

A COMPARISON OF MILLER ANALOGIES TEST SCORES WITH  
UNDERGRADUATE AND GRADUATE GRADE-POINT  
AVERAGES OF GRADUATE STUDENTS

APPROVED:

*Sidney Hamilton*

Major Professor

*Harold D. Holloway*

Minor Professor

*Luane Kingery*

Dean of the School of Education

*Robert B. Touloucy*

Dean of the Graduate School

A COMPARISON OF MILLER ANALOGIES TEST SCORES WITH  
UNDERGRADUATE AND GRADUATE GRADE-POINT  
AVERAGES OF GRADUATE STUDENTS

THESIS

Presented to the Graduate Council of the  
North Texas State University in Partial  
Fulfillment of the Requirements

For the Degree of

MASTER OF SCIENCE

By

Don Matlosz, B. S.

Denton, Texas

January, 1969

## TABLE OF CONTENTS

|                                       | Page |
|---------------------------------------|------|
| LIST OF TABLES . . . . .              | iv   |
| CHAPTER                               |      |
| I. INTRODUCTION . . . . .             | 1    |
| Statement of the Problem              |      |
| Review of the Literature              |      |
| Statement of the Hypotheses           |      |
| II. METHOD . . . . .                  | 21   |
| Subjects                              |      |
| Procedure                             |      |
| Description of Instrument             |      |
| III. RESULTS AND DISCUSSION . . . . . | 30   |
| Results                               |      |
| Discussion                            |      |
| IV. SUMMARY. . . . .                  | 40   |
| BIBLIOGRAPHY. . . . .                 | 45   |

## LIST OF TABLES

| Table   | Page |
|---|------|
| I. Multiple and Single Predictions of Graduate School Success Using a Number of Variables .....                                   | 8    |
| II. Average <u>Miller Analogies Test</u> scores of Graduate Students at Syracuse University for Different Honor Point Ratios..... | 27   |
| III. Reliability Coefficients for Four Forms of the <u>Miller Analogies Test</u> .....  | 28   |
| IV. Coefficients of Correlation for <u>Miller Analogies Test</u> Scores and Grade-Point Averages.....                             | 30   |
| V. Average <u>Miller Analogies Test</u> Scores for Students for Grade-Point Averages in Graduate School.....                      | 31   |
| VI. Mean, Standard Deviation and $\bar{t}$ for Male and Female Subjects - Summary of Results.....                                 | 32   |
| VII. Mean, Standard Deviation and $\bar{t}$ for Guidance and Elementary Education Students - Summary of Results.....              | 33   |
| VIII. Mean, Standard Deviation and $\bar{t}$ for Subjects in Age Groups of 20-35 and 36-60 - Summary of Results.....              | 33   |

## CHAPTER I

### INTRODUCTION

Most universities and colleges annually face the problem of selecting from applicants those who they feel will ultimately succeed in graduate school. Several criteria are used in choosing the prospective students. The student's academic record, cumulative grade-point average over the last sixty hours in the area in which the student majored, assessments by the faculty, and examinations such as the Miller Analogies Test and Graduate Record Examination are used in determining eligibility.

Administrators and researchers in the area of evaluation and selection have pointed out the desirability of accurate assessment of a student's potential success in graduate school. A high scholastic index does not guarantee that a student will be able to think independently or that he will be able to adjust to an environment other than the university or college which he attended, as many employers have discovered in hiring only those students with the highest grades. However, investigations have found that a high positive correlation exists between academic achievement in school and success following graduation (2, 4, 5, 6, 7, 17, 19).

Both the Graduate Record Examination and Miller Analogies Test are utilized in measuring scholastic aptitudes at the graduate school level.

Fryer concluded that college scholarship is without doubt, the best single predictor of professional school scholarship (9, p. 438). The Miller Analogies Test, which is being used widely, is also considered an excellent predictor of graduate school success. If Fryer's statement and the predictive powers of the Miller Analogies Test hold true, then it should not be too unreasonable to expect that fairly high correlations exist between Miller scores of graduate school students and their respective grade-point averages on both the undergraduate courses and courses in the graduate school.

Psychological texts state that in overall intelligence males and females are equal (16, 17). However, it is indicated that females usually surpass males in tasks involving verbal abilities and social relationships, in perceiving details quickly and accurately, and in making rapid accurate manual movements. On the other hand, males surpass females generally in tests which involve spatial, numerical and mechanical ability. (14)

The Miller Analogies Test is a test based upon many diversified areas of subject matter, including social sciences, literature, mythology, mathematics, vocabulary, and general information.

Individual differences, not primarily gender, are the primary reasons for variations in the level of intellectual functioning. Therefore, it should seem reasonable to assume that there would be no significant difference between the performance levels of male and female subjects on the Miller Analogies Test.

Since students, aged nineteen through sixty-five years, are pursuing training on the graduate level, and since most graduate schools use the Miller Analogies Test or Graduate Record Examination as a criterion in selecting students, the question arises as to the effect of increasing age upon test scores. To answer this question adequately, the experimental design would involve securing longitudinal data on a large number of individuals.

Although Wechsler's findings are not universally accepted, there seems to be evidence to justify the claim that after the age of twenty-five our intellectual capacities are definitely on the decline, and just as in the case of physical strength, continue to fall off progressively with age (25). Therefore, it shall be assumed that subjects thirty-five years of age or younger may score significantly higher on the Miller Analogies Test than subjects who are thirty-six years of age or older.

Students majoring in the field of psychology are introduced to testing techniques early in their studies. Specific courses in psychology are centered around the methodology and techniques of administering individual, group, intelligence, personality and aptitude tests. Most psychology students are "test wise", that is, quite familiar with the mechanics and design of psychological tests.

In establishing the norms for the Miller Analogies Test, Miller (15) found that psychology majors were the highest scoring group. Doppelt (7) found the mean score of a group of psychology students taking the

Miller Analogies Test to be superior to the mean scores of the other two groups involved in the study (Science students and Non-science students).

In the present study, subjects enrolled in the field of Guidance and Personnel Administration had a more expansive background in psychology than those doctoral candidates enrolled in the field of Elementary Education. This review of the student's records, therefore, produced information which led to the rationale that doctoral candidates enrolled in the field of Guidance and Personnel Administration would attain a mean score on the Miller Analogies Test which was superior to the mean score of Elementary Education doctoral candidates.

#### Statement of the Problem

The main purpose of the present study was to investigate the relationship between the Miller Analogies Test scores of graduate school students and their undergraduate grade-point averages, as well as to find the relationship between Miller Analogies Test scores and the grade-point averages in graduate school of these students.

A secondary purpose of this study was to investigate the relationship between each of the following variables: sex, age, curriculum and the performance levels of the subjects.

#### Related Literature

A study by Weber (23) was based upon the records of three hundred nineteen graduate students who completed their undergraduate work at



Northwestern University. The correlation between undergraduate grade-point averages and graduate grade-point averages for the group was .62. This prediction of graduate academic success from undergraduate marks is better than generally obtained in predicting undergraduate success from high-school grade-point averages. The prognosis might be different for graduate students who completed their undergraduate work at other schools.

Undergraduate grade-point averages in the major fields showed no closer relationship for prediction purposes than undergraduate marks as a whole. The coefficient of correlation between undergraduate marks achieved in the major field and graduate grade-point averages was .62.

The intelligence-test scores and graduate grade-point averages correlated .31, which is not high enough to be of much value in predicting graduate school success. It should be considered, however, that this correlation was probably reduced because of the fact that the intelligence tests were taken when the students were freshmen and were restricted to students who had received their undergraduate and graduate grades in the same institution.

In his report of the investigation of high-level tests done by the University of Pittsburgh's "Division of Psychological Services", Jenson (11) stated the necessity of developing improved methods of selecting graduate students. The council concentrated mainly upon comparing and combining the scholastic predictive value of these tests with the

customarily used undergraduate grade-point average for different groups of first-year graduate students.

Prior to this investigation, a number of researchers had investigated the predictive value of certain variables. Weber (23), Crawford (6), and Petersen (22) agreed that the overall undergraduate grade-point average was the best overall predictor of graduate school grades, even though all recognized the hazards of using this method of prediction. Petersen and Weber found that comprehensive tests of educational attainment such as the Graduate Record Profile Tests, when used with total undergraduate grade-point average, tended to improve prediction of graduate student grade-earning ability. Weber also found no intelligence test which could equal the predictive power found with the Graduate Record Examinations. A relatively accurate predictive method was found to be a subjective appraisal of graduate student potential based on quality point average and Graduate Record Examination scores. However, this means was proven less dependable than statistically acquired predictions. The criterion normally used for measuring success in graduate school was graduate grade-point average. There was no decisive evidence in favor of alternative criteria used in some of the studies. As yet, the best, or the only, success criterion has not been discovered; so at this time, regarding this study, scholastic achievement represents an intermediate, if not immediate, criterion for predicting graduate school success, rather than an ultimate measure of it.

This investigation was essentially an applied study. It was designed to evaluate the efficiency with which a battery of standardized tests of high-level aptitude type, and overall undergraduate grade-point average would forecast the success of groups of first-year graduate students in education, English, chemistry, and psychology.

The sample departmental groups consisted of one hundred seventy-seven students from education, sixty-two from English, forty-nine from chemistry, and forty-seven from psychology. All of the graduate work had been taken in one field of study at the University of Pittsburgh.

The predictive variables consisted of (a) overall Undergraduate Quality Point Average (UQPA), (b) Miller Analogies Test, (c) the Iowa Mathematical Aptitude Test; and (d) the Cooperative Reading Comprehension Test, Higher Form, C2. The reading test yields vocabulary, speed of comprehension, level of comprehension, and total reading scores. These tests were selected because of the wide range of aptitudes they are intended to measure.

Graduate Quality Point Average computed on twelve to thirty hours of graduate credit was employed as the measure of scholastic success in graduate study.

Summarizing the findings, for all groups the most dependable predictors of Graduate Quality Point Average in each group are shown in Table I, page 8. It is apparent that the predictive weight of the tests

TABLE I  
SINGLE AND MULTIPLE PREDICTORS OF  
SUCCESS IN GRADUATE SCHOOL

| Group      | Single Predictions | Multiple Predictions |
|------------|--------------------|----------------------|
| Education  | L or T             | AL<br>UAL<br>UAML    |
| English    | U or S             | US<br>UVS<br>UAVS    |
| Chemistry  | M or U             | UM<br>UAM<br>UAMS    |
| Psychology | M or T             | MT<br>UMT<br>UMVS    |

\*G - Graduate Quality Point Average  
 U - Undergraduate Quality Point Average  
 A - Miller Analogies Test  
 M - Iowa Mathematical Aptitude Test  
 V - Cooperative Reading Vocabulary  
 S - Cooperative Reading Speed  
 L - Cooperative Reading Level  
 T - Cooperative Reading Total

and UQPA varies from one group to another. This seemed especially true for single predictions. Mathematical Aptitude was the only test which had the best single predictive power for more than one group. UQPA appeared twice, first for English, and second for chemistry. Speed, Level and Total Reading tests appeared once each. The best

single tests for the education group seemed to be reading level or total; for English, UQPA or reading speed; for chemistry, math aptitude or UQPA, and for psychology, Mathematical aptitude or Reading Total. Vocabulary had very little predictive significance for any of the groups involved.

In this study, as in most others of this design, it was discovered that maximum efficiency in predicting scholarship was not greatly enhanced by using more than two or three of these predictive variables.

In general, the GQPA of first-year graduate students in education, English, and psychology can be predicted accurately two times out of three within an error of about three-tenths of a quality point in either direction of the best predicted Graduate Quality Point Average.

One of the early investigations into the predictability of success of the Miller Analogies Test was conducted by Cook (4). In this study at the University of Minnesota, a battery of tests, given to seven hundred eighty-eight students during 1940-1941, contained the following tests: Miller Analogies Test, the Educational Information and Application; the Co-operative Survey Test in Mathematics, the Cooper English Examination. A profile of the student's scores, based on grade school norms, was furnished his advisor, to enable the advisor to interpret as best he could from a knowledge of the nature of the tests. It was hoped that the battery might ultimately prove valuable for the following purposes:

- (1) partial basis for admission to graduate school;
- (2) basis for preliminary advice to the student regarding which courses he might pursue with prof it;
- (3) in indicating whether the plan to be followed in qualifying for the master's degree should emphasize research or course work;
- (4) as a partial basis for awarding assistanceship;
- (5) as a partial basis for encouraging capable students as early as possible, in order to give them a special attention they warrant;
- (6) in making it possible to know more intimately the great number of graduate students who are on the campus only during the summer sessions.

A preliminary study was undertaken to determine the value of the tests for the purposes indicated. Three criteria for success in graduate work were considered.

The following factors were decided upon as being of most importance in the respective values of the tests:

- (1) honor point ratio based on better grades in all graduate courses;
- (2) numerical scores of students in various graduate courses, the scores being those upon which the final better grades in these courses were based;
- (3) numerical scores on graduate comprehensive examinations administered to both M. A. and Ph. D. candidates at the end of their course in the fields of educational psychology, educational administration, secondary education, and elementary education.

The chief value of the tests is that they enable staff members to identify superior students immediately upon entrance into the graduate school, thus enabling them to give appropriate counsel.

The results indicated the relationship between success in graduate courses in the College of Education and scores on the four aptitude tests, which are 1) Miller Analogies Test, .48; 2) Education (Cook), .60; 3) Co-operative Math (Form P), .33; and 4) Co-operative English (Form OM), .43.

The test scores are sufficiently indicative of success in graduate courses to warrant their continual use, not only as a measure of general aptitude, but also of extreme strength or weakness in fields of English or mathematics.

In the main, it can be said that most of the outstanding graduate students who are worthy of being encouraged to do advanced graduate work were found in the sixteenth percentile on three of four tests (mathematics). At the other extreme, among students ranking below the sixteenth percentile on three of four tests will be found most of the failures, those students incapable of profiting from graduate work.

In a study by Platz, McClintock, and Katz (20), subjects were one hundred twenty-four students who entered the graduate program at the University of Michigan during the years 1951-1955. Information at the time of entrance included the student's total grade point average, the undergraduate grade point average in science and mathematics, the undergraduate grade point average in psychology, Miller Analogies Test score, and objective comprehensive examination in psychology. The measures of success included grade point average in graduate courses, marks on doctoral preliminary examination, faculty ratings of professional contribution and potential scientific controls. The results showed the best predictor of marks in graduate courses was undergraduate grade point average in science courses; it was also shown that the undergraduate grade point average in science courses best predicted preliminary examination grades. The Miller Analogies Test significantly predicted graduate course grades and was the best predictor of potential scientific contributions of the student.

Cureton (5) used the Miller Analogies Test and the Columbia Intelligence Scale CAUD. Two criterion measures were used to check the validity of these two tests. First was grade point average in all courses taken in graduate school at the University of Tennessee. The second criterion was an overall rating completed by each member of the department who was teaching graduate courses in psychology.



It was concluded from the data that for the population of which the group may be considered a sample, the fifty-minute Miller Analogies Test is at least as valid as the Columbia Intelligence Scale CAUD, which requires about four hours to complete. The chances are about nineteen to one that its validity for predicting average rating is above .28 and that these validities may be as high as .75 and .82 respectively.

Glaser (10) examined the merits of a battery of tests in data obtained at the University of Indiana School of Medicine, in order to evaluate the predictive efficiency of this battery of tests as an aid in selection of medical students. The criterion employed in this study was achievement in the first year of medical school as indicated by the general grade average at the end of the first year. The Miller Analogies Test, Differential Aptitude Test, Army General Classification Test, the Minnesota Multiphasic Personality Inventory, and the United States Armed Forces Institute Subject Examinations were administered to one hundred fifty students at the beginning of their first year of medical school. Results indicated that there was a high degree of intercorrelations between tests in the battery and the single-test validities. The particular tests which correlated highest with the criterion were the United States Armed Forces Institute Subject Examinations for general educational development, Interpretation of Reading Materials in Natural Science with a correlation coefficient of .48 and Miller Analogies Test with an  $r$  of .33. The multiple correlation of these two tests with the

criterion was .50. The Minnesota Multiphasic Personality Inventory showed no relationship to achievement in the first year of medical school.

In a study by Patterson and Waters (22), pooled ratings of fifty-one Minnesota Ph.D. students in psychology were obtained from graduate faculty members on the following three characteristics: (a) intellectual competence, (b) research capacity, and (c) administrative clinical or service competence. The students also took the Miller Analogies Test. The results showed that there is a moderate but statistically significant positive correlation between Miller Analogies Test raw scores (plus .50) and pooled ratings of intellectual competence of the fifty-one Ph.D. students in psychology at the University of Minnesota.

In general, the findings of these studies indicated that the Miller Analogies Test has moderate success in predicting academic potential in graduate school.

In reviewing the literature regarding the Miller Analogies Test, no studies or information were found on the differences in performance levels of male and female subjects.

Doppelt (7) investigated the effect of differences in curriculum upon performance levels of the Miller Analogies Test. Five thousand three hundred eleven college seniors and graduate students who had taken Form G of the Miller Analogies Test were utilized as subjects.

Three major field categories were used to classify these students. Students studying the physical and biological sciences, mathematics, engineering and related major fields were included in the Science group; students majoring in social studies, languages, arts, etc., were put in the non-science category. The relatively homogeneous category of psychology majors was the last group studied.

The ninety-four items on the Miller Analogies Test were separately classified as either "science" or "non-science" items.

Doppelt found that the psychology majors surpassed the performance levels of both the science and non-science groups on "science" items and "non-science" items. Their performance, therefore, proved superior, in all aspects, to the other two groups on all items of the Miller Analogies Test.

In a study by Osborne and Sanders (18), an attempt was made to ascertain the effect of age increment upon the various profile test scores of the Graduate Record Examination.

The subjects were one thousand eight hundred seven graduate students attending the University of Georgia between February, 1946, and February, 1952. Data were obtained from official records of the University of Georgia.

Group means were computed for four-year intervals for all seven Graduate Record Examination profiles. Significance of the mean differences from group to group were computed.

The rate of decline with age varied among the different subject matter areas. Mean science scores, especially biology, fell off sharply with age increment, while social science, literature, and fine arts scores held up well throughout the entire nineteen years through sixty-five years range.

### Statement of Hypotheses

The following hypotheses were investigated:

(1) There will be a significant positive correlation between the Miller Analogies Test scores of students and their grade-point averages in graduate school.

(2) There will be a significant positive correlation between undergraduate grade-point averages and averages in graduate school courses.

(3) There will be no significant difference in performance on the Miller Analogies Test between a group composed exclusively of female subjects and a separate group composed exclusively of male subjects.

(4) A group of doctoral candidates enrolled in the field of Counseling and Personnel Administration will attain significantly greater scores on the Miller Analogies Test than a group of doctoral candidates enrolled in an Elementary Education curriculum.

(5) Those doctoral candidates ranging in age from thirty-five years of age and under will score significantly higher on the Miller Analogies Test than the group of doctoral candidates who are thirty-six years of age and older.

## CHAPTER BIBLIOGRAPHY

1. Berdie, R. F., and W. L. Layton, "Predicting Success in Law School", Journal of Applied Psychology, 36 (August, 1951) 60-63.
2. Chancey, H., and H. M. Fowler, "Estimating Graduate School Success: A Comparison of Predictions Made from the Undergraduate Record and from the Graduate Record Examination Profile at Harvard University", New York: Carnegie Foundation for the Advancement of Teaching, 1940.
3. Cockrum, L. V., "Predicting Success in Training for the Ministry", Religious Education; 47 (May-June, 1952) 198-202.
4. Cooke, W. W., "Predicting Success of Graduate Students in a College of Education", School and Society; 56 (September, 1942) 192-195.
5. Cureton, E. E., L. W. Cureton and R. Bishop, "Prediction of Success in Graduate Study at the University of Tennessee", American Psychologist; 4 (August, 1949) 361.
6. Crawford, A. B., "The Prediction of Success in Graduate School", New York: Graduate Record Examination, 1940.
7. Doppelt, J. E., "Difficulty and Validity of Analogies Items in Relation to Major Field of Study", Journal of Applied Psychology, 35 (August, 1951), 39-40.
8. Feldman, M. J., "A Comparison of Miller Analogies Test Scores", Journal of Counseling Psychology, 5 (Summer, 1958), 149-150.
9. Fryer, D. H., "Professional Psychology", Fields of Psychology; New York: Van Nostrand; 1949.
10. Glaser, R. "Predicting Achievement in Medical School", Journal of Applied Psychology, 35 (August, 1951) 272-274.
11. Jenson, R. E., "Predicting Scholastic Achievement of First Year Graduate Students", Educational and Psychology Measurements, XII (September, 1953) 322-329.

12. Kelly, E. L. and D. W. Fiske, The Prediction of Performance in Clinical Psychology, Ann Arbor: University of Michigan Press, 1951.
13. McConnell, P. "Changes in Scores on the Psychological Examinations of American Council of Education from Freshman to Senior Year", Journal of Educational Psychology, XXV (January, 1934), 66-69.
14. Miller, W.S., Miller Analogies Test Bulletin No. 2, The Psychological Corporation, New York, New York, December, 1948.
15. Miller, W.S., Miller Analogies Test Manual, 1960 Revision, New York: The Psychological Corporation, 1960.
16. Morgan, C. T., Introduction to Psychology, New York: McGraw-Hill Company, 1956.
17. Munn, N., Introduction to Psychology, Boston, Mass.: Houghton-Mifflin, Company, 1962.
18. Osborne, R. T. and W. Sanders, "Differential Decline in Graduate Record Examination Scores with Age", Journal of Geriatric Psychology, 87 (April, 1955), 309-316.
19. Petersen, S. C., The Measurement and Prediction of Scholastic Achievement on the Graduate Level. New York: Graduate Office (1943). A Ph. D. Dissertation from the State University of Iowa.
20. Platz, A., C. McClintock and D. Katz, "Undergraduate Grades and the Miller Analogies Test on Prediction of Graduate Success", American Psychologist, 14 (June, 1959), 285-289.
21. Tully, G. E., "Screening Applicants for Graduate Studies with the aptitude Test of the Graduate Record Examination," Colleges and Universities, 38 (February, 1963), 51-60.
22. Watters, G. V. and D. G. Paterson, "Miller Analogies Test Scores and Ratings of Ph. D. 's in Psychology", American Psychologist, 8 (February, 1955), 89-90.
23. Weber, J., et al., "Success in Graduate School," Journal of Higher Education, XIII (January, 1942), 19-24.

24. Webster, E.C., Alexander Winn, and J.A. Oliver, "Selection Tests for Engineers", Personnel Psychology, 4 (Winter, 1951), 339-345.
25. Wechsler, D., The Range of Human Capacities, New York: William and Wilkins, 2nd Edition, 1952.



## CHAPTER II

### METHOD

#### Subjects

The subjects for the present study consisted of doctoral candidates who had taken Form L of the Miller Analogies Test between January of 1965 and February of 1967. A total of three hundred twenty-two student's records were examined.

The group of three hundred twenty-two subjects was composed of two hundred fifty-five male and sixty-seven female students. The age range of the subjects was from twenty years through sixty years.

From this original sample of subjects, the test scores of a group of fifty subjects from the School of Education at North Texas State University were used to investigate the first hypothesis. The remainder of the sample (two hundred seventy-two subjects) could not be utilized in the testing of this hypothesis, as they had either enrolled in another school, were not accepted by North Texas State University, had enrolled in North Texas State University and dropped out of school shortly thereafter, or had not fulfilled the two-semester residency requirement set for the examination of the hypothesis.

In testing hypothesis two, the transcripts of the group of fifty subjects were examined to ascertain grade-point averages in both graduate and undergraduate schools. The mean Miller Analogies Test score for this group was 49.82.

For hypothesis three, the three hundred twenty-two subjects were divided into two groups, designating those who were female doctoral candidates (sixty-seven subjects) and the two hundred fifty-five male doctoral candidates who had submitted their Miller Analogies Test scores to North Texas State University. The mean Miller score for males was 43.17; the Mean Miller score of female subjects was 41.50.

To attempt to examine hypothesis four, it was necessary to examine the Miller Analogies Test scores of subjects enrolled in these different fields of study. The subjects (N = 41) who were doctoral candidates enrolled in the Guidance and Personnel Administration at North Texas State University consisted of seven females and thirty-four males. The mean Miller Analogies Test score for this group was 44.09.

A second group of subjects (N = 26), who were enrolled in the Elementary Education department at North Texas State University, consisted of eighteen male and eight female doctoral candidates. The mean Miller Analogies Test score for this group was 41.24.

In investigating the fifth hypothesis, two hundred ninety-two subjects were used from the original sample of three hundred twenty-two. The remaining thirty subjects could not be included in this aspect of this study, as the information on these students' records was incomplete.

One group of the two hundred ninety-two subjects included one hundred seventy-one subjects who were thirty-five years of age and younger. The second group of one hundred twenty-one subjects were thirty-six years of age and older. The mean age for the first group of subjects was 29.02 years. The second group's mean age was 42.11 years. The mean Miller Analogies Test score for the first group was 45.88; for the second group, the mean Miller score was 43.22.

#### Procedure

The transcripts of the students utilized in this study were obtained with the permission of the registrar of North Texas State University. Undergraduate grade-point averages were ascertained by adding the number of quality points earned by each student, and dividing the sum by the number of hours attempted and completed by that student.

All of the students had earned the master's degree. The grades reviewed were on the doctoral level. Another criterion of eligibility was at least a two-semester residency requirement.

Grade-point averages were computed in the same manner for doctoral candidates as for undergraduate students.

The individual Miller Analogies Test scores were compared with the individual undergraduate grade-point averages, and the grade-point averages on the doctoral level, respectively. This analytic procedure yielded indices of the efficiency with which the Miller Analogies Test scores and the grade-point averages are able to predict eventual success in graduate school.

Pearson product-moment coefficients of correlation were computed to determine the relationship between Miller Analogies Test scores and (1) graduate school grade-point averages, and (2) undergraduate grade-point averages of the subjects studied.

The t test was used to determine if the mean differences between the six groups composed of Guidance and Personnel Administration majors and Elementary Education majors, males and females, and subjects aged twenty through thirty-five years and subjects aged thirty-six years and older was significant. The .05 level of confidence were utilized as the initial level of statistical significance.

#### Description of Instrument

Miller Analogies Test. --This test consists of one hundred analogy items based on many diversified areas of subject matter, including chemistry, social sciences, literature, mythology, mathematics,

vocabulary, and general information. The test items require the recognition of relationships, rather than a display of knowledge. The analogies appear in standard form, A: B: C: D. The examinee selects the one possibility out of the four which will complete the analogy.

The Miller Analogies Test (MAT) was developed to measure scholastic aptitude at the graduate school level. It represents one of the earliest efforts in the field of selection at that level, since it has been in use in one form or another since 1926. The first of present forms (Form G) was developed at the University of Minnesota, and was originally restricted to use at that institution. However, the test soon won favorable attention from a number of other institutions. The presently available forms of the MAT are in widespread use for the selection of graduate students in universities and of high-level personnel in business and government (2, p. 3).

This test is designed for use as one factor in the selection of Graduate students in universities. To keep it useful, it is important that it be completely unavailable except to authorized personnel. The Miller Analogies Test is released only under a plan of licensing whereby certain universities undertake to control the use of the test in their own institution and nearby areas. . . (2, p. 3).

The test is printed in a four-page booklet, with full instructions on the cover page. The test booklet is reusable. The subject records his answers on a special IBM answer sheet which can be either machine or hand scored. The time allowed for testing is fifty minutes (2, p. 3).

The Miller Analogies Test is primarily a power test; the contribution of speed to a student's score is of negligible importance. . . . The correlation between 50- and 90- minute scores was .99. . . . The subject's score is the number of his right answers. The maximum possible score is 100 (2, p. 3).

The validity of the Miller is determined primarily by its ability to predict success in graduate school study. Sources of evidence are represented. . .by direct measures of relationship between test scores and graduate school success as defined by honor point ratios (or variants thereof), satisfactory completion of course work, or subject-field examinations. . .(2, p. 10).

A study of the validity of the Miller was presented by W.W. Cooke (1), showing correlations with numerical grades in graduate courses in a school of education, with a range of coefficients from .14 ( $n = 100$ ) to .78 ( $n = 12$ ).

Another technique used to assess the validity of the Miller is to note the ability of the test to discriminate between groups who are believed to differ in the ability measured by the test. (2) Table II presents mean scores of the Miller Analogies Test in relation to honor point ratios.

Reliability data for the Miller Analogies Test are presented in Table III. The two kinds of evidence shown in Table III are as follows: (1) Odd-even correlation coefficients for each of the four forms, and (2) correlations between Forms G and H, Forms H and J, and Forms J and K.

Both sources of evidence demonstrate that the Miller Analogies Test is reliable enough so that the test user may have considerable confidence in the stability of the scores earned by individual students.

TABLE II  
 AVERAGE MILLER ANALOGIES TEST SCORES OF GRADUATE  
 STUDENTS AT SYRACUSE UNIVERSITY FOR  
 DIFFERENT HONOR POINT RATIOS

| Number | Honor Point Ratio | Mean Score |
|--------|-------------------|------------|
| 14     | 2.9 - 3.0         | 72.6       |
| 30     | 2.7 - 2.8         | 64.4       |
| 27     | 2.5 - 2.6         | 59.4       |
| 47     | 2.3 - 2.4         | 53.4       |
| 29     | 2.1 - 2.2         | 47.6       |
| 18     | 1.9 - 2.0         | 44.7       |
| 26     | Below 1.9         | 42.7       |

\*From a study by Milton Hahn and Stanley Ostrom of Syracuse University.

TABLE III  
 RELIABILITY COEFFICIENTS FOR THE FOUR  
 FORMS OF THE MILLER ANALOGIES TEST

| Method of Determining Reliability | Form    | r   | Mean           | SD           | Population   |
|-----------------------------------|---------|-----|----------------|--------------|--|
| Odd-even                          | G       | .93 | 63.7           | 16.0         | 387 graduate students                                  |
| Odd-even                          | H       | .94 | 47.9           | 16.9         | 334 seniors and graduate students at nine institutions |
| Odd-even                          | H       | .93 | 53.6           | 14.0         | 426 employed engineers                                 |
| Odd-even                          | J       | .94 | 48.3           | 17.4         | 220 seniors and graduate students at nine institutions |
| Odd-even                          | J       | .92 | 52.5           | 16.6         | 324 seniors and graduate students at six institutions  |
| Odd-even                          | K       | .93 | 51.7           | 16.4         | 312 seniors and graduate students at six institutions  |
| Alternate form                    | G vs. H | .89 | G55.7<br>H54.2 | 15.9<br>16.0 | 681 seniors and graduate students                      |
| Alternate form                    | H vs. J | .88 | H49.6<br>J49.8 | 17.4<br>17.9 | 554 seniors and graduate students                      |
| Alternate form                    | H vs. J | .85 | H57.3<br>J58.5 | 16.0<br>15.9 | 135 graduate students                                  |
| Alternate form                    | J vs. K | .87 | J54.5<br>K54.5 | 16.2<br>16.8 | 636 seniors and graduate students                      |

\*NOTE: All odd-even coefficients have been corrected by Spearman-Brown formula.



## CHAPTER BIBLIOGRAPHY

1. Cook, W.W., "Predicting Success of Graduate Students in a College of Education", School and Security, 56 (September, 1942), 192-195.
2. Miller, W.S., Miller Analogies Test Manual 1960 Revision, New York: The Psychological Corporation.

## CHAPTER III

### RESULTS AND DISCUSSION

The coefficient of correlation between Miller Analogies Test scores and the undergraduate grade-point averages was .67, as shown in Table IV. This was significant with 48 degrees of freedom at the 1 percent level.

TABLE IV

COEFFICIENTS OF CORRELATION FOR MILLER ANALOGIES TEST  
SCORES AND GRADE-POINT AVERAGES

|                          | Undergraduate | Graduate |
|--------------------------|---------------|----------|
| Group I<br>Miller Scores | .67           | .74      |

The  $r$  of .74 between the Miller Analogies Test scores and the grade point averages of the students in graduate school was significant at the 1 percent level of confidence.

Therefore, the assumptions that college scholarship is an indication of possible success in graduate school and the assumption of the selectivity of the Miller Analogies Test made at the outset of this study were confirmed. It was assumed that a high correlation existed between Miller Analogies Test scores of certain students and their undergraduate

grade-point averages. This was confirmed in finding an  $r$  of .67, as stated above. It was further stated that a similarly high correlation was believed to exist between the Miller Analogies Test scores and the grade-point averages of the students in graduate school. This is confirmed in finding an  $r$  of .74.

TABLE V

AVERAGE MILLER ANALOGIES TEST SCORES FOR STUDENTS  
FOR GRADE-POINT AVERAGES IN GRADUATE SCHOOL

| Grade-Point Average | Number | Mean Miller Score |
|---------------------|--------|-------------------|
| 2.9 - 3.0           | 10     | 55.6              |
| 2.7 - 2.8           | 11     | 52.7              |
| 2.5 - 2.6           | 11     | 42.0              |
| 2.3 - 2.4           | 3      | 49.0              |
| 2.1 - 2.2           | 2      | 42.5              |
| 1.9 - 2.0           | 8      | 46.5              |
| below 1.9           | 5      | 47.4              |

The discrepancies in any individual cases between Miller scores and Graduate School grade-point averages were brought more sharply into focus (Table V) by finding a mean Miller score for students whose average grades were closely related. For example, ten students were found to have a grade-point average between 2.9 and 3.0, and their Miller

scores averaged 55.6. It is of interest to note from Table V that the largest number of students is found in the range of averages 2.7 to 2.8, which approximates the mean Miller score for the total number on a grouped frequency distribution (49.82).

Hypothesis III of the study was tested, using a  $t$  test technique. The results of this analysis and presented in Table VI.

TABLE VI

MEAN, STANDARD DEVIATION AND  $t$   
FOR MALE AND FEMALE SUBJECTS  
SUMMARY OF RESULTS

|         | Mean Score | S D   | SEM   | $t$  |
|---------|------------|-------|-------|------|
| Females | 41.50      | 10.12 | 1.244 | .937 |
| Males   | 43.17      | 11.31 | .707  |      |

The obtained  $t$  value of .937 for sex scores did not reach levels of statistical significance. Thus, Hypothesis III was confirmed.

In order to test Hypothesis IV the data were analyzed by the  $t$  test. The results of this analysis are shown in Table VII. As shown in Table VII, the variable investigated, difference in performance levels between Guidance and Counseling students and Elementary Education students, did not reach a critical statistical level of significance. Thus, Hypothesis IV was not confirmed. Miller Analogies

TABLE VII  
 MEAN, STANDARD DEVIATION AND  $t$  FOR GUIDANCE  
 AND ELEMENTARY EDUCATION STUDENTS

|                         | Mean Score | SD    | SEM  | $t$  |
|-------------------------|------------|-------|------|------|
| Guidance                | 44.09      | 9.27  | 1.43 | 1.48 |
| Elementary<br>Education | 41.24      | 10.86 | 2.17 |      |

Test scores did not differ significantly between Guidance and Counseling students and Elementary Education students.

Hypothesis V of this study was tested, using the  $t$  test. Results of this analysis are presented in Table VIII.

TABLE VIII  
 MEAN, STANDARD DEVIATION AND  $t$  FOR SUBJECTS  
 IN AGE GROUPS OF 35 AND UNDER, 36 AND OLDER  
 SUMMARY OF RESULTS

| Ages         | Mean Scores | SD    | SEM    | $t$  |
|--------------|-------------|-------|--------|------|
| 35 and under | 45.88       | 10.33 | .792   | 20.4 |
| 36 and older | 43.22       | 12.30 | 1.1254 |      |

The obtained  $t$  value of 20.4 for age scores was significant, at better than the .001 level of probability as shown in Table VIII.

Thus, Hypothesis V was confirmed. Those subjects who were between the ages of twenty and thirty-five years did attain significantly greater scores on the Miller Analogies Test than a group of subjects ranging in age from thirty-six through sixty years.

### Discussion

Following a review of the results listed in Table IV, it would appear that the Miller Analogies Test would be a good indicator of academic success in graduate school. Both coefficients of correlation yielded high positive relations. There are two possible explanations of why the correlation between the graduate course grades and the Miller Analogies Test score is higher than the correlation between undergraduate course grades and Miller Analogies Test scores.

First, grade-point averages based upon one hundred twenty semester hours or more will probably yield a lower grade-point average. In general, these same students took between nine semester hours and thirty semester hours during a two semester period. Taking fewer hours and a schedule that essentially consisted of course work in one major field would enable these students to achieve higher grade point averages.

Secondly, the grading system used in graduate school differs from the grading system utilized in undergraduate course work. Generally speaking, the grade range in graduate school is from "A"

to "C". The fifty cases reviewed in this investigation produced not one "D" or "F". In undergraduate work, the grading range is from "A" to "F". Students must also take an assortment of courses in order to meet degree requirements specified by the school. For various reasons a student may have difficulty in a certain subject and perform poorly; his grade in this subject mirrors his dislikes of subject matter, biased opinion of the teacher, or other extenuating circumstances. In graduate school, a student specializes in one area of study. On the doctoral level, a student has possibly had fifty hours or more in one specialized field. The homogeneity of the subject matter makes it easier to assimilate. In a grading system which puts emphasis on the grade, or subject-centered approach, as opposed to student-centered approach, a graduate student is required to make an "either-or" grade ("A" or "B") in order to "Pass" a course on the Graduate School level. Fine discriminations in a larger continuum must, of necessity, be contained in an "either-or" grade. It is possible using the grade-centered approach, to hold classes for an entire semester without administering a single written examination. This is unheard of in undergraduate school; testing is the principal criterion used in assessing a student's merit.

The instructor is then confronted with the task of assigning grades of the "either-or" variety. Hence, based upon this assumption one

can expect grades based on purely subjective judgements will correlate well with an objective, standardized tool, such as the Miller Analogies Test.

Another possible approach to this problem would be to find the relationships between Miller scores and grades acquired in the early stages of graduate school, possibly on the Master's level. This would give all students a certain consistency in having their graduate grades based on an equal amount of hours studied. This uniformity of hours could be achieved in selecting only those students who had received a certain status in the graduate school.

In general, the knowledge of a Miller score alone is not believed to be sufficient evidence on which to accept or to reject an application to the graduate school. In the opinions of most researchers and administrators, whatever else the Miller Analogies Test measures, the score does give an indication to the "pattern of thinking", which is fairly well set for an individual student by the time he enters graduate work. In the studies which have been previously reported, no information is listed showing scores lower than 20; in an article prepared by Kelly and Fiske (11), some successful doctorates scored as low as 30, although the evidence for doctoral candidates is more highly in favor of those who make higher scores. Perhaps the study of the Miller Analogies Test should begin with the thesis that a student who makes a score of 30, and subsequently receives a doctorate has found and developed to an



unusual degree a "pattern of thought" which is unique for him.

The third hypothesis, that there would be no significant difference in performance on the Miller Analogies Test between groups composed of male post-graduate students and female post-graduate students was confirmed. There was no significant difference in regard to this variable. It might be added that no previous studies could be found that dealt with performances on the Miller Analogies Test with regard to sex.

The results obtained through a t test do not seem to indicate support for the fourth hypothesis, namely that students enrolled in the Guidance and Counseling department would attain significantly greater scores on the Miller Analogies Test than a group of students enrolled in the Elementary Education department.

There is a study which disagrees with the findings of the present study: Doppelt's (1) study on performance by different major fields who had taken the Miller Analogies Test. He found that psychology majors made higher scores than any other group. The majority of the Guidance students used in this study had a more expansive background in psychology than the Elementary Education group. A possible explanation offered by Doppelt is that psychology majors are generally more "test oriented" than other students, and perhaps this is of some assistance in obtaining higher scores.

There is a possibility, also, that if the two groups had been the same size (Guidance - N = 41, Elementary Education - N = 26) there would have been a significant difference in the mean scores of the two groups. The standard error of the mean decreases as the number of subjects increases.

The fifth hypothesis, that candidates who were aged thirty-five years and under should score significantly higher on the Miller Analogies Test than the group composed of post graduate students who were thirty-six years of age and older, was confirmed. The results of the t test seem to indicate support of this hypothesis.

There is a possibility that, on a power test such as the Miller Analogies Test, younger subjects (20-35 years of age) would fare better than the older subjects (36-60 years of age). There was no other experimentation in this area to be found to support this hypothesis. Generally speaking, it has been stated that the human organism reaches a point of intellectual maturity and then depreciates year after year. Of course, due to limited scientific investigation using the Miller Analogies Test in this area and individual difference variables, this statement cannot be supported with any favorable degree of confidence. The apparent differential decline in Miller Analogies Test scores with age also could be due to inadequate sampling techniques.

The problem of investigating intelligence and aptitudes, and understanding their relationship to age offers a wide field of investigation.

## CHAPTER BIBLIOGRAPHY

1. Doppelt, J.E., "Difficulty and Validity of Analogies Items in Relation to Major Field of Study," Journal of Applied Psychology, 35 (August, 1951), 39-40.

## CHAPTER IV

### SUMMARY

The purpose of this study was to determine the value of the Miller Analogies Test for measuring scholastic aptitude of fifty doctoral candidates at North Texas State University. The method used was to correlate Miller Analogies Test scores with grade-point averages earned in both undergraduate and graduate school courses.

The secondary purpose of the study was to investigate the effects of sex, age, and curriculum differences upon performance levels of the subjects.

Five hypotheses were tested in this study.

Hypothesis I stated there will be a significant relationship between the Miller Analogies Test scores of certain graduate students and their grade-point averages.

Hypothesis II stated that a positive correlation should result from a comparison between the grade-point averages earned in graduate school courses and the grade-point averages based upon undergraduate courses.

Hypothesis III stated that there will be no significant difference in performance levels in the Miller Analogies Test between a group

composed exclusively of female subjects and a separate group composed exclusively of male subjects.

Hypothesis IV stated that the group of doctoral candidates enrolled in a Guidance and Counseling curriculum will attain significantly greater scores on the Miller Analogies Test than a group of doctoral candidates enrolled in an Elementary Education curriculum.

Hypothesis V stated that those doctoral candidates who were thirty-five years of age or younger would score significantly higher on the Miller Analogies Test than a group of doctoral candidates who were over thirty-six years of age.

The subjects for the present study consisted of doctoral candidates who had taken Form L of the Miller Analogies Test between January of 1965 and February of 1967. A total of three hundred twenty-two student's records were used in this study.

The three hundred twenty-two subjects were composed of two hundred fifty-five male students and sixty-seven female students. The range of the subjects involved in this study is twenty years through sixty years of age.

One group of fifty subjects from an original sample of three hundred twenty-two subjects was selected from the School of Education at North Texas State University. Their transcripts were examined to ascertain grade-point averages in both undergraduate and graduate schools. The mean Miller Analogies Test score for this group was 49.82.

Two groups of subjects, female doctoral candidates (sixty-seven subjects in the group), and male doctoral candidates (two hundred fifty-five subjects in the group) submitted their Miller Analogies Test scores to North Texas State University. The mean Miller Analogies Test score for the group of female subjects was 41.50; the mean Miller Analogies Test score for the group of male subjects was 43.17.

Forty-one doctoral candidates enrolled in the Guidance and Counseling Department achieved a mean Miller Analogies Test score of 44.09.

A second group of twenty-six subjects who were enrolled in the Elementary Education department at North Texas State University attained a mean Miller Analogies Test score of 41.24.

One group of subjects (one hundred seventy-one) who were between the ages of twenty and thirty-five years of age submitted their Miller Analogies Test scores to North Texas State University. The second group of subjects considered here ranged between the ages of thirty-six and sixty years. There were one hundred twenty-one subjects in the second group.

The mean Miller Analogies Test score for the first group was 45.88, for the second group, 43.22.

Transcripts of fifty students were examined to ascertain undergraduate grade-point averages and graduate grade-point averages. The individual Miller Analogies Test scores were compared with the individual

undergraduate grade-point averages and grade-point averages on the doctoral level. Coefficient of correlation was utilized to determine the relationship between Miller Analogies Test scores and undergraduate grade-point averages and graduate school grade-point averages.

The t test was used to determine if the mean differences between the six groups, into which the three hundred twenty-two subjects in this study were divided, was significant.

All hypotheses were confirmed in this study, with the exception of one. Hypothesis I was confirmed; a coefficient of correlation (Pearson r) of .74 was attained. Hypothesis II was confirmed and a coefficient of correlation (Pearson r) of .67 was attained. Hypothesis III did not reach a level of significance, as was predicted. Hypothesis IV obtained results did not seem to indicate supporting evidence that students enrolled in the Guidance and Counseling department score significantly higher on the Miller Analogies Test than students enrolled in the Elementary Education department. Hypothesis V was significant beyond the  $p=.001$  level of probability.



## BIBLIOGRAPHY

### Books

- Kelly, E. L. and D. W. Fiske, The Prediction of Performance in Clinical Psychology, Ann Arbor: University of Michigan Press, 1951.
- Miller, W. S., Miller Analogies Test Manual, 1960 Revision, New York: The Psychological Corporation, 1960.
- \_\_\_\_\_. Miller Analogies Test Bulletin No. 2; The Psychological Corporation, New York, New York, December, 1948.
- Morgan C. T., Introduction to Psychology, New York: McGraw-Hill Company, 1956.
- Munn, N., Introduction to Psychology, Boston, Mass.: Houghton-Mifflin, Company, 1962.
- Wechsler, D., The Range of Human Capacities, New York: William and Wilkins, 2nd Edition, 1952.

### Articles

- Berdie, R. F., and W. L. Layton, "Predicting Success in Law School", Journal of Applied Psychology, 36 (August, 1951), 60-63.
- Chancey, H., and H. M. Fowler, "Estimating Graduate School Success: A Comparison of Predictions Made from the Undergraduate Record and from the Graduate Record Examination Profile at Harvard University", New York: Carnegie Foundation for the Advancement of Teaching, 1940.
- Cockrum, L. V., "Predicting Success in Training for the Ministry", Religious Education, 47 (May-June, 1952), 198-202.

- Cooke, W.W., "Predicting Success of Graduate Students in a College of Education", School and Society, 56 (September, 1942), 192-195.
- Cureton, E.E., L.W. Cureton and R. Bishop, "Prediction of Success in Graduate Study at the University of Tennessee", American Psychologist, 4 (August, 1949), 361.
- Crawford, A.B., "The Prediction of Success in Graduate School", New York: Graduate Record Examination, 1940.
- Doppelt, J.E., "Difficulty and Validity of Analogies Items in Relation to Major Field of Study", Journal of Applied Psychology, 35 (August, 1951), 39-40.
- Feldman, M.J., "A Comparison of Miller Analogies Test Scores", Journal of Counseling Psychology, 5 (Summer, 1958), 149-150.
- Fryer, D.H., "Professional Psychology", Fields of Psychology, New York: Van Nostrand; 1940.
- Glaser, R., "Predicting Achievement in Medical School", Journal of Applied Psychology, 35 (August, 1951), 272-274.
- Jenson, R.E., "Predicting Scholastic Achievement of First Year Graduate Students", Educational and Psychology Measurements, XIII (September, 1953), 322-329.
- McConnell, P., "Changes in Scores on the Psychological Examinations of American Council of Education from Freshman to Senior Year", Journal of Educational Psychology, XXV (January, 1934), 66-69.
- Osborne, R.T. and W. Sanders, "Differential Decline in Graduate Record Examination Scores with Age", Journal of Geriatric Psychology, 87 (April, 1955), 309-316.
- Platz, A., C. McClintock and D. Katz, "Undergraduate Grades and the Miller Analogies Test on Prediction of Graduate Success", American Psychologist, 14 (June, 1959), 285-289.
- Tully, G.E., "Screening Applicants for Graduate Studies with the Aptitude Test of the Graduate Record Examination", Colleges and Universities, 38 (February, 1963), 51-60.

- Watters, G. V. and D.G. Paterson; "Miller Analogics Test Scores and Ratings of Ph.D.'s in Psychology", American Psychologist, 8 (February, 1955), 89-90.
- Weber, J., et al. "Success in Graduate School", Journal of Higher Education, XIII (January, 1942), 19-24.
- Webster, E. C., Alexander Winn, and J. A. Oliver, "Selection Tests for Engineers", Personnel Psychology, 4 (Winter, 1951), 339-345.

Ph.D. Dissertation

- Petersen, S. C., "The Measurement and Prediction of Scholastic Achievement on the Graduate Level," unpublished Ph.D. dissertation, Department of Psychology, State University of Iowa, 1943.