowledge and to refrain from rash statement. The effect upon total scores resulting from the various attitudes of individuals could be, with the extremes of temperament operating, within the limits of chance distributions of score, to indicate some essentials of achievement. Refinements of scoring serve to distinguish between those who secure and maintain high rank, and those who would naturally slip through the sieve because of lesser ability or amount of training. Still further refinements of method for securing test results, which would take better account of different personal reactions to the situation, may be well within the realm of possibility. Assurance of greater reliability in answers to true-false questions apparently can be achieved to a very useful extent by the device of indications of confidence. The multiple factors underlying student reaction to the situation, if known, might contain information of great importance for the estimation of an individual's probable success in a specific type of education.