ACHIEVEMENT IN HIGH SCHOOL COMPUTER SCIENCE: A PREDICTION STUDY

A THESIS

PRESENTED TO

THE FACULTY OF GRADUATE STUDIES

THE UNIVERSITY OF MANITOBA

IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE DEGREE

MASTER OF EDUCATION

BY

ARTHUR RICHARD GEITH

MAY, 1975

ACHIEVEMENT IN HIGH SCHOOL

COMPUTER SCIENCE: A PREDICTION STUDY

by

ARTHUR RICHARD GEITH

A dissertation submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the requirements of the degree of

MASTER OF EDUCATION

ACKNOWLEDGMENTS

This study would not have been possible without the active cooperation of several individuals. The writer wishes to acknowledge the assistance given by Vern Barrett, George Dorward, and John Fast in collecting the data. Mark Bishop and Dr. B. Johnston assisted in the statistical analysis. Above all, the writer is grateful to his wife Dianne for typing the manuscript.

ABSTRACT

The purpose of this study was to identify variables which were significant predictors of achievement in Computer Science 205, an eleventh grade course taught in Manitoba's high schools. The value of the study lay in its establishing a base of knowledge which would be of use to teachers, counsellors, and administrators in their attempts to help students decide whether they should or should not enroll in Computer Science 205.

The study investigated three subtests of the Differential Aptitude Tests, the Otis-Lennon Mental Abilities Test, the Grade Ten Average, and the Mathematics 100 Score as predictors of achievement in Computer Science 205. The sample consisted of 91 students from three Winnipeg high schools. The analysis involved stepwise regression at the .05 level of significance.

Of the variables investigated, only the Mathematics 100 Score was found to be a significant predictor of the criterion. This variable accounted for about 23 per cent of the variance in the Computer Science scores.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

LIST OF APPENDIXES

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

Computer science has come of age as an academic subject at the high school level only within the last ten to fifteen years. This new discipline differs from other subjects in that few high school administrators, counsellors, or teachers are trained in it or are familiar with the unique demands which it makes of a student.

The problem undertaken in this study was to establish criteria for predicting student achievement in computer science at the grade eleven level. An underlying objective was to develop a base of knowledge about such criteria which would assist administrators, guidance counsellors, and teachers in their efforts to channel students into or out of computer science courses in high school.

Specifically, this study was designed to determine whether there exist criteria locally and readily available, which could be used to predict student achievement in grade eleven computer science. The statistical technique used was forward stepwise regression. In this analysis technique a regression equation was generated which contained only the significant independent variables. The success of the study hinged upon identification of one or more of these significant independent variables.

The significance of the study

High school computer science is a very recent curriculum innovation in Manitoba. Most guidance counsellors, teachers, and school administrators were not trained in computer science and hence probably have not been fully sensitive to the unique demands which this discipline makes of a student. Similarly, these decision-makers in our school system at the time of the study did not know which of the available criteria were significant predictors of student achievement in computer science. As an illustration, the writer was aware that some Winnipeg high schools would admit any grade ten major stream student to the eleventh grade computer science course while other high schools demanded achievement at the A or B level in the university entrance courses. This controversy was not a local phenomenon. Caroline Beckner reported:

> One dedicated high school business teacher, Morris Miller of Martin Van Buren High School, New York City, views data processing as a social need. He is conducting an experiment. Until recently, only the ultra-bright math students had been permitted to study unit-record, programming and computer concepts in this high school. But Mr. Miller is convinced that an average student with only basic math can learn the concepts as easily as the superior math student. He is working with such a class now and declares that any average ability student who is interested in learning data processing can do so. His approach coincides with actual practice in the business world, where it is typical for more than half of the data processing department to be made up of former average students.[1]

---

[1]Caroline Beckner, "Who Should Study Data Processing?," Business Education World, XLVIII (April, 1968), 10-11.

The study can therefore be justified in two ways: 1) it should provide information to educators useful to them in counselling and streaming students; 2) it should enable students to make a better decision on their own if the results of the study are interpreted for them by skilled counsellors.

Definitions

Some terms were used again and again in this study. Recurring terms are defined below.

1. Computer Science. In this study "Computer Science" will refer to the eleventh grade course designated as Computer Science 205 by the Curriculum Branch of Manitoba's Department of Youth and Education. Any noncapitalized reference will refer to the overall discipline of computer science.
2. Achievement. This term will refer to the numerical percentage grade assigned to a student who has completed a specific high school course. Depending on the context, the term may also refer to a student's average performance in a specific course.
3. Predictor Variables. This term will refer to the independent variables which were investigated in this study.
4. Criterion. This term will refer to the dependent variable (Computer Science achievement) investigated in this study.

Assumptions

It was assumed that a student's achievement as defined in a course or grade was in fact a valid measure of his level of performance in comparison to the performance of members of his peer group. It was further assumed that the population from which the sample was taken was similar to and representative of the population in other Winnipeg high schools. It was assumed finally that the schools sampled were similar to other Winnipeg high schools with regard to practices involving grading procedures.

Limitations

The study had three obvious limitations which must be pointed out.

The predictors used constituted a limitation. They were chosen because they represented statistics readily available in a modern high school with adequate testing procedures and could therefore be of real value to counsellors and teachers who wished to use the results of this study. Many other predictors could have been added but they would not have been useful in the above mentioned sense.

A second limitation lay in the sampling procedure. The statistics were drawn from only three of about twenty-five high schools which offered Computer Science. One must therefore be cautious in extrapolating the results of this study to other high schools not included in the study.

Finally, the predictors used in this study were in no way capable of taking into account the many complex factors such as health, personality, motivation, home life, etc. which are important contributors to academic achievement.

Chapter 2

THE REVIEW OF THE LITERATURE

In this study the writer made use of various mental measurements as predictors. It seemed appropriate therefore to review briefly the historical development of mental tests. This was followed by an extensive review of the many complex variables which influence academic achievement. Finally the writer undertook an intensive review of the literature relevent to prediction in the field of computer science.

An Historical Overview of Mental Measurements

The origins of modern day testing date back to the nineteenth century. Not much progress can be reported prior to the work done by Alfred Binet, who probably ranks as the foremost genius in psychological measurement. His intelligence scale, published in 1911, incorporated three features:[1]

1) the use of age standards;

2) the kinds of mental functions brought into play;

3) the concept of general intelligence.

---

[1] Lewis M. Terman, The Measurement of Intelligence (Cambridge: The Riverside Press, 1916), pp. 40-42.

The intelligence quotient of a child could be expressed by the simple formula

$$I\,Q = \frac{\text{Mental Age}}{\text{Chronological Age}} \times 100$$

and this device was employed for the first time in 1916 by Terman[2] in his revision of Binet's test. Terman's adaption and revision of Binet's scale appeared as the Stanford-Binet intelligence test. It was defective at both extremes of the age spectrum, the sampling was inadequate, and some items were unsatisfactory due to low validity and scoring difficulties. Also, the instructions for administration and scoring of the test were lacking in precision. No alternative form of the test existed for use in retesting. Nevertheless it was a milestone in the measurement of intelligence because it incorporated Binet's concepts of age standards, mental functions, and general intelligence in a test suited to the American scene.

Terman and Merril, in 1937, completed a ten year revision of the Stanford-Binet test. It retained the concepts of age standards and intelligence as a trait which develops with age. The test was a distinct improvement over the 1916 version in that many of the problems of sampling, extremes,

---

[2]Lewis M. Terman and Maud A. Merril, Stanford-Binet Intelligence Scale (Boston: The Riverside Press, 1960), p. 6.

validity, etc., had been solved to a large extent.[3]

The Stanford studies of intelligence have continued up to the present time and have set a standard of excellence in mental measurements.

Arthur S. Otis, under the direction of Lewis M. Terman, began the development of large group intelligence tests before World War I. When war broke out the test materials were placed at the army's disposal for adaption to the problem of selecting soldiers for various army roles. The test became known as the Army Alpha test. This use of the Army Alpha popularized the testing movement and after the war was finished many tests were developed, among them the Dearborn, Haggerty, Miller, National, Otis, Pressey, Terman and Thorndike tests.[4] The significance of the Army Alpha test was that for the first time, a measuring device was being used in an important noneducational setting. After the war such test instruments came into widespread use in American industry to assist in the selection of clerical and production personnel.[5]

Despite the fact that the global type of tests

---

[3]Francis N. Maxfield in The 1938 Mental Measurements Yearbook ed. by Oscar Krisen Buros (New Brunswick: Rutgers University Press, 1938), p. 113.

[4]G. M. Ruch and George D. Stoddard, Tests and Measurements in High School Instruction (Chicago: World Book Company, 1927), p. 3.

[5]James D. Linden and Kathryn W. Linden, Tests on Trial (Boston: Houghton Mifflin Company, 1968), p. 16.

developed by Binet and Otis continue to be of value up to the present day, rumblings of dissent began early in this century. Among others, Edward L. Thorndike, L. L. Thurstone and J. P. Guilford argued to varying degrees the existence of many different kinds of intelligence. Guilford, for example, has speculated that as many as one hundred and twenty specific abilities may exist.[6] For this reason much work has been done developing tests which assess specific traits such as verbal, numerical, abstract, spatial and mechanical abilities. Tests which assess attitudes and emotions have also been developed.

Buros classified tests which were in print at the time (1961) into fifteen different categories. His classification as set forth in Table 1 is of interest because it illustrates in quantitative terms the amount of interest shown by researchers in various areas of mental measurements.[7] Personality and character tests were the most important classification in that 306 tests fell into this category. Vocational tests and general intelligence tests were next in importance with 287 and 238 tests in each category respectively. Tests to measure specific aptitudes were relatively fewer in number, and only 20 multi-aptitude batteries had been developed.

---

[6] Ibid.

[7] Oscar Krisen Buros, ed., Tests in Print (Highland Park, N. J. : The Gryphon Press, 1961), p. xvii.

Table 1

In Print Tests By Major Classifications

Classification	Number of tests	Percentage of total
Personality and Character	306	14.4
Vocations	287	13.5
Intelligence	238	11.2
Miscellaneous	233	11.0
Mathematics	198	9.3
English	192	9.0
Reading	159	7.5
Social Studies	113	5.3
Science	106	5.0
Foreign Languages	92	4.3
Sensory-Motor	55	2.6
Business Education	53	2.5
Achievement Batteries	45	2.1
Fine Arts	29	1.4
Multi-Aptitude Batteries	20	0.9
Total	2126	100.0

Reading Buros' Mental Measurements Yearbooks gives one a clear impression of how much research has gone into the development of mental tests. These tests appear to encompass intellective, personality and sociological variables.

Review of Independent Variables

The writer surveyed extensively the entire range of variables which have been considered as determinants of academic performance, in order to provide a conceptual framework for his own research. The variables fall into three broad categories: intellective, personality, and sociological.

Lavin,[8] referring to literature reviews by Cronbach (1949) and Henry (1950), reported correlations between college level ability tests and college grade point averages (GPA) to be in the range from 0.50 to 0.55 Moreover, recent findings center on a correlation of 0.50 with a range from 0.30 to 0.70. At the high school level the comparable correlation is typically about 0.60, the higher value being accounted for by the fact that high school students vary more widely in ability.

Of all the variables used as predictors the one which is usually reported to be the best predictor of

---

[8] David E. Lavin, The Prediction of Academic Performance (New York; Russell Sage Foundation, 1965), p. 51.

academic performance is the grade point average.[9] The following reports are typical of the voluminous research done in this area.

Klugh and Bierly (1959) tested the validity of the School and College Ability Test (SCAT) against a criterion of first semester freshman grade point average (GPA) at Alma College.[10] Correlations between these two variables ranged from r = .51 to r = .67 for men and women tested over two years. A further objective of this study was to determine whether addition of the high school GPA would significantly improve the prediction. Using this additional variable the multiple correlations ranged from R = .661 to R = .782. The researchers concluded that the SCAT and high school GPA were of roughly equal value in predicting freshman performance.

Sharp and Pickett(1959) investigated the General Aptitude Test Battery (GATB) as a predictor of college achievement.[11] The sample consisted of 262 students who had taken the GATB in their senior year of high school. The GATB consists of nine aptitude subtests: 1) Intelligence

---

[9] Ibid., p. 51.

[10] H. E. Klugh and R. Bierly, "The School and College Ability Test and High School Grades as Predictors of College Achievement," Educational and Psychological Measurement, XIX (1959), 625-626.

[11] H. C. Sharp and L. M. Pickett, "The General Aptitude Test Battery as a Predictor of College Success," Educational and Psychological Measurement, XIX (1959), 617-623.

2) Verbal Aptitude, 3) Numerical Aptitude, 4) Spatial Aptitude, 5) Form Perception, 6) Clerical Perception, 7) Motor Coordination, 8) Finger Dexterity, and 9) Manual Dexterity. Correlations between the GATB subtests and college GPA were found to range from .037 to .444. A t-test for significant difference was computed between the means of the successful and unsuccessful groups for each of the subtests. The researchers concluded that subtests 1, 2, 3, 4, and 5 were of some value as predictors of academic success in engineering.

Durnall (1954) investigated the effectiveness of the widely used Miller Analogies Test (MAT) as a predictor of academic success at the graduate level in education.[12] The sample consisted of 153 students who had completed at least thirty hours of graduate work at Oregon State College. The MAT scores were correlated with GPA's in Education and a correlation of 0.21 was obtained. A correlation between the MAT scores and the rank order of the GPA'S resulted in a correlation of 0.50. The conclusion reached was that some relationship existed between MAT scores and GPA scores in Education at the graduate level but that the degree of relationship was only moderate.

Ability measures are, according to Lavin, the best single type of predictor. However, they account for less

---

[12] E. J. Durnall Jr. "Predicting Scholastic Success for Gruaduate Students in Education," School and Society, LXXX (1954), 107.

than half of the variance in academic performance.[13] This has led many researchers to pursue personality factors as predictors of academic performance. Some variables are related to motivational states; anxiety, interest in a field of study, and so on. A second type of variable relates to what Lavin terms "style". Examples are measures of extroversion, dominance and impulsiveness. A third type of variable relates to manifestations of pathology. Classic examples of test instruments in this area are the Minnesota Multiphasic Personality Inventory and the Rorschach. Some interesting examples of research relating personality to academic performance were reviewed.

Weiss, Wertheimer and Groesbeck (1959) studied the influence of motivation on achievement.[14] Sixty male psychology students at the University of Colorado wrote the Edwards Personal Preference Schedule (PPS), McClelland's Picture Story Test and the locally constructed Academic Aptitude Test (AAT). Simple and multiple correlations were computed using the university grade point average as the criterion. It was found that the combination of motivation and ability measures was a powerful predictor of academic achievement with R = .68.

---

[13] Lavin, The Prediction of Academic Performance, p. 64.

[14] P. Weiss, M. Wertheimer, and B. Groesbeck, "Achievement Motivation, Academic Aptitude, and College Grades," Educational and Psychological Measurement, XIX (1959), 663-666.

Merrill and Murphy (1959) studied the personality traits of over-achieving college students at the University of Utah.[15] About 100 students were divided into an over-achieving group and a normally-achieving group on the basis of predicted GPA and GPA achieved after one quarter. These students were then required to take the PPS. The over-achievers were found to be more dominant and less autonomous; more deferent, less exhibitionistic, less affiliative, less concerned with change, and more enduring than the normal achievers. Other variables on the PPS did not differentiate significantly between the two groups.

Beach (1960) studied the combined effect of student sociability and method of instruction on achievement.[16] Students enrolled in an advanced psychology course at the University of Michigan were tested on Guilford's Inventory of Factors Sociability scale and placed into four learning situations; lecture, discussion group with instructor, autonomous small groups, and independent study. The less sociable students were found to achieve more than the sociable students in the lecture situation. Conversely, in the small group situation the sociable students outperformed the less sociable students.

---

[15] R. M. Merrill and D. T. Murphy, "Personality Factors and Academic Achievement in College," Journal of Counselling Psychology, VI (1959), 207-210.

[16] L. R. Beach, "Sociability and Academic Achievement in Various Types of Learning Situations," Journal of Educational Psychology, LI (1960), pp. 208-212.

Cooper (1955) investigated the value of the Rorschach Test as a predictor of academic performance at San Franscisco State College.[17] The Rorschach is a clinical instrument used to investigate underlying personality dynamics. This test was found to be poor in the prediction of academic achievement. Similar results were obtained by Quinn (1957) using another clinical instrument, the Minnesota Multiphasic Personality Inventory (MMPI).[18]

In Lavin's review of the literature, he found the research findings related to personality factors rather disappointing.[19] The picture presented by the research was confused and often contradictory. An equally difficult set of predictors consists of the sociological determinants to which a member of society is exposed. Examples are his socio-economic status, sex, religious background, and age. Again, a sampling of the research done in this area was conducted.

Schoonover (1959) studied three questions:[20] 1) Is there a relationship between a child's ordinal position and

---

[17] J. G. Cooper, "The Inspection Rorschach in the Prediction of Academic Success," Journal of Educational Research, XLIX (1955), pp. 275-283.

[18] S. B. Quinn, "Relationship of Certain Personality Characteristics to College Achievement," Dissertation Abstracts, 17: 809, 1957.

[19] Lavin, The Prediction of Academic Performance, p. 111.

[20] S. M. Schoonover, "The Relationship of Intelligence and Achievement to Birth Order, Sex of Sibling, and Age Interval," Journal of Educational Psychology, L (1959), 143-146.

his mental test performance? 2) Does a sibling's sex influence the mental test performance of a child? 3) What relationship exists between age interval and degree of resemblance in mental achievement of siblings? The sample consisted of 59 sibling pairs whose statistics were drawn from the University Elementary School at the University of Michigan. Schoonover found that older and younger siblings performed equally well on mental tests and no relationship existed between age interval and average score difference on the mental tests. However, in every case, siblings with brothers were found to perform better on mental tests than siblings with sisters. The rationale presented for this finding is that a male sibling is more aggressive than a female sibling and hence is more stimulating.

Northby (1959) made the well known observation that girls perform better than boys throughout all grades.[21] In a study of 12,826 Connecticut high school students he ranked boys and girls into decile columns. He found 930 girls or 13.2 per cent of the girls scored in the top decile. Of the boys, only 366 or 6.3 per cent did as well. At the bottom end of the scale a reverse relationship existed. These figures were cited as clear evidence that girls were academically superior to boys. However, Northby notes that when boys and girls are compared on the basis of mental tests the

---

[21] A. S. Northby, "Sex Differences in High School Scholarship," <u>School and Society</u>, LXXXVI (1958), 63-64.

differences are relatively negligible.

Coster (1959) drew a sample of 878 cases from 3,000 pupils in nine Indiana High Schools in an attempt to determine the distinctive characteristics of three socio-economic groups.[22] Each student was classified as high, middle, or low income according to a "House and Home" scale and subjected to a 27 item attitudinal questionnaire. Coster found that the high socio-economic students differed significantly from the other two groups in that more of them planned to go to university and they were twice as likely to achieve A and B averages as the lower income students.

Tibbetts (1955) investigated the effect of family relationships on the achievement of high school pupils.[23] The hypothesis was that high achieving boys and their parents would describe family relationships in similiar terms more frequently than would underachievers and their parents. A second hypothesis was that the former families would express more satifaction with these relations than would the latter. These hypotheses were both verified. The high achieving students were found to identify closely with their parents and were motivated by a desire to please them.

---

[22] J. K. Coster, "Some Characteristics of High School Pupils From Three Income Groups," Journal of Educational Psychology, L (1959), 55-62.

[23] J. R. Tibbetts, "The Role of Parent-Child Relationships in the Achievement of High School Pupils: A Study of the Family Relationships Associated With Underachievement and High Achievement," Dissertation Abstracts, 15; 232, 1955.

The parents of high achievers were characterized as thoughtful, understanding, and warm. Apparently, there exists a relationship between the quality of home-life and student achievement.

Carter (1956) investigated the effect of age on achievement at the elementary level.[24] His concern was to determine whether children admitted to the Austin, Texas school system prior to the age of six were at a disadvantage in terms of academic performance. Fifty underage children were compared with fifty normal children. The study controlled sex, and intelligence. It was found that the underage children were indeed at a disadvantage in academic achievement. Moreover, this disadvantage persisted throughout the elementary school years. According to this study 87 per cent of the underage children did not perform as well as the normal age children.

In summary, prediction of academic performance based on previous academic achievement and tests of mental ability appears to be well founded. However, researchers have attempted to improve upon predictions based solely on intellective factors. This has pushed the research into the area of personality and societal influences with some

---

[24]L. B. Carter, "The Effect of Early School Entrance on the Scholastic Achievement of Elementary School Children in the Austin Public Schools," Journal of Educational Research, L (1956), 91-103.

positive results. Much of the latter research is contradictory, however, and the variables involved are difficult to qualify.

Prediction Studies in Computer Science

The writer conducted an intensive review of the literature pertaining to prediction studies in computer science. Some of the studies reviewed naturally had an educational setting. Others originated in the data processing industry, usually as attempts to improve the efficiency with which computer programmers were selected for training. Summaries of the research in this area are arranged in chronological order.

Rowan (1957) reported on the Rand Corporation's experience in the selection of programmer personnel.[25] In 1952, research into the predictive capabilities of thirty differents tests narrowed the field down to two: the Thurstone Primary Mental Abilities Test (PMA), and the Thurstone Temperament Schedule (TTS). Using supervisory ratings as the criterion, a multiple correlation was calculated using multiple regression analysis. The sample consisted of Rand employees involved in computer programming. The specific subtests found to be important predictors of programming competence were those labelled "Verbal Meaning","Reasoning",

[25] T. C. Rowan, "Psychological Tests and Selection of Computer Programmers," Journal of the Association for Computing Machinery, IV (1957), 348-353.

and "Emotional Stability". The first two of these subtests were designed to measure important components of general intelligence. The multiple correlation yielded by these three subtests was R = .54. Subsequent attempts to improve the predictive capability of the test battery up to 1956 were unsuccessful.

McNamara and Hughes (1961) in a review of the literature, pointed out their development of the Programmer Aptitude Test (PAT) under the auspices of the I B M corporation.[26] The PAT is a one hour paper and pencil test consisting of three parts: 1) Number Series, 2) Figure Analogies, and 3) Arithmetic Reasoning. When the PAT became compromised through overuse, these two researchers developed an alternate version known as the Revised Programmer Aptitude Test (RPAT). The RPAT is an 86 item multiple choice, paper and pencil test consisting of three subtests with the same names as the PAT subtests. McNamara and Hughes reported a 1958 study wherein PAT scores and supervisor ratings were correlated at 0.44, significant at the .05 level. A previous study of a similar sort in New York yielded an r = .36, also significant at .05. Better correlations were also reported. A study by Upshall and Riland (1958) yielded a Spearman rho of 0.61 in a study of the relationship between PAT scores and six-point supervisor ratings. Other significant

[26] W. J. McNamara and J. L. Hughes, "A Review of Research on the Selection of Computer Programmers," Personnel Psychology. XIV (1961), 39-51.

predictors found by Upshall and Riland were the Brown-Carlsen Listening Comprehension Test; $r = .60$, and the Edwards Personal Preference Schedule. No correlation was reported for the latter.

Oliver and Willis (1963) researched the validity of the Programmer Aptitude Test in predicting final grades in a computer programming course at Southern Illinois University.[27] The sample consisted of 27 students who completed the course during the summer quarter of 1962. Students who registered but did not complete the course were excluded from the study. The PAT was administered at the beginning of the course and the results were not made available to the instructor until after he had assigned the final grades. Validity coefficients were computed between the three subtests of the PAT and the final course grade. These were: Number Series, 0.65; Figure Analogies, 0.70; and Arithmetic Reasoning, 0.40. The total validity coefficient was 0.78 and all results were significant at the .001 level. The researchers concluded that for the small sample involved, the PAT had demonstrated high predictive validity.

Peres and Arnold (1963) performed a study designed to give some insight into those performance factors considered

---

[27] T. C. Oliver and W. K. Willis, "A Study of the Validity of the Programmer Aptitude Test," Educational and Psychological Measurement, XXIII (1963), 823-825.

to be important in the programming profession.[28] Twenty-three programmers and six supervisors were asked to write an essay describing the best programmer they had known. The essays were analysed and 218 statements regarding programmer behaviour were extracted. Careful attention was paid to the choice of phraseology and vocabulary. By combining similar items the number of descriptive statements was reduced to 140. These statements were then sorted into 12 categories representing predicted performance factors. The 140 statements were presented to the 29 subjects with instructions to rate the best and worst programmers they had ever known on an applicability scale of 1 to 5. The 58 check lists were subjected to analysis consisting of the determination of Pearson product-moment correlation coefficients and a modified central factor analysis by the Wherry-Winer method for factoring large numbers of items. The analysis revealed six specific factors related to programming competence. The competent programmer was characterized as having the following qualities: 1) personal maturity and stability; 2) good interpersonal relations; 3) good communication skills; 4) thoroughness and dependability; 5) professional competence; and 6) job interest and zeal.

The Computer Personnel Research Group (1964) reported

---

[28]Sherwood H. Peres and Phil H. Arnold, "Identifying Programmer Behaviour," <u>Datamation</u>, (January, 1963), 40.

research undertaken for the United States Air Force.[29] The study was designed to investigate the relationship of rated job performance to the cognitive abilities, vocational interests, and biographical background of computer programmers. A test battery composed of the Programmer Aptitude Test, the Strong Vocational Interest Blank (SVIB), a Personal Background Data Form, and the Test of Sequential Instructions (TSI) was administered to 534 programmers from 24 companies. The sample was divided into two subsamples: 233 were classified as business programmers and 301 were classified as scientific programmers. The TSI was a research instrument which measured the subject's ability to do several tasks simultaneously, an ability considered to be necessary for programming.

Analysis of the data showed that the PAT and TSI scores correlated more highly with the scientific subsample performance ratings than with the business subsample performance ratings. Correlations computed on an individual company basis showed wide fluctuations (-.67 to .90). This may be accounted for by the fact that six of the participating companies had used the PAT to select the programmers involved in the study. The supervisor rankings then forced previously

---

[29] Robert N. Reinstedt et al, Computer Personnel Research Group Performance Prediction Study, Memorandum RM - 4033-PR, a report prepared for the U. S. Air Force Project Rand, Santa Monica, California, (March, 1964).

acceptable employees into a poor or fair category, causing a bias against the PAT as a prediction instrument. The SVIB findings indicated that successful programmers tend to be scientific, professionally-oriented, aesthetic persons. Both samples showed a high interest in mathematics. As a group the programmers were found to be relatively young as compared to other professions and to have relatively few years of experience. The scientific subsample was more highly educated than the business subsample. Eighty-six per cent of the scientific subsample participants had at least one degree as compared to two-thirds of the business subsample.

Biamonte (1965) reported fascinating research into the effect of attitudes on the learning of computer programming.[30] A sample consisting of 201 students enrolled in introductory and advanced programming courses at New York University was subjected to attitudinal and intelligence testing. Each student was required to write the California F-scale, a measure of authoritarianism; Rokeach's Dogmatism scale, a measure of closed-mindedness; and McClosky's Conservatism scale, a measure of the extent to which a person accepts the tenets of conservative social and political philosophy. In addition to these attitudinal tests each

---

[30] A. J. Biamonte, "A Study of the Effect of Attitudes on the Learning of Computer Programming," Proceedings of the Third Annual Computer Personnel Research Conference (Silver Spring, Maryland: Computer Personnel Research Group, June 17-18, 1965), pp. 68-74.

student was required to write an intelligence test; the 12 minute, 50 item, paper and pencil Wonderlic Personnel Test. The criterion against which these test scores were correlated was the grade point score achieved in the course. Intercorrelations ranging from $r = .288$ to $r = -.182$ were found between the Wonderlic and attitudinal scales. The multiple correlation between all independent variables and the grade point score was found to be $R = .531$ but the grade point score was correlated negatively with each of the attitudinal scales. A correlation of 0.518 was found between the grade point score and the Wonderlic. An analysis of covariance was performed to suppress intelligence as a contributing variable.

Biamonte concluded that possession of a high level of intelligence was reflected in a personality profile which was nonauthoritarian, nondogmatic and nonconservative. Also, highly intelligent persons were found to be good programmers. Thus the attitudinal scales could be considered to be indirect predictors of programming competence through the common factor of intelligence.

Hollenbeck and McNamara[31] (1965) investigated the validity of the Computer Usage Company Programmer Aptitude Test (CUCPAT) developed by Opler[32] in 1963. The CUCPAT is

---

[31] G. P. Hollenbeck and W. T. McNamara, "CUCPAT and Programming Aptitude," Personnel Psychology, XVIII (1965), 101-106.

[32] Ascher Opler, "Testing Programming Aptitude," Datamation, (October, 1963), 28-31.

a computer based test requiring a computer and computer operation for about 45 minutes of testing time. Twenty seven college level persons, 9 women and 18 men, constituted the sample. Predictor variables were the CUCPAT and the Data Processing Aptitude Test (DPAT) scores. The DPAT is an instrument designed by IBM for internal use only and is composed of 40 letter series items, 30 figure analogy items and 25 arithmetic items. The criterion was programming performance during a 26 week training program. Evaluations were made on the basis of observed achievement, personal interviews and student-teacher discussions. The CUCPAT did not stand up under scrutiny; the investigators found a rather low correlation (0.35) between the CUCPAT and the well established DPAT. Furthermore the CUCPAT-criterion r was found to be nonsignificant at the .05 level.

Seiler (1965) reported research involving the General Aptitude Test Battery (GATB) as a predictor of success in four automatic data processing occupations: 1) business systems analyst, 2) business programmer, 3) engineering and scientific programmer, and 4) console operator.[33] The GATB yields 9 aptitude scores: 1) General Learning Ability, 2) Verbal Aptitude, 3) Numerical Aptitude, 4) Spatial Aptitude, 5) Form Perception 6) Clerical Perception,

---

[33] Joseph Seiler, "Abilities for ADP Occupations," Proceedings of the Third Annual Computer Personnel Research Conference (Silver Spring, Maryland: Computer Personnel Research Group, June 17-18, 1965), pp. 52-59.

7) Motor Coordination, 8) Finger Dexterity, and 9) Manual Dexterity. GATB test scores and supervisor ratings were obtained for each of 475 workers employed by 30 companies in Wisconsin and California. Correlations were computed between each of the GATB aptitudes and the supervisor ratings. Analyses showed that four aptitudes; General Learning Ability, Verbal Aptitude, Numerical Aptitude, and Spatial Aptitude were required for performing the job duties of a business systems analyst, a business programmer, and the engineering and scientific programmer. Only the first three of these aptitudes were required for the console operator.

Stalnaker (1965) reported research at the Georgia Institute of Technology in which the Watson-Glaser Critical Thinking Appraisal was investigated as a predictor of programming performance.[34] The Watson-Glaser is a 99 item, 40 minute test designed to measure five elements of critical thinking: 1) Inferential Reasoning, 2) Recognition of Assumptions, 3) Deductive Reasoning, 4) Interpretation, and 5) Evaluation of Arguments. The sample consisted of 108 Georgia Tech students involved in the ALGOL programming language course. The Watson-Glaser scores were correlated

---

[34] Ashford W. Stalnaker, "The Watson-Glaser Critical Thinking Appraisal as a Predictor of Programming Performance," Proceedings of the Third Annual Computer Personnel Research Conference (Silver Spring, Maryland: Computer Personnel Research Group, June 17-18, 1965), pp. 75-77.

with the ALGOL scores for two of four subsamples. An item analysis showed that 95 per cent of the variation in answers was contained in only 54 questions. Furthermore, 14 additional items showed two levels of response, the choice between the two being statistically random. A revised scoring key was devised based only on the significant items and correlations were computed between the revised Watson-Glaser scores and the ALGOL scores of the other two subsamples. Correlations of 0.76 and 0.81 were computed. In his conclusion, Stalnaker did not appear overly enthusiastic about the Watson-Glaser as a predictor of programming performance. He speculated that the useful items were merely reflecting the important relationship between general intelligence and programming aptitude. In other words, an intelligence test might have been equally valid as a predictor of programming performance.

Dugan (1966) delivered a paper in which he discussed the achievement of graduates of predominantly negro colleges as compared to the achievement of graduates of northern integrated colleges on the Programmer Aptitude Test.[35] In comparing the two groups of graduates it was found that the graduates of negro colleges had more difficulty with the PAT than did the graduates of northern schools. Indeed, the PAT score which ranked the negro college graduates at the 82nd percentile of their group was equivalent to a ranking at the

---

[35]R. D. Dugan, "Current Problems in Test Performance of Job Applicants," Personnel Psychology, XIX (1966), 18-24.

24th percentile at the northern integrated colleges. Dugan suggested that the PAT did not cut across cultural barriers as a predictive device since many southern graduates did go on to become successful computer company employees despite low scores on the PAT.

Gordon and Denis (1966) reported the experience of the Bell Telephone Laboratories in testing a battery of predictors against supervisor ratings of 52 Computer Service Operators (CSO's).[36] The CSO is a job classification one level below that of programmer. A person classified as a CSO performs a service role; mounting tapes, console monitoring, paper and card handling, malfunction troubleshooting, etc. Many CSO's go on to become programmers. The test battery selected consisted of the Programmer Aptitude Test, the Logical Analysis Device (LAD), the Punched Card Machine Operator Test (PCMOT), the Work Inventory (Motivation Survey) and the Wonderlic Personnel Test. The LAD is a hands-on testing device which requires the testee to use logical reasoning to turn on console lights. The test ostensibly measures the logical faculty necessary for successful programming. The PCMOT is a 12 minute paper-pencil test designed for use in the operation of IBM card punch equipment. It is divided into two parts: letter-digit substitution,

---

[36] Bruce F. Gordon and Richard A. Denis, "Characteristics and Performance Predictors of 7094 Computer Service Operators," Proceeding of the Fourth Annual Computer Personnel Research Conference (Los Angeles: Computer Personnel Research Group, 1966), pp. 96-106.

and name checking. The Motivation Questionnaire is a 69 item test arranged in triads. The testee has to choose one of three equally socially acceptable responses. However, one response is always central to task completion.

The scores from the test battery were correlated with supervisor ratings. The conclusion reached was that none of the tests were capable of predicting CSO performance as measured by supervisor ratings.

Seiler (1967) reported that a survey of the industry indicated that wide-spread satisfaction existed re the ability of current test instruments to select competent computer programmers.[37] About ten employers reported that a score of A or B on the Revised Programmer Aptitude Test or its successor, the Aptitude Test for Programmer Personnel (ATPP) was a strong predictor of success in programmer training. One company reported that 45 out of 50 testees who scored a combination of 80 points on the Wonderlic and RPAT went on to become excellent programmers on the job. Seiler concluded his survey by stressing the need for the development of culture-free predictors of programming competence. The problem with such tests however, was that they were also likely to be "validity-free". He quoted Anastasi's

---

[37] Joseph Seiler, "Survey of Validation Studies on Computer Personnel Selection Instruments," Proceedings of the Fifth Annual Computer Personnel Research Conference, (Computer Personnel Research Group, 1967), pp. 43-51.

opinion regarding this problem: "A better solution is to choose test content in terms of its criterion relevence and then investigate the effect of moderator variables."

Howell, Vincent, and Gay (1967) used the Robot test as a criterion variable and several other variables as predictors: age, education, Civil Service grade, supervisory ratings, the PAT, the Color Naming Test, the General Aptitude Test Battery and the Federal Service Entrance Examination (FSEE).[38] The Robot test was developed by the Bureau of the Census and has face validity in that it provides a sample of programming activity analagous to problem solving on a computer. The FSEE consists of three parts: 1) Verbal, 2) Numerical, and 3) Abstract Reasoning. Pearsonian intercorrelations were computed among all variables. Multiple correlations were obtained with F tests computed for the significance of each added variable. The study showed that parts 2 and 3 of the PAT and part 2 of the FSEE were of value as predictors with multiple correlations of $R = .60$ and $R = .71$ on two different samples.

Mayer and Stalnaker (1967) reported on the state of the art of computer personnel selection proceedures in

---

[38] M. A. Howell, J. W. Vincent and R. A. Gay, "Testing Aptitude for Computer Programming," <u>Psychological Reports</u>, XX (1967), 1251-1256.

North America.[39] The computer Personnel Research Group had just sponsored a survey in which 483 American and 98 Canadian firms participated. The number of employees of these firms totalled 23,000 in the United States and 1,000 in Canada. It was found that 68 per cent of the American firms and 72 per cent of the Canadian firms used test instruments to select personnel. The practice among these firms was to use two basic types of test: intelligence and aptitude. The most commonly used intelligence test was the 12 minute, paper and pencil, spiral Wonderlic Personnel Test. Sixty American and 7 Canadian firms used the Wonderlic. In comparison only 13 firms used the next most popular test, the Thurstone Test of Mental Alertness. The overwhelmingly favored aptitude test was the Programmer Aptitude Test. This test was used by 282 American and 67 Canadian firms. The next most important aptitude test was the Federal Service Entrance Exam but only 13 firms used it. The number of firms who depended on personality or interest tests was negligible.

Bauer, Mehrens, and Vinsonhaler (1968) conducted a computer programming prediction study at Michigan State

---

[39] David B. Meyer and Ashford W. Stalnaker, "Computer Personnel Research - Issues and Progress in the 60's," Proceedings of the Fifth Annual Computer Personnel Research Conference (Computer Personnel Research Group, 1967), pp. 6-41.

University.[40] Undergraduates enrolled in a computer programming course constituted the sample. Achievement in the programming course was correlated with scores from the Aptitude Test for Programmer Personnel, the Strong Vocational Interest Blank (SVIB), the College Qualification Test (CQT), and grade point averages (GPA's) attained in all previous courses at Michigan State. The best predictors in descending order were the GPA ($r = .68$), the CQT Numerical ($r = .53$), the ATPP Total ($r = .51$), the CQT Total ($r = .49$) and the ATPP Part III ($r = .48$). Interestingly, the SVIB Programmer was found to be a relatively weak ($r = .33$) predictor of the criterion.

A stepwise addition multiple regression analysis was computed on the data. The results showed an R of .81 when the GPA, ATPP Part II, and Computer Programmer Scale of the SVIB were used in combination as predictors. Without the SVIB contribution the other two variables yielded an R of .76.

Hunt and Randhawa (1968) noticed that a bimodal distribution existed in the achievement scores of computer science students in a second year course at the University of Saskatchewan.[41] In an attempt to identify some of the

---

[40]Roger Bauer, William A. Mehrens, and John F. Vinsonhaler, "Predicting Performance in a Computer Programming Course," Educational and Psychological Measurement, XXVIII (Winter, 1968), 1159-1164.

[41]D. Hunt and B. S. Randhawa, "Relationships Between and Among Cognitive Variables and Achievement in Computational Science," Educational and Psychological Measurement, XXVIII, No. 3, (1968), 921-928.

cognitive abilities which separated the high achievers from the low achievers, 119 students enrolled in Computer Science 253 were subjected to a battery of tests from the Kit of Reference Tests for Cognitive Factors; namely, Seeing Problems, Object-Number, Figure Classification, Locations, Letter Sets, Thing Categories and Hidden Patterns, along with five subtests of the Watson-Glaser Test of Critical Thinking (CT). The criterion was achievement in Computer Science 253.

A stepwise regression was performed to reduce the number of possible predictors to four. The sample was then divided into high and low achievement groups and a one way multivariate analysis of variance was performed. It was found that 23.1 per cent of the variance was accounted for by four predictors; Locations, Hidden Patterns and Seeing Problems from the Reference Tests and Deduction from the CT test. Furthermore, it was found that these four tests also discriminated between high and low achievers in Computer Science 253.

Correnti (1969) studied the value of the Aptitude Test for Programmer Personnel, the Survey of Study Habits and Attitudes (SSHA- Form C), and the State University of New York Admissions Examination (SUAE) as predictors of performance in a first year college level programming course.[42]

---

[42] R. J. Correnti, "Predictors of Success in the Study of Computer Programming at Two-Year Institutions of Higher Learning," Dissertation Abstracts, 30: 3718A, 1969.

Other predictors used were the grade point average and the parental occupational level as determined by Roe's classification. The sample was composed of 261 students enrolled at four New York State colleges. Regression equations were developed and checked in a cross-validation. The predictors were found to be significant indicators of programming success but not to the degree of certainty required to justify acceptance or rejection of a student applying for a course in programming. Nevertheless students could justifiably be classified according to ability on the basis of the predictors used.

Miller (1970) studied the Programmer Aptitude Test in conjunction with high school scores and the Semantic Differential Test (SDT).[43] The SDT is a test designed to reveal a person's area of interest. The sample consisted of 50 male and female students enrolled in the Vocational Computer Operations Program at Garinger Senior High School in Charlotte, North Carolina. Miller computed correlations between the subtests of the PAT, the SDT, the students' grade point averages, and the scores achieved in the high school computer science course. The data was analysed using multiple regression procedures. Correlations between the PAT subtests and the computer science course scores ranged

---

[43] Robert L. Miller, "A Study of the Relationship Which Exists Between Achievement in High School Vocational Data Processing Programs and the Programmer Aptitude Test" (Master's thesis, North Carolina State University, 1970), pp. 26-27.

from $r = .45$ to $r = .56$, indicating significance at the .01 level. The SDT had correlations in the order of $r = .20$ with the PAT, the grade point average, and the computer science course scores. The conclusion that the student grade point average was the best predictor of success ($r = .76$) in the computer science course was not surprising.

Walker and Markham (1970) administered the Revised Programmer Aptitude Test and the Otis Higher Form B intelligence test to 86 Information Science students at the New South Wales Institute of Technology.[44] A second group of 297 students was given the RPAT and the International Computers and Tabulators Computer Programming Aptitude Test (ICT). The ICT consists of four subtests: 1) Letter Series, 2) Arithmetical Reasoning, Number Series and Code Deciphering, 3) Coding According to Instructions, 4) Verbal Analogies, Vocabulary, Logical Reasoning and Verbal Series. This 148 item test is apparently more complex than the 86 item RPAT which consists of only three subtests: 1) Number Series, 2) Figure Analogies, and 3) Arithmetical Reasoning.

The objective of this research was to develop norms for Australian candidates and make comparisons among the various tests and the high school scores of the candidates. It was found that American norms for the RPAT were not appropriate to the Australian scene. The ICT and RPAT, both

---

[44]Elaine Walker and S. J. Markham, "Computer Programming Aptitude Tests," Australian Psychologist, V (March, 1970), 52-58.

of which purport to measure programming aptitude had a total correlation of only 0.49. Furthermore, the correlation between equivalent subtests, the RPAT subtest 3 and the ICT subtest 2, was only 0.51. All other subtest correlations were lower. Correlations between the high school mathematics, English, and average scores and the two aptitude test scores were very low. Further research was found to be necessary before either test could be adapted meaningfully to the Australian scene.

A prediction study by Carol Ann Alspaugh (1970), attempted to define the profile of a successful university computer science student.[45] Predictor variables employed were the Thurstone Temperament Schedule, the Programmer Aptitude Test, the Watson-Glaser Critical Thinking Appraisal, the School and College Aptitude Test, and a mathematics background measure. The developed profile of the successful computer science student was as follows: nonimpulsive, strong mathematics course background, strong numerical ability, low sociability, nondominant, and reflective.

The above study provides valuable insight into the personality traits of the successful computer science student with two reservations: 1) most of the variance was not

---

[45] Carol Ann Alspaugh, "A Study of the Relationships Between Student Characteristics and Proficiency in Symbolic and Algebraic Computer Programming" (Doctor's thesis, University of Missouri, 1970), pp. 56-60.

explained, and 2) the mathematics background measure is a measure of the number of courses taken and may therefore be a measure of age or of survival ability rather than of mathematical ability as is implied.

John Alspaugh (1971), in a study designed to investigate the effects of differences in grade placement upon programming achievement and programming aptitude for high school seniors as compared to college juniors and seniors, found correlations between the Ohio Psychological Examination (Ohio), the Programmer Aptitude Test and high school achievement scores.[46] His samples were rather small; fourteen high school students and twenty three college students. The Ohio is a general intelligence test administered to Missouri high school students during the late junior or early senior years. The correlations found included $r = 0.46$ between the Ohio and high school achievement and $r = 0.57$ between the PAT and high school achievement. At the college level the corresponding correlations were -.06 and 0.36. In both cases the PAT had a higher correlation with the grade point achievement than did the Ohio.

This intensive review of prediction studies in the computer science field permitted some general conclusions to be drawn. Most of the research has been conducted by

[46] John W. Alspaugh, "The Relationship of Grade Placement to Programming Aptitude and Fortran Programming Achievement," Journal for Research in Mathematics Education, January, 1971, pp. 44-48.

industrial firms and post secondary institutions of learning. In industry, the tests known as PAT, RPAT, DPAT and ATPP were all developed by IBM corporation and are overwhelmingly favoured as selection instruments in the hiring of computer personnel. These tests are largely cognitive in nature. Some research has been conducted in an attempt to explore personality and sociological variables. The research in this area has been interesting and worthwhile but a definitive study has yet to be done.

Methods of Prediction

In this study the need existed for a statistical method which would consider several independent variables $X_1, X_2, \ldots X_n$, a dependent variable Y, and fit these variables into an equation of the form $Y = B_1X_1 + B_2X_2 + \ldots B_nX_n$. The technique required also had to be capable of making appropriate statistical significance checks to ensure that nonsignificant variables were not included in the final equation. Draper and Smith [47] outline the possible approaches to the regression equation: a) all possible regressions, b) backward elimination, c) forward selection, and d) stepwise regression. Each of these will be briefly explained.

---

[47] N. R. Draper and H. Smith, Applied Regression Analysis (New York: John Wiley and Sons, Inc., 1968), pp. 167-172.

All possible regressions. This is an extremely cumbersome procedure because it involves the fitting of every possible regression equation. If there are k independent variables there are 2^k equations which must be examined at the second step of the procedure. While the researcher may be confident he has examined all possibilities, the procedure is wasteful of human effort and computer time. Many of the possible equations could be dismissed at once without this laborious technique.

Backward elimination. This approach is much better than the first one in that it attempts to cut through to the best regression without examining every possibility. The procedure starts by computing a regression equation containing all variables. A partial F-test value is computed for each variable as if it were the last variable to enter the equation. The lowest F-value variable is compared with a predetermined significance level and accepted or rejected. The procedure then examines each of the other variables in a similar manner. At the conclusion the fitted regression equation contains only those variables which make a significant contribution to the prediction.

Forward Selection. This procedure is the mirror image of backward elimination in that it starts with one independent variable and builds up the equation until all independent variables have been considered. The order of inclusion depends upon the magnitude of the partial correlation coefficient. At each step the latest variable's partial F-value is compared with a predetermined value and accepted or rejected. As soon as the most recently entered variable has a nonsignificant F-value, the procedure is terminated.

Stepwise regression. This procedure, despite its different name, is in fact an improvement of the forward selection procedure. The crucial difference is that every previously entered variable is re-examined in terms of its F-value as if it had been the most recently entered variable. This procedure, which may be characterized as forward selection with a backward glance, takes into account the interrelatedness which exists among many predictor variables. Any variable found to be nonsignificant is eliminated from the equation regardless of its point of entry.

It is clear from the above discussion that the stepwise regression procedure is best suited to the problem of curve fitting. Draper and Smith concur. Therefore, the writer chose to use this procedure for his analysis.

Summary

The literature is immense in the area of mental measurements. Accordingly the writer broke the literature review into three manageable parts: an historical overview, an extensive and nonexhaustive survey of the many variables which affect academic performance, and an intensive review of research pertinent to the writer's inquiry. No research using the writer's research design, variables, or similar sample was found.

Finally, the statistical methods appropriate to the writer's study were reviewed and justification was given for the choice of stepwise regression.

Chapter III

THE DESIGN OF THE STUDY

In this chapter the research design and related considerations are reported in detail. The independent and criterion variables are discussed. Background to the Differential Aptitude Test (DAT) and the Otis-Lennon intelligence test is provided. Justification is offered to support the choice of predictors. The sample and population are discussed. The method used in recording the data and converting it to numerical form is reported. A step by step explanation of stepwise regression is supplemented by a discussion of all the statistical problems and issues encountered in analysing the data. Finally, reference is made to use of a regression equation on a subsample.

The Variables Used in the Study

The criterion variable was achievement in Computer Science. As noted in Chapter 1, the reference is to Computer Science 205 currently authorized as a high school course in Manitoba. Six predictor variables were chosen:

1) The Differential Aptitude Test Verbal Percentile;
2) The Differential Aptitude Test Numerical Percentile;

3) The Differential Aptitude Test Abstract Percentile;

4) The Otis-Lennon Intelligence Quotient;

5) The Grade Ten Average;

6) The Mathematics 100 Score.

The review of the literature did not uncover a study which used the same predictors although some of the studies reviewed did involve high school grades, intelligence tests, and aptitude tests. The writer's choice of variables can therefore be justified on the basis that a new area of research is being explored. Furthermore, the writer wanted to choose variables whose statistics were readily available in a modern high school with a well developed diagnostic program. Each variable will now be examined and justified in more detail.

<u>The Differential Aptitude Tests</u>

As noted in the historical overview of this study, the concept of a global measure of intelligence has had many critics. These critics asserted that tests should be created which were capable of assessing the various facets of mental ability. It was important to know what specific aptitudes a student had for purposes of counselling, vocational training, and career selection. One of the best known test batteries developed for these purposes is the Differential Aptitude Test battery.

In 1947, the Psychological Corporation published the first version of the Differential Aptitude Tests (DAT) under the joint authorship of Bennet, Seashore and Wesman. The tests purported to differentiate between and provide a score for each of the following: 1) verbal reasoning, 2) numerical ability, 3) abstract reasoning 4) clerical speed and accuracy, 5) space relations, 6) spelling, 7) grammar, and 8) mechanical reasoning. All of the eight subtests of the DAT are power tests except for clerical speed and accuracy. A power test measures the strength rather than the speed or accuracy which a person has in a specific area. Alternate forms are available and the entire battery can be administered in six class periods, preferably over a short span of days. It is possible to subtest selectively without regard to the entire battery.[1]

In a review of the DAT, Quereshi acknowledges that "Reviewers in the past have justifiably dubbed the DAT as the 'best' available instrument of its kind."[2] He, Carroll, Horrocks, and Schoonover are however aware of certain problems with the DAT, such as its inability to take account of factors such as motivation and the apparent interrelatedness of the variables. The writer took the position in his research that despite the difficulties associated with the DAT, its status as the "best" test of its kind made it worthy of consideration as a predictive instrument.

Of the eight DAT subtests, the writer chose three: 1) Verbal Reasoning, 2) Numerical Ability, and 3) Abstract Reasoning. The other five subtests had no obvious face

---

[1] John E. Horrocks and Thelma I. Schoonover, Measurement For Teachers (Columbus, Ohio: Charles E. Merrill Publishing Company, 1968), p. 386.

[2] M. Quereshi in The Seventh Mental Measurements Yearbook, Vol. II, ed. by Oscar Krisen Buros (Highland Park, N. J.: The Gryphon Press, 1972), p. 1049.

validity as predictors of Computer Science and were therefore discarded. The Verbal subtest was chosen because verbal ability is commonly regarded as an important element of intelligence and therefore ought to be a predictor of Computer Science, a subject which takes the full measure of a student's intellectual capacity. The Numerical subtest was chosen because computer science is very similar to mathematics in that both disciplines demand a capacity to handle symbolism. The Abstract subtest was chosen because many of the algorithms which a student in Computer Science is required to develope demand abstract thought on a high level.

The choice of three of the DAT subtests was soundly based on considerations of reviews by other writers, face validity, availability, and the nature of the intellectual capacities demanded of a student enrolled in Computer Science.

Otis-Lennon Mental Ability Tests

These tests(alternate forms) follow in Binet's tradition in so far as they are conceived to be broadly based measures of general intellectual ability.[3] This does not mean that an innate capacity is being measured. The authors clearly point out that "...the Otis-Lennon tests are measures of learned or developed abilities in the broadest sense.[4]

[3]Arthur S. Otis and Roger T. Lennon, Technical Handbook (New York: Harcourt,Brace and World, Inc., 1969),p.8.

[4]Ibid., p. 7.

These tests can be administered to groups, require no more than sixty minutes for testing, and are capable of assessing intelligence from kindergarten to grade twelve. To help prevent misinterpretation by persons untrained in assessment, the results are reported as percentiles and stanines as well as the usual deviation intelligence quotient.[5]

It should be noted that Arthur S. Otis, deceased, one of the authors of this test, had a great deal of experience in testing. His credits included the Otis Group Intelligence Scale, Otis Self-Administering Tests of Mental Ability, Otis Employment Tests, Otis Quick-Scoring Mental Ability Tests, many journal articles and several books.[6] His experience dated back to the original Army Alpha tests. Knowledge of this background makes the following assessment not unexpected:

> The construction and norming of this test bespeaks adherence to the highest level of current standards. Sophisticated authorship apparently backed by the publisher's determination to spare no effort or expense has resulted in a product of exceptional merit.[7]

The Otis-Lennon Mental Ability Test was found to be one of the best intelligence tests available and for this reason to be worthy of inclusion in the writer's study.

---

[5] Ibid., pp. 4-23.

[6] Arthur S. Otis and Roger T. Lennon, Manual for Administration (New York: Harcourt Brace and World, Inc., 1967), p. 5.

[7] John E. Milholland in The Seventh Mental Measurements Yearbook, I, p. 690.

Grade Ten Average

The writer chose the Grade Ten Average as a predictor of Computer Science achievement. Lavin reviewed the research of Swenson and others and came to this conclusion:

> Of all the measures used in these prediction batteries, the one that consistently emerges as the best single predictor is the high school average or high school rank.....Cronbach reported one study in which correlations between intelligence and grades were .55. Travers found that correlations between intelligence and grades ran between .50 and .75 on the eighth- to tenth-grade levels.[8]

It is not surprising that past academic performance should be a good predictor of future academic performance. Since the literature showed the high school average to be an excellent predictor of future academic performance, the writer chose to include the Grade Ten Average as a predictor in his study.

Mathematics 100 Score

At the time of the study there were four courses in mathematics available to Manitoba grade ten students: Mathematics 100 for students of university caliber; Mathematics 101 for average students who would likely proceed to a community college; Mathematics 102 for students who wanted business mathematics training; and Mathematics 103 for vocational and industrial students.

In the course of preparing the research design of

---

[8] Lavin, The Prediction of Academic Performance, p. 52.

this study the writer became aware that several local high schools had adopted the practice of using Mathematics 100 as a criterion for admission to Computer Science in the following year. The teachers, counsellors, and administrators had evidently come to the conclusion that a strong mathematics background was essential to success in Computer Science. Students who enrolled in the other available grade ten mathematics courses were not considered to have a sufficiently strong mathematics background.

The writer, on the basis of his own classroom experience, agreed with other teachers that a relationship probably did exist between achievement in Computer Science and achievement in Mathematics 100. Accordingly, the Mathematics 100 variable was selected as a predictor in order that the relationship could be explored.

The Sample and Population

The sample of 91 students was drawn from three suburban high schools: Garden City Collegiate, West Kildonan Collegiate, and River East Collegiate over a period of two years. The students of six different Computer Science teachers were involved in the study. Every student who took Computer Science at West Kildonan Collegiate and Garden City Collegiate during that two year period was included in the study providing that all of the DAT, Otis-Lennon, Grade Ten Average, and Mathematics 100 scores were available for that individual. River East Collegiate was sampled for only one year because enrollments in Computer Science were larger there and therefore more student data were returned to the writer in the first sampling. It was the writer's intention to retain some balance among the three high schools with regard to the number of student statistics drawn from each school. Table 2 contains information re the sample, population and the number of teachers involved at each of the three schools.

The writer did not claim that the sample was a scientific random sample of the schools involved. It appeared however, that the students upon whom the research was based were typical academic students who were attending public high schools similar to others in terms of staffing, funding, size, and curriculum. No factors were observed which intervened to differentiate these students from other academically oriented students.

Recording the Data

The data was recorded on the Computer Science Achievement Prediction Data Form found in Appendix A. The form was constructed in a manner such that data were not likely to be entered incorrectly. At West Kildonan Collegiate and River East Collegiate it was not possible for the writer to record the data personally. He therefore carefully explained the research and the form to each of the other teachers involved. The co-operating teachers were instructed to include every student, even if some of the scores were missing. In this way the writer was able to monitor the number of students who eventually had to be excluded from the study, as well as the

Table 2

The Sample and Population in Relation to the Number of Teachers and the Schools Involved in the Prediction Study

School	Number of students in sample	Number of Computer Science teachers	Number of students in school
River East Collegiate	40	2	1,000
Garden City Collegiate	25	2	900
West Kildonan Collegiate	26	2	700
Totals	91	6	2,600

reasons for exclusion.

At River East Collegiate the grading system was the traditional percentage system and all data were recorded in numerical form. At the other two schools a letter grade system was in use. Accordingly, the Computer Science, Mathematics 100, and Grade Ten Average scores had to be converted from letter grades to percentages. This was done with the aid of a conversion table supplied by the school. Thus, if a student's Mathematics 100 score was a letter grade of B and the conversion table (See Appendix B) specified a B range from 70 to 79 inclusive, a rounded midpoint integral score of 75 was recorded for that student. This conversion was in harmony with current high school practice.

The letter grade average posed a special problem. Here, the procedure was to convert each academic letter grade mark (guidance and physical education scores were excluded) to its equivalent percentage. The percentages were then averaged to obtain the Grade Ten Average in numerical form.

No difficulties were found in recording the DAT or Otis-Lennon scores because these were simply numerical in nature and no conversion was necessary.

The Analysis Technique

The justification for an analysis based on stepwise regression was presented in Chapter II. The actual regression was performed by the University of Manitoba's IBM 370 computer using the computer program described by Mitchell and hereafter referred to as the "SPSS package".[9] A summary of the regression procedure is provided.

Step 1. The most highly correlated independent variable was entered into regression with the criterion variable.

Step 2. Using partial correlation coefficients, the next independent variable most highly correlated with the criterion was entered into regression.

Step 3. Each variable previously selected was re-examined under a partial F-test to determine whether it still made a significant contribution to the prediction. If significant, the variable in question was retained and if not, it was rejected.

Step 4. A new regression equation was generated using least squares.

Step 5. The procedure outlined in steps 2 to 4 was performed recursively until all independent variables had been examined.

Using the regression procedure an equation of the form $Y = B_1X_1 + B_2X_2 + \ldots B_nX_n + C$ was developed. Y represents the predicted Computer Science score, B_i where i = 1,6 represent the regression coefficients, and X_i where i = 1,6 represent the predictor variables. C is the regression constant.

---

[9] W. C. Mitchell, "Multiple Regression Analysis: Subprogram Regression", Statistical Package for the Social Sciences, ed. N. Nie et al (New York: McGraw - Hill Inc., 1970), pp. 174-195.

Some Statistical Considerations

The writer found the SPSS package to be lacking in documentation. This made interpretation of the computer output rather difficult. This section of the study was included to explain some points of interest in understanding the F-test and stepwise regression.

First of all, statistical inference can be broken down into two main categories: parametric and nonparametric inference. Parametric tests are computed from sample statistics such as the arithmetic mean ($\overline{X}$), the variance (S^2), and the correlation coefficient (r). Each of these statistics is an estimate of a population parameter. This explains why these tests are referred to as "parametric". Common parametric tests are those associated with the t-distribution and the z- or F-distribution. Nonparametric tests are not computed from estimates of population parameters. Instead, these tests are evaluated in terms of frequencies, proportions or ranks drawn from the sample. Examples of nonparametric tests are those associated with the binomial distribution, the Poisson distribution, and the chi-square distribution.

The writer's research design involved stepwise regression, which in turn involved a parametric test: the F-test. Unlike the t-test, which involves comparisons between means, the F-test makes a comparison between two variances. A simple ratio represents the F level in many cases:

$$F = \frac{\text{variance between classes}}{\text{variance within classes}}$$

where the variance within classes is usually assumed to be homogeneous.

The concept of variance can be developed quickly.[10] The simplest approach is through the standard deviation (a measure of dispersion) of a finite population:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N}}$$

where: σ is the standard deviation:

X_i is the variate;

μ is the population mean; and

N is the number of variates in the population.

Squaring both sides of the formula yields the variance

$$\sigma^2 = \frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N}$$

Both formulas presented here are intuitively meaningful. It should be remembered however, that research typically deals with a sample rather than a population, in which case a different formula for the variance should be used.

---

[10] J. E. Hill and A. Kerber, Models, Methods and Analytical Procedures in Education Research (Detroit: Wayne State University Press, 1967), p. 167.

To conceptualize the F-ratio as used in stepwise regression, in geometric terms, consider Figure 1. In Figure 1 consider V_1 to be the variance contribution due to using a slope; V_2 to be the variance of observations about the regression line; and $\bar{Y}$ to be the average y-value of the plotted points. The F-ratio may be interpreted[11] to be:

$$F = \frac{\text{Variance contribution due to using a slope}}{\text{Variance of observations about the regression line}}$$

An interesting theoretical point considered was whether hypotheses, null or otherwise, should be stated explicitly since F-tests are conducted at each step of the regression. No statistics text reviewed by the writer dealt with this point specifically. Rather the authors seemed to avoid the problem. The writer inferred that the statisticians' viewpoint could be summarized as follows:

1) The hypotheses are stated implicity prior to computation of the F-ratios; and

2) Since this type of analysis usually involves many F-tests it would be exceedingly cumbersome to set up an explicit hypothesis for each test.

Accordingly, the stepwise regression research design used in this study did not involve the explicit statement of an hypothesis.

Another consideration dealt with concerned the

---

[11] Explanation by Dr. B. Johnston, Statistics Department, University of Manitoba, August 1, 1974.

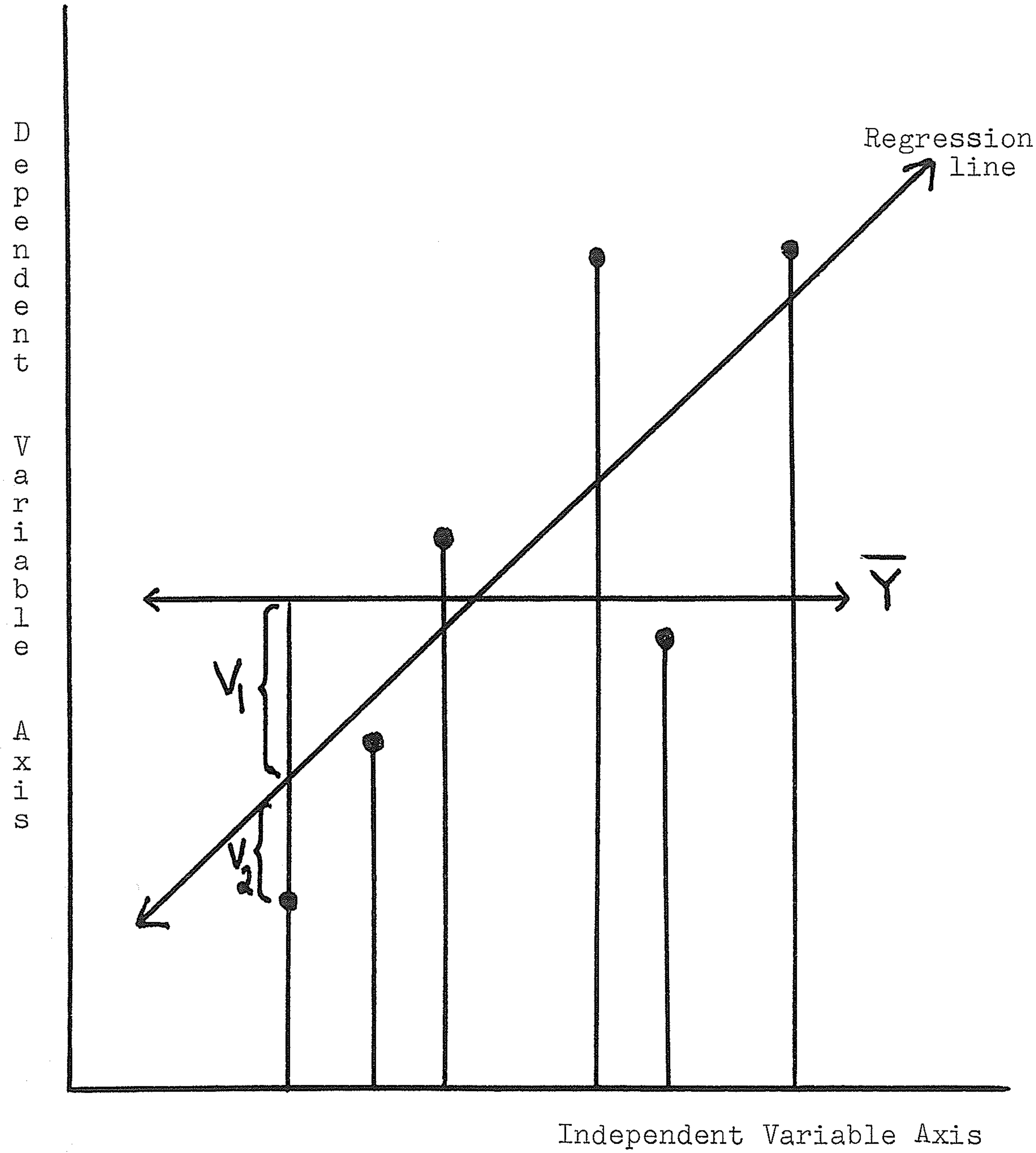


Figure 1

Geometric Interpretation of the F-Ratio in Regression:

$$F = \frac{V_1}{V_2}$$

problem of one-tailedness as opposed to two-tailedness. Johnston stated that the F-tests used in stepwise regression were one-tailed.[12] Ferguson agrees, stating that in the analysis of variance the F-ratio is one-tailed, and the F tables found in any standard reference are also one-tailed.[13]

An interesting problem encountered involved the degrees of freedom of each step of the regression. The specific question which had to answered was "Are the degrees of freedom the same for the F level of the regression equation and the partial F computed at each step?" In answer to the question, each partial F was calculated with one degree of freedom at any step of the regression but the F level of the regression equation was tested on the basis that a degree of freedom was added at each successive step of the regression.[14]

A final issue encountered was the level of significance which a researcher sets for his research design. There are two approaches to the problem. Traditionally, researchers have chosen a level of significance appropriate to their research, done the research, and accepted the inferential result of the chosen level of significance. In reading the literature, a more recent trend became apparent. Some researchers now take a "fishing trip" approach in that they do

---

[12] Ibid., Dr. B. Johnston.

[13] George A. Ferguson, Statistical Analysis in Psychology and Education (New York: McGraw - Hill Book Company, 1959), p. 183.

[14] Ibid., Dr. B. Johnston.

not preselect the significance level. Instead, they do the research, perform the significance tests, and report results at whatever level of significance is generated by the data. This latter approach may be regarded as less rigorous than the first but its supporters would claim that significant results, at whatever level, are not ignored. Faced with a choice between these two alternatives, the writer chose the conservative approach and preset the level of significance at .05. He compromised, however, by deciding to report any results which were significant at a higher level of significance.

The problems outlined above posed difficulties to the writer. A discussion of them was considered necessary to compensate for lack of sufficient documentation in the SPSS package used in the computerized execution of the regression.

Application of a Regression Equation

After the regression equations had been developed, a subsample of 27 students was drawn randomly from the total sample. The nonnormalized regression equation was used to predict each student's Computer Science score. The predicted and actual scores were compared on the basis of correlation, means and standard deviations. No test of significance was done because the subsample was drawn from the original sample. In this context a significance test would have been redundant. The results of this application are reported in

Chapter IV.

Summary

In this chapter the research design was explained. Consideration was given to the variables, the sample, and the population. The method of recording the data was reported. The statistical technique and related problems were explored. Reference was made to use of the final regression equation on a subsample.

As noted before, the actual computations were performed by computer. This did not absolve the writer of the necessity to become familiar with the statistical concepts involved. The relevent statistical considerations were reported in this chapter.

Chapter IV

THE FINDINGS OF THE PREDICTION STUDY

In this chapter the statistical results are presented. The variables are summarized in terms of range, mean, standard deviation, kurtosis and skewness. A simple correlation matrix for all variables is presented, as well as partial correlations between the dependent and independent variables. The regression is traced through step by step in terms of the multiple correlation, per cent of variance accounted for and significance. Inferences are drawn. The regression equations are developed and interpreted.

Characteristics of the Predictors and Criterion

In this study 91 students were used as a sample to obtain the representative statistics. The characteristics of the independent and dependent variables are set forth in Table 3. The kurtosis and skewness figures in Table 3 are included to give an indication of how the frequency curve for each variable differed from the normal curve. Evidently all of the variables were negatively skewed and all except the DAT Abstract Percentile and Computer Science Score were platykurtic to some extent. However, no extreme variations likely to invalidate the study were present. A graphical analysis showed the skewness and kurtosis to be undetectable visually. Therefore the graphs were omitted from this report.

Table 3

Population Sample Range, Mean, Standard Deviation, Kurtosis and Skewness for Each of the Six Independent Variables and One Dependent Variable

Variable number	Variable	Range	Mean	Standard deviation	Kurtosis	Skewness
1	DAT Verbal Percentile	3-99	65.264	27.255	-0.571	-0.765
2	DAT Numerical Percentile	10-99	71.934	23.447	-0.462	-0.779
3	DAT Abstract Percentile	15-100	77.055	20.146	+0.676	-1.139
4	Otis-Lennon IQ	92-144	117.352	11.008	-0.319	-0.042
5	Grade Ten Average	51-95	72.538	9.337	-0.320	-0.009
6	Math 100 Score	50-98	74.516	11.847	-0.737	-0.159
7	Computer Science Score	15-100	70.824	18.754	+0.325	-0.798

Table 4

Simple Correlation Matrix for all Variables

	DAT Verbal Percentile	DAT Numerical Percentile	DAT Abstract Percentile	Otis-Lennon IQ	Grade Ten Average	Math 100 Score	Computer Science Score
DAT Verbal Percentile	1.000						
DAT Numerical Percentile	0.470	1.000					
DAT Abstract Percentile	0.382	0.382	1.000				
Otis-Lennon IQ	0.654	0.672	0.433	1.000			
Grade Ten Average	0.521	0.373	0.223	0.547	1.000		
Math 100 Score	0.480	0.382	0.324	0.504	0.773	1.000	
Computer Science Score	0.295	0.350	0.286	0.378	0.470	0.479	1.000

The Correlation Matrix

The simple correlation matrix for all variables is presented in Table 4. An examination of this table shows that Computer Science and Math 100 are moderately correlated ($r = .479$). Computer Science and the Grade Ten Average are also moderately correlated ($r = .470$). The correlations between Computer Science and the other predictors are relatively low.

Another interesting observation is that there exist some relatively high correlations among the predictor variables. For example, the Otis-Lennon IQ is correlated with the DAT Numerical Percentile at $r = .672$. The Stepwise regression procedure derives its power by virtue of being able to discount such inter-relatedness among the predictor variables in the course of generation of the regression equation.

The Stepwise Regression

The stepwise regression findings are presented in Tables 5, 6, and 7. Table 5 shows the order in which the independent variables were entered into the regression equation. The order of entry depended upon the descending order of magnitude of the partial correlations between the predictor and criterion variables. In this manner the order of entry of the predictors was: Math 100, DAT Numerical Percentile, Grade Ten Average, DAT Abstract Percentile, DAT Verbal Percentile, and finally, the Otis-Lennon IQ. Table 5

Table 5

Partial Correlation Coefficients Between Each of the Predictor Variables and the Criterion Variable

Step number	Variable entered	Partial correlation coefficient*
1	Math 100 Score	0.479
2	DAT Numerical Percentile	0.206
3	Grade Ten Average	0.158
4	DAT Abstract Percentile	0.120
5	DAT Verbal Percentile	-0.054
6	Otis-Lennon IQ	0.027

*The simple correlation coefficient was given for the first variable entered into regression.

shows the partial correlation coefficients.

A comparison of Table 4 with Table 5 was rather interesting. For example, in Table 4 the correlation between Computer Science and the Grade Ten Average was rather good ($r = .470$). But in Table 5 the same two variables have a partial correlation of only 0.158. This low partial correlation caused the Grade Ten Average variable to be entered into regression in the third step rather than in the second step as one might surmise from a casual examination of Table 4. It is evident from this example that the partial correlation is of more importance than the simple correlation in considering the predictive capability of a variable.

Examination of columns 3 and 4 of Table 6 showed that the variables entered in steps 1 and 2 of the regression appeared to make relatively important contributions to the multiple correlation ($R = .479$ and $R = .512$) as well as to the per cent of variance accounted for (22.903 per cent and 26.180 per cent). The contribution made by the Grade Ten Average was rather marginal in that the per cent of variance accounted for rose only to 28.026 per cent. The contribution made by the last two variables, the DAT Verbal Percentile and the Otis-Lennon IQ, was apparently rather negligible.

Columns 5 and 6 of Table 6 were of some limited use in determining whether a regression equation based on the six independent variables would be statistically significant. At each step of the regression, the F level of the regression equation was higher than the required .05 level of significance.

Table 6

Stepwise Regression Findings for Six Independent Variables and One Dependent Variable

Step Number	Variable entered	Multiple R	Per cent of variance accounted for	F level of regression equation	Required .05 F level
1	Math 100 Score	0.479	22.903	26.438	3.96
2	DAT Numerical Percentile	0.512	26.180	15.604	3.11
3	Grade Ten Average	0.529	28.026	11.292	2.72
4	DAT Abstract Percentile	0.539	29.058	8.807	2.48
5	DAT Verbal Percentile	0.541	29.262	7.032	2.33
6	Otis-Lennon IQ	0.541	29.313	5.806	2.21

On this basis an incautious researcher might have inferred that all six predictors should have been included in the regression equation. Judgement was withheld, however, until consideration had been given to Table 7.

As the stepwise regression proceeded a partial F was calculated for each new variable being entered into the regression equation. The significance of the most recently entered variable was established by comparing the computed partial F value with a table of F values, using one degree of freedom. This was done at the .05 and .01 levels of significance and the results are summarized in Table 7.

Table 7 clearly shows that only one variable, Math 100, made a significant contribution to the prediction of Computer Science scores. The computed F level is 26.438 with only 6.96 being required for a .01 level of significance. Thus, while the other five predictors made contributions to the accounted for variance and the multiple correlation, they were rejected because they did not contribute significantly.

The Regression Equations

The findings summarized in Table 7 were used to develop regression equations. Variables found to be of no significance were excluded. Two equations were developed at the .01 level of significance, one for nonnormalized data and one for normalized data.

Since only one variable was found to be significant

Table 7

Partial F Findings for the Six Independent Variables in Regression

Step number	Variable entered	Partial F
1	Math 100 Score	26.438*
2	DAT Numerical Percentile	3.907**
3	Grade Ten Average	2.232
4	DAT Abstract Percentile	1.252
5	DAT Verbal Percentile	0.245
6	Otis-Lennon IQ	0.060

*Significant at the .01 level.

**Not significant.

the equation for nonnormalized data was of the type $Y = B_6X_6 + C$ where Y was the predicted Computer Science score on a scale from 0 to 100; B_6 was a coefficient which could be interpreted as the slope of a straight line; X_6 was the student's Math 100 score on a scale from 0 to 100; and C was a constant which could be interpreted as the y-intercept of the straight line. B_6 and C were computed and fitted into the equation which then became:

$$Y = 0.758\,X_6 + 14.369$$

The standard error for this equation was found to be 16.559 with 22.903 per cent of the variance accounted for. The standard error may be regarded as the typical error in prediction.[1]

The equivalent normalized equation, which would generate a Computer Science standard score was of the type

$$Y = \beta_6 \frac{(X_6 - \overline{X_6})}{S.\ D.}$$

where Y was the predicted Computer Science score in standard form; β_6 was a coefficient; X_6 was the student's Math 100 score on a scale from 0 to 100; $\overline{X}_6$ was the average Math 100 score; and S. D. was the standard deviation of the Math 100 variable. Using the values for $\overline{X}_6$ and S. D. reported in Table 3, the standardized equation was found to be:

[1]Mitchell, "Multiple Regression Analysis," p. 185.

$$Y = \frac{0.479\ (\ X_6 - 74.516\ )}{11.847}$$

Note that no regression constant C exists in the standardized regression equation. This is always the case because when all the independent variables take on their mean value, the predicted value is the mean value of the dependent variable.[2] Another interesting point to note is that the beta coefficient, $\beta = 0.479$ is also the simple correlation between Computer Science and Math 100.

Interpretation of a Regression Equation

Consider the nonstandardized equation

$$Y = 0.758\ X_6 + 14.369$$

where Y is the predicted Computer Science score on a scale from 0 to 100 and X_6 is the student's Math 100 score on the same scale. The standard error of this equation was found to be 16.559 with 22.903 per cent of the variance accounted for. A teacher or counsellor using this equation for counselling students would have to be very cautious. A student whose Math 100 score generated a prediction of 70 in Computer Science would have a 68 per cent chance of scoring in the range of 70 ± 17 where 17 is the rounded standard error. His score would be anywhere from 53 to 87 but the most probable score would be 70.

---

[2] Mitchell, op. cit., p. 177.

The standardized prediction might be used as follows. Suppose a student's predicted Computer Science score were +1. This could be interpreted to mean that he would probably score as well as or better than 84 per cent of his fellow students, subject to the same uncertainties as before.

In order for a regression equation to be meaningful and useful to a student, the counsellor must be prepared to explain fully the meaning of the prediction. To use the equation in any other way would be of doubtful value.

Comparison of Predicted and Actual Computer Science Scores

As noted in Chapter III, part of the research design was to draw a subsample of 27 students from the original sample and do a comparison between the actual and predicted Computer Science scores. The subsample findings are recorded in Tables 8 and 9. Means, Standard Deviations and the Correlation between the Actual and Predicted Scores are reported.

No statistical analysis of the comparison between the actual and predicted Computer Science scores which involved a significance test was attempted because such an analysis would have been redundant. It could only have had significance if a new sample had been drawn from the population. Such a sample was not available to the writer and hence the research design did not include a test for significant difference between the actual and predicted Computer Science scores. The data in Tables 8 and 9 were set forth

Table 8

Actual and Predicted Computer Science
Scores of Students Comprising the Subsample

Student number	Math 100 Score	Actual Computer Science Score	Predicted Computer Science Score
4	75	83	71.22
6	75	67	71.22
10	67	85	65.15
12	63	63	62.12
13	63	67	62.12
14	95	77	86.38
17	98	100	88.65
18	94	95	85.62
21	85	100	78.80
23	76	53	71.98
25	85	95	78.80
27	90	96	82.59
28	92	79	84.10
30	78	86	73.49
33	84	64	78.04
37	79	71	74.25
38	73	92	69.70
45	83	87	77.28
51	65	78	63.64
59	50	61	52.57
60	64	66	62.88
61	64	50	62.88
62	86	63	79.56
65	79	88	74.25
73	75	31	71.22
83	63	85	62.12
87	63	63	62.12

Table 9

Means, Deviations and Correlation
Of and Between the Actual and Predicted
Computer Science Scores

Variable	Mean	Standard deviation	Correlation
Actual Computer Science score	75.741	16.997	
			0.474
Predicted Computer Science score	72.314	9.244	

so that an intuitive understanding of how well the regression equation predicted performance could be gained. It is the writer's opinion that the regression equation performed about as well as one could expect since it was based on 22.903 per cent of the variance accounted for.

Summary

The findings of the study were reported in this chapter. The variables were summarized in Table 3. A correlation matrix was presented. The stepwise regression was reported and certain inferences regarding significance were drawn. The regression equations were explained in terms of their possible school use. Finally, a regression equation was used to predict the Computer Science scores of a 27 student subsample and some descriptive statistics regarding the comparison of actual and predicted Computer Science scores were presented.

Chapter V

SUMMARY AND CONCLUSIONS

In this chapter, a review of the study is conducted with a summary of the findings presented in Tables 3 to 9. Conclusions are drawn and recommendations are made for further research.

Review of the Study

This study arose out of the writer's personal interest in mental testing. The writer perceived mental tests, supplemented by data such as the academic average, as devices which would enable a valid prediction of future achievement in Computer Science to be made. The writer also wanted to establish practical criteria to be used by high school personnel in counselling students re the advisability of studying Computer Science.

The research design was set up so as to involve variables represented by statistics available in the school records; the DAT scores, the Otis-Lennon IQ, the Grade Ten Average and the Math 100 Score. These variables were used as predictors of a criterion variable, Computer Science achievement. A stepwise regression at the .05 level of significance was performed and the findings were reported in Chapter IV.

The inferences to be drawn from Chapter IV suggest that the writer expected too much from mental testing. The relatively low correlations which exist between the mental test variables and the Computer Science Score may be attributable to the fact that the sample was rather homogeneous. This homogeneity is shown in Table 3 where the Otis-Lennon IQ is 117 with a standard deviation of 11.008. The norms are 100.6 and 15.7 for this particular test.[1] Perhaps, if the sample had been truly heterogeneous, one or more of the mental test variables would have been significant.

Another possibility investigated to account for the moderate correlations reported in this study was whether the implicit assumption of continuous variables was seriously violated by converting the letter grade scores to midpoint scores. Dr. Johnston confirmed that the conversion procedure used was conceptually quite acceptable.[2] This possibility was therefore ruled out.

The study was a success in the sense that, of the independent variables investigated, a judgement could be rendered re their relative importance as predictors of achievement in Computer Science.

---

[1]Arthur S. Otis and Roger T. Lennon, Technical Handbook (New York: Harcourt, Brace and World, Inc., 1969), p. 19.

[2]Ibid., Dr. B. Johnston, January 13, 1975.

Major Findings

The statistical procedure generated information about simple correlations, partial correlations, multiple correlations, per cent of variance accounted for and the significance of each variable entered. These findings may be summarized as follows:

1) The simple correlation matrix showed moderate to strong correlations between many of the variable-pairs.

2) The partial correlations computed in the step-wise regression ranged from a moderately strong 0.479 to a very weak 0.027.

3) The multiple correlation rose from 0.479 to 0.541 during the six steps of regression.

4) The variance accounted for rose from 22.903 to 29.313 per cent during the six steps of regression.

5) Only the first regression step yielded results significant at the .05 level. These results were also significant at the .01 level.

6) The Math 100 variable was found to be a very significant predictor of achievement in Computer Science. No other predictor variable made a significant contribution to the prediction at the .05 level of significance.

7) The regression equation, when tested on the subsample, appeared to work reasonably well according to the descriptive statistics shown in Table 9.

Major Conclusions

Any conclusions drawn must be rather tentative since only 22.903 per cent of the variance was accounted for with the one significant predictor variable. With this caveat, the following major conclusions appear to be warranted:

1) Of the independent variables investigated only achievement in Mathematics 100 was found to be a significant predictor of achievement in Computer Science.

2) The other variables investigated as predictors of Computer Science achievement failed completely to live up to expectations.

Recommendations for Further Study

This study was able to identify only one variable which functioned significantly as a predictor of Computer Science achievement. Moreover, the per cent of variance accounted for was rather low. The writer therefore concluded that further research should be undertaken. The research design envisaged by the writer would be an attempt to improve the prediction. In addition to the Math 100 variable, measures of personality and sociological measures could be included as predictors. Other grade ten subjects could conceivably function as predictors. Another avenue which could be explored is the influence which teachers have on the prediction. Finally, some measure of computer science aptitude currently used in industry should be investigated.

Summary

This study was designed to identify criteria which would be of practical value in helping school administrators, guidance counsellors, and teachers to counsel students re the advisability of enrolling in Computer Science 205, a course currently offered in Manitoba high schools. The study was a success in that one important predictor of Computer Science achievement was identified: achievement in Mathematics 100, the most rigorous mathematics course offered at the grade ten level in Manitoba.

SELECTED BIBLIOGRAPHY

SELECTED BIBLIOGRAPHY

A. Books

Buros, Oscar Krisen, ed. Tests in Print. Highland Park, N. J.: The Gryphon Press, 1961.

Buros, Oscar Krisen, ed. The 1938 Mental Measurements Yearbook. New Brunswick: Rutgers University Press, 1938.

Buros, Oscar Krisen, ed. The Seventh Mental Measurements Yearbook. Highland Park, N. J.: The Gryphon Press, 1972.

Cronbach, L. J. Essentials of Psychological Testing. New York: Harper and Brothers, 1949.

Draper, N. R., and Smith, H. Applied Regression Analysis. New York: John Wiley and Sons, Inc., 1968.

Ferguson, George A. Statistical Analysis in Psychology and Education. New York: McGraw-Hill, Inc., 1959.

Guilford, J. P. Fundamental Statistics in Psychology and Education. Third ed. New York: McGraw-Hill Book Co., 1956.

Hill, J. E., and Kerber, A. Models, Methods and Analytical Procedures in Education Research. Detroit: Wayne State University Press, 1967.

Horrocks, John E., and Schoonover, Thelma I. Measurement For Teachers. Columbus, Ohio: Charles E. Merrill Publishing Company, 1968.

Keeping, E. S. Introduction to Statistical Inference. Princeton: D. Von Nostrand, 1962.

Lavin, David E. The Prediction of Academic Performance. New York: Russell Sage Foundation, 1965.

Linden, James D., and Linden, Kathryn W. Tests on Trial. Boston: Houghton Mifflin Company, 1968.

SELECTED BIBLIOGRAPHY (continued)

Mitchell, W. C. "Multiple Regression Analysis: Subprogram Regression." Statistical Package for the Social Sciences. ed. by N. Nie, D. H. Bent, and C. H. Hull. New York: McGraw-Hill, Inc., 1970.

Otis, Arthur S., and Lennon, Roger T. Technical Handbook. New York: Harcourt Brace and World, Inc., 1969.

Otis, Arthur S., and Lennon, Roger T. Manual for Administration. New York: Harcourt Brace and World, Inc., 1967.

Ruch, G. M., and Stoddard, George D. Tests and Measurements in High School Instruction. Chicago: World Book Company, 1927.

Terman, Lewis M. The Measurement of Intelligence. Cambridge: The Riverside Press, 1916.

Terman, Lewis M., and Merril, Maud A. Stanford-Binet Intelligence Scale. Boston: The Riverside Press, 1960.

B. Periodicals

Alspaugh, John W. "The Relationship of Grade Placement to Programming Aptitude and Fortran Programming Achievement." Journal for Research in Mathematics Education, January, 1971. pp. 44-48.

Bauer, R. S., Mehrens, W. A., and Vinsonhaler, J. F. "Predicting Performance in a Computer Programming Course." Educational and Psychological Measurement, XXVIII (1968), 1159-1164.

Beach, L. R. "Sociability and Academic Achievement in Various Types of Learning Situations." Journal of Educational Psychology, LI (1960), 208-212.

Beckner, Caroline. "Who Should Study Data Processing?" Business Education World, XLVIII (April, 1968), 10-11.

Carter, L. B. "The Effect of Early School Entrance on the Scholastic Achievement of Elementary School Children in the Austin Public Schools." Journal of Educational Research, L (1956), 91-103.

Cooper, J. G. "The Inspection Rorschach in the Prediction of Academic Success." Journal of Educational Research, XXXXIX (1955), 275-283.

SELECTED BIBLIOGRAPHY (continued)

Coster, J. K. "Some Characteristics of High School Pupils from Three Income Groups." Journal of Educational Psychology, L (1959), 55-62.

Dugan, R. D. "Current Problems in Test Performance of Job Applicants." Personnel Psychology, XIX (1966), 18-24.

Durnall, E. J.,Jr. "Predicting Scholastic Success for Graduate Students in Education." School and Society, LXXX (1954), 107.

Hollenbeck, G. P., and McNamara, W. J. "CUCPAT and Programming Aptitude." Personnel Psychology, XVIII (1965), 101-106.

Howell, M. A., Vincent, J. W., and Gay, R. A. "Testing Aptitude for Computer Programming." Psychological Reports, XX (1967), 1251-1256.

Hunt, D., and Randhawa, B. S. "Relationship Between and Among Cognitive Variables and Achievement in Computational Science." Educational and Psychological Measurement, XXXIII, 3(1968), 921-928.

Klugh, H. E., and Bierly, R. "The School and College Ability Test and High School Grades as Predictors of College Achievement." Educational and Psychological Measurement, XIX (1959), 625-626.

McNamara, W. J.,and Hughes, J. L. "A Review of Research on the Selection of Computer Programmers." Personnel Psychology, XIV (1961), 39-51.

Merrill, R. M.,and Murphy, D. T. "Personality Factors and Academic Achievement in College." Journal of Counselling Psychology, VI (1959), 207-210.

Northby, A. S. "Sex Differences in High School Scholarship." School and Society, LXXXVI (1958), 63-64.

Oliver, T. C., and Willis, W. K. "A Study of the Validity of the Programmer Aptitude Test." Educational and Psychological Measurement, XXIII, 4 (1963), 823-825.

Opler, Ascher. "Testing Programming Aptitude." Datamation, October, 1963. pp. 28-31.

Peres, Sherwood H., and Arnold, Phil H. "Identifying Programmer Behavior." Datamation, January, 1963. p. 40.

SELECTED BIBLIOGRAPHY (continued)

Rowan, T. C. "Psychological Tests and Selection of Computer Programmers." Journal of the Association for Computing Machinery, IV (1957), 348-353.

Schoonover, S. M. "The Relationship of Intelligence and Achievement to Birth Order, Sex of Sibling, and Age Interval." Journal of Educational Psychology, L (1959), 143-146.

Sharp, H. C., and Pickett, L. M. "The General Aptitude Test Battery as a Predictor of College Success." Educational and Psychological Measurement, XIX (1959), 617-623.

Walker, Elaine, and Markham, S. J. "Computer Programming Aptitude Tests." Australian Psychologist, V (March, 1970), 52-58.

Weiss, P., Wertheimer, M., and Groesbeck, B. "Achievement Motivation, Academic Aptitude, and College Grades." Educational and Psychological Measurement, XIX (1959), 663-666.

C. Miscellaneous

Alspaugh, Carol Ann. "A Study of the Relationships Between Student Characteristics and Proficiency in Symbolic and Algebraic Computer Programming." Doctor's thesis, University of Missouri, 1970.

Bennet, George K., Seashore, Harold G., and Wesman, Alexander G. Differential Aptitude Tests. Form L specimen set. New York: The Psychological Corporation, 1968.

Biamonte, A. J. "A Study of the Effect of Attitudes on the Learning of Computer Programming." Proceedings of the Third Annual Computer Personnel Research Conference. Silver Spring, Maryland: Computer Personnel Research Group, June 17-18, 1965. pp. 68-74.

Correnti, R. J. "Predictors of Success in the Study of Computer Programming at Two-Year Institutions of Higher Learning." Dissertation Abstracts, 30: 3718A, 1969.

Gordon, Bruce F., and Denis, Richard A. "Characteristics and Performance Predictors of 7094 Computer Service Operators." Proceedings of the Fourth Annual Computer Personnel Research Conference. Los Angeles: Computer Personnel Research Group, 1966. pp. 96-106.

SELECTED BIBLIOGRAPHY (continued)

Johnston, Dr. B. Statistics Department, University of Manitoba. Explanation of statistical procedures. August 1, 1974.

Miller, Robert L. "A Study of the Relationship Which Exists Between Achievement in High School Vocational Data Processing Programs and the Programmer Aptitude Test." Master's thesis, North Carolina State University, 1970.

Meyer, David B., and Stalnaker, Ashford W. "Computer Personnel Research - Issues and Progress in the 60's." Proceedings of the Fifth Annual Computer Personnel Research Conference. Computer Personnel Research Group, 1967. pp. 6-41.

Otis, Arthur S., and Lennon, Roger T. Otis-Lennon Mental Ability Test. Intermediate level, Form J specimen set. New York: Harcourt, Brace and World, Inc., 1967.

Quinn, S. B. "Relationship of Certain Personality Characteristics to College Achievement." Dissertation Abstracts, 17: 809, 1957.

Reinstedt, Robert N., et al. Computer Personnel Research Group Programmer Performance Prediction Study. Memorandum RM-4033-PR, a report prepared for the U. S. Air Force Project Rand, Santa Monica, California, (March, 1964).

Seiler, Joseph. "Abilities for ADP Occupations." Proceedings of the Third Annual Computer Personnel Research Conference. Silver Spring, Maryland: Computer Personnel Research Group, June 17-18, 1965. pp. 52-59.

Seiler, Joseph, "Survey of Validation Studies on Computer Personnel Selection Instruments." Proceedings of the Fifth Annual Computer Personnel Research Conference. Computer Personnel Research Group, 1967. pp. 43-51.

Stalnaker, Ashford W. "The Watson-Glaser Critical Thinking Appraisal as a Predictor of Programming Performance." Proceedings of the Third Annual Computer Personnel Research Conference. Silver Spring, Maryland: Computer Personnel Research Group, June 17-18, 1965. pp. 75-77.

Tibbetts, J. R. "The Role of Parent-Child Relationships in the Achievement of High School Pupils: A Study of the Family Relationships Associated with Underachievement and High Achievement of High School Pupils." Dissertation Abstracts, 15: 232, 1955.

APPENDIX A

Computer Science Achievement

Prediction Data Form

COMPUTER SCIENCE ACHIEVEMENT PREDICTION DATA FORM

Student Initials											
Differential Aptitude Verbal Percentile											
Differential Aptitude Numerical Percentile											
Differential Aptitude Abstract Percentile											
Otis-Lennon Intelligence Quotient											
Grade Ten Average											
Mathematics 100 Score											
Mathematics 101 Score											
Computer Science 205 Score											

APPENDIX B

Letter Grade to Numerical Range Conversion Table

GARDEN CITY COLLEGIATE

ACHIEVEMENT KEY

LETTER GRADES COMPARED TO PERCENTAGES

A+	90% and over
A	80 to 89%
B	70 to 79%
C	56 to 69%
D	50 to 55%
F	Less than 50%
I	Incomplete subject. Student is behind in work and unable to assess his progress at this time.

J. M. Yakmission,

Principal.

APPENDIX C

Letter Requesting Specimen Sets of the Mental Tests

August 7, 1974.

Guidance Centre,
1000 Yonge Street,
TORONTO, Ontario.
M4W 2K8

Dear Sirs,

Please send two specimen sets as identified on the enclosed order form. The specimens are needed for a master's thesis currently under way. The authorization for this order was given by Professor A. M. McPherson, Department Head, Department of Curriculum, Faculty of Education, University of Manitoba.

Thank you for your cooperation. A certified cheque is enclosed.

Yours truly,

A. Geith,

WINNIPEG, Manitoba.
R

Encl. 2

APPENDIX D

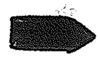
Five Sample Items From The Otis-Lennon Mental Abilities Test, Intermediate Level, Form J

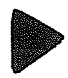
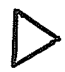
Mark the one best answer for each question in the test. Do not guess blindly, but you may mark an answer even when you are not perfectly sure that it is right.

J-1. A school without pupils is like a book without

a) pictures b) pages c) soft covers

d) hard covers e) an index

J-9.  is to  as  is to

a)  b)  c)  d)  e) 

J-21. Mary has a piece of clay shaped like a block. If she rolls it into a ball, what will happen to the weight and volume of the clay?

a) Weight will decrease but volume will increase.

b) Weight will increase but volume will decrease.

c) Both will increase.

d) Both will decrease.

e) Both will remain unchanged.

J-25. John is facing north. He walks forward 1 block, turns to his right, walks 2 blocks, and then turns to his right again. The direction he is now facing is

a) east b) west c) north d) south

e) impossible to find out from what we are told

J-46. Choose the word that best completes this sentence:

Because the sign warning about the swift river was ________ by the boys, they were nearly drowned.

a) unchecked b) followed c) recognized

d) unheeded e) painted

APPENDIX E

Five Sample Items From the Numerical Subtest of the Differential Aptitude Tests, Form L

You are to pick out the correct answer.....If you are not sure of an answer, mark the choice which is your best guess.

		ANSWER	
L-N-4.	Multiply $\begin{array}{r} 2.04 \\ \underline{\times .75} \end{array}$	A B C D E	1.5300 153.0 1530 15300 none of these
L-N-12.	Divide $.04 \div 4.036$	A B C D E	1.009 10.9 10.09 100.9 none of these
L-N-18.	Add 3 lbs. 3 oz. 6 lbs. 7 oz. 7 lbs. 5 oz. 11 lbs. 1 oz.	A B C D E	28 lbs. 16 oz. 28 lbs. 27 lbs. 16 oz. 18 lbs. none of these
L-N-33.	Cube root $\sqrt[3]{\frac{1}{8} \times \frac{125}{64}}$	A B C D E	5/8 375/512 2 1/2 15 5/8 none of these
L-N-39.	What one number can replace both question marks? $\frac{6.25}{?} = \frac{?}{16}$	A B C D E	4 10 16 50 none of these

APPENDIX F

Five Sample Items From the Verbal Subtest of the Differential Aptitude Test, Form L

You are to pick out words which will fill the blanks so that the sentence will be true and sensible. If you are not sure of an answer, mark the choice which is your best guess.

L-V-2. ..... is to cavalry as foot is to .....

A. horse ______travel
B. horse ______infantry
C. horse ______yard
D. cemetery ___yard
E. horse ______armory

L-V-18. ..... is to childhood as adolescence is to.....

A. infantry ______adultery
B. infancy _______maturity
C. infantry ______intelligence
D. infancy _______adultery
E. health ________intelligence

L-V-28. ..... is to horse as bray is to .....

A. neigh _______donkey
B. hoof ________donkey
C. saddle ______wagon
D. hoof ________wagon
E. hoof ________pony

L-V-41. ..... is to physician as secretary is to.....

A. doctor ______office
B. nurse _______executive
C. doctor ______stenographer
D. medicine ____office
E. medicine ____executive

L-V-50. ..... is to diamond as circle is to.....

A. gold ______round
B. square ____oval
C. shape _____round
D. cube ______round
E. square ____round

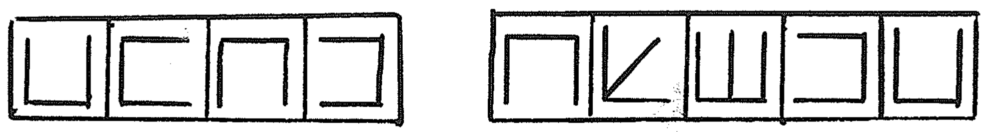
APPENDIX G

Five Sample Items From the Abstract Subtest of the Differential Aptitude Tests, Form L

You are to select the one figure from among the Answer Figures which belongs next in the series. If you are not sure of an answer, mark the choice which is your best guess.

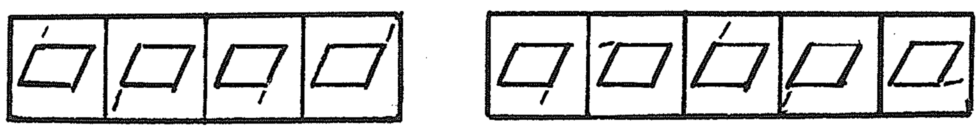
PROBLEM FIGURES ANSWER FIGURES

L-A-6.



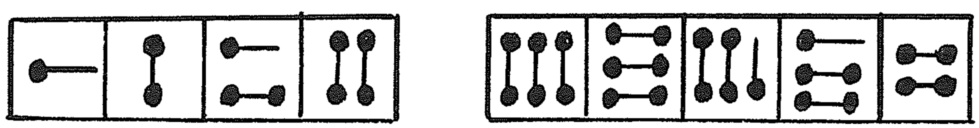
A B C D E

L-A-15.



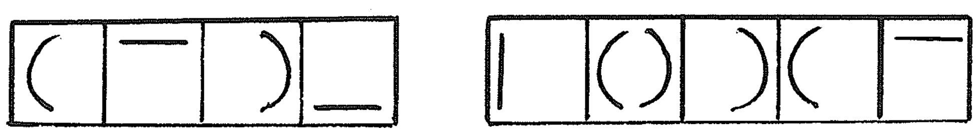
A B C D E

L-A-23.



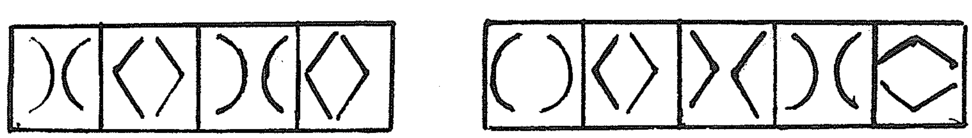
A B C D E

L-A-37.



A B C D E

L-A-46.



A B C D E