1965

The Raven Coloured Progressive Matrices, Metropolitan Readiness and Detroit First Grade Intelligence Tests as Predictors of Achievement in Primary Grades

Robert Harold Wirt
Central Washington University

Follow this and additional works at: http://digitalcommons.cwu.edu/etd

Part of the Educational Assessment, Evaluation, and Research Commons, and the Educational Methods Commons

THE RAVEN COLOURED PROGRESSIVE MATRICES,

METROPOLITAN READINESS AND DETROIT FIRST

GRADE INTELLIGENCE TESTS AS PREDICTORS OF

ACHIEVEMENT IN PRIMARY GRADES

———

A Thesis

Presented to

the Graduate Faculty

Central Washington State College

———

In Partial Fulfillment

of the Requirements for the Degree

Master of Education

———

by

Robert Harold Wirt

July 1965

APPROVED FOR THE GRADUATE FACULTY

_____

Eldon E. Jacobsen, COMMITTEE CHAIRMAN

_____

H. L. Anshutz

_____

Howard B. Robinson

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION AND PROBLEM

Group tests of mental ability or talent are used in
the primary grades of the public schools for two main
purposes: (1) to assess readiness of children for learning
certain kinds of material and (2) to facilitate the
necessary grouping of youngsters according to ability
within a subject area. The Metropolitan Readiness Test
(MRT) and the Detriot First Grade Intelligence Test
(DFGIT) are two instruments commonly used to fulfill
these aims. The value of testing in the public school
setting is often expressed through some measure of their
relative efficiency in predicting success, which in the
school setting would be academic performance or achievement.

I. THE PROBLEM

Statement of the problem. This study was designed
to investigate the extent to which the Raven Coloured
Progressive Matrices Test (RPM) correlates with academic
performance; and, when used with the MRT, the extent to
which it increases the efficiency of predicting school
achievement. Of principal interest were primary students
in Grades One, Two, and Three.

Also investigated were the relationships among the

above-named measures and their relation to the criterion, teacher judgment or rating.

Reliability measures were computed for the RPM scores and for the Teacher Ratings. Further investigation of the relationship between RPM scores and academic achievement was done with students on an individual basis.

II. DEFINITION OF TERMS USED

Academic achievement. Achievement is defined as the teacher's evaluation of each student's academic performance without regard to estimated intelligence or intellectual potential.

Teacher ratings. The teachers were instructed to rank their students on a six point scale and then place the students in rank order according to actual achievement.

Concurrent validity. The new or different test (RPM) is compared with existing procedures or tests (MRT and DFGIT).

Predictive validity. The students' expected academic achievement is predicted on the basis of test scores.

The chapter immediately following will review the literature with emphasis on the RPM. The procedure used in

collecting the data will be described in Chapter Three, followed by statistical methods and the obtained results. The results and their implications are discussed in Chapter Five, followed by a summary chapter.

CHAPTER II

BACKGROUND

A review of the literature revealed few studies of direct relevance to the aims of the present investigation with the RPM as the instrument of primary concern.

The author (5) of the RPM states:

> The scale as a whole can be described as a test of observation and clear thinking. Each problem in the scale is really the mother or source of a system of thought, while the order in which the problems are presented provides the standard training in the method of working.

> The way in which the test is presented, the fact that it is untimed, and the group of figures from which the choice has to be made, have been chosen to insure that success depends only upon the person's present capacity for intellectual activity.

Green and Ewert (2:139-42) found that scores on the RPM correlated to about the same degree with intelligence tests which stress verbal elements as with those stressing nonverbal aspects. They inferred the RPM cannot be considered a test of nonverbal reasoning ability, but instead should be considered a test of fairly complex intellectual processes.

In a study by Martin and Wiechers (4:143-4) high correlations (.91, .84, .83) with Full Scale, Verbal and Performance I.Q's, respectively, were obtained between the RPM and the WISC scales. Subjects were 100 school children between 9 and 10 years of age. Because of these

high correlations and the ease and speed of administration, the authors conclude that the RPM will find more extensive use in the clinical testing of children.

In an unpublished study by Wilkins and Wirt (6) stability coefficients were computed for grades one, three, and six, with r's of .761, .795, and .781, respectively. These are based on N's of 41, 21, and 23. The test-retest method was used with a time interval of one month. The relationships between teachers' estimates of the students' status in class regarding mental ability were compared with RPM scores. A tetrachoric r of .399 was obtained with an N of 201. From this study it was concluded:

> The RPM is suited to the purpose of efficient and time saving appraisal of mental ability among children of early school age, primarily in grades one through four. The secondary usefulness is in the direction of locating students who through inconsistencies between RPM scores, teacher appraisal, or other observations, present problems of learning that are secondary to other adjustment difficulty either of behavioral or organic origin.

The aforementioned studies lead to the inference that the RPM may be effective for use in the public schools. If this is true, it would give an efficient (but different, in that it is nonverbal) and additional measure of present and potential intellectual functioning.

The MRT is used as a validity check of academic

standing because of its common use in the schools and its high correlation with achievement. Anastasi (1:475-7) includes the MRT among the best known reading readiness tests. She reports validation studies against subsequent achievement test scores in reading, with the validity coefficients ranging from the .50's to the .80's. Correlations within the same range have been found between the readiness tests and tests of general intelligence for the primary grades.

CHAPTER III

PROCEDURE

For this study one self-contained classroom from
Grades One, Two, and Three was utilized. The classes were
in the Vale Building, Cashmere, Washington. The teachers
of their respective grades were instructed to rate each
student on a six point scale--excellent, above average,
good, fair, below average, and poor. They then placed the
students in rank order according to their academic achieve-
ment and standing at that time. This procedure was carried
out during the month of December, 1961, and repeated
during the month of May, 1962.

The RPM was administered in December, 1961, by this
investigator with the assistance of a proctor. The test
was administered to small groups of five or six students.
The RPM was again administered in May, 1962, without the
assistance of a proctor, since familiarity of the students
with the general procedure facilitated the administration.

Scores on the Detroit First Grade Intelligence and
the Metropolitan Readiness Tests were obtained from the
students' cumulative record folder.

Of the original number of students tested in
December (N=66), 61 remained in the Cashmere School through
May. Consequently, this study has an N of 61 which is

distributed as follows: Grade One, 17; Grade Two, 21; and Grade Three, 23. The size of the sample[1] and the nature of the data led to the decision to use Spearman's rank-difference correlation method. All test data and the Teacher Ratings were ranked from low to high, assigning low the rank of 1.

---

[1]When samples are small, a common procedure applied to regular data in the place of the product-moment method is the rank-difference method of Spearman. It is conveniently applied when the number of pairs, or N, is less than 30 (3:310).

CHAPTER IV

RESULTS

The results are presented by grade level with the exception of reliability coefficients for the RPM scores and Teachers' Rating, which are reported for all three grades in Table I. The stability coefficients were computed by means of the rank-difference method. The ranks used were obtained by the test-retest method as described in Chapter III. The reliabilities for the RPM are all .87 or higher and are significant at or beyond the one per cent level of confidence. Teacher Ratings are similarly high, again reaching the one per cent level. The RPM reliabilities are consistent with those reported earlier in the study by Wilkins and Wirt (6).

I. GRADE ONE

Interrelationships reported for Grade One (Table II) are the highest of the three grades. The rho coefficients are significant at the .01 confidence level with the exception of RPM scores with Teacher Ratings and RPM with DFGIT. The primary concern was to determine the predictive validity of RPM with Teachers' Rating of achievement as the criterion. Here the rho of .498 failed to reach significance.

Of secondary interest was the extent to which RPM shows concurrent validity with other predictive measures, namely, MRT and DFGIT. In this instance the earlier and more widely known instruments, DFGIT and MRT, show greater predictive efficiency than RPM, correlating with Teacher Ratings .596 and .824 respectively. Both are significant beyond the one per cent level of confidence. The concurrent validity as shown by the interrelationship of RPM with MRT, rho=.686, is significant at the one per cent level of confidence. This validity is not shown with RPM and DFGIT as the rho is only .318.

Once knowing the various intercorrelations it was hypothesized that RPM might be assessing factors somewhat different from other predictors and might contribute to a multiple correlation. The difference, adding RPM to MRT was negligible. Again using the two more traditional measures we find a similar negligible difference, with MRT proving to be as effective alone as with the addition of any other measures.

II. GRADE TWO

The correlations reported for Grade Two (Table III) are the lowest of the three grades. However, the rho coefficients exceed the .05 level of confidence with the

exception of RPM and Teacher Rating. The predictive
validity of RPM, rho=.151, is not significant.

The predictive efficiency of DFGIT and MRT with
Teacher Rating is significant with rho's of .644 and .471
respectively. Using RPM and MRT as multiple predictors of
academic success does not yield a significant correlation
(multiple correlation=.385). The multiple correlation,
.624, of MRT and DFGIT is significant, but does not
appreciably increase the efficiency of prediction when
compared to using these tests as independent predictors.

III. GRADE THREE

Interrelationships reported for Grade Three
(Table IV) are significant at the one per cent level of
confidence with the exceptions of RPM with DFGIT,
significant at the five per cent level, and RPM with
Teacher Rating. The predictive validity of RPM with
Teacher Rating again fails to reach the five per cent level
of confidence, rho=.357.

As in Grade One the predictive validity of MRT
with Teacher Rating and DFGIT with Teacher Rating is
significant, as is the concurrent validity of RPM with
MRT and RPM with DFGIT. Also, the multiple predictors of
academic success are significant at the one per cent level

of confidence, but the MRT proves to be as effective alone
as with the addition of any other measures.

TABLE I

RELIABILITY COEFFICIENTS (rho) FOR TEACHER
RATINGS AND RAVEN PROGRESSIVE MATRICES SCORES

GRADE ONE:	RPM - - - - - - - - .872*
	Teacher Rating - - .870*
GRADE TWO:	RPM - - - - - - - - .905*
	Teacher Rating - - .952*
GRADE THREE:	RPM - - - - - - - - .891*
	Teacher Rating - - .851*

*Significant at .01 level of confidence.

TABLE II

GRADE ONE

RHO CORRELATION COEFFICIENTS AMONG TEACHER
RATINGS, RAVEN PROGRESSIVE MATRICES, METRO-
POLITAN READINESS TEST, AND DETROIT FIRST
GRADE INTELLIGENCE TEST

		2	3	4
1.	TEACHER	.498	.824*	.596*
2.	RPM		.686*	.318
3.	MRT			.755*

MULTIPLE CORRELATIONS: CRITERION-TEACHER RATING

RPM and MRT - - - - - - - - - .829*

MRT and DFGIT - - - - - - - - .825*

*Significant at .01 level of confidence

TABLE III

GRADE TWO

RHO CORRELATION COEFFICIENTS AMONG TEACHER
RATINGS, RAVEN PROGRESSIVE MATRICES, METRO-
POLITAN READINESS TEST, AND DETROIT FIRST
GRADE INTELLIGENCE TEST

		2	3	4
1.	TEACHER	.151	.471**	.644*
2.	RPM		.513**	.468**
3.	MRT			.474**

MULTIPLE CORRELATIONS: CRITERION-TEACHER RATING

RPM and MRT - - - - - - - - .483

MRT and DFGIT - - - - - - - .671**

*Significant at .01 level of confidence.
**Significant at .05 levle of confidence.

TABLE IV

GRADE THREE

RHO CORRELATION COEFFICIENTS AMONG TEACHER
RATINGS, RAVEN PROGRESSIVE MATRICES, METRO-
POLITAN READINESS TEST, AND DETROIT FIRST
GRADE INTELLIGENCE TEST

		2	3	4
1.	TEACHER	.357	.706*	.575*
2.	RPM		.665*	.519**
3.	MRT			.609*
4.	DFGIT			

MULTIPLE CORRELATIONS: CRITERION-TEACHER RATING

RPM and MRT - - - - - - - - .722**

MRT and DFGIT - - - - - - - .729**

*Significant at .01 level of confidence.
**Significant at .05 level of confidence.

CHAPTER V

DISCUSSION AND IMPLICATIONS

One major limitation of this study is the small N when the total sample is divided by grade level for the correlation computations. However, the high reliability (stability) coefficients obtained over a six month period may partially offset this limitation.

The rank order method of evaluating the students' relative standing in their class has some merit as the teachers were able to make their judgments without being verbally descriptive.

The interrelationships reported in this study between the RPM and MRT are in accord with other similar studies reported in Chapter Two. The MRT and DFGIT also correlated with achievement in the primary grades to the degree expected on the basis of correlations reported in other studies.

The relationship of main concern in this study, RPM and academic achievement, was found to be relatively low. All of these correlations were positive, but failed to reach the .05 level of confidence. When used with the MRT as a multiple predictor, the RPM adds little or nothing. This is also true for DFGIT when added to the MRT. For these reasons it appears that the RPM is not suited for

regular use as a group test in those primary grades similar to the Cashmere sample for the purpose of predicting academic achievement.

The concurrent validity of RPM with tests used to predict academic success, MRT and DFGIT, is significant. On the basis of these findings and other studies reporting high correlations between commonly used verbal tests and the RPM, the predictive validity of the RPM in this study is suprisingly low. Because of this the relationship between RPM and academic standing was investigated on an individual basis with selected subjects of this study.

Six students, two from each grade, were selected for individual study. The two students from each grade level were selected according to the greatest deviation between Teacher Rating and the RPM results. The deviation was in terms of a higher RPM rank than Teacher Rank. The students were given the ten regularly administered subtests, five verbal and five performance, of the Wechsler Intelligence Scale for Children (WISC).

These cases are reported individually, followed by some observations of the group as a whole.

Case A was a male, six years and seven months of age, in Grade One. The Teacher Rank was 1, lowest in the class, and RPM rank was 7.5. The obtained WISC IQ scores

were Verbal=89, Performance=107, and Full Scale=97.   The
full scale score would not indicate this youngster to have
the lowest academic standing in the class, although this
verbal score would suggest some difficulty with academic
work.

Case B was a male, seven years and six months of
age, in Grade One.   The Teacher Rank was 2 and the RPM rank
was 16.   In a class of 17, this places case B next to the
lowest academically and next to the highest on the RPM.
The obtained WISC IQ scores were Verbal=106, Performance=120,
and Full Scale=114.   Although the verbal score is signifi-
cantly lower than the performance score, case B appears to
be an underachiever.

Case C was a female, nine years and one month of
age, in Grade Two.   The Teacher Rank was 1 and the RPM rank
was 20.   The obtained WISC IQ scores were Verbal=75,
Performance=101, and Full Scale=86.   Since the verbal score
is in the borderline category of intelligence, the low
academic standing would be expected.   However, the
significantly higher performance score suggests some uniqe
learning problems.

Case D was an eight year old female in Grade Two.
The Teacher Rank was 5 and the RPM rank was 18.   The
obtained WISC IQ scores were Verbal=103, Performance=117

and Full Scale=110. On the basis of these scores one would not expect this student to be achieveing in the lower one-fourth of her class. Student D appears to be an underachiever.

Case E was a female, nine years and ten months of age, in the Third Grade. The Teacher Rank was 2 and the RPM rank was 21.5. The obtained WISC IQ scores were Verbal=87, Performance=107, and Full Scale=96. While on the basis of the verbal score, student E may not be a significant underachiever, the wide difference between verbal and performance results does suggest this student has unique learning problems.

Case F was a nine year old male in the Third Grade. The Teacher Rank was 6 and the RPM rank was 17. The obtained WISC IQ scores were Verbal=104, Performance=132, and Full Scale=119. On the basis of these scores one would not predict academic standing in the lower one-fourth of his class. It appears student F is a significant underachiever.

Some common factors regarding these six students are apparent. They are all performing academically in the lower one-fourth of their respective classes. All of them obtained considerably higher rank on the RPM than that given by the teacher. The difference in IQ points

obtained on the verbal and performance portions of the WISC are all significant, and in each case the verbal score is the lower of the two. Considerable variation is noted within the group. Three students (B,D,F) were above average on the full scale IQ score and could be described as significantly underachieving. Two (A,E) were near average on the full scale score, but the verbal results were low enough to suggest probable difficulty with academic achievement. Student C could be expected to be low academically, but the difference of 26 IQ points between verbal and performance suggests that she may have unique learning problems.

Actually, all six students have unique learning problems because of the significant variation of intellectual skills. Case C is a classic example of a student with deficient verbal skills and average nonverbal skills, as measured by the Wechsler Intelligence Scale for Children. Inspection of the nonverbal results (Block Design Scaled Score=14) suggests the deficiencies are other than constitutional. This student does not technically fit in Special Education classes for retarded children, as state standards are currently defined. However, minimal academic achievement, similar to the mentally retarded, would be predicted on the basis of the verbal test results.

It can not be inferred that this youngster has the capacities for assimilation of academic skills under regular classroom procedures at the present time. Continued enrollment in the regular classroom is all that is available and in all probability this will result in continued failure.

Frequently the background information reveals pre-school experiences and environmental conditions not conductive to the acquisition of academic or verbal skills. These may include low socio-economic factors, low educational level of parents, broken marriages, frequent moves, etc. The background and experiences of Case C included enough of these factors to be detrimental to the development of verbal skills.

In the six cases studied the RPM results approximate the nonverbal rather than the verbal scores. This factor may have contributed significantly to the low relationship between RPM and academic achievement.

The results obtained in this study suggest that the RPM is not specifically suited for predicting academic success in the primary grades. It does correlate moderately with other commonly used instruments such as the MRT or DFGIT and was expected to add predictive efficiency in a multiple correlation. However, the results

do not justify the time and effort involved.

The RPM may be useful for screening students with unique learning problems in a way the more verbal tests could not and in determining which students need a more thorough evaluation of mental abilities. Further investigation of the RPM as a screening instrument for students with academic problems seems warranted.

CHAPTER VI

SUMMARY

This study investigated the extent to which the Raven Coloured Progressive Matrices Test correlates with academic achievement and, when used with the Metropolitan Readiness Test, whether it would affect or increase the adequacy of predicting academic achievement. Teacher Ratings and RPM scores were obtained twice over a six month period for 61 students in the primary grades. The reliability (stability) coefficients were significant at the .01 level of confidence.

Scores on the Detroit First Grade Intelligence and Metropolitan Readiness Tests were obtained from the students' cumulative record folder. Rho correlations were reported by grade for the following:

Teacher Ratings and RPM

Teacher Ratings and MRT

Teacher Ratings and DFGIT

RPM and MRT

RPM and DFGIT

MRT and DFGIT

The correlations were positive and all were statistically significant with the exception of Teacher Rating and RPM, the relationship of primary concern in this study.

Multiple correlations were computed for each grade using RPM and MRT scores with Teacher Rating as the criterion and MRT and DFGIT scores with Teacher Rating as the criterion. With the exception of the second grade there was little difference between the multiple predictors and all were significant at the .01 level of confidence.

This study did not show the RPM as a valid predictor of academic achievement in the Cashmere primary grades. Neither did it show an increase in the efficiency of prediction when used with the MRT or even compared to MRT alone.

Other studies reported in Chapter II show that the RPM correlates significantly with verbal tests and in this study the correlation between RPM and MRT was high. In view of this the low correlation between the RPM and achievement was definitely not hypothesized. This low relationship was investigated by examining six of the subjects, using the Wechsler Intelligence Scale for Children. All of these subjects had verbal scores significantly lower than performance scores. It was inferred, besides being low achievers, that they presented unique learning problems.

The RPM may be useful in screening for underachievers and students with unusual or unique learning problems.

Further investigation of these aspects seems warranted.

BIBLIOGRAPHY

BIBLIOGRAPHY

1. Anastasi, Anne. _Psychological Testing_. New York: The Macmillan Company, 1954.

2. Green, Meredith W., and Ewert, Josephine C. "Normative Data on Progressive Matrices (1947)," _Journal of Consulting Psychology_, 19 (1955), pp. 139-142.

3. Guilford, J. P. _Fundamental Statistics in Psychology and Education_. McGraw-Hill Book Company, 1950.

4. Martin, Anthony W., and Wiechers, James E. "Raven's Colored Progressive Matrices and the Wechsler Intelligence Scale for Children," _Journal of Consulting Psychology_, 18 (1954), pp. 143-144.

5. Raven, J. C. _Guide to Using the Coloured Progressive Matrices_. H. K. Lewis and Company, London, Revised Order 1956.

6. Wilkins, J. W. Jr., and Wirt, Robert H. "Clinical Merit of the Raven Progressive Matrices," Unpublished Paper, State Department of Research Conference, University of Washington, November, 1961.