DOCUMENT RESUME

ED 063 489 VT 015 493

AUTHOR Lokan, Janice J.; Halpern, Gerald
TITLE Differential Validities For Shop Courses.
INSTITUTION Ottawa Board of Education, Ontario. Research Centre.
SPONS AGENCY Department of Manpower and Immigration, Ottawa (Ontario).; Ontario Dept. of Education, Toronto.
PUB DATE Mar 72
NOTE 49p.; Paper presented at the annual meeting of American Educational Research Association (Chicago, Illinois, April 1972)

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Career Choice; Educational Research; Foreign Countries; Noncollege Preparatory Students; *Occupational Guidance; *Predictive Validity; Program Improvement; *Shop Curriculum; *Vocational Education
IDENTIFIERS *Canada

ABSTRACT

In a vocational high school in Ottawa, Canada, with an enrollment of approximately 750 students, 25 different shop courses are available. Students must take six shop courses during the first year, then specialize in two during the second year. During the first 2 years of operation, between two-thirds and three-fourths of those enrolled requested changes in their program. In order to reduce the number of changes, this study sought to provide information, in the form of differential validities, derived from a suitable battery of tests, which would be useful in identifying those occupations most likely to offer rewarding careers to non-academic individuals. Three separate cohorts of students, each containing 350 students, were used in the study of a double cross-validation design containing four phases. It was concluded that differences in academic performance and performance in shop areas can be predicted with some success for both sexes. (GEB)

DIFFERENTIAL VALIDITIES FOR SHOP COURSES

Paper presented at Annual Meeting
of American Educational Research Association
Chicago, April 1972

Janice J. Lokan

Gerald Halpern

Research Centre, Ottawa Board of Education
March, 1972

DIFFERENTIAL VALIDITIES FOR SHOP COURSES

This paper is an abbreviated account of an extensive study which has been described fully in several reports published by the Research Centre of the Ottawa Board of Education, Ottawa, Ontario, Canada.

The reports are:

Lokan, J. J., "Differential Validities for Shop Courses: Progress Report", Research Report 69-06, June 1970.

(Describes development and selection of aptitude measures included in predictor battery).

Halpern, G. and Lokan, J. J., "Differential Validities for Shop Courses: Second Progress Report", Research Report 69-07, July 1970.

(Describes development of instruments used to measure aspects of vocational interests and occupational preferences).

Lokan, J. J., "Differential Validities for Shop Courses: Final Report", Research Report 70-05, April 1971.

(Gives a brief overview of the study, and presents all results derived with the validation sample).

Lokan, J. J., "Differential Validities for Shop Courses: An Explanation of Purposes and Results", Research Memorandum 71-03, October 1971.

(Gives a brief discussion of the study, intended for lay readers).

A further report, incorporating cross-validation results, is in preparation.

Dr. Robert L. Linn, of the Educational Testing Service, acted as Statistical Consultant for the study.

The study was funded by the Ontario Department of Education, and the Canada Department of Manpower and Immigration.

DIFFERENTIAL VALIDITIES FOR SHOP COURSES

Introduction

Researchers have now been grappling for more than half a century with the idea that a person's future performance, and perhaps also his future satisfaction, can be predicted on the basis of his present characteristics and behaviour. A vast number of studies with the aim of making such predictions has been conducted. Several statistical techniques have been suggested for handling the kinds of prediction and classification problems that are encountered in educational and vocational guidance settings. Yet, to date, no great breakthrough in improving the accuracy of prediction in these settings has occurred. This state of affairs has led Goldman (1972) to express the view that test scores have little to offer in counselling.

The school in which the study described in this paper was carried out, however, represents a case for the more optimistic view that test scores _can_ provide information that is useful to counsellors and students alike. The school is a vocational high school where students who are thought to have minimal chances of success in any regular high school curriculum undergo a two- or four-year shop-oriented programme. Half of each day, however, is devoted to academic work at an appropriate level. Total enrolment at the school is usually about 750 students. Twenty-five shop courses are available; each student must take six in his or her first year, before specialising in two shops from among these six during his or her second year. Thus, assuming that it is desirable for the students' shop specialty training to be in the area in which they will later find work, crucial decisions concerning possible future occupations for them need to be made at the grade nine level.

In its first two years, the school operated without a general testing programme. During that time a large number of students, somewhere between two-thirds and three-quarters of those enrolled, requested changes (often more than one, and at more than one stage during the year) in the programmes of shop options that they had selected at the beginning of the year. The Guidance Department at the school felt that some of these changes arose because many of the students came from limited experiential backgrounds (most fathers were in semi-skilled or unskilled occupations, or were unemployed or non-existent; about 30% of the students came from families on welfare). The students probably knew very little about their own abilities beyond the fact that they had experienced constant failure at elementary school. Thus it seemed that a comprehensive series of tests, provided that they were at a suitable level of difficulty, could be of real assistance in placing the students in appropriate training courses.

Aim of the study

The over-all aim of the study was to provide information, in the form of differential validities derived from a suitable battery of tests, which would be useful in identifying those occupations most likely to offer rewarding careers to non-academic individuals.

Following the suggestions of French (1955) and Horst (1957) that comparative information is of particular use in guidance, stress was placed on the differential prediction of success in some types of training programmes rather than in others. It was felt that the range of courses offered at the school was broad enough to warrant attempts at differential prediction (this is discussed further in Section 2).

1. Related Research

Only a small number of validity studies employing techniques of differential prediction have so far been reported in the literature, though the techniques themselves have been under discussion for many years (e.g. Brogden, 1946; Mollenkopf, 1950; Thorndike, 1950; Horst, 1954). The main reason for this is probably the large number of problems associated with differential prediction, discussed by Wesman and Bennett (1951), Kelleher (1969), and Norris and Katz (1970). However, attempts to overcome these problems have increased in recent years, as the value of comparative prediction in guidance has become more widely recognized.

As with most absolute prediction studies, the differential prediction studies reported have been concerned largely with college or potential college students. The predictor measures tried have ranged from the large, diversified set of 42 measures of aptitude, interest and personality "factors" used by French (1961), to the limited set of scores in four areas of the ACT tests investigated by Cole (1969). Biographic information has also been included in some analyses (Lunneborg and Lunneborg, 1966; Lunneborg, 1968). Differences between grades in several college major fields have been the criteria. By and large, differential correlations found in these studies have been low, not exceeding 0.40. Verbal - quantitative or Humanities - Science differences could, not surprisingly, be predicted best. None of the studies cited above reported results separately by sex.

Two large-scale diffenrential prediction studies have been carried out with high school students (French, 1964; Norris and Katz, 1970). Students at all ability levels in the upper grades of high school were included in both studies, but in the differential prediction of marks, attained at Grade 12 or Grade 13 levels, only academic subject fields were considered as criteria. Shop grades were used for trade school students in French's study, but only for absolute prediction. In both studies results were presented separately by sex. Again, the majority of the differential correlations found were less than .40, though a few exceeded this value, and one or two in each study exceeded .50.

A second strand of relevant research relates to the non-academic nature of the students who were the subjects of this study. In the last few years there has been a great deal of concern regarding possible middle-class culture bias in the verbal, academically-oriented, testing programmes currently in use. Much effort is being devoted to exploring methods of assessment which will not be discriminatory against disadvantaged groups. In most research articles the word "disadvantaged" has had ethnic connotations. However, any group of low-achieving non-academic students, for whom existing guidance batteries are unsuitable, warrants the development of appropriate assessment measures. The work of Freeberg (1969) with disadvantaged adolescents in New York City proved particularly valuable to the present study.

2. Requirements for Successful Differential Prediction

Differential prediction is the prediction of differences between performance on pairs of criteria. The most commonly used method for calculating differential correlation coefficients is due to Mollenkopf (1950), who derived the following formula relating predicted and actual differences:

- - - - - - - - - - -

Insert from p. 1 of Appendix here

- - - - - - - - - - -

It can be seen from this equation that the multiple R's for the absolute prediction of the two criteria should be high. Likewise, the correlation between predictions, r_{a*b*}, should be low. "It is thus the goal of differential prediction to get good predictions of each of the criteria, predictions which are at the same time as independent of each other as possible" (Lunneborg, 1968, p298).

Since the quantities r_{a*b*} and r_{ab} tend to be closely related (Norris and Katz, 1970), it follows that the correlation between actual criterion scores, r_{ab}, should not be too high. This is in agreement with the common sense idea that if criterion variables are highly related, the real differences between them will be small, and difficult to predict. In differential prediction the proportion of variation in the predictors and criteria which is unique becomes important. The chief reason why most differential prediction results have so far been fairly low is thought to be that, in the academic areas studied, the criteria, and also many of the predictors, share too much common variance (Cole, 1969).

In the present study it was hoped that the diversity of shop courses offered, and the planned inclusion in the predictor battery of several types of tests, would provide favourable circumstances for differential prediction.

3. The Study

Design, Methods and Data Collection

To achieve the over-all objective, the study was planned to be longitudinal. In addition to the usual validation and cross-validation procedures, a "pilot" stage was necessary so that tests suitable for the non-academic population in question could be identified or developed. Altogether three separate cohorts of students, comprising all first -year students entering the school in three successive years, were involved in the study. Each cohort contained about 350 students, of whom about two-thirds were boys.

For the double cross-validation design selected (Mosier, 1951), four main phases[1] were delineated:

I: (1968): Pre-testing and selection of predcitor battery-Cohort 1.

[1] A fifth phase, in which the validation sample is being followed through its first year of work experiences, is currently in progress, but does not form part of this report.

II:	Sept. 1969 to Aug. 1970:	Validation (or "derivation") phase - Cohort 2 (hereafter referred to as "validation sample" (VS) or "Sample 1").
III:	Sept. 1970 to July 1971:	Cross-validation phase - Cohort 3 (hereafter referred to as "cross-validation sample" (CVS) or "Sample 2").
IV:	Aug. to Oct. 1971:	Double cross-validation phase, Samples 1 & 2.

Details of the composition of Samples 1 and 2 are shown in Figure 1.

- - - - - - - - - - - - -
Figure 1 about here
- - - - - - - - - - - - -

Phase 1. Where possible it was initially intended to use existing instruments as predictors. Many published tests were each administered to subgroups of about 40 students, who had been selected by sampling methods to ensure representativeness. It soon became apparent that, because of the nature of the student population involved, several tests would have to be constructed for the study. The nature of the population can perhaps be understood from an examination of Figure 2, which shows the distribution by reading grade level of both the VS and CVS. A requirement for admission to the school is that the student must be 15 years of age, yet the average reading grade level is about 5.7. The mean IQ (non-verbal) of both samples was about 86. Published

- - - - - - - - - - - - -
Figure 2 about here
- - - - - - - - - - - - -

tests for average students of their age group would have yielded very restricted ranges of scores, scores which would in most instances not be valid anyway because the students would not have been able to read the test items. Some tests for lower age groups were also tried, but were usually found to be too difficult or too long. (Most of the students come to the school with poor motivation for test-taking, and are characterized by short attention spans).

The selected predictor battery included measures covering scholastic, verbal, mathematical, clerical and mechanical aptitude, eye-hand co-ordination, general motor ability, vocational interests and occupational preferences. A list of the tests included is shown in Table 1. Specially constructed tests are indicated in this table, and some reliability indices are provided. Published tests which were tried but not selected for the battery are shown in Table 2.

Tables 1 and 2 about here

The suitability of the new and modified tests in terms of a) appropriate difficulty level and b) potential for use in prediction was assessed in part by considering the distributions of scores. In addition, item analyses were carried out in Phase II on the two Mathematics tests and the Filing test.

A brief description of the non-published predictor tests is included here. Further details are given in the Research Reports cited at the beginning of this paper.

1. Highland Park Mathematics: Tests basic operations with whole numbers, some with fractions; a few items testing simple concepts; measurement items; a few one-step problems. Items are at Grade 6 level or less.
   45 items. No time limit, but intended for use in one class period.

2. Filing: Intended to test ability to alphabetize and file correctly.
   Item format: file of five names, one name to be filed. Each file used for two items, to minimize reading. Most names are common names of one or two syllables. Variations in order of last and first names, and in positions of differentiated letters within last names.
   24 items. Time allowed: 6 minutes.

3. Object Drawing: Intended to measure eye-hand co-ordination. Items require students to draw a specified object through given dots, to trace over a given line, or to draw a line between given lines. Several items contain shapes to be copied. Scoring requires an overlay of tolerance regions, and a list of criteria to be followed.
   27 items. Time allowed: 5 minutes.

4. Vocational Interest Inventory: (adapted with permission from Freeberg, 1969)
   Item format: pictorial illustration of typical job task, accompanied by simple description of task. Drawings and lettering clear and dark. Items to be rated on 4-point scale, ranging from "Pretty bad - I couldn't take doing it" to "This is great - just the kind of thing I would like to do".
   Separate forms for boys and girls. Four scale scores provided for each sex (the scales were generated from factor analyses of item responses in Sample 1)

The scales are:

White Collar/Clerical
Blue Collar/Aesthetic & Technical
Personal Service/Personal Service
Outdoor/Low Level Occupations (incl. Outdoor)

Boys: 30 items; Girls: 28 items. No time limit (ten minutes is generally sufficient)

5. Self Location of Traits: Intended as a self-report instrument for students to express their preferences for various tasks within the Data-People-Things hierarchies. Item format: simple description of job task, e.g. "Find numbers in one place and copy them in another".
Students responded by marking on a three point scale whether or not they would like to do each task on a job.
26 items: time required, about 5 minutes.

This instrument was considered to be experimental only, and since results from the analyses carried out did not support the hypothesized hierarchies, no meaningful scores could be derived from this instrument. It was therefore excluded from the battery.

6. Preference Record Form: Intended to supply implicit selection information for use in correcting for range restriction (Linn, 1967). Listed all shop courses, asked students to rank in order the 6 they would most like to take, then in order the 6 they would least like to take. This was a complicated task for the students, as well as producing confounding with sex preferences in the "dislikes" section (due to faulty design of the instrument). This instrument was also excluded from the battery.

Phases II and III. In terms of procedure, these two phases paralleled each other. In Phase II, the predictor battery was administered in September and October to the validation sample, make-up testing extended into November, and criterion data was collected for this sample at the end of the school year. In Phase III the same procdures were carried out the following year with the cross-validation sample, except that both sets of measures in this phase were restricted on the basis of analyses carried out in Phase II. A mishap occurred in that the Motor Ability scores for the CVS were discarded by the school before we had been able to record them. Since this test requires individual administration, it was not practical to give the test again.

Mean scores and standard deviations on the predictor tests are shown for the two samples in Tables 3 and 4. In general these indicate the samples to be fairly comparable, though the CVS boys were significantly lower in Reading and

- - - - - - - - - - - - - - - -

Tables 3 and 4 about here

- - - - - - - - - - - - - - - -

Mathematics. Intercorrelations among the predictors are shown for the two samples in Tables 5 and 6. The overall pattern of intercorrelations is very similar, with the one exception that the Mechanical Reasoning test correlated considerably higher with most other tests for boys in the CVS than it did in the VS.

- - - - - - - - - - - - - - - -

Tables 5 and 6 about here

- - - - - - - - - - - - - - - -

All tests were administered under standardized conditions. To compensate as much as possible for the generally low reading level of the students every effort was made to see that they understood test directions. The maximum number of students at any one session was about 80, and, for most sessions, one proctor for approximately every ten students was present. Directions for all tests were read aloud as the students followed the relevant sections of their test papers. For the three Clerical Aptitude tests the procedure of including a complete practice page, called "Part I" of the test but not scored, was adopted, since it was felt in the pilot sessions that many students lost valuable testing time through not being sure of what they were required to do. All items in the SLOT Profile were read aloud while the students worked through them. Throughout all sessions both the attitude of the students and the general testing conditions were good.

Most of the tests were scored by hand. For the Reading, IPAT Intelligence, DAT Mechanical Reasoning and Maths IXF tests the students used separate answer sheets. For all other tests their responses were written directly on the test booklets, following the findings by Clark (1968) that slow learners made significantly more errors when using answer sheets than when writing answers on their test papers. In Phase II, responses from the Highland Park Mathematics test, the Filing test, the Vocational Interest Inventory (VII), the Preference Record Form (PRF) and Self Location of Traits (SLOT) were key-punched and transferred to tape for further analyses.

Criterion data comprised marks in all six academic subjects and in all shop courses taken during the first year. All marks were expressed as percentages. An over-all academic average was computed for each student. Since differential prediction cannot successfully be achieved for highly similar criteria, and since it was assumed that the 25 "different" shops would involve considerable overlap in the abilities and skills required to succeed in them, it was proposed that the shops should be clustered in Phase II so that similar shops would be considered together. Initial plans for establishing shop clusters by multiple discriminant analysis had to be abandoned. A requirement of MDA is that the categorization of subjects must be unique, therefore "shop success" could not be used for categorization, since most students were successful in more than one shop. As an alternative, the selection of "favorite shop" was considered. However, most of the students were then distributed among the few most popular shops, leaving many shops with insufficient numbers for MDA to be legitimately carried out. It was therefore decided that intercorrelations among shop grades should be the major factor in determining the shop clusters, though this also was partly unsatisfactory because some pairs of shops had only small numbers of students taking both. The full list of shops available at the school, and the allocation of these shops to clusters, are shown in Tables 7 and 8. Generally speaking the clusters are in accordance with common ideas about the nature of the

- - - - - - - - - - - - - - -
Tables 7 and 8 about here
- - - - - - - - - - - - - - -

shops themselves. For each student an average grade per cluster was computed.

Since most students took at least one shop from each cluster, the problems of range restriction and bias due to self-selection (e.g. see Wesman and Bennett, 1951) were not a serious issue in this study.

The number of students in each sample with scores in each criterion cluster is shown separately by sex in Table 9. In Tables 10 and 11 the intercorrelations between pairs of criteria are shown. Many of the correlation coefficients are moderately high, despite the attempts to differentiate shop areas by means of clustering.

Tables 9, 10 and 11 about here

In the validation phase scores derived from rating scales of satisfaction with shop courses were also analyzed. Item intercorrelations showed that motivator/hygiene dimensions of satisfaction (Herzberg, et al, 1959) could not meaningfully be distinguished for this sample of students. Satisfaction scores used as criteria were a simple 5-point rating of over-all satisfaction, and motivator/ hygiene ratings added together. For each student an average satisfaction per cluster was computed, based on his ratings of the shops that he took.

In the validation sample complete predictor data was obtained for 192 boys and 95 girls (if he missed only one test a student was considered to have a complete record). Criterion data was collected for 172 of these boys and 86 of the girls. In the cross-validation sample complete predictor data (on a reduced battery) was obtained for 202 boys and 140 girls, and criterion data was collected for 171 of these boys and 120 of the girls. The reader is referred again to Figure 1 for a diagram showing the composition of the samples.

Phase IV. The procedures followed in this double cross-validation stage involved statistical analyses only. These will be discussed in the next section.

4. Results

Intercorrelation matrices for all predictors and criteria were computed separately by sex. Validities for the satisfaction measures were mostly low and somewhat random in nature, and so further analyses were restricted to the academic average and the four shop averages. Table 12 shows a complete list of all predictors and criteria that were considered in the validation phase. Intercorrelations between the restricted predictor and criterion lists are shown for boys in both samples in Table 13, and for girls in both samples in Table 14. In general, the patterns of correlation coefficients are similar.

Tables 12, 13 and 14 about here

In Phase II, stepwise regression analyses were done separately by sex for each criterion in turn, adding tests from the pool of predictors as long as R^2 increased by at least .01. The resulting combinations of predictor tests, standardized regression weights and multiple R's are presented in Table 15 for boys and Table 16 for girls.

Tables 15 and 16 about here

The single most useful test for this sample was the Highland Park Mathematics test, which had large weights for academic averages for both sexes, and moderate weights for several shop averages. Some of the clerical tests and the interest scales were also useful, as was the specially developed "Object Drawing" test. In general, the girls' results were slightly more predictable than the boys'. The highest value of R obtained was 0.75, for girls' academic average, and the lowest was 0.32, for girls' shop miscellaneous.

In order that Phase IV, the double cross-validation stage, could be carried out, the absolute validities for Sample 1 were re-computed, considering only scores on the restricted battery of predictor tests taken by both Sample 1 and 2[2]. The resulting beta weights and multiple R's for the boys and girls

[2] The Object Drawing test was also omitted, since several items were modified and the scoring system was changed to yield greater variance.

in Sample 1 are shown in Tables 17 and 18. It can be seen that some reduction

- - - - - - - - - - - - - - - - -
Tables 17 and 18 about here
- - - - - - - - - - - - - - - - -

occurred in all of the R's, though this was generally only slight. These tables also show the results of the cross-validation phase, when the Sample 1 weights were used to predict the criterion scores for the sample 2 students. Tables 19 and 20 show the double cross-validated R's, when sets of weights were derived in Sample 2 and then applied back to Sample 1. The over-all absolute validity results can best be seen from the summary shown in Table 20a. With one or two

- - - - - - - - - - - - - - - - - -
Tables 19, 20 and 20a about here
- - - - - - - - - - - - - - - - - -

exceptions, notably shops cluster B for boys, the cross-validated and double cross-validated multiple correlations held up very well. All are high enough to be of at least marginal value for absolute prediction, and several are much higher. Once the regression equations for each criterion for each sex, based on the selected predictor tests shown in Tables 17 to 20, had been set up, two scores were predicted for each criterion for each student in both samples (one from the weights derived in his own sample, the other from the weights derived in the other sample). Two sets of intercorrelations for each sex in each sample were then computed among pairs of predicted criteria (the r_{a*b*} terms in Mollenkopf's equation). These are shown in Table 21 for boys, and in Table 22 for girls. For completeness, the actual criterion intercorrelations (r_{ab} terms) are also shown in these tables (see also Tables 10 and 11).

- - - - - - - - - - - - - - - - - -
Tables 21 and 22 about here
- - - - - - - - - - - - - - - - - -

The predicted criterion intercorrelations are very consistent for the two samples when the same set of weights is used, and show some consistency across sets of weights. Many of the values are quite high, a usual, though undesirable, finding in differential prediction studies (Norris and Katz, 1970).

Finally, two sets of differential validity coefficients were computed for each sex in each sample, using the two sets of r_{a*b*} values and the two sets of absolute validities in Mollenkopf's equation. These results are shown in Table 23 for boys and in Table 24 for girls. Again, the results are very

Tables 23 and 24 about here

consistent for the two samples when the same set of weights is applied, and show similarities across sets of weights. The differential correlations computed with Sample 1 weights are generally higher than those computed with Sample 2 weights, and appear to show that differences in criterion performance are more predictable for girls than for boys. Differential correlations computed from Sample 2 weights, however, show no clear superiority in results for either sex. It is a familiar finding in absolute prediction studies that girls are more predictable than boys, but so far evidence in differential prediction studies is conflicting. For example, Norris and Katz (1970) found differential prediction of course marks to be better for girls than for boys in Grade 13, but better for boys than girls in Grade 12.

Despite the relatively high values of r_{a*b*} (shown in Tables 21 and 22) and the moderately high r_{ab} values in Sample 2, most of the cross-validated and double cross-validated differential correlations are at least as good as, or slightly better than, values reported in other studies. The tendency for r_{ab} and r_{a*b*} values to vary together is shown in Figure 3, which was plotted from the set of results computed for Sample 1 using the regression weights shown in Tables 15 and 16. The scatter plot in Figure 4 shows that the governing factor in the magnitude of the differential validities may not be the values of r_{a*b*} and r_{ab} as such, but rather the proximity of each r_{a*b*} to the r_{ab} between the same pair of criteria. A plot of the same two quantities, using data from p 40

Figures 3 and 4 about here

and p 44 of their report, from Norris and Katz' study shows a similar relationship.

This seems to indicate that accuracy in the prediction of criterion scores is as important to successful differential prediction as having criteria which can be differentiated.

In practical terms, considering the differential validities shown in Tables 23 and 24, it appears that differences in academic performance and performance in shop areas can be predicted with some success for both sexes. Among the shop areas for girls, Typing (shops A) can be differentiated moderately well from all other clusters, and Personal Grooming (shops B) can be differentiated from Domestic tasks (shops C). For boys it appears that differences between performance in pairs of shop clusters cannot be predicted very well, with the possible exception of Mechanical tasks (shops A) compared with both Construction tasks (shops C) and Miscellaneous (shops D). According to French (1964), differential validities should be assessed in the same way as absolute validities, but bearing in mind that differential prediction is more difficult. Thus, even differential correlations in the .30's may lead to statements of students' relative chances of success in different areas which could be of some use in guidance.

REFERENCES

Brogden, H. E. An approach to the problem of differential prediction. Psychometrika, 1946, 11, 139-154.

Clark, C. A. The use of separate answer sheets in testing slow-learning pupils. Journal of Educational Measurement, 1968, 5, 61-64.

Cole, N. S. Differential validity in the ACT tests. ACT Research Report No. 30. Iowa City: American College Testing Program, 1969.

Freeberg, N. E. Assessment of disadvantaged adolescents: a different approach to research and evaluation measures, Research Bulletin RB-69-41. Princeton, New Jersey: Educational Testing Service, 1969.

French, J. W. The logic of and assumptions underlying differential testing. Proceedings ETS Invitational Conference on Testing Problems, 1955, reproduced in Anastasi, A. (Ed.) Testing problems in perspective Washington, D. C.: American Council on Education, 1966, p. 321-330.

French, J. W. Comparative prediction and success in college major fields. Part II: Pooling and analyses of results and conclusions. Research Bulletin 61-7. Princeton, N. J.: Educational Testing Service, 1961.

French, J. W. Comparative prediction of high-school grades by pure-factor aptitude, information and personality measures. Educational and Psychological Measurement, 1964, 24, 321-329.

Goldman, Leo. Tests and counselling: the marriage that failed. Measurement and Evaluation in Guidance, 1972, 4, 213-220.

Herzberg, F., Mausner, B., and Snyderman, B. B. The motivation to work. New York: Wiley, 1959.

Horst, P. A technique for the development of a differential prediction battery. Psychological Monographs, 1954, No. 380.

Horst, P. Differential prediction in college admissions. College Board Review, 1957, 33, 19-23.

Kelleher, E. J. Differential prediction for non-random subgroups. American Educational Research Journal, 1969, 6, 633-644.

Linn, R. L. Range restriction problems in the validation of a guidance test battery. Research Bulletin 67-8, Princeton, N. J.: Educational Testing Service, 1967.

Lunneborg, C. E. Use of factor scores in differential prediction of academic success. Journal of Educational Measurement, 1968, 5, 297-300.

Lunneborg, P. W. and Lunneborg, C. E. The differential prediction of college grades from biographic information, Educational and Psychological Measurement, 1966, 26, 917-925.

Mollenkopf, W. G. Predicted differences and differences between predictions. Psychometrika, 1950, 15, 409-417.

Mosier, C. I. Symposirum: The need and means of cross-validation. I: Problems and designs of cross-validation. Educational and Psychological Measurement, 1951, 11, 5-11.

Norris, L. and Katz, M. R. The Measurement of Academic Interests. Part II: The predictive validities of the Academic Interest Measures. College Board Research and Development Report 70-71, No. 5, and ETS Research Bulletin 70-67. Princeton, N. J.: Educational Testing Service, 1970.

Thorndike, R. L. The problem of classification of personnel. Psychometrika, 1950, 15, 215-235.

Wesman, A. G. and Bennett, G. K. Problems of differential prediction. Educational and Psychological Measurement, 1951, 11, 265-272.

APPENDIX: TABLES AND FIGURES

Contents

Mollenkopf's Formula for Differential Validities.

One of the more frequently used formulae for expressing differential alidity is due to Mollenkopf (1950), and was the one used in the present study. rom the formula the validity of a battery in predicting a difference between two riteria, a and b, can be found. If d is the observed difference between performnce on the two criteria (d = a - b) and d* is the difference between predicted erformance on the same criteria (d* = a* - b*), then the validity of the predicted ifference is obtained from the formula:

$$R_{d*d} = \sqrt{\frac{R^2_{a*a} + R^2_{b*b} - 2R_{a*a}\, R_{b*b}\, r_{a*b*}}{2(1 - r_{ab})}}$$

here R_{a*a} or R_{b*b} is the validity of the battery for predicting criterion a or b

r_{a*b*} is the correlation between predicted criteria

r_{ab} is the correlation between actual criteria.

HIGHLAND PARK DIFFERENTIAL VALIDITIES STUDY: VALIDATION AND CROSS VALIDATION OF PREDICTOR BATTERY

Schematic diagram, showing composition of groups, and time periods.

I. VALIDATION (or "derivation") PHASE

September 1969

All students (N=359) enrolled in the first year programme on 30th Sept., 1969, designated as the validation group (includes [illegible] "failures" from previous years)

October 1969

Predictor battery administered, Validation sample established as the 287 students who wrote all, or all but one, predictor test (includes 21 of the "failures")

Excluded

New Canadians	16
Left before end of first term	25
Not enough predictor data	28
Promoted to second year	2
Omitted in error	1
	72

287
+ 72
359

June 1970

Criterion data (end of year grades and satisfaction) collected for the 258 students from the validation sample for whom grades were available.

Excluded

Left school before end of year	29

258
+ 29
287

See next page for further details of this group.

II. CROSS-VALIDATION PHASE

September 1970

All students (N=394) enrolled in the first year programme on 30th Spt, 1970, designated as the cross-validation group (includes 23 "failures" from previous years)

October 1970

Predictor battery administered, cross-validation sample established as the 342 students who wrote all, or all but one predictor test (includes 18 of the "failures")

Excluded

New Canadians	24
Left school before end of first term	14
Not enough predictor data	14
	52

342
+ 52
394

June 1971

Criterion data (end of year grades) collected for the 291 students from the cross-validation sample for whom grades were available.

Excluded

Left school before end of year	50
No grades available	1
	51

291
+ 51
342

further follow-up of these students is not includec in the frame workof the present study.

Figure 1

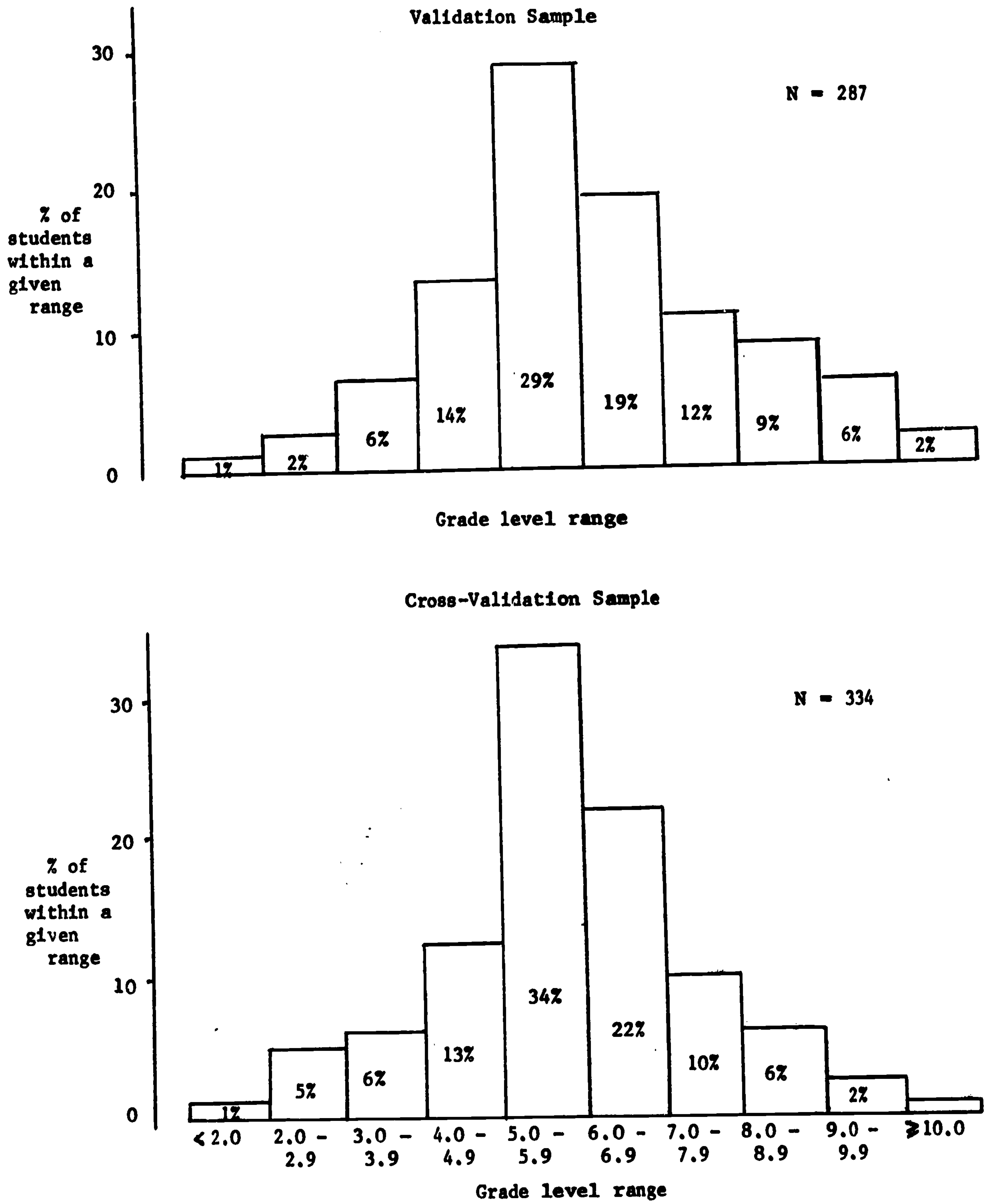

Figure 2: Distribution by sample of reading levels, as measured by Nelson Reading Test, Form A.

Table 1

Predictor Tests Used With Validation Sample

Content Area	Test
SCHOLASTIC APTITUDE	I. P. A. T. "Culture Fair" Intelligence Form A.
VERBAL APTITUDE	Nelson Reading, Form A.
MATHEMATICAL APTITUDE	Maths IXF (Ottawa Board of Education) Highland Park Mathematics*
CLERICAL APTITUDE	Number Comparisons (Personnel Name Comparisons (Research Institute) Filing*
MECHANICAL APTITUDE	Mechanical Reasoning (D. A. T)
MANUAL DEXTERITY	Object Drawing*
VOCATIONAL INTERESTS	Vocational Interest Inventory*
OCCUPATIONAL PREFERENCES	Self Location of Traits Profile* Preference Record Form*
MOTOR ABILITY	Western Motor Ability - Boys only (University of Western Ontario)

* Instruments developed specifically for this study.

Reliability Indices

Maths IXF : 0.93 (K-R 20)
Highland Pk. Math : 0.91 (K-R 20)
Filing : 0.88 (K-R 20 on first thirteen items, stepped up by Spearman-Brown formula for a test of 24 items- gives an estimate of the reliability if the whole test had been done under power conditions)

VII Scales	Boys	Girls	
Clerical	0.74	0.79	(α coefficients)
Service	0.73	0.67	
Outdoor	0.68	0.78	
Technical	0.81	0.77	

Table 2

Published tests tried in pilot stages of the study, and reasons for their exlusion from the predictor battery

Area	Test	Reason for exclusion
Scholastic Aptitude	Henmon-Nelson (Houghton -Mifflin)	Too verbal for use with group with low reading skills.
Verbal Aptitude	Gates-MacGinitie (Teachers College Columbia Press)	Yielded similar results to the Nelson test, but not as easy to administer and score.
Mathematical Aptitude	Metro. Achievement Arith-Advanced	Much too difficult
	Metro. Achievement Arith-Intermediate (Harcourt Brace)	Rather difficult. Too long for students with short attention span.
Clerical Aptitude	Short Tests of Clerical Ability (S. R. A.)	
	Language	Too difficult
	Arith, Parts I & II	Too difficult
	Checking and Coding	Could have been used; were excluded because of overlap with the selected tests in this area.
Manual Dexterity	Object Completion (Psychometric Affiliates)	Rather easy; seemed more related to spatial perception than motor co-ordination
	Purdue Pegboard (S. R. A.)	Impractical, because it requires individual administration. Would be valuable in situations where fine eye-hand coordination is important.
Vocational Interests	Geist Picture Interest Inventory (Western Psych. Services)	The relatively detailed, fine line drawings were confusing to many of the students.
	Minnesota Voc. Int. Inv. (Psych. Corporation)	Much too long, vocabulary level much too high
Occupational Preferences	Gordon Occupational Checklist (Harcourt Brace)	Too long and difficult. Too much of its content beyond the realm of the students' experiences

Table 3

Mean Score and Standard Deviation by Sex on Predictor Tests

Sample 1 (Validation)

Total number of boys in sample = 192
Total number of girls in sample = 95

Test	BOYS N	BOYS Mean	BOYS S. D.	GIRLS N	GIRLS Mean	GIRLS S. D.
Highland Pk. Math.	186	26.1	8.2	92	20.8	8.5
Filing	191	8.8	5.4	94	7.8	5.6
Number Comparisons	192	29.2	8.1	95	30.3	9.3
Name Comparisons	186	27.3	9.0	95	28.7	12.2
Object Drawing	187	13.9	5.9	93	13.8	6.2
VII Clerical	188	21.7	6.1	93	19.2	4.5
VII Service	188	11.6	3.4	93	17.7	3.9
VII Outdoor	188	8.0	3.1	93	13.0	3.8
VII Technical	188	20.8	6.4	93	15.2	7.5
DAT Mech. Reas.	187	36.8	8.5	92	29.6	6.3
Reading (raw score)	192	72.8*	22.2	95	68.1*	20.8
I.P.A.T. Intelligence (raw score)	190	24.8**	6.4	92	22.4**	7.0
Math IXF[(a)]	171	11.6	4.2	86	9.7	4.1
Motor Ability[(b)]	159	46.5	14.8			

* Equivalent to grade levels of 5.9 and 5.7 for boys and girls respectively

** Equivalent to IQ's of 87 and 83 for boys and girls respectively

(a) & (b) No make-up testing of absentees on these tests was attempted, hence the lower N's

(b) Not administered to girls

Table 4

Mean Score and Standard Deviation by Sex on Predictor Tests

Sample 2 (Cross-Validation)

Total number of boys in sample = 202
Total number of girls in sample = 140

	BOYS			GIRLS		
Test	N	Mean	S.D.	N	Mean	S.D.
Highland Pk. Math.	191	21.8	8.9	134	18.0	7.7
Filing	198	10.4	4.6	138	10.5	5.4
Number Comparisons	200	31.9	8.4	134	33.6	9.3
Name Comparisons	200	27.8	8.2	136	29.1	8.1
Object Drawing	199	28.5	7.6	132	26.6	8.0
VII Clerical	193	19.6	5.9	132	17.5	5.1
VII Service	193	11.5	3.4	132	17.3	6.4
VII Outdoor	191	8.0	3.3	132	12.1	3.9
VII Technical	191	20.8	5.2	132	14.1	4.6
DAT Mech. Reas.[(a)]	174	36.4	9.0			
Reading (raw score)	197	62.2*	23.4	137	63.0*	22.7
I.P.A.T. Intelligence (raw score)	180	24.2**	5.7	132	21.5**	5.8

(a) Not administered to girls

* Equivalent to grade levels of 5.4 and 5.5 for boys and girls respectively

** Equivalent to IQ's of 86 and 82 for boys and girls respectively

Table 5

Intercorrelations of Predictor Tests - VALIDATION SAMPLE[(a)]

		1	2	3	4	5	6	7	8	9	10	11	12	13	14
Maths 1XF	1		0.52	0.41	0.21	0.25	0.14*	0.02*	0.07*	0.05*	-0.08*	0.09*	0.32	0.25	0.30
Highland P.Ma.	2	0.52		0.61	0.33	0.35	C.28	0.07*	0.08*	0.09*	-0.06*	0.12*	0.38	0.51	0.60
Filing	3	0.39	0.55		0.27	0.24	0.19	0.07*	0.16	0.21	0.03*	0.12*	0.21	0.46	0.43
Number Compa.	4	0.22	0.39	0.42		0.61	0.16	0.22	-0.02*	-0.07*	-0.10*	-0.02*	0.17	0.32	0.24
Name Compar.	5	0.15*	0.40	0.28	0.45		0.23	0.19	-0.01*	-0.01*	-0.12*	0.07*	0.18	0.35	0.20
Object (Drawing)	6	0.14*	0.30	0.36	0.40	0.26		0.18	-0.05*	-0.11*	-0.07*	0.11*	0.24	0.22	0.23
Motor Ability	7								-0.03*	0.12*	0.08*	0.26	0.30	0.10*	0.27
VII Clerical	8	0.10*	-0.08*	-0.06*	0.07*	0.07*	0.07*			0.42	0.29	0.24	-0.16	0.03*	0.01*
VII Service	9	-0.10*	-0.30	-0.16*	-0.22	-0.00*	-0.11*		0.24		0.47	0.43	-0.20	0.01*	0.09*
VII Outdoor	10	-0.11*	-0.34	-0.13*	-0.14*	-0.11*	-0.12*		0.05*	0.36		0.41	-0.02*	-0.12*	0.01*
VII Technical	11	-0.06*	-0.26	-0.01*	-0.07*	0.03*	-0.02*		-0.08*	0.17*	0.27		0.08*	0.01*	0.19
DAT/MR	12	0.36	0.29	0.27	0.22	0.25	0.10*		0.03*	-0.11*	0.11*	0.08*		0.17	0.43
Reading	13	0.39	0.53	0.53	0.30	0.26	0.20*		-0.01*	-0.11*	-0.24	-0.01*	0.32		0.30
I. Q.	14	0.40	0.67	0.55	0.55	0.42	0.35		-0.03*	-0.28	-0.34	-0.12*	0.37	0.49	

(a) Results for males above diagonal; results for females below diagonal.

* These correlation coefficients are not significantly different from zero at the .05 level.

Table 6

Intercorrelations of Predictor Tests - CROSS-VALIDATION SAMPLE[(a)]

Test		1	2	3	4	5	6	7	8	9	10	11	12	13	14
Maths IXF[(b)]	1														
H. Pk Math	2			0.51	0.25	0.42	0.26		-0.02*	-0.04*	-0.22	-0.02*	0.53	0.56	0.51
Filing	3		0.54		0.25	0.28	0.17		0.01*	0.13*	0.01*	0.10*	0.50	0.45	0.42
Number Compa.	4		0.44	0.45		0.66	0.16		-0.06*	-0.06*	-0.02*	0.06*	0.26	0.26	0.23
Name Compa.	5		0.46	0.48	0.75		0.20		-0.05*	0.02*	-0.08*	0.06*	0.42	0.39	0.22
Object Drawing	6		0.30	0.40	0.40	0.36			0.12*	-0.02*	-0.06*	0.10*	0.42	0.15	0.31
Motor Ability[(b)]	7														
VII Clerical	8		0.09*	0.10*	0.20	0.23	0.00*			0.44	0.14*	0.28	-0.04*	-0.07*	-0.10*
VII Service	9		-0.03*	-0.04*	-0.02*	0.02*	-0.03*		0.21		0.34	0.37	0.08*	-0.05*	-0.04*
VII Outdoor	10		-0.03*	-0.03*	0.06*	0.07*	-0.11*		0.20	0.17*		0.12*	-0.03*	-0.09*	-0.14*
VII Technical	11		-0.16*	-0.14*	-0.04*	-0.05*	-0.02*		0.27	0.15*	0.51		0.22	-0.05*	0.02*
DAT/MR[(c)]	12													0.47	0.56
Reading	13		0.59	0.57	0.50	0.57	0.44		0.10*	-0.14*	-0.08*	-0.12*			0.39
IQ	14		0.55	0.52	0.51	0.60	0.48		0.13*	0.15*	0.02*	0.00*		0.53	

(a) Results for males above diagonal; results for females below diagonal
(b) Scores on these tests not available for cross-validation sample
(c) Not administered to girls in cross-validation sample

* These correlation coefficients are not significantly different from zero at .05 level

Table 7.

Shop courses available at the school.

Art
Auto Body
Auto Service
Carpentry
Drafting
Industrial Sewing
Dry Cleaning
Electrical Repair
Food Services
Graphic Arts (Printing)
Home Management
Horticulture
Hospital Care
Machine Shop
Music
Painting and Decorating
Personal Grooming (Hair Dressing)
Retailing
Sheet Metal
Small Engines
Trowel Trades
Typing and Office Practice
Upholstery
Welding

Shops for boys only:	For girls only:	For boys and girls:
Auto Body	Home Management	Art
Auto Service	Personal Grooming	Dry Cleaning
Building Maintenance		Food Services
Carpentry & Millwork		Graphic Arts
Drafting		Horticulture
Electrical Repair		Hospital Care
Machine Shop		Industrial Sewing
Sheet Metal		Instrumental Music
Small Engines		Painting & Decorating
Trowel Trades		Retailing
Welding		Typing & Office Practice
		Upholstering

Table 8

Allocation of Shops to Clusters

BOYS Cluster	BOYS Shop	GIRLS Cluster	GIRLS Shop
A	Auto Body	A	Typing & Office Practice
	Auto Service		
	Electrical Repair		
	Graphic Arts	B	Personal Grooming
	Machine Shop		
	Small Engines	C	Dry Cleaning
			Food Services
B	Art		Home Management
	Drafting		Industrial Sewing
	Instrumental Music		
	Painting and Decorating	D	Art
	Retailing		Graphic Arts
	Upholstering		Horticulture
			Hospital Care
C	Carpentry and Millwork		Instrumental Music
	Sheet Metal		Painting and Decorating
	Trowel Trades		Retailing
	Welding		Upholstering
D	Building Maintenance		
	Dry Cleaning		
	Food Services		
	Horticulture		
	Hospital Care		
	Industrial Sewing		
	Typing & Office Practice		

Table 9

Frequency of Students by Criterion, Sample and Sex

Criterion	Number of students with scores Validation Sample Boys	 Girls	 Cross-Validation Sample Boys	 Girls
Academic grade average	168	85	169	119
Shop grade average				
- Cluster A	148	74	155	103
- Cluster B	140	66	135	48
- Cluster C	140	75	128	111
- Cluster D	97	82	95	118

Table 10

Intercorrelations of Selected Criteria - BOYS
(Sample sizes in parentheses)

I. Validation Sample

Variable		Academic average 1	2	3	4
Shops A average:	2	0.56 (147)			
Shops B average:	3	0.48 (139)	0.22 (119)		
Shops C average:	4	0.61 (140)	0.37 (134)	0.52 (111)	
Shops D average:	5	0.36 (96)	0.39 (78)	0.09 (91)	0.19 (71)

II. Cross-Validation Sample

Variable		Academic average 1	2	3	4
Shops A average:	2	0.65 (155)			
Shops B average:	3	0.54 (135)	0.42 (123)		
Shops C average:	4	0.53 (128)	0.53 (120)	0.51 (97)	
Shops D average:	5	0.42 (95)	0.53 (81)	0.36 (75)	0.41 (61)

Table 11

Intercorrelations of Selected Criteria - GIRLS
(Sample sizes in parentheses)

I. Validation Sample

Variable		Academic average 1	2	3	4
Shops A average:	2	0.54 (73)			
Shops B average:	3	0.45 (65)	0.35 (60)		
Shops C average:	4	0.69 (74)	0.49 (69)	0.36 (64)	
Shops D average:	5	0.49 (81)	0.49 (71)	0.26 (63)	0.45 (72)

II. Cross-Validation Sample

Variable		Academic average 1	2	3	4
Shops A average:	2	0.59 (103)			
Shops B average:	3	0.38 (48)	0.41 (41)		
Shops C average:	4	0.[illegible]6 (111)	0.59 (96)	0.53 (42)	
Shops D average:	5	0.70 (118)	0.64 (102)	0.46 (48)	0.56 (111)

Table 12

Complete List of Predictor and Criterion Variables used with Validation Sample

<u>Predictors</u>

1.	Mathematics IXF	
2.	Highland Park Mathematics	
3.	Filing	
4.	Number Comparisons	
5.	Name Comparisons	
6.	Object Drawing	
7.	Motor Ability (boys only)	
	Vocational Interest Inventory (VII)	
8.	White Collar - Boys;	Clerical - Girls
9.	Service - Boys;	Service - Girls
10.	Outdoor - Boys;	Low Level Occupations - Girls
11.	Blue Collar - Boys;	Aesthetic/Technical - Girls
12.	DAT Mechanical Reasoning	
13.	Reading	
14.	IQ (non-verbal)	
15.	Preference Record Form (PRF)	Average for Shop Cluster A
16.	Preference Record Form (PRF)	Average for Shop Cluster B
17.	Preference Record Form (PRF)	Average for Shop Cluster C
18.	Preference Record Form (PRF)	Average for Shop Cluster D

<u>Criteria</u>

1. Academic grade average
2. Shop grade average, Cluster A
3. Shop grade average, Cluster B
4. Shop grade average, Cluster C
5. Shop grade average, Cluster D
6. Overall satisfaction (item 9 only)
7. Total satisfaction (sum of items 1 to 9)
8. Motivator/hygiene satisfaction (sum of items 1 to 8)
9. Shop satisfaction average, Cluster A
10. Shop satisfaction average, Cluster B
11. Shop satisfaction average, Cluster C
12. Shop satisfaction average, Cluster D

Table 13

Intercorrelations of Predictors with Selected Criteria - BOYS

(Sample sizes in parentheses)

I. Validation Sample

Crit. \ Pred.	2	3	4	5	6	8	9	10	11	14	12	13	7
Ac. av.	0.59 (164)	0.41 (167)	0.22 (168)	0.22 (164)	0.29 (165)	0.21 (168)	0.08 (168)	0.00 (168)	0.02 (168)	0.39 (167)	0.19 (165)	0.18 (168)	0.14 (139)
Shops A	0.31 (144)	0.30 (147)	0.21 (148)	0.19 (144)	0.28 (145)	0.09 (148)	0.14 (148)	-0.02 (148)	-0.01 (148)	0.33 (147)	0.33 (145)	0.17 (148)	0.29 (124)
Shops B	0.25 (137)	0.15 (139)	0.22 (140)	0.20 (136)	0.22 (137)	0.12 (140)	0.09 (140)	0.04 (140)	-0.05 (140)	0.25 (139)	0.05 (137)	0.02 (140)	0.13 (115)
Shops C	0.25 (136)	0.17 (139)	0.13 (140)	0.14 (136)	0.22 (138)	0.09 (140)	0.12 (140)	-0.07 (140)	0.08 (140)	0.19 (139)	0.21 (138)	-0.01 (140)	0.14 (120)
Shops D	0.22 (97)	0.17 (96)	0.26 (97)	0.21 (95)	0.15 (97)	0.21 (97)	0.08 (97)	0.07 (97)	0.11 (97)	0.12 (97)	0.04 (95)	0.20 (97)	-0.00 (75)

II. Cross-Validation Sample

Crit. \ Pred.	2	3	4	5	6	8	9	10	11	14	12	13
Ac. av	0.60 (162)	0.42 (164)	0.27 (167)	0.36 (167)	0.30 (166)	0.04 (164)	0.03 (164)	-0.11 (163)	-0.01 (164)	0.41 (154)	0.46 (152)	0.49 (166)
Shops A	0.30 (149)	0.34 (150)	0.13 (153)	0.24 (153)	0.27 (153)	-0.04 (152)	0.03 (152)	0.00 (151)	0.14 (152)	0.34 (143)	0.44 (143)	0.29 (152)
Shops B	0.35 (130)	0.28 (130)	0.17 (133)	0.23 (133)	0.24 (134)	0.13 (131)	0.10 (131)	-0.09 (130)	0.20 (131)	0.22 (123)	0.42 (125)	0.23 (132)
Shops C	0.29 (123)	0.16 (127)	0.24 (128)	0.24 (128)	0.21 (125)	0.14 (124)	-0.02 (124)	-0.20 (123)	0.08 (124)	0.16 (114)	0.27 (117)	0.11 (126)
Shops D	0.10 (91)	0.20 (93)	0.23 (94)	0.22 (94)	0.20 (93)	0.12 (93)	0.00 (93)	0.15 (92)	0.05 (93)	0.19 (84)	0.21 (81)	0.15 (93)

Table 14

Intercorrelations of Predictors with Selected Criteria - GIRLS
(Sample sizes in parentheses)

I. Validation Sample

Crit. \ Pred.	2	3	4	5	6	8	9	10	11	14	12	13
Ac. av.	0.71 (81)	0.45 (83)	0.34 (85)	0.49 (85)	0.27 (85)	-0.11 (85)	-0.16 (85)	-0.11 (85)	-0.29 (85)	0.51 (83)	0.23 (83)	0.31 (85)
Shops A	0.33 (72)	0.36 (72)	0.48 (74)	0.49 (74)	0.28 (74)	-0.15 (74)	-0.36 (74)	-0.15 (74)	-0.31 (74)	0.41 (73)	0.25 (72)	0.27 (74)
Shops B	0.31 (64)	0.16 (65)	0.19 (66)	0.20 (66)	0.31 (66)	-0.10 (66)	0.03 (66)	-0.11 (66)	-0.16 (66)	0.17 (65)	0.19 (65)	0.16 (66)
Shops C	0.51 (72)	0.32 (73)	0.28 (75)	0.32 (75)	0.15 (74)	-0.15 (75)	-0.13 (75)	-0.10 (75)	-0.09 (75)	0.34 (74)	0.18 (73)	0.25 (75)
Shops D	0.25 (79)	0.18 (80)	0.11 (82)	0.19 (82)	0.08 (82)	-0.15 (82)	-0.11 (82)	-0.17 (82)	-0.03 (82)	0.23 (80)	0.09 (80)	0.15 (82)

II. Cross-Validation Sample

Crit. \ Pred.	2	3	4	5	6	8	9	10	11	14	12	13
Ac. av.	0.65 (114)	0.47 (116)	0.37 (113)	0.42 (115)	0.44 (114)	0.07 (115)	0.00 (115)	-0.06 (115)	-0.08 (115)	0.47 (113)		0.54 (117)
Shops A	0.32 (100)	0.36 (100)	0.34 (98)	0.40 (99)	0.38 (99)	0.16 (99)	0.01 (99)	-0.09 (99)	-0.07 (99)	0.38 (97)		0.41 (101)
Shops B	0.41 (45)	0.09 (47)	0.23 (45)	0.17 (45)	0.31 (45)	0.14 (45)	0.22 (45)	0.04 (45)	0.12 (45)	0.32 (45)		0.17 (45)
Shops C	0.41 (107)	0.38 (108)	0.23 (105)	0.35 (107)	0.41 (106)	0.11 (109)	-0.07 (109)	-0.15 (109)	-0.04 (109)	0.37 (106)		0.45 (109)
Shops D	0.41 (113)	0.35 (115)	0.28 (112)	0.41 (114)	0.32 (113)	0.08 (114)	-0.03 (114)	0.00 (114)	0.03 (114)	0.37 (112)		0.50 (116)

Table 15

Standardized Regression Weights and Multiple Correlations (Boys)*

Validation Sample, Complete Predictor Battery

Criterion	Predictors	Standardized Regression Weights	Multiple Correlation
Academic Grade Average	H. Pk. Math	.5964	.6417
	Object Drawing	.2008	
	VII - White Collar	.1497	
	Reading	-.1541	
Shop Grades, Cluster A	DAT - MR	.1488	.4935
	Filing	.1811	
	Motor Ability	.1830	
	Object Drawing	.1768	
	IQ	.1203	
Shop Grades, Cluster B	H. Pk. Math	.2307	.3826
	Object Drawing	.1803	
	Reading	-.1965	
	Number Comparisons	.1774	
	VII - White Collar	.1219	
Shop Grades, Cluster C	H. Pk. Math	.3117	.3976
	Object Drawing	.1850	
	Reading	-.1818	
	Motor Ability	.1017	
	VII - Outdoor	.1587	
	VII - Service	-.1194	
Shop Grades, Cluster D	Maths IXF	.2144	.3993
	Number Comparisons	.2211	
	VII - White Collar	.1961	

* Variables added to regression equation as long as increment to squared multiple R was at least .01.

Table 16

Standardized Regression Weights and Multiple Correlations (Girls)*

Validation Sample, Complete Predictor Battery

Criterion	Predictors	Standardized Regression Weights	Multiple Correlation
Academic Grade Average	H. Pk. Math	.6332	.7484
	Maths IXF	.1716	
	Name Comparisons	.1912	
	Reading	-.1461	
	VII - Aesth./Tech.	.1122	
Shop Grades, Cluster A	Name Comparisons	.3227	.6420
	VII - Aesth./Tech.	-.2580	
	Number Comparisons	.2527	
	Maths IXF	.2196	
	VII - Clerical	-.1494	
	H. Pk. Math	-.1789	
Shop Grades, Cluster B	H. Pk. Math	.3241	.4471
	Object Drawing	.2498	
	VII - Aesth./Tech.	.1637	
	DAT - MR	.1496	
	VII - Service	-.1146	
	IQ	-.1377	
Shop Grades, Cluster C	H. Pk. Math	.4534	.5498
	Maths IXF	.2019	
	VII - Low Level Occupations	.1174	
Shop Grades, Cluster D	H. Pk. Math	.1861	.3249
	VII - Service	-.1774	
	VII - Low Level Occupations	.1470	
	IQ	.1452	

* Variables added to regression equation as long as increment to squared multiple R was at least .01.

Table 17

Standardized Regression Weights (β) and Multiple Correlations* for Restricted Predictor Battery - BOYS

Sample 1 (Validation Sample)

Criterion	Predictors	β	R**	Sample 1 weights applied to Sample 2
Academic average	H. Pk. Math	.6312	.61(.64)	R = .54
	VII-White Collar	.1282		
	Reading	-.1293		
Shop grades, Cluster A	DAT-MR	.2666	.41(.49)	R = .44
	Filing	.2295		
	VII-Blue Collar	.1062		
Shop grades, Cluster B	H. Pk. Math	.2737	.33(.38)	R = .28
	Number Comp	.1859		
	Reading	-.1751		
Shop grades, Cluster C	H. Pk. Math	.2851	.32(.40)	R = .32
	Reading	-.1682		
	DAT-MR	.1315		
Shop grades, Cluster D	Number Comp	.2276	.36(.40)	R = .26
	VII-White Collar	.2014		
	H. Pk. Math	.1134		

* Stepwise analyses were limited to three steps, since previous analyses showed little improvement in prediction when more than three variables were used.

** Values shown in brackets are those derived from the full predictor battery.

Table 18

Standardized Regression Weights (β) and Multiple Correlations* for Restricted Predictor Battery - BOYS

Sample 2 (Cross-Validation Sample)

Criterion	Predictors	β	R	Sample 2 weights applied to Sample 1
Academic average	H. Pk. Math	.4104	.64	R = .51
	Reading	.2142		
	DAT-MR	.1249		
Shop grades, Cluster A	DAT-MR	.3326	.46	R = .39
	Filing	.1600		
	Name Comp	.0593		
Shop grades, Cluster B	DAT-MR	.3256	.46	R = .18
	VII-White Collar	.1489		
	H. Pk. Math	.1714		
Shop grades, Cluster C	H. Pk. Math	.2436	.37	R = .25
	Number Comp	.1884		
	VII-White Collar	.1516		
Shop grades, Cluster D	Number Comp	.1857	.32	R = .23
	VII-Service	.1613		
	DAT-MR	.1614		

* Stepwise analyses were limited to three steps, since previous analyses showed little improvement in prediction when more than three variables were used.

FILMED FROM BEST AVAILABLE COPY

Table 19

Standardized Regression Weights (β) and Multiple Correlations* for Restricted Predictor Battery - GIRLS

Sample 1 (Validation Sample)

Criterion	Predictors	β	R**	Sample 1 weights Applied to Sample 2
Academic average	H. Pk. Math	.6641	.73(.75)	R = .66
	Name Comp	.1599		
	VII-A/T	.1168		
Shop grades, Cluster A	Name Comp	.2968	.60(.64)	R = .35
	VII-A/T	-.2364		
	Number Comp	.2622		
Shop grades, Cluster B	H. Pk. Math	.3670	.37(.45)	R = .45
	VII-A/T	.2080		
	VII-Service	-.1261		
Shop grades, Cluster C	H. Pk. Math	.5395	.53(.55)	R = .40
	VII-LL	.1018		
	VII-Service	-.0809		
Shop grades, Cluster D	H. Pk. Math	.2696	.30(.32)	R = .39
	VII-Service	-.1760		
	VII-LL	.1291		

* Stepwise analyses were limited to three steps, since previous analyses showed little improvement in prediction when more than three variables were used.

** Values shown in brackets are those derived from the full predictor battery.

Table 20

Standardized Regression Weights (β) and Multiple Correlations* for Restricted Predictor Battery - GIRLS

Sample 2 (Cross-Validation Sample)

Criterion	Predictors	β	R	Sample 2 weights applied to Sample 1
Academic average	H. Pk. Math	.4806	.69	R = .66
	Reading	.2099		
	Filing	.0915		
Shop grades, Cluster A	Reading	.2125	.47	R = .46
	Name Comp	.2047		
	Filing	.1446		
Shop grades, Cluster B	H. Pk. Math	.4468	.50	R = .28
	VII-A/T	.2104		
	VII-LL	.1652		
Shop grades, Cluster C	Reading	.3018	.50	R = .42
	H. Pk. Math	.2314		
	VII-Service	-.1227		
Shop grades, Cluster D	Reading	.3167	.54	R = .22
	Name Comp.	.1592		
	H. Pk. Math	.1539		

* Stepwise analyses were limited to three steps, since previous analyses showed little improvement in prediction when more than three variables were used.

Table 20a

Comparison of Multiple R's in Cross-Validation and Double Cross-Validation

Summary of Tables 17 to 20

	BOYS[a]				GIRLS[b]			
	C - V		D C - V		C - V		D C - V	
Criterion	(1,1)*	(1,2)	(2,2)	(2,1)	(1,1)	(1,2)	(2,2)	(2,1)
Academic average	.61	.54	.64	.51	.73	.66	.69	.66
Shops A average	.41	.44	.46	.39	.60	.35	.47	.46
Shops B average	.33	.28	.46	.18	.37	.45	.50	.28
Shops C average	.32	.32	.37	.25	.53	.40	.50	.42
Shops D average	.36	.26	.32	.23	.30	.39	.54	.22

* Indicates Sample 1 weights used in Sample 1, etc.

a The number of boys in each sample was approximately 200.

b The number of girls in each sample was approximately 100.

Table 21

Intercorrelations of Predicted and Actual Criteria - BOYS
Sample 1 (Validation) and Sample 2 (Cross-Validation)

I. Correlations between Predicted Criterion Scores ($r_{a^*b^*}$ terms)

(a) Using Sample 1 weights (Sample 1 below diagonal, Sample 2 above diagonal)

	1	2	3	4	5
1. Acad. av.	--	.53	.76	.90	.59
2. Shops A	.54	--	.39	.63	.43
3. Shops B	.81	.46	--	.79	.69
4. Shops C	.83	.65	.80	--	.43
5. Shops D	.67	.45	.71	.39	--

(b) Using Sample 2 weights (Sample 1 below diagonal, Sample 2 above diagonal)

	1	2	3	4	5
1. Acad. av.	--	.78	.83	.75	.45
2. Shops A	.69	--	.89	.59	.68
3. Shops B	.77	.89	--	.75	.58
4. Shops C	.80	.51	.67	--	.58
5. Shops D	.49	.66	.63	.65	--

II. Correlations between Actual Criterion Scores (r_{ab} terms)
(Sample 1 below diagonal, Sample 2 above diagonal)

	1	2	3	4	5
1. Acad. av.	--	.65	.54	.53	.42
2. Shops A	.56	--	.42	.53	.53
3. Shops B	.48	.22	--	.51	.36
4. Shops C	.61	.37	.52	--	.41
5. Shops D	.36	.39	.09	.19	--

Table 22

Intercorrelations of Predicted and Actual Criteria - GIRLS
Sample 1 (Validation) and Sample 2 (Cross-Validation)

I. Correlations between Predicted Criterion Scores (r_{a*b*} terms)

(a) Using Sample 1 weights (Sample 1 below diagonal, Sample 2 above diagonal)

	1	2	3	4	5
1. Acad. Av.	--	.52	.89	.94	.81
2. Shops A	.64	--	.19	.43	.37
3. Shops B	.88	.32	--	.81	.82
4. Shops C	.95	.54	.85	--	.86
5. Shops D	.77	.45	.82	.86	--

(b) Using Sample 2 weights (Sample 1 below diagonal, Sample 2 above diagonal)

	1	2	3	4	5
1. Acad. Av.	--	.81	.76	.93	.91
2. Shops A	.83	--	.50	.83	.95
3. Shops B	.70	.48	--	.59	.61
4. Shops C	.93	.83	.53	--	.93
5. Shops D	.91	.96	.54	.93	--

II. Correlations between Actual Criterion Scores (r_{ab} terms)
(Sample 1 below diagonal, Sample 2 above diagonal)

	1	2	3	4	5
1. Acad. Av.	--	.59	.38	.66	.70
2. Shops A	.54	--	.41	.59	.64
3. Shops B	.45	.35	--	.53	.46
4. Shops C	.69	.49	.36	--	.56
5. Shops D	.49	.49	.26	.45	--

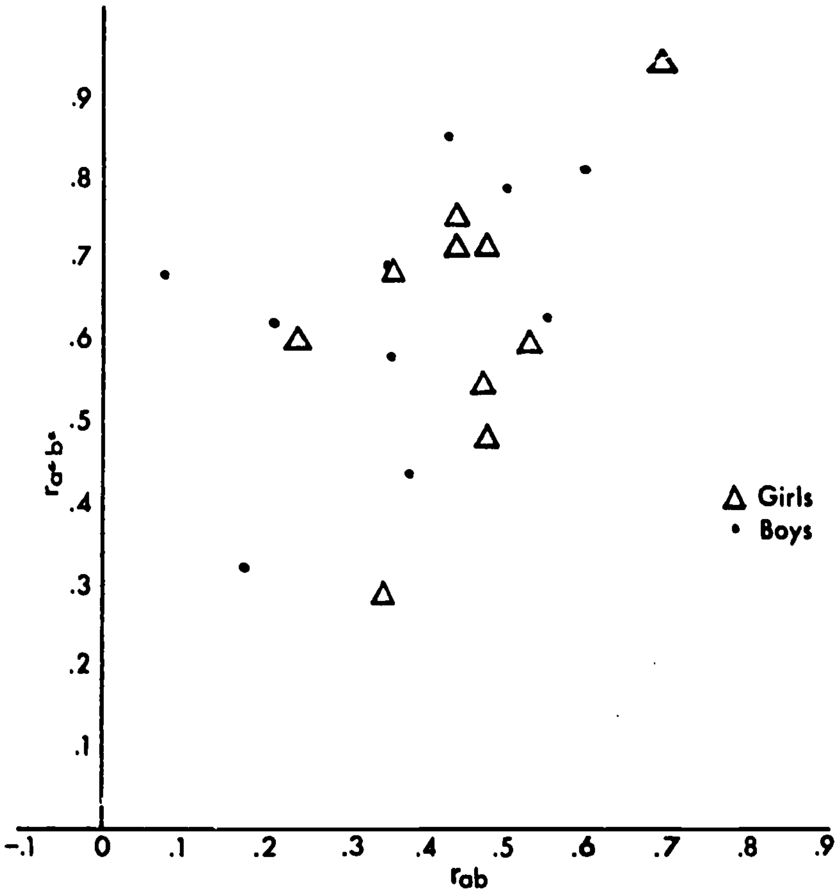

FIG. 3: RELATIONSHIP BETWEEN r_{ab} AND $r_{a^*b^*}$

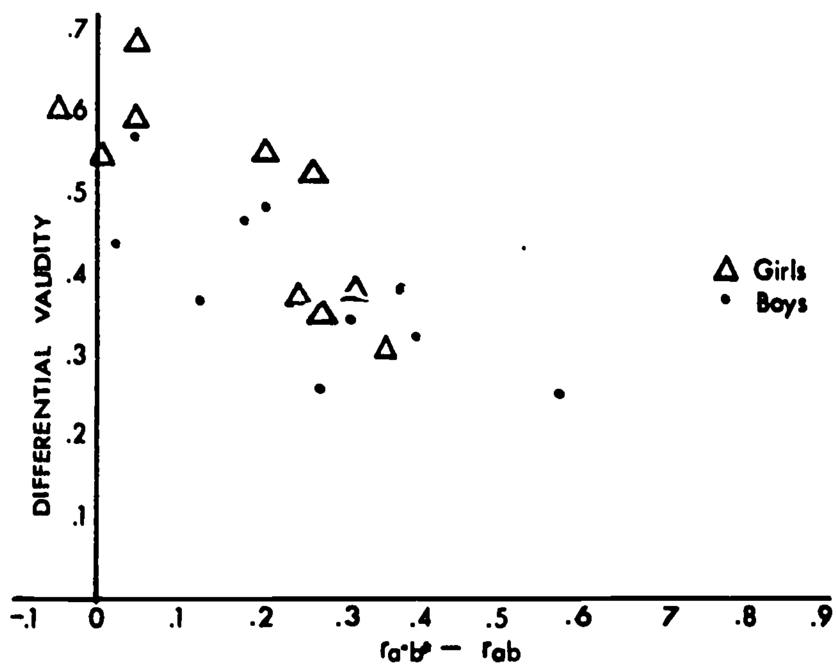

FIG. 4: RELATIONSHIP BETWEEN DIFFERENCE IN $r_{a^*b^*}$ AND r_{ab} AND DIFFERENTIAL VALIDITY

Table 23

Differential Validities (Rd*d) - BOYS
Sample 1 (Validation) and Sample 2 (Cross-Validation)

(a) Using Sample 1 weights (Sample 1 below diagonal, Sample 2 above)

	1	2	3	4	5
1. Acad. av.	--	.6[illegible]	.44	.36	.46
2. Shops A	.55	--	.39	.34	.43
3. Shops B	.39	.31	--	.21	.24
4. Shops C	.44	.28	.21	--	.34
5. Shops D	.41	.37	.19	.29	--

(b) Using Sample 2 weights (Sample 1 below diagonal, Sample 2 above)

	1	2	3	4	5
1. Acad. av.	--	.49	.38	.45	.53
2. Shops A	.49	--	.21	.40	.35
3. Shops B	.40	.17	--	.31	.34
4. Shops C	.46	.37	.36	--	.29
5. Shops D	.49	.32	.27	.23	--

Table 24

Differential Validities (Rd*d) - GIRLS
Sample 1 (Validation) and Sample 2 (Cross-Validation)

(a) Using Sample 1 weights (Sample 1 below diagonal, Sample 2 above)

	1	2	3	4	5
1. Acad. av.	--	.73	.39	.35	.66
2. Shops A	.60	--	.59	.67	.66
3. Shops B	.42	.52	--	.32	.20
4. Shops C	.36	.54	.26	--	.33
5. Shops D	.53	.53	.17	.30	--

(b) Using Sample 2 weights (Sample 1 below diagonal, Sample 2 above)

	1	2	3	4	5
1. Acad. av.	--	.45	.40	.34	.39
2. Shops A	.41	--	.45	.32	.20
3. Shops B	.47	.43	--	.47	.44
4. Shops C	.36	.28	.43	--	.21
5. Shops D	.30	.16	.41	.19	--