A WRITTEN TRADE TEST: MULTIPLE CHOICE METHOD

By J. CROSBY CHAPMAN, Western Reserve University; and
HERBERT A. TOOPS, Teachers College, Columbia University

The Army Trade Test Division was, on a *priori* grounds, extremely sceptical of written trade examinations. For this reason, in the theoretical examination, an oral test was used. The oral method has proved its power to distinguish, with a fairly high degree of certainty, between novices, apprentices, and journeymen, and to a less degree between journeymen and so-called experts. The possibilities and limitations of one type of written test form the subject of this article. In connection with the work of this Division an opportunity was given of trying out a written trade examination constructed on the multiple-choice plan. There are many advantages in being able to give group examinations of this kind. Among these are: (1) conservation of examiner's time, (2) reduction of scoring time and increase of scoring accuracy through the use of stencils, etc., (3) elimination of subjective judgments, (4) the possibility of rapidly eliminating the trade "bluffer," leaving only those who show evidence of some knowledge for further consideration or examination by means of oral, performance, or picture tests.

ASSEMBLY OF QUESTIONS

In conference with expert tradesmen, sixty-two questions were prepared, each question having four possible answers, of which one was correct. Nearly all the questions of the army oral test for Bricklayers, partly collected by the authors, were adapted to this form of examination, and the remainder were secured from other sources. Trade terms were used almost exclusively in the four answers from which the choice was to be made, on the supposition that a person might be able to recognize a trade term from among a number of terms not applicable to the trade and yet have no real knowledge of its significance. A part of the test is reproduced.

A WRITTEN TRADE TEST: MULTIPLE CHOICE METHOD 359

BRICKLAYERS' TEST*

Directions:

In the sentences below, there are four choices for a correct sentence. Only one of the four choices is correct. Draw a line under the *one* choice which makes the best sentence. If you are not sure, *guess*; an omitted answer will count as a wrong answer in determining your score. The first two sentences, A and B, are answered correctly as a sample for you. Read them carefully, then go on to question 1, question 2, and so on to the end of the list.

Samples:

- A. A half of a brick is called a: chunk, block, heel, bat
 B. Fire-bricks are laid in: concrete, cement, fire-clay, mortar.

Questions:

1. The top course of stone on a wall is called coping, bond-stone, clip-course, capstone.
2. A brick set on end is called. upright, soldier, row-lock, stud.
3. Before plumbing up a corner, you should lay: 3-courses, 6-courses, 9-courses, 12-courses.
4. A brick that is set on the narrow edge is called: stretcher, oarlock, row-lock, header.
5. In coming to a height, if there is a course of brick difference in the level, you would call it a: haunch, filler, line-level, hog.
6. Filling-in the space between a front and back course of brick is called: slushing-up, tempering, tuck-pointing, plastering.
7. Bricks laid across the wall reaching to the face are called: stretchers, headers, layers, fillers.
8. A fire stop around a flue is formed by a coping, skew-back, corble, indent.
9. If you have no metal ties in plain bond work, the front work of a building is tied-in with: king-course, clip-course, projection-course, belt.
10. In backing a 4-inch ashlar stone wall, you would tie it with: tie-joint, binder, brace, anchor.
11. A brick cut in half lengthwise is called. filler, soap, bat, chunk
12. To keep the line level in the middle of a long wall you use: level, plumb-line, square, trigger.
13. In cutting a brick you would use a: chisel-bar, striker, brick-set, cutter.

*These eighteen questions were selected from the original set of sixty-two, which constituted the examination. This selection was made on the basis described in the text.

14. The distance above the frame for a lintel should be: $\frac{1}{4}$ -inch, $\frac{3}{4}$ -inch, 1-inch, 1 $\frac{1}{2}$ -inches
15. An ordinary stone sill should be bedded on the: front, back, solid, two ends.
16. To locate a building you use: water-level, sighter, transit, line-level.
17. A skutch is used in cutting: cement, glazed-brick, stone, fire-brick.
18. Another name for a blind header course is: belt-course, clip-course, dental-course, hidden-course

The answers to the above questions with the percentages of experts passing each are as follows:

1. Coping—98%	7. Headers—98%	13. Brick-set—100%
2. Soldier—98%	8. Corble—85%	14. $\frac{1}{4}$ -inch—88%
3. 3-courses—88%	9. Clip-course—100%	15. Two-ends—93%
4. Rowlock—100%	10. Anchor—85%	16. Transit—88%
5. Hog—100%	11. Soap—85%	17. Fire-brick—100%
6. Slushing-up—98%	12. Trigger—93%	18. Clip-course—93%

THE SUBJECTS OF THE EXPERIMENT

The examination was given without time limit to 101 persons, distributed as follows: 35 novices without trade experience, 4 apprentices, and 62 trade union members. Of the novices, 29 were soldiers in an army training school, being instructed in other trades at the time, but inexperienced in the bricklayers' trade. These men were mechanically inclined, and above average in intelligence and education. It is obvious, therefore, that the questions were subjected to much more intelligent guessing than ordinary conditions, under which they would be employed, would furnish. Of the other six novices, 3 were stenographers who had several days experience, immediately preceding this test, in taking and transcribing stenographic notes relating to the general subject of bricklaying. The other 3 novices were clerical workers in the Department of Psychology of Carnegie Institute of Technology. These subjects were well informed concerning the technique employed in assembling such tests, and used various methods of "suggestion, inspection, and elimination" in efforts to make as high scores as possible.

Owing to the great scarcity of apprentices, only four could be examined. Their length of trade experience ranged from one month to five years. In view of experience and performance in the army oral test, the lowest apprentice can be con-

A WRITTEN TRADE TEST: MULTIPLE CHOICE METHOD 361

sidered little better than novice, while the best apprentice perhaps deserves to be ranked as a low journeyman.

The tradesmen, being all members of the union, more than fulfil the minimum entrance requirements in point of years of trade experience and skill. On the basis of their own written statements in regard to years of trade experience, type of work, and foremanship ability, these 62 men were divided by the authors into the two following groups. (a) 36 journeymen (hereafter called J.) or men of medium amount of skill and knowledge; (b) 26 experts (hereafter called E), men of superior skill, knowledge and foremanship ability. This rather arbitrary classification was made in an effort to determine from the results of the examination whether a set of questions could be so selected as to differentiate not only non-tradesmen from tradesmen, and apprentices from skilled men, but also journeymen from journeymen experts, so divided

PRESENTATION OF RESULTS

Four points were allowed for each correct answer, there were no partial credits. Four points was selected as an arbitrary value because of army custom. A graph for each question was drawn, showing the average score made by novices, apprentices, journeymen and experts. Samples of such graphs are shown in fig. 1. Since little was known as to the most desirable combination of questions to make up the best set for an examination, five methods of combining varying numbers of questions were tried; only two of these methods are here described.

Method I is the entire set of 62 questions.

Method II is a selection, from the original 62, of 18 questions, chosen on the basis of the combined judgments of three persons experienced in similar selection for the army oral trade tests. The requirements for a question in the army tests were: (1) increasing average scores from class to class, novices to experts; (2) a fairly low average novice score; (3) a comparatively high average expert score. The ideal type of question for the choice-method is shown by the straight line of fig. 1.

The distribution of scores by Method I are given in Graph I. By inspection, arbitrary lines of demarcation of N-A, and A-J may be made, corresponding to the *number of questions correct* as follows:

Fig.1. Some Graphs of Questions.

Novice score	0-30
Apprentice score	31-42
Skilled tradesmen	43-62

It can be seen at once that the test does not differentiate between the journeymen and expert classes. Either the division into these two classes by judgments, based on the men's written statements, has nothing to do with the men's ability to answer this form of question, or the test as given is not sufficiently difficult to differentiate among skilled tradesmen. The homogeneity of the skilled tradesmen class as to scores can be easily seen from Graph I, where the expert distribu-

A WRITTEN TRADE TEST: MULTIPLE CHOICE METHOD 363

tion has been added to the journeyman distribution by means of the dotted squares

With the two arbitrary critical lines just mentioned, only three skilled union men out of 62, or less than 5 per cent, are displaced out of this larger class into the apprentice class. We can assume that all union men have had four years trade experience and possess a certain minimum requirement of ability. The apprentice, with only one month's experience, is here rated with the novices, as is only fair under the conditions. His guessing ability is as poor as the poorest novice, from which we can assume that he has learned next to no trade terms in one month of trade experience. Our own judgment is to the effect that intelligence is a better single measure of trade skill than are years of trade experience, after a certain minimum time for learning has elapsed.

By chance alone, the poorest novice should theoretically be able to answer 15 of the 62 questions correctly (one out of every four). Here, only two novices fail to come up to this standard, and they fail by only a single question each, both answering 14 questions correctly. The average novice answers a little better than 19 questions.

If the object is merely to separate skilled tradesmen from the apprentices, or the tradesmen of experience from the "trade bluffers," the complete set of questions of Method I is a remarkably accurate instrument for so doing. For this purpose, with the persons tested, it is at least 95 per cent efficient.

Method II, which is the selection of the best 18 questions according to the usual army criteria, gives us, with the arbitrary lines of demarcation adopted in Graph II, a slightly better separation of the three classes—novices, apprentices, and skilled union workmen. The one apprentice, with one month of trade experience, is still placed with the novices, but is now better than the average novice, and only one skilled union tradesman (a journeyman) is rated as an apprentice. To make the separations indicated, the test is 98 per cent efficient. This would probably be lowered if we had a fair sampling of apprentices.

Nevertheless, in determining the efficiency of a test for industrial purposes, it seems desirable to weigh the classes according to their frequency of occurrence in industry; if there is only one apprentice for every 25 skilled men in industry, an employment manager would be unable to make many mistakes by hiring apprentices under the impression that they were skilled men.

In industry, also, it seems desirable to measure the efficiency of test methods by taking into account the amount of damage caused by misplacement of a man by the test, some such measure as "dollar displacements," as suggested by R. W. Kelly. For it is a greater industrial mistake to give a journeyman or expert rating to an apprentice than it is to rate a novice as an apprentice; this may be seen from a consideration of the relative amounts of responsibility and wages involved.

From Graph II, Method II, we note that there is a slightly better differentiation of J. and E. than in Method I, but not sufficient to justify any arbitrary line of division between the two classes. If a line were drawn at 60-64, only two experts would be classed as J. while over half of the J. would be classed as E.

The final selection of questions, used in Method II and percentages of experts answering each correctly, are given in the

A WRITTEN TRADE TEST: MULTIPLE CHOICE METHOD 365

preceding sample. The average scores and average deviations of the trade classes in both sets of questions are shown in Table I.

TABLE I
AVERAGE SCORES AND AVERAGE DEVIATIONS THEREFROM, BY METHODS
AND BY TRADE CLASSES

Trade Class	Method		No. of Cases
	I	II	
Novices	78 3 ±12 5	12 5 ±5 1	35
Apprentices	128 5 ±37 5	37.0 ±8 5	4
Journeymen	206.4 ±12 7	63 3 ±5 9	36
Experts	207 4 ±14 3	66 8 ±3 9	26
Possible Score	248	72	Total 101

SUMMARY

(1) A written trade examination, based on the multiple choice answer method, was given, without time limit, to 35 novices, 4 apprentices, and 62 journeymen

(2) Two combinations of questions were tried—the complete set of 62 and a specially selected series of 18.

(3) Both combinations, with a very high degree of accuracy, separate the novices, apprentices and journeymen, but neither of them differentiate between the journeymen and the so-called experts, although the eighteen-question set is slightly more efficient.

(4) In the absence of extreme language difficulty, the written group examination can be used with great advantage to (1) eliminate the bluffer, (2) to separate the apprentice of average attainment from the journeymen.

The authors acknowledge the help accorded them by Messrs. T. J. Kirby, R. Pintner, and R. M. Reinhold, all members of the Army Trade Test Division.