Proceedings of the Iowa Academy of Science

Volume 61 | Annual Issue Article 49

1954

Development and Preliminary Evaluation of a Test of Mechanical Ingenuity

Harold Martinek
Iowa State College

Development and Preliminary Evaluation of a Test of Mechanical Ingenuity

By Harold Martinek

Introduction

There are many phases of industry that need the maximal use of high level personnel. Perhaps the most important of these is engineering, particularly machine design engineering. Although there are many machine designers, only a small percentage can be described as truly creative; these few individuals, however, are among the most important men in industry whether the nation is at war or at peace.

Some men may have been "born creative", but they probably did not reach their maximum usefulness until they had received experience or training or both. If industry could predict who is potentially creative without waiting some indefinite period for this creativity to show itself, the potentially creative designer could receive special training and begin producing earlier and better.

Creative ability has been measured to some extent by various tests. Guilford, Wilson, and Christensen (1) have used the method of factor-analysis to define factors which may be important to creative ability. However, it is not known if these measured factors actually discriminate the creative from the non-creative person since one does not necessarily achieve validity by looking at pure factors.

The General Electric Company (4) has developed selection techniques which also lack demonstrated empirical validity. Their selection program consists of Test assignments, interviews, college grades, and a three and one-half miniature-task type of test. Although the program looks good, the only real evidence of validity that is offered is that there is a demand for the graduates of their training program and that the final standings of these graduates are said to correlate to some extent with their standing upon selection.

The present study, which represents one aspect of an Office of Naval Research project, employs the cross-sectional approach. The general procedure of the project was: 1, to find out as much as possible about the characteristics of creative machine designers; 2, to establish certain working hypotheses about types of

tests which would discriminate creative from non-creative designers; 3, to develop as many tests consistent with these hypotheses as possible; 4, to empirically validate the tests in such a manner that the hypothesized discrimination could be affirmed or denied.

Method

Rationale of the Test

The test reported on in this paper, the Power Source—Apparatus Test, is a miniature-task type of test intended to measure in context. It was not designed to measure a specific ability, but rather to measure the complex of those abilities which are needed to design a mechanism for a particular purpose.

In nine of the ten items of the test, the testee is required to design a mechanism which will produce one type of motion from another. That is, given a powering motion, the testee is to design a mechanism which will produce some other specified motion. For example, one problem without its accompanying drawing is as follows:

PROBLEM 8—To get motion B from motion A by means of a suitable mechanism. Give as many solutions as you can in 10 minutes.

MOTION A—Rotary motion of shaft A which has its axis on the X coordinate.

MOTION B—Square motion of a rod. Rod B moves in a square path which lies in the XY plane.

Almost any machine design problem will yield examples of this type which must have confronted the designer at some time. One would expect the creative designer to be more adept at thinking of mechanisms to solve this type of problem since this is actually the definition of a creative machine designer.

In the tenth problem the testee is presented with a picture of a simple mechanism and asked to give as many uses for this mechanism as he can in ten minutes. From one point of view, the creative designer simply finds new uses for common machine elements. It thus seems logical to hypothesize the validity of such an item.

Scoring

Like most open-ended tests, it would seem that the scoring would be subjective and difficult. This is not true of one of the two scoring methods employed. This method, hereafter called the number of responses method, consists simply of counting the number of solutions that the testee gave. The quality of the solution is not considered; credit is even given for incomplete solutions.

The more difficult and subjective method of scoring, hereafter called the number of workable solutions method, consists of counting the solutions which meet certain scoring criteria. These scoring criteria were developed by considering just which types of mechanisms need what details in order to work.

Without knowledge of the testee's classification, two graduate students in psychology independently scored the test using both methods. There was little disagreement when they used the number of responses method. Any differences represented errors in counting. However, when they used the number of workable solutions method, there was an average of 30 per cent disagreement. Although this is fairly high, it should be noted that this was the first time that either of the scorers had used the method and that they could hardly be considered as "experienced". Many of the disagreements were outright errors on the part of the scorers rather than differences in interpretation of the scoring criteria. Moreover, as the section on results will show, even this subjective scoring method works quite well.

Administration of the Test

The sample consisted of 70 engineers employed in nine companies. Of this, 36 were classified as being creative and 34 as being non-creative by their supervisors. Presumably, the classification was made on the basis of the following definition:

> The creative machine designer is defined as one who has demonstrated the ability to comprehend the nature of a design problem, and to produce a novel, ingenious, or original solution in the form of a total, functional, and practical mechanism. Creativity, in this sense, does not necessarily involve the conception of an entirely new principle, but does involve the combination of existing principles or mechanisms in such a way as to produce new and unique solutions to previously unsolved problems.
>
> The non-creative designer is defined as one whose major function is to work out the details of a design; that is, the engineer who does not produce original ideas, but who works out the routine problems of what materials to use, and who smoothes out the design according to established practices.

The above definition and a letter explaining the project were sent to various companies. Those that agreed to assist in the project were scheduled for testing during June, July, and August of 1953.

RESULTS

Item Selection

The main purpose of this study was to select items which discriminated between creative and non-creative groups. The sig-

nificance of the difference between the mean ranks of the two groups on each item was evaluated by using the H-test (3). Since this is a non-parametric statistic, its assumptions are easily met.

Only in item number eight were there no significant differences at the 10 per cent level between the mean ranks of the two groups if both methods of scoring are considered. The 10 per cent level was decided on since errors of omission of good items are more serious than errors of inclusions of poor items, at least at this stage, and since interest is actually in the discrimination of the total test rather than of single items.

Reliability

The uncorrected, odd-even reliability of the test was .72; the corrected coefficient is .84. These values are for the number of responses scoring method and were computed on the total test rather than on the selected items only. Further estimates of reliability were not considered on this sample.

Validity

Kendall's tau correlation coefficient (2), a non-parametric statistic, was used to estimate validity since the underlying assumptions of this statistic are easily met. Using this coefficient, the correlation between the criterion and the combined score of both scoring methods for the total test is .35. Similar values were found for each scoring method taken separately.

Since these validity coefficients are significant at the 1 per cent level, the hypothesis that the test does not discriminate between the two groups is rejected. Moreover, since the correlations are fairly high and since they were computed using all of the items rather than just the items which significantly discriminated between the two groups, it is believed that the Power Source-Apparatus Test will be useful in industry when refined by item-selection.

Discussion

Of considerable importance in the evaluation of a test is the Criterion. Although supervisors' ratings may be subjective, they were the only practical criterion available. While it is possible that the supervisors did not select the *S*'s on the basis of creativity, it is quite doubtful. If they did select on some extraneous variables they must have done so consistently since the test does discriminate and since no differences were found between companies, i.e., a chi-square tnalysis showed that, in respect to the test scores, there was no interaction between classification by company and

classification by criterion, and that the proportion of responses made by the *S*'s of each company is consistent throughout all companies. It is unlikely that all of the supervisors selected *S*'s on the basis of the same extraneous variable.

In conclusion, it is felt that the Power Source-Apparatus Test does measure creativity in machine design and that, following cross-validation, it will be useful in both educational and industrial selection.

References Cited

1. Guilford, J. P., Wilson, R. C., and Christensen, P. R. 1952. A Factor-Analytic Study of Creative Thinking. U. of S. California. Reports from the Psychology Laboratory. No. 8.
2. Kendall, Maurice G. 1948. Rank Correlation Methods. Charles Griffen and Company Limited, London.
3. Kruskal, W. H. and Wallis, W. A. 1952. Use of Ranks in One-criterion Variane Analysis. J. of Am. Stat. Assn. 47: 583-621.
4. Schmall, W. H. 1952. The General Electric Company. Creative Engineering Program. Private Communications. The General Electric Company.

DEPARTMENT OF PSYCHOLOGY
IOWA STATE COLLEGE
AMES, IOWA