ON THE USE OF OBJECTIVE EXAMINATIONS

C. H. COOMBS

University of Michigan

Introduction

IT seems to be a common experience of individuals taking objective tests to feel confident about eliminating some of the wrong alternatives and then guess from among the remaining ones. This procedure is usually encouraged, as the odds are in the individual's favor. A procedure for administering and scoring objective tests which will discriminate levels of partial knowledge and will yield a greater variance of test scores than the conventional procedure for the same number of items is proposed here.

There is a psychological rationale underlying this proposed procedure. It is assumed that partial knowledge exhibits itself in recognizing some of the wrong answers. Complete knowledge, knowing what is right, is equivalent to knowing everything that is wrong. Misinformation is distinguished from partial information in the individual's belief that the right answer is wrong. Objective tests constructed with one right alternative and the others wrong can be administered and scored so as to provide on each item, a scale from complete misinformation through several degrees of partial information to complete information.

Suggested Procedure

Individuals taking the test should be instructed to cross out all the alternatives which they consider wrong. An individual who knows the right answer will cross out all the wrong alternatives. An individual not knowing the right answer may, however, have sufficient partial knowledge to recognize that one or more of the wrong alternatives is wrong. Such an individual should be instructed to cross out those alternatives he believes to be wrong but *not* to guess among the remaining

ones. He should, in effect, exhibit just what knowledge he has and is sure of. The scoring procedure used is such that the odds are against him if he guesses.

The weights used in the scoring procedure are dependent on the number of wrong alternatives. Each wrong alternative crossed out is worth a unit of credit. But if the right alternative is crossed out this is given a weight of 1–k where k is the total number of alternatives. Thus the possible scores and response patterns for a 4 alternative item are given in Table 1. It is evident that a four alternative item could yield a scale from

TABLE 1

Response Patterns and Scores on a Four Alternative Item

Alternatives	Response Patterns and Scores**															
a*									X	X	X	X	X	X	X	X
b	X	X	X		X				X	X	X		X			
c	X	X		X		X			X	X		X		X		
d	X		X	X			X		X		X	X			X	
Item score	+3	+2			+1			0		−1			−2			−3

* Correct alternative.
** An X indicates that the alternative is crossed out.

−3 to +3, a seven point scale, instead of the usual two point scale at +1 and −⅓.

It is interesting to note that this procedure constitutes a generalization of the true-false test scored R–W and without scoring omissions.

General Comments

1. The labor involved in scoring is at least doubled, as the test will have to be scored for the number of right alternatives crossed out and the number of wrong alternatives crossed out.

2. No correction for guessing is involved. Each individual sets up his own standard of assurance. This source of error is not present in procedures which require the individual to answer every item. In the latter case random error variance is substituted for the above standard-of-assurance variance. The random error variance, however, would tend to decrease with increasing number of items. The standard-of-assurance variance

would not. In this respect it is of interest to note the article by Cronbach (1) comparing multiple multiple-choice (variable number of correct alternatives, and subject is instructed to mark all the correct ones) with multiple true-false (variable number of correct alternatives, and subject is instructed to mark each alternative true or false). These bear a certain resemblance to the procedure proposed here, but only superficially. Both the multiple multiple-choice and the multiple true-false are essentially indistinguishable from a true-false test. One of the virtues of the conventional multiple choice is to reduce the probability of getting an item right by guessing and thereby reducing the error variance. In the multiple multiple-choice and the multiple true-false this component of the error variance is increased.

3. This procedure is applicable only to items with one right alternative and the remaining alternatives wrong. However, it would probably not yield anything different in a test such as an arithmetic test in which an individual would *compute* an answer and then cross out all the other alternatives.

4. Indices of item difficulty and discrimination could be constructed by generalization of many of the existing indices.

5. As four or five alternative items would have a seven or nine point scale respectively, product moment correlations between items, or between items and appropriate criteria, are feasible.

6. The first time students are exposed to this method they should be given a clear account of the procedure, including a copy of Table 1. The danger of guessing should be pointed out to them. A small scale tryout of the procedure at the University of Michigan indicated that the majority of the students were favorably inclined.

7. Ultimate evaluation of the procedure is dependent on empirical and comparative studies of reliability and validity for various purposes.

REFERENCE

1. Cronbach, L. J. "An Experimental Comparison of the Multiple True-False and Multiple Multiple-Choice Tests." *Journal of Educational Psychology.* XXXII (1941), 533-543.