CORRELATES OF THE MILLER ANALOGIES TEST WITH THE WONDERLIC PERSONNEL TEST AND THE SHIPLEY-INSTITUTE OF LIVING SCALE

BY

STEPHANIE FISHER ROBERTS

CORRELATES OF THE MILLER ANALOGIES TEST WITH THE WONDERLIC PERSONNEL TEST AND THE SHIPLEY-INSTITUTE OF LIVING SCALE

A Research Paper

Presented to

the Graduate Council of

Austin Peay State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

in Education

---

by

Stephanie Fisher Roberts

Summer 1973

To the Graduate Council:

I am submitting herewith a Research Paper written by Stephanie Fisher Roberts entitled "Correlates of the Miller Analogies Test with the Wonderlic Personnel Test and the Shipley-Institute of Living Scale." I recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Arts in Education, with a major in Counseling and Guidance.

Major Professor

Accepted for the Council:

Dean of the Graduate School

ACKNOWLEDGEMENTS

The author wishes to express sincere appreciation to Dr. John D. Martin, Associate Professor of Psychology, Austin Peay State University, who suggested the problem and who aided and counseled her during the course of the study; to Dr. Elizabeth Stokes, Department of Psychology, and Dr. Donald Lambert, Department of Education, for their cooperation in the completion of the study. Gratefulness is also extended to Dr. Peter Minetos and Dr. Stephen Davis for their assistance in the computation and analyses of statistical data.

The author wishes to thank her husband for helping in every way during the study; and her parents, Mr. and Mrs. George Fisher, and brother, Mickey, for their patience and understanding.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION TO THE PROBLEM

Development of the Miller Analogies Test (MAT) was initiated some 47 years ago by W. S. Miller as an aid in selecting students for graduate work at the University of Minnesota. It represents one of the earliest efforts to measure scholastic aptitude at this academic level. As a result of interest on the part of other institutions, the presently available forms of the MAT are now in widespread use for the selection of graduate students in universities and of high-level personnel in business and government.

The data from studies of the MAT reported in the literature are often difficult to analyze and interpret because of variations in graduate school programs. Miller points out that validity discrepancies occur from school to school, from course to course, and from year to year. These variations may result from differences in the number of applicants for a school or department, disparity of grading procedures, and the relevance of the ability measured by the test to the qualities required by courses for which it is predicting.

The correlations (.92-.95) of odd and even half-scores for each of the six forms of the test and correlations (.95-.90) between pairs of alternate forms of the test demonstrate that the MAT is sufficiently reliable to generate confidence in the stability of scores earned by individual examinees.

A considerable body of information is available on the correlation of the MAT and other tests. A portion of the coefficients of correlation as included in the 1970 revision of the test manual are as follows: Graduate Record Examination - Verbal (.80), National Teacher Examinations - Commons (.71), Terman Concept Mastery Test (.73), Ohio State University Psychological Test (.77), College Entrance Examination Board Scholastic Aptitude Test - Verbal (.79).

The continued utilization of the Miller Analogies Test in the graduate selection process has initiated a considerable amount of research which has added significantly to the over-all understanding of the instrument. When a test achieves such distinction and widespread use in so notable a decision-making procedure, it is essential that persons using it be cognizant of its potential as well as its limitations.

Some researchers believe that MAT scores can be used effectively in the prediction of academic success in graduate study. On the other hand, many proclaim that the evidence in favor of its predictive value is inconclusive.

Thumin and Boernke (1966) cite the varying and occasionally low (.20-.50) correlation coefficients in some studies relating MAT scores to graduate grade point averages. Various researchers have used similar statistics as proof of its inadequacy as a predictive tool. In contradiction to this point, Payne and Tuttle (1966) designate the cause of validity coefficient variations in MAT scores and graduate grade point averages to be, not a discrepancy in the test, but a lack of uniform grading practices or departmental bias.

Miller's validity studies of his own test at the University of Minnesota, Syracuse University, and the University of Illinois have been reviewed extensively by Durnall (1954). He notes that the correlation coefficients obtained seemed to justify Miller's conclusion that the MAT is a valuable tool for the prediction of success in graduate education. Other studies, as reviewed by Mehrabian (1969), also provide support for the predictive validity of the MAT in the selection of qualified graduate students.

Dailey (1959) describes the MAT as a "well constructed test of general academic scholarship potential with a difficulty pitched at a high graduate student level." He attributes to the Miller Analogies Test a greater validity than is probable for other criteria of academic graduate scholarship because of its loading with highly profound subject matter. Humphreys (1965) portrays the MAT as a

"good" test with plenty of "top" which has high reliability and fairly substantial predictive validity for the criterion of graduate school success.

According to Jenson (1953), most researchers agree that "the" predictive success criterion for graduate study has not yet been identified. He reminds educators that scholastic achievement is not an ultimate measure of graduate success, but only an immediate or intermediate one. The ambiguity in the value and usefulness of the MAT in graduate education has been attributed by Smith (1957) to the grading criteria of different institutions. He describes these criteria as being at times unreliable and inadequate. Most important, however, is the fact that these evaluative standards may vary not only among institutions, but also within the various departments of a single university. Smith's defense of the MAT as a predictive tool rests on his assertion that the instrument continues to relate to other high level tests. Hyman (1957) supports this assertion with his own postulation that the MAT possesses sufficient common variance with the other "verbal" paper and pencil tests so as to maintain a degree of discrimination ability for the purpose of graduate student selection.

In a study of the relationships between creativity and academic achievement, Bentley (1966) found that MAT scores correlated with all mental operation categories,

although the relationship was less for divergent thinking and evaluative categories. He disclosed that no differences were found between MAT scores and creative scores in predicting academic achievement.

In a review by Willingham (1965), it was noted that systematic analyses of what constitutes successful graduate work were largely nonexistent. He concluded that a "hodge-podge" predictor such as the MAT may provide the best available answer to an inadequate question.

The present study was designed to determine the degree of correlation between the MAT and the Wonderlic Personnel Test, and the degree of correlation between the MAT and the Wonderlic Personnel Test, and the degree of correlation between the MAT and the Shipley-Institute of Living Scale; to add to the research previously conducted on the usefulness of the MAT; and to determine if the MAT could be utilized as a reliable measure of intellectual aptitude of prospective graduate students.

CHAPTER II

METHOD

THE SAMPLE

The sample employed in this study was composed of graduate students currently enrolled in graduate psychology and education courses at Austin Peay State University, Clarksville, Tennessee. The sample consisted of a total of 38 subjects of which 26 were females and 12 were males.

DESCRIPTION OF THE INSTRUMENTS

The Miller Analogies Test was developed to measure scholastic aptitude at the graduate school level. It consists of 100 analogy items in verbal form based on many areas of knowledge: vocabulary, literature, social sciences, chemistry, biology, physics, mathematics, and general information. The items are arranged in order of difficulty. The subject's score is the number of his right answers. The maximum possible score is 100. The time allowed for testing is 50 minutes.

The Miller Analogies Test is now available in four regular forms: J, K, L, and M. A fifth form, R (previously Form H), is available for re-examination.

Two well-established tests of mental ability were selected as the criteria in the study. The researcher wished to determine the degrees of correlation among these instruments and the Miller Analogies Test.

The Wonderlic Personnel Test was chosen because of its ease in administration and its accuracy in the measurement of mental ability levels. The instrument was so named in order to decrease the hesitancy and fear generally associated with psychological tests.

The Wonderlic Personnel Test (Wonderlic) is practically self-administered and requires only 12 minutes. All the necessary instructions are given on the first page and sample questions indicate clearly to the examinee the type of questions he will encounter on the test. The fifty items on each form constitute the examination and are answered by the subject without interruption. Included are analogies, analysis of geometric figures, arithmetic problems, disarranged sentences, sentence parallelisms with proverbs, similarities, logic, definitions, judgment, direction-following using clerical items, spatial relations, etc. The final score is the number of questions correct.

In a comparative study of three mental ability tests, Thumin (1970) disclosed favorable comparisons among the Mental Dexterity Test (MDT), the Otis Self-Administering Test of Mental Ability, and the Wonderlic Personnel Test. Other correlational studies between the Wonderlic

and the Otis Self-Administering Test have yielded coefficients ranging from .81 to .87.

The Wonderlic Personnel Test has also been shown to correlate highly with Aptitude G: General Learning Ability of the General Aptitude Test Battery. Droege (1972) cites the similarity of item content for the Wonderlic and the GATB as an indication that the two instruments measure essentially the same abilities.

The Shipley-Institute of Living Scale (Shipley), also known as the Shipley-Hartford and the Shipley Conceptual Scale, was designed as an aid in detecting mild degrees of intellectual impairment in individuals of normal original intelligence. It may also be used as a test of intelligence. The test is composed of a vocabulary test which includes forty multiple-choice items, and an abstraction test, which is composed of twenty completion items. Both tests yield maximum scores of forty; the abstraction test raw score is multiplied by two.

Shipley (1940) obtained reliability coefficients of .87 for the vocabulary test, .89 for the abstraction test, and .92 for the two combined in a study of 322 army recruits.

Wiens and Banaka (1940) administered the Shipley-Institute of Living Scale and the Wechsler Adult Intelligence Scale to 140 subjects, including all major psychiatric classifications, at the Oregon State Hospital. From the

obtained correlation coefficient of .80, the investigators concluded that the Shipley may be used in an institutional setting as a substitute for the WAIS when economy of time in administration is desired.

In similar hospital studies of the relationship between the Shipley-Hartford and the WAIS, conclusions of the investigators were in agreement with that of Wiens and Banaka. Mack (1970) studied patients admitted to a state psychiatric hospital during a seven-month period. His comparison of group test estimates of intelligence revealed the Shipley-Hartford to be the best measure for approximating WAIS Full Scale IQ. Paulson and Lin (1970), in a study of 219 psychiatric patients, found that valid estimates of WAIS IQ could be obtained from the Shipley-Hartford total raw score.

During an investigation of suitable instruments for estimating intelligence, Bartz and Loy (1970) found the Shipley-Hartford to be an excellent screening test for IQ. Wahler and Watson (1962) concluded that a Shipley estimate can be sufficient when IQ is a minor variable in a decision complex.

ADMINISTRATION AND SCORING

The Wonderlic Personnel Test and the Shipley-Institute of Living Scale were administered in groups by the researcher. Each instrument was scored according to

the manual of directions. Miller Analogies Test scores were obtained from the Graduate Office, Austin Peay State University. All names and corresponding test scores were coded by Dr. John D. Martin in order to maintain anonymity.

CHAPTER III

RESULTS

The Pearson Product-Moment technique was employed to compute the correlation coefficients. In the statistical analyses of data, raw scores were used for each instrument. Table I summarizes the correlations for all measures.

TABLE I

CORRELATIONS BETWEEN TESTS

Tests	r
1. Won.-S.H.Abs.	.32*
2. Won.-S.H.Voc.	.23
3. Won.-MAT	.29
4. S.H.Abs.-S.H.Voc.	.01
5. S.H.Abs.-MAT	.43**
6. S.H.Voc.-MAT	.66**

*Significant beyond the five percent level of significance.

**Significant beyond the one percent level of significance.

CHAPTER IV

DISCUSSION

The correlation coefficients obtained in the study were somewhat lower than the majority of coefficients reported in the literature. This may be explained in terms of lack of ideal testing conditions and lack of sophistication of the examiner. Under these and other obvious limitations of the study, correlation coefficients between the MAT and the two sections of the Shipley-Hartford were still significant beyond the .01 level. To the extent that these correlations are positive and significant, the instruments are said to be measuring essentially the same mental operations.

Significant correlations between the Miller Analogies Test and the Shipley-Hartford indicate that the latter may be substituted for the MAT when economy of time in administration and scoring necessitate. The degree of correlation between the two tests substantiates the assessment potential of both as predictors of mental ability levels.

The MAT and the vocabulary section of the Shipley are verbal measures of intelligence; the Shipley abstraction test is primarily verbal also. This similarity of item

content may partially account for the significant correlations between these tests, and for the lack of significant correlation between the MAT and the Wonderlic Personnel Test which has a more varied content.

Until "success" in graduate school performance has been specifically defined and generally accepted, and an instrument has been devised to predict aptitude for such achievement, the Miller Analogies Test may well remain as one of the most appropriate measures for screening graduate school applicants.

CHAPTER V

SUMMARY

The present study was designed to determine the degree of correlation between the Miller Analogies Test and the Wonderlic Personnel Test, and the degree of correlation between the MAT and the Shipley-Institute of Living Scale. The study was also designed to determine if the MAT could be utilized as a reliable measure of intellectual aptitude of prospective graduate students.

The selection of the criteria instruments--the Wonderlic Personnel Test and the Shipley-Institute of Living Scale--was based on their reliability and validity, their extensive usage in educational and institutional settings, and their similarity to the MAT.

The correlation coefficients obtained in the study between the Miller Analogies Test and the criteria instruments ranged from .29 to .66. Correlations between the MAT and Shipley Vocabulary, and the MAT and Shipley Abstractions were significant beyond the .01 level.

In light of these correlations, it is concluded that the MAT is a valid measure of intellectual functioning which may be utilized effectively in the selection of graduate students.

REFERENCES

Bartz, W. and D. Loy. The Shipley-Hartford as a brief I.Q. Screening Device. Journal of Clinical Psychology, 1970, 26 (1), 74-75.

Bentley, J. Creativity and Academic Achievement. Journal of Educational Research, 1966, 59 (6), 269-272.

Dailey, J. in The Fifth Mental Measurements Yearbook. Oscar K. Buros (ed.). The Gryphon Press, Highland Park, New Jersey: 1959.

Durnall, E. Predicting Scholastic Success for Graduate Students in Education. School and Society, 1954, 80, 107.

Droege, R. in The Seventh Mental Measurements Yearbook. Oscar K. Buros, (ed.). The Gryphon Press, Highland Park, New Jersey, 1972.

Humphreys, L. in The Sixth Mental Measurements Yearbook. Oscar K. Buros, (ed.). The Gryphon Press, Highland Park, New Jersey, 1965.

Hyman, S. The Miller Analogies Test and University of Pittsburgh Ph.D's in Psychology. American Psychologist, 1957, 12, 35-36.

Jenson, R. Predicting Scholastic Achievement of First-year Graduate Students. Educational and Psychological Measurement, 1953, 13, 322-329.

Mack, J. A comparative study of group test estimates of WAIS verbal, performance, and full scale I.Q's. Journal of Clinical Psychology, 1970, 26 (2), 177-180.

Mehrabian, A. Undergraduate ability factors in relationship to graduate performance. Educational and Psychological Measurement, 1969, 29, 409-419.

Miller Analogies Test Manual. New York: The Psychological Corporation, 1970.

Payne, D., and C. Tuttle. The predictive relationship of the Miller Analogies Test to objective and subjective criteria of success in a graduate school of education. Educational and Psychological Measurement, 1966, 26, 427-430.

Paulson, M., and T. Lin. Predicting WAIS IQ from Shipley-Hartford Scores. Journal of Clinical Psychology, 1970, 26 (4), 453-461.

Shipley-Institute of Living Scale Manual. Hartford, Connecticut: The Institute of Living, 1967.

Shipley, W. A self-administering scale for measuring intellectual impairment and deterioration. Journal of Psychology, 1940, 9, 371-377. Cited in Shipley-Institute of Living Scale Manual. Hartford, Connecticut: The Institute of Living, 1967.

Smith, L. Correlates of the Miller Analogies Test. School and Society, 1957, 85, 286-287.

Thumin, F. Comparative Study of Three Mental Ability Tests. Journal of Industrial Psychology, 1970, 5 (1), 1-7.

_____, and C. Boernke. Miller Analogies Test performance as related to age, sex, and academic field. Psychological Reports, 1966, 19 (3, pt. 1), 751-754.

Vickers, D. Correlates of the Slosson Intelligence Test with the California Short-Form Test of Mental Maturity, The Quick Word Test, The Wide Range Vocabulary Test, and the Shipley-Institute of Living Scale. Unpublished Master's Thesis, Austin Peay State University, 1970.

Wahler, H., and L. Watson. A comparison of the Shipley-Hartford as a Power Test with the WAIS Verbal Scale. Journal of Consulting Psychology, 1962, 26 (1), 105.

Wiens, A., and W. Banaka. Estimating WAIS IQ from Shipley-Hartford Scores: A Cross-Validation. Journal of Clinical Psychology, 1960, 16 (4), 452.

Willingham, W. in The Sixth Mental Measurements Yearbook. Oscar K. Buros, (ed.). The Gryphon Press, Highland Park, New Jersey, 1965.

Wonderlic Personnel Test Manual. Northfield, Illinois, 1961.