A New Approach to Programmer Aptitude Testing

Charles J. Testa
University of Maryland

During the decade of the 1960's many organizations used aptitude tests to screen computer programmer applicants.(1) The results of these tests were claimed to be valid and reliable indicators of programmer potential. As such, they were often the sole criterion on which employment decisions were based. Although the content of available tests varied, most tests were oriented towards determining an individual's ability by using measures such as letter series, number series, and arithmetic reasoning. In more recent years test batteries have been developed which include aptitude, intelligence, and interest tests to provide a more thorough examination of the applicant. Inherent in this testing process, however, is a cultural and language bias. Recent state and federal legislation has forced employers to carefully examine personnel evaluation procedures to guard against this problem. Research is now being directed toward the development of tests which can correctly be used to assist in the selection process.

In a series of articles which first appeared in the October, 1971 edition of Computers and Automation, MacDonald(2) cited the need for a non-verbal, non-mathematical aptitude test which was not culture, occupation, or background limited. The author suggested the use of an untimed pictorial reasoning test to meet these requirements and offered the hope that this test would provide a valid measure of a person's capacity to perform certain functions, i.e., observe, adapt, reason, program a computer, etc. Thus, the author implicitly questioned the appropriateness of number series, arithmetic reasoning, geometric figures, etc. as measures of programmer aptitude.

A copy of the test was included in the article. Respondents were requested to provide answers to the twenty pictorial reasoning items and also furnish survey data. Included in the survey data were questions about an individual's capability in certain fields, i.e., computer programming, systems analysis, and management. Certain difficulties with this approach have been aptly noted by MacDonald. First, the "correct" answers on which the test is scored are inconclusive. It may not be possible to prove that only one correct answer exists. Also the procedure of respondents assessing their own capabilities is fraught with bias.

In spite of these restrictions, a study of responses has indicated a degree of "correlation" between test score and opinion of programming capability. The results are shown in Table 1 where it can be seen that about 80% of the good or excellent programmers have test scores higher than 12. Thus, the test does appear to distinguish traits such as good programming. However, it must be noted that this approach does not actually measure aptitude, but rather "job performance."

METHOD AND PROCEDURE

The present study attempted to find a predictor which would have a logical relationship to programmer aptitude.(3) An analysis of the programming task revealed perceptual and cognitive factors which seemed related to Witkin's concept of perceptual style.(4) Programming requires an ability to perceive the whole and a concomitant ability to proceed from the general to the particular. After conducting numerous studies, Witkin concluded that a person's manner of perceiving does not easily change and represents an engrained feature of his psychological being -- his "perceptual style." According to Karp, individuals can be classified on a field dependent/field independent continuum depending upon their ability to extract a figure from an embedding context.(5) Field independent people are better able to perform this task.

It has been demonstrated by Goodenough et al(6) that an individual's performance in the perceptual test represents the nature of his functioning in other areas of life, i.e., his dealings with symbolic representations. Thus, it was hypothesized that performance on the perceptual test would predict programmer aptitude.

The embedded figures test (EFT) developed by Witkin examines the manner in which a part is perceived within a larger field. It requires an individual to locate a simple figure which is perceptually obscured in a larger complex figure. The test utilized eight simple figures and twelve complex figures as shown in Figure 1. Also included are the simple and complex figures used for a practice trial before actual testing. To further obscure the simple figures, complex figures were colored using a design described in Figure 1. Since one of the complex figures (A-2) was sufficiently difficult in its original form, it was left uncolored.

The instructions to the subject were written down and presented essentially as follows: "I am going to show you a series of colored designs. After you examine each of these designs I will show you a simpler figure which is contained in that larger design. You will then be given the larger design again, and your job will be to locate the simple figure in it. Let us go through a practice trial to show you how it is done."(7) The score for the test is the average time taken to locate each figure. This test score provides a measure of perceptual style for each subject.

The 34 students who participated in the study were selected on a volunteer basis from an introductory course in computer programming at the University of Maryland. None of these students had any previous programming experience.

The COBOL programming language taught in the course is comprised of English-like statements and sentences. The name COBOL is an acronym for "Common Business Oriented Language." Thus, it is a language used primarily for processing large volumes of business data.

During the semester three tests and three projects were given to students to ascertain capability in COBOL. However, project grades are strongly influenced by factors such as motivation, outside assistance, etc., which somewhat negate their use as "true" measures of capability. Thus, for purposes of this study only test grades were utilized.

The embedded figures test (EFT) was administered to all 34 subjects. Scores on the EFT were then correlated with average test grades.

RESULTS AND DISCUSSIONS

The frequency distribution of individual scores and the mean and standard deviation for the EFT are shown in Table 2 and presented in graphic form in Figure 2. A scatter diagram of EFT scores versus class grades (Figure 3) revealed a linear relationship. Correlation analysis proved that the relationship was significant beyond the .001 level of confidence as shown in Table 3. The negative correlation indicates that individuals who had high EFT scores tended to score lower in class tests.

According to Witkin, performance on the EFT reveals a person's tendency to function at a more differentiated or less differentiated level in other domains. The use of the EFT as a tool to assess this differentiation has certain advantages. First, perceptual functions can be measured by objective procedures. Second, the EFT is based upon the universal perceptual function of disembedding. Thus, it is amenable for testing people of widely different mental levels and socioeducational backgrounds.[8] It is difficult to compare reported studies such as MacDonald's[2] with the current study because of differences in environmental conditions, program languages, subjects, and techniques used to measure capability. Nevertheless, it is interesting that a test (EFT) not purposely designed to measure specific programming capacities had such a high correlation with course grades. Clearly, the EFT must have tapped some of the characteristics of the programming task.

CONCLUSIONS

Test scores on the EFT correlated highly with class test grades. Field independent people tended to score higher in class tests than field dependent people. In fact, an apparent threshold exists at an EFT score of 809 above which subjects are not well suited to programmer training. This finding is in concord with Witkin's belief that perception per se is resistant to training.

The current study has developed a predictor of success in training which appears to be at least as effective as those previously described and which can be administered in a shorter period of time. In addition, the approach adopted has the benefit of being free of limitations imposed by cultural or educational background. However, further validation is required before the EFT can be used as a screening technique for programmer aptitude.

REFERENCES

1. Dickman, R. A., and Lockwood, J. Computer personnel research group: 1966 survey of test use in computer personnel selection. Proceedings of the 4th Annual Computer Personnel Research Conference, 1966, 15-27.

2. MacDonald, N. Pictorial reasoning tests and aptitudes of people. Computers and Automation, October 1971.

3. Cross, E. M. The behavioral styles of computer programmers. Proceedings of the 8th Annual Computer Personnel Research Conference. 1970, 61-82

4. Witkin, H. A., R. Dyk, H. F. Faterson, D. R. Goodenough, and S. A. Karp. Psychological Differentiation. New York: Wiley and Sons, 1962.

5. Karp, S. A. Field dependence and overcoming embeddedness. Journal of Consulting Psychology, 1963, 27, 294-302

6. Goodenough, D. and S. A. Karp. Field dependence and intellectual functioning. Journal of Abnormal and Social Psychology, 1961, 63, 241-246

7. Witkin, H. A., P. K. Oltman, E. Raskin, and S. A. Karp. A Manual for the Embedded Figures Test. Palo Alto, Calif.: Consulting Psychologists Press, Inc., 1971, p. 16

8. Karp, S. A., L. Silberman, and S. Winters. Psychological differentiation and socioeconomic status. Perceptual and Motor Skills, 1969, 28, 55-60

Table 1

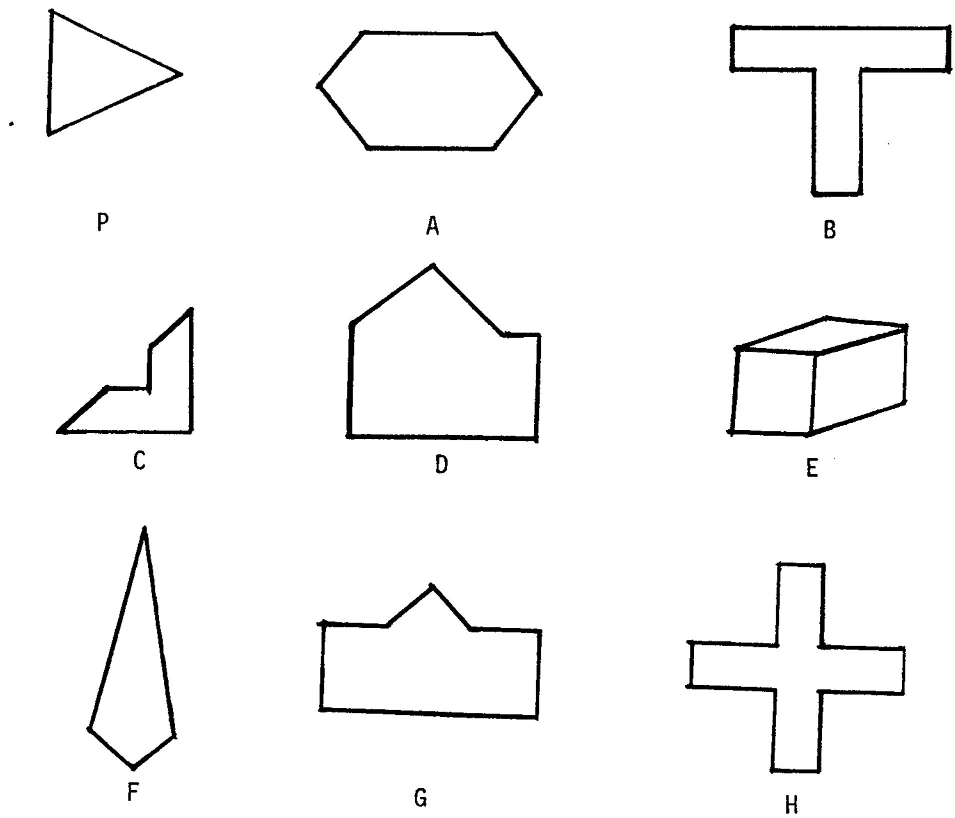
Frequency of Scores - 129 Readers of C & A Who Classify Themselves as "Good" or "Excellent" Programmers

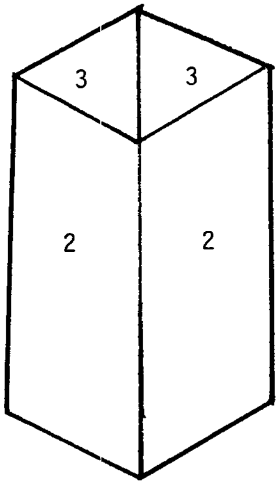
Score	Frequency	Score	Frequency
0	0	11	6
1	0	12	31
2	0	13	25
3	0	14	24
4	1	15	13
5	2	16	8
6	1	17	2
7	0	18	0
8	1	19	0
9	7	20	0
10	8		

FIGURE 1
Simple and Complex Figures Used in the
Embedded-Figures Test

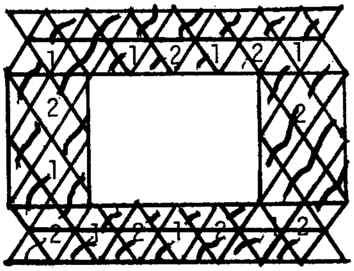
The simple figures are designated by a letter; the complex figures are designated by a letter and a number, the letter corresponding to that of the simple figure which it contains. Figures P and P-1 are the practice figures.

The specific colors used in each complex figure are represented by numbers; and wherever necessary the area covered by a given color is indicated by wavy lines radiating from the number. Figure A-2 remained uncolored. The colors to which the numbers refer are as follows: 1-red, 2-blue, 3-orange, 4-yellow, 5-brown, 6-dark green.

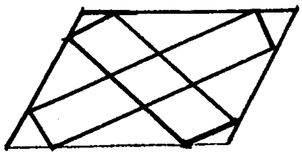


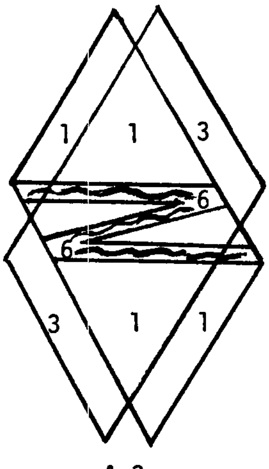


P-1

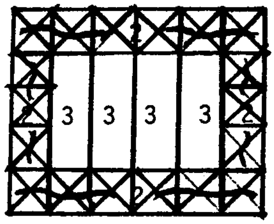


A-1

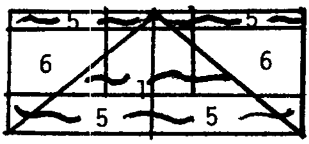

A-2



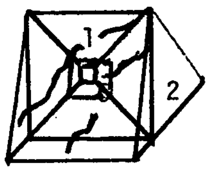
A-3



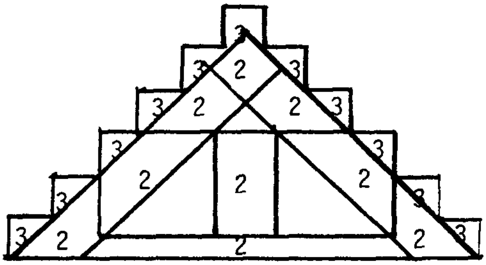
B-1



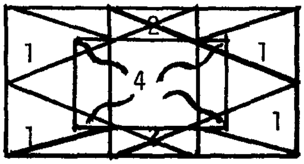
C-1



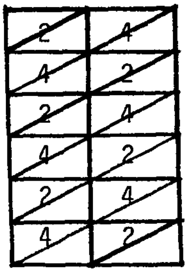
C-2



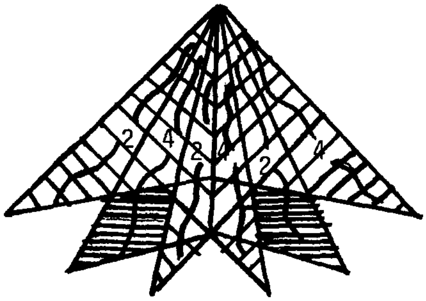
D-1



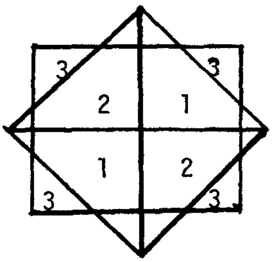
E-1



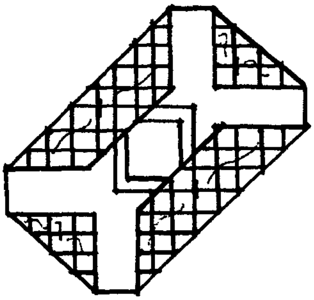
E-2



F-1



G-1



H-1

Table 2

Distribution of Scores and the Mean and Standard Deviation for the EFT

Total Time (Seconds)	Frequency
90-269	3
270-449	9
450-629	5
630-809	8
810-989	5
990-1169	1
1170-1349	1
1350-1529	0
1530-1709	2

Mean (Seconds/Item)	54.5
Standard Deviation	29.4

Figure 2
Subject Performance on the EFT

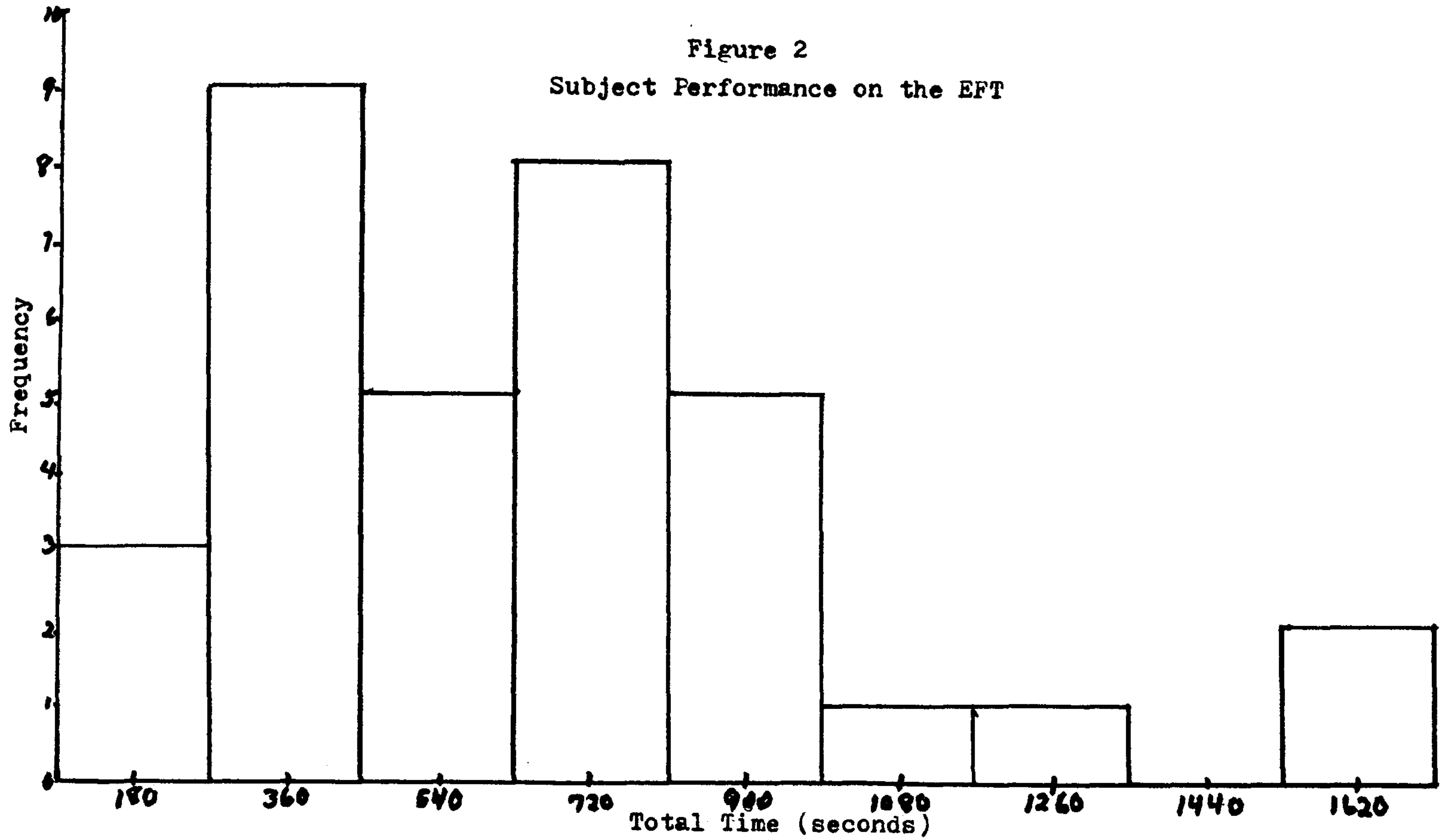


Figure 3
Relationship Between Perceptual
Style and Glass Grades

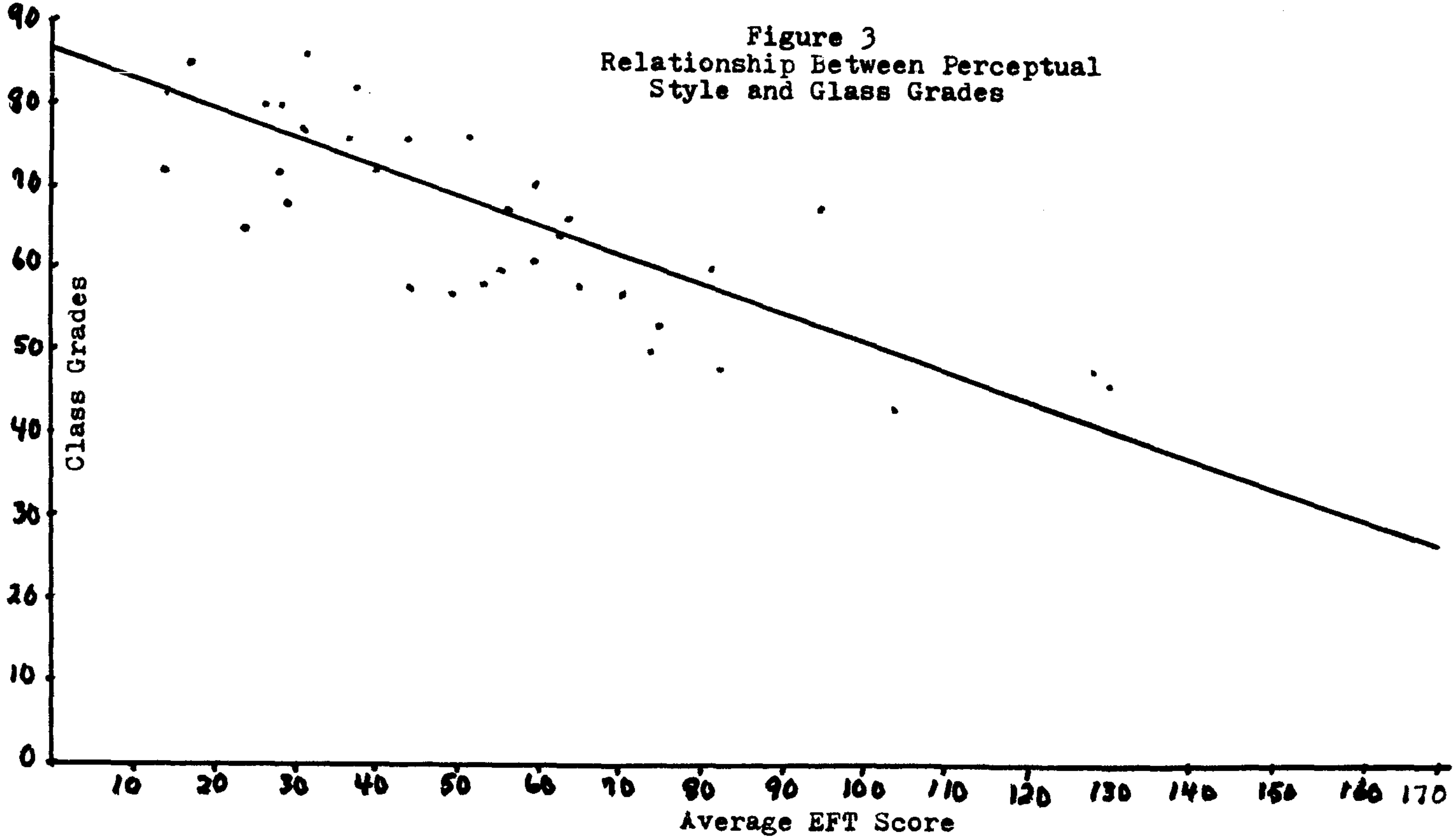


Table 3

Correlation of EFT scores with course test grades

Test	Group Size	Corelation
EFT	N = 34	-.80***

***$p < .001$