This is to certify that the

thesis entitled

"A Comparison of Three Methods of
Mathematics Placement for College Freshmen"

presented by

Mary A. Bone

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Higher Education

Major professor

Date 2/13/81

O-7639

# A COMPARISON OF THREE METHODS
# OF
# MATHEMATICS PLACEMENT FOR COLLEGE FRESHMEN

By

Mary Adams Bone

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Administration and Higher Education

1981

ABSTRACT

# A COMPARISON OF THREE METHODS OF MATHEMATICS PLACEMENT FOR COLLEGE FRESHMEN

By

Mary Adams Bone

An experiment was designed to compare the effectiveness of three methods of mathematics placement at Lake Superior State College in Sault Ste. Marie, Michigan.

Nineteen-eighty incoming freshmen were randomly assigned to one of three treatment groups. Placement was based on one of the following:

(1) Performance on a mathematics placement test written by LSSC faculty member Dr. Gary Thesing. The experiment was cross-validating for the cutoff scores employed.

(2) Performance on the American College Testing Program's Mathematics subtest (ACTM) and the Cooperative Mathematics Test, Algebra II (CAT II). Norms had been cross-validated earlier.

(3) The recommendation made by a mathematics professor, following an individual conference with the student, whose high school transcript, ACT profile and application

for admission were made available beforehand.

Measurement was based on fall term 1980 mathematics grades. "Success" was defined as a grade of A, B or C; "unsuccessful" students were those receiving a D, F or N. The subjects placed in Intermediate Algebra, a remedial course, were contrasted with those placed in college algebra and in calculus. A "log likelihood test," similar to the chi-square test, was used to analyze the resulting three-way contingency table. A second analysis was performed with "A" grades excluded.

Both analyses were statistically significant. One-degree-of-freedom tests, most of which were significant at the one-half of one percent critical level, revealed differences between testing and advising for the criterion variables "success" and "course level." Confidence intervals, obtained from the normal approximation to the binomial distribution, indicated that students receiving faculty advice tended to be placed at a higher level of mathematics than those placed by tests and that they were less likely to be successful. There were no interaction effects.

Instructors judged their mathematics students as being at the right course level, too high or too low. Analysis of these ratings produced a statistically significant difference between counseling and testing at the one-half of one percent critical level. Students rated as "should have taken a lower-level course" occurred in greater proportion in the group placed by faculty members.

No significant differences were found between the two testing methods for any of the dependent variables; however, the "success" effect was significant at the one-half of one percent level when students who accepted a testing recommendation were compared with those who rejected advice. Confidence intervals confirmed that the former group was much more likely to be successful.

A "good test" for mathematics placement was defined as one having a cross-validated success rate of 70 percent or more at a particular institution. It was concluded that the Thesing Placement Test is a good test for Lake Superior State College and that the ACTM-CAT II combination is probably a good testing method there.

Adoption of Dr. Thesing's test for mathematics placement at LSSC is recommended; non-statistical considerations indicate that it is preferable to the combination of the standardized instruments. Continuing reliability and validity studies are also suggested. Means of inducing students to accept the test's prognoses would be beneficial.

Mary Adams Bone

Further conclusions, partially substantiated by the literature of "clinical versus statistical prediction," are that members of a college mathematics faculty will place students at a higher level of beginning mathematics than a good actuarial method will, and that the students placed by faculty are less likely to succeed.

Genuine comparative experiments in mathematics placement should be carried out more often. The phenomenon of "clinical versus statistical placement" may deserve further investigation.

Dedicated to the memory of Miss Myrtle Elliott

February 10, 1891 - January 25, 1980

Teacher and Principal, Sault Ste. Marie Public Schools

1914 - 1956

## ACKNOWLEDGMENTS

To Dr. Van C. Johnson, the chairman of my guidance committee, my major advisor and steadfast supporter.

To the other members of my committee at Michigan State University: Dr. Walter Johnson, Dr. William Mehrens and Dr. John Wagner.

To Dr. Kenneth Shouldice, President of Lake Superior State College, for approving this endeavor; to Dr. Harry Pike, Vice-President for Student Services, for supervising it, serving on my committee, and editing the manuscript.

To the Lake Superior State College mathematics faculty: Professors Bernard Arbic, Thomas Mickewich, Galen Harrison, Gerald Samson, Gary Thesing and Paul Wilson; and to Dr. Reino Hakala, formerly of that faculty, for the generous donation of their time and effort.

To the capable and creative administrators at Lake Superior State College: Mr. Wayne Bohm, Director of Computer Services; Mr. Duane Graham, Registrar; Mr. James Honkanen, Dean of Admissions; Dean Emeritus C. Ernest Kemp, former Head of the Department of Earth Science, Mathematics and Physics; Mr. John Truckey, Director of Counseling and Testing; and their staff members.

To Dr. David Behmer and Dr. Susan Ratwik of Lake Superior State College, and Dr. Larry Lezotte of Michigan State University, for methodological expertise.

To my parents, Ruth and Paul Adams, for bibliographic and financial aid.

To my older son, Michael Yeaton, for technical and psychological assistance when I needed both.

To my patient, loving and helpful family: Ken, Elizabeth and Reggie Bone, and Jenny Yoas. Thank you.

TABLE OF CONTENTS


Chapter Page

List of Tables................................vii

List of Figures............................... x

I. The Problem............................... 1

Significance of the Problem................ 2
Design of the Experiment................... 15
Organization of the Study.................. 16

II. A Review of the Literature................. 17

Placement and Prediction of Success in Mathematics................................ 17
Reviews from the Mental Measurements Yearbooks............................. 17
ACT Reports to Lake Superior State College................................ 21
Shana'a's Survey, Review and Results.... 23
Review of Other Literature.............. 26
Prediction of Success in College........... 39
Clinical Versus Statistical Prediction..... 41
Summary.................................... 46

III. Design and Methodology..................... 50

Design of the Experiment................... 51
Experimental Procedures.................... 53
General Description..................... 53
Avoiding Reactive Arrangements.......... 59
Random and Other Assignments............ 67
Norming the TPT and the ACTM-CAT II Combination.............................. 71
Measurement................................ 81
The Independent Variables............... 81
Difficulties............................ 85
The Dependent Variables................. 92
Rationale............................... 94
Criterion Contamination.................100

Chapter Page

III. Design and Methodology

Analysis....................................103
Contingency Tables and the Log
Likelihood Test..........................103
Data Based on Grades.....................105
Data Matrix..............................105
Preliminary Analysis.....................106
The Fundamental Analysis.................107
Further Analysis if the Null
Hypothesis Is Rejected...................108
Further Analysis if the Null
Hypothesis Is Not Rejected...............109
The Variable Y...........................111
Data Based on Instructors' Judgments....111
Analysis of Naming Versus Coding........112
Confidence Intervals.....................114
Summary.....................................115

IV. Statistical Analysis.........................117

Data Based on Grades........................117
The Data Matrix..........................117
Preliminary Analysis.....................120
The Fundamental Analysis.................124
The Variable Y...........................128
A Final Consideration....................130
Data Based on Instructors' Ratings..........131
The Data Matrix..........................131
Statistical Analysis of A................134
Criterion Contamination.....................137
Naming Versus Coding........................137
Confidence Intervals........................139
Success Rates for Placement Methods.....139
Remedial Placement Rates.................143
Rate of Placement Acceptance.............144

V. Summary and Recommendations..................145

Conclusions.................................149
Discussion..................................150
The Tests................................150
Clinical Versus Statistical Placement...153
Recommendations.............................161

Chapter Page

Appendix A: Supplementary Material..............162

Appendix B: Statistical Calculations............172

Bibliography....................................184

Footnotes.......................................190

LIST OF TABLES


| Table | | Page |
|---|---|---|
| 1 | Mathematics Placement Standards at Lake Superior State College, 1978 | 6 |
| 2 | Mathematics Placement Standards at Lake Superior State College, 1979 | 8 |
| 3 | Proposed Guidelines for Use of the Thesing Placement Test in 1980 | 11 |
| 4 | Validity Data for the Thesing Placement Test (1979) | 73 |
| 5 | Guidelines for Use of the ACTM and CAT II in 1980 | 75 |
| 6 | Validity Data for Standardized Tests (1975-77) MA 100 Intermediate Algebra | 77 |
| 7 | Validity Data for Standardized Tests (1976 and 1977) MA 111, MA 121, MA 181 - College Algebra | 78 |
| 8 | Validity Data for Standardized Tests (1975-77) MA 132 Calculus with Analytic Geometry I | 79 |
| 9 | Comparison of Validity Data for the TPT with that of the ACTM-CAT II Combination | 80 |
| 10 | The Data Matrix M | 118 |
| 11 | Treatment Group-Outcome Contingency Table | 121 |
| 12 | Testing Methods-Outcome Contingency Table | 122 |
| 13 | Placement Method-Outcome Contingency Table | 122 |

| Table | | Page |
|---|---|---|
| 14 | Adjusted Treatment Group-Outcome Contingency Table | 123 |
| 15 | Treatment Group-Course Level-Success Contingency Table | 124 |
| 16 | Testing Methods-Course Level Contingency Table | 125 |
| 17 | Placement Method-Course Level Contingency Table | 126 |
| 18 | Course Level Effect with MA 090 Included | 126 |
| 19 | Testing Method-Success Contingency Table | 127 |
| 20 | Placement Method-Success Contingency Table | 128 |
| 21 | Treatment Group-Course Level-Success Contingency Table (A Grades Omitted) | 129 |
| 22 | Acceptance-Success Contingency Table | 130 |
| 23 | Acceptance-Success Contingency Table (A Grades Omitted) | 131 |
| 24 | The Data Matrix A | 133 |
| 25 | Instructors' Ratings for Students Who Rejected Placement Advice | 133 |
| 26 | Treatment Group-Ratings Contingency Table | 134 |
| 27 | Testing Method-Ratings Contingency Table | 135 |
| 28 | Placement Method-Ratings Contingency Table | 135 |
| 29 | Acceptance-Ratings Contingency Table | 136 |
| 30 | Ratings and Grades Assigned by Students' Placement Advisors | 138 |

| Table | | Page |
|---|---|---|
| 31 | Nomenclature-Outcome Contingency Table | 139 |
| 32 | Thesing Placement Test Difficulty Data (1979) | 162 |
| 33 | Thesing Placement Test Discrimination Data Subjects Ranked by TPT Score (1979) | 163 |
| 34 | Thesing Placement Test Discrimination Data Subjects Ranked by ACTM Score (1979) | 164 |
| 35 | Reduced Nomenclature-Outcome Contingency Table | 176 |
| 36 | Contingency Table for the Fundamental Analysis | 177 |
| 37 | Success-Course Level Contingency Table | 177 |
| 38 | Contingency Table for the Fundamental Analysis (A Grades Omitted) | 180 |
| 39 | Success-Course Level Contingency Table (A Grades Omitted) | 180 |
| 40 | Testing Method-Course Level Contingency Table (A Grades Omitted) | 181 |
| 41 | Placement Method-Course Level Contingency Table (A Grades Omitted) | 181 |
| 42 | Testing Method-Success Contingency Table (A Grades Omitted) | 182 |
| 43 | Placement Method-Success Contingency Table (A Grades Omitted) | 182 |

LIST OF FIGURES


| Figure | | Page |
|---|---|---|
| 1 | Mathematics Placement Form | 55 |
| 2 | Treatment Group-Outcome Contingency Table Format | 106 |
| 3 | Treatment Group-Course Level-Success Contingency Table Format | 107 |
| 4 | Treatment Group-Grades Contingency Table Format | 110 |
| 5 | Treatment Group-Instructors' Judgments Contingency Table Format | 112 |
| 6 | Nomenclature-Outcome Contingency Table Format | 113 |
| 7 | Two by Two Contingency Table | 173 |
| 8 | Two by Three Contingency Table | 173 |

# CHAPTER I

## THE PROBLEM

What placement method should a post-secondary school use to help its incoming freshman students choose an appropriate beginning mathematics class? By "appropriate" is meant a course which is suited to the student's curriculum, ability and experience, so that a reasonable amount of effort, neither trivial nor gargantuan, would result in a grade of "C" or better.

There are various options in curricula; for example, one calculus course for mathematics majors, another for engineering students, still another for those in a business curriculum. In addition to these options, some students will require remedial work; others will be ready for graduate courses.

The number of methods used to place college freshmen in mathematics courses is diverse. Some colleges rely on a professional counseling staff to place students, a staff guided by such tools as high-school grades and standardized test scores. In other schools, similar counseling is done by members of the mathematics faculty. In some cases, assignment is made by an actuarial method and no individual judgment is involved.

In this dissertation, three methods of mathematics placement are compared experimentally. They are:

(1) placement based on performance on a local testing instrument;

(2) placement based on scores obtained on two standardized, commercially available tests;

(3) placement recommended by members of a mathematics faculty following individual conferences with students.

## SIGNIFICANCE OF THE PROBLEM

The problem has great personal significance to the author, because of experiences both in placing students and in being placed - sometimes for the better and once very definitely for the worse - in mathematics courses.

Shana'a has ably stated the case for the importance of mathematics placement:

> ...the increasing horizons of knowledge demand efficiency of educational effort. Yet one cannot enter indiscriminately at an advanced level into a field of study, for many academic areas require an ordered, systematic conquering of concepts and techniques.
>
> The placement of a student in a course which proves to be a review of previous experiences can result in boredom and a loss of interest by him in the entire field. Placement at too difficult a level is equally frustrating and unrewarding, and can generate an acute dislike for the subject. If the student thus placed fails, not only will he probably feel his time and effort have been wasted, but he still may lack the needed knowledge and skills to

> even repeat the course successfully. In addition, the psychological effects of failure are frequently detrimental and combined with the resulting grade point deficiency may cause lost opportunities and drop-outs by capable young people.
>
> Either of the foregoing extremes in placement ... can result in circumstances which prove to be a disservice not only to the individual, but to his classmates, his instructor, and the academic field under study. Educational institutions ... must make maximum utilization of available instructional facilities. The improperly-placed students consume the same physical facilities, the same teaching time, and the same instructional effort--in fact, often more of the latter--as the student capable of maximum profit from the course. This is a drain on our limited educational resources.
>
> In the field of mathematics, the problem of correct placement takes on special urgency. Mathematics is a fabric consisting of various interwoven and dependent branches of study. The student's success, therefore, in a mathematics course is often dictated as much by his former studies in this field as by the new material presented in the course.[1]

Few involved in mathematics placement would dispute this description. More controversial than the need for mathematics placement are the questions: How should it be done? Who should do it?

At Lake Superior State College in Sault Ste. Marie, Michigan, the question has a special significance. The college is oriented strongly towards science, technology and business; most students are involved with a mathematics sequence at some point in their curricula. Educators there have never been completely satisfied with their placement procedures.

From 1973 through 1977, mathematics placement was based on the American College Testing Program's Mathematics and Composite scores and the Educational Testing Service's Cooperative Mathematics Test, Algebra II. (These tests will be abbreviated as ACTM, ACTC and CAT II, respectively, in what follows. Lake Superior State College will usually be denoted as LSSC).

The major placement problem at LSSC is distinguishing students who should elect MA 100, "Intermediate Algebra,"

> An algebra course for those students who have not had at least 1½ years of high school algebra or who need a refresher course in algebra. Elementary operations, first degree equations, one unknown, products and factoring, algebraic fractions, exponents and radicals, quadratic equations, functions and systems of equations.2

from those who are prepared for more advanced work. There are three courses at LSSC in which students ordinarily begin a college-level mathematics sequence:

> MA 111 (for biology, business, social science, etc.) Freshman Mathematics I.
> Sets, analytic geometry topics, functions and graphs (including exponential, logarithmic and trigonometric), logarithms, probability, with problem solving applications of these concepts.

> MA 121 (for mathematics, chemistry, engineering, and students wanting a rigorous calculus course) Pre-Calculus Mathematics.
> Basic theory of functions including polynomial, exponential, logarithmic and trigonometric functions. Inequalities, topics from analytic geometry and plane trigonometry.

MA 181 (for technology) Applied Mathematics I. A course in algebra and trigonometry covering elementary operations, functions and graphs, fractions, systems of linear equations, quadratic equations, exponents and radicals, trigonometric functions, graphs of trigonometric functions, vectors and logarithms.3

Although they are not of equal difficulty and two cover some trigonometry, these courses will be loosely and collectively termed "college algebra."

A secondary problem at LSSC is determining which students may bypass college algebra and safely attempt MA 132, "Calculus and Analytic Geometry I." The 1973-77 standards for placement were:

(1) For students scoring at or below the 30th percentile on ACTM and CAT II and below the 50th percentile on ACTC, MA 100, Intermediate Algebra, was the recommended course. Students who had not taken Algebra II in high school and Canadian students scoring at or below the 30th percentile on CAT II were also recommended for MA 100.

(2) Students whose performance exceeded these criteria but who did not achieve percentile scores of 90 or above on all three tests: ACTM, ACTC and CAT II, were considered as appropriately placed in college algebra.

(3) For a student whose scores on all three tests were in the 90th percentile or above, MA 132, Calculus with Analytic Geometry I, was recommended, provided the student had taken three and one-half or four years of high school mathematics, including trigonometry.[4]

A belief had grown in the LSSC mathematics department that the CAT II "over-predicted" success in mathematics, because it was somewhat out-of-date and the material covered was easier than the content of MA 100.[5] For that reason, Dr. Gary Thesing, then department chairman, prepared in 1977 a placement examination based on the department's behavioral objectives for Intermediate Algebra. By using it as part of the final examination in some sections of MA 100, preliminary norms were obtained in 1978. The test was administered to freshmen during their 1978 pre-orientation sessions at the college and faculty advisors were supplied with the following placement guidelines.

TABLE 1

MATHEMATICS PLACEMENT STANDARDS AT LAKE SUPERIOR STATE COLLEGE, 1978

| Score (out of 50) | Recommendation |
|---|---|
| 0-19 | Strongly recommend Intermediate Algebra |
| 20-29 | Recommend Intermediate Algebra |
| 30-39 | Recommend college algebra |
| 40-50 | Recommend Calculus with Analytic Geometry I if the student has an adequate trigonometry preparation[6] |

Student performance in 1978 indicated that Dr. Thesing's test and norms might be "under-predicting" mathematical success; that perhaps a significant number of students were placed in Intermediate Algebra when in fact one of the college algebra courses was appropriate.[7]

Test responses for the 415 freshmen for whom American College Test scores were available were then analyzed.[8] The mean score was 20.29; the standard deviation, 8.60; the median score, 18.72; and the mode, 16 (29 students). The percentage of students who didn't complete the test was 27.47, indicating that the test should be shortened and/or the 50-minute time period lengthened before calculating a split-half reliability coefficient. The Pearson product-moment correlation coefficient of the test with the ACTC was 0.64; with ACTM, 0.78.

Data were then ranked in descending order by test score and a detailed item analysis was performed, comparing the number of students in the upper third who chose each "distracter" with the number in the lower one-third who did so.[9] Difficulty, discrimination and their indices were also computed. The same analysis was done after arranging the data in descending order by ACTM score, and again after ordering by ACTC. The analysis pointed to the desirability of revising a number of items or eliminating them from Dr. Thesing's test; this was done. Dr. Thesing also shortened the test to 45 items and extended the time period to an hour. Other modifications included the listing of distracters vertically, rather than horizontally, in ascending order for numbers (in ascending order by degree for polynomials, etc.) and in general attempting to follow the recommendations of Mehrens and Lehmann.[10] This revision of the original Thesing test will be denoted TPT in what follows; if the original Thesing test is referenced, that will be made plain.

There was not time after the revision of the test to determine norms for mathematics placement in 1979. A "seat of the pants" recommendation was made by Dr. Thesing, based on his experience with the first version of the instrument.

TABLE 2

MATHEMATICS PLACEMENT STANDARDS AT LAKE SUPERIOR STATE COLLEGE, 1979

| Score (out of 45) | Recommendation |
|---|---|
| 0-14 | Strongly recommend Intermediate Algebra |
| 15-24 | Recommend Intermediate Algebra |
| 25-34 | Recommend college algebra |
| 35-45 | Recommend Calculus with Analytic Geometry I if the student has an adequate trigonometry preparation[11] |

An analysis similar to that done in 1978 was done for the 394 students who had provided the college with their ACT scores. The mean score on the TPT was 21.18;

the standard deviation, 9.77; the median score, 19.29; and the mode, 14 (28 students). This time, 96.70 percent of those taking the test were able to complete it and a split-half reliability coefficient of 0.92 was obtained by finding the Pearson product-moment correlation coefficient of scores on odd-numbered questions with those on even-numbered items (0.8581) and applying the Spearman-Brown correction. The standard error of measurement was then calculated; it was 2.70.

The correlation coefficient of the TPT with ACTC was 0.62; with ACTM, 0.79. Discrimination indices ranged from 21 to 80 for all items but one; that had a discrimination index of 11.8 and on close examination did not appear to need revision. A summary of discrimination and difficulty data may be found in Appendix A.

Equally encouraging information was revealed when fall term 1979 grades in MA 100, MA 111, MA 121, MA 181 and MA 132 were examined. More than 84 percent of the students scoring 14 or more on the TPT passed MA 100 with a grade of C or higher, or "succeeded." Conversely, more than 60 percent of those whose score fell below 12 received a D or F or withdrew from the course (N). Just over half of the students scoring 13 and 14 (10 out of 17) were successful.

In the college algebra group of courses, more than

85 percent of those scoring 28 or more succeeded; more than 80 percent of those scoring 25 or above did so. Below 22, less than 28 percent were successful, while those scoring 22 or 23 had a two-thirds chance (six out of nine cases) of achieving a grade of C or better. The small number (32) of observations available for Calculus with Analytic Geometry I showed that 79 percent of those scoring above 35 succeeded, while neither of the two students scoring below 33 did so (their scores were 29 and 31). Students who scored 34 or 35 had a two-thirds chance of success: three people obtained each score and two succeeded.

The mathematics department had concurrently been developing a "Basic Mathematics" course to meet the needs of students whose mathematical background was insufficient for MA 100. A pilot project covering arithmetic operations, exponents, proportions, percentages, graphing and an introduction to algebra[12] was offered in winter term 1979-80. The college made plans to offer the course each quarter, beginning in fall term 1980.

The availability of the Basic Mathematics course, combined with the 1979 validity data for the TPT, suggested the following guidelines for mathematics placement at Lake Superior State College.

TABLE 3

PROPOSED GUIDELINES FOR USE OF THE THESING PLACEMENT TEST IN 1980

| Score (out of 45) | Recommendation |
|---|---|
| 0-11 | Strongly recommend Basic Mathematics |
| 12-13 | Recommend Basic Mathematics |
| 14-22 | Strongly recommend Intermediate Algebra |
| 23-24 | Recommend Intermediate Algebra |
| 25-33 | Recommend college algebra |
| 34-45 | Recommend Calculus with Analytic Geometry I if the student has an adequate trigonometry preparation[13] |

Despite the promise of placement accuracy held by the test, a number of questions remained. Even if cross-validation in 1980 were to show the same degree of success, one might reasonably ask, "Would placement not have been equally effective had the ACTM and CAT II been used? With the plethora of data available for them, surely a cross-validated set of norms could be developed to meet the standard of the TPT." And in fact this proved to be the case.

An important question frequently asked at LSSC and at other institutions is whether members of the mathematics department should not advise incoming students on the appropriate mathematics course. Kurtz, asking

"Who should place college freshmen in mathematics?" weighed the qualifications of the professional counseling bureau against those of the mathematics faculty and found for the latter:

> ...it is the belief of this writer that the advisement of freshmen students in mathematics should be done by the student's adviser in mathematics. It seems that the mathematics adviser ... is in a better position to provide adequately for these students in the course he teaches.14

LSSC is certainly a suitable setting for such advising. It is a residential college serving approximately 2,000 students. Mathematics classes are small: 35 students or fewer except in some sections of MA 100, which may be as large as 80. In those courses, however, four student tutors assist the instructor. All the full-time members of the mathematics faculty have taught at the college for seven years or longer; thus, all are familiar with the students and their abilities. All are dedicated teachers.

No clear-cut picture emerges from the literature of mathematics placement. Kurtz' viewpoint, often expressed by teachers of mathematics (including this author), may be the best one, but it is supported only by opinion. The literature of "clinical versus statistical prediction" suggests that statistical methods of placement could be at least as effective as advisors.

If so, then the actuarial method, which "takes less time, less effort and - no minor point - can be entrusted to lower paid personnel possessing much less skill" is to be preferred.[15]

Shana'a studied "the placement program for freshmen entering their initial college mathematics courses at the University of Oklahoma" and concluded that

> The ACTM appears to be the best single variable for use as a placement guideline. This conclusion, however, is probably reflecting an inherent bias toward this variable resulting from its use as the primary guideline for the original placement of our sample in their [sic] mathematics courses.
>
> Discriminant functions dependent on the ACTM, the ACTC, the high school mathematics grade point average, and the number of semesters of high school mathematics are of value in distinguishing membership in different mathematics courses at the five percent level of significance. However, they do not prove significantly better than the ACTM at this level as placement tools.16

In contrast, Morgan employed discriminant analysis for mathematics placement at an Ohio community college and found that "a significant loss ensues" when any of the four variables: CAT II score; high school mathematics grade point average; "number of years of academic high school mathematics;" and age; are deleted.[17]

Michigan State University, like LSSC, uses a local placement test for mathematics. A recommendation modest in the extreme was provided for that instrument by Professor Douglas Hall: "The statisticians came around a few years ago and looked at our data. They told us that our test was lousy and the ACTM was worthless." He added, meaningfully, "Of course, that's not fair, because we do placement on the basis of our test."[18]

Shana'a and Hall have pinpointed the essential weakness in the majority of mathematics placement studies: "a single group is studied only once."[19]

> ...such studies have such a total absence of control as to be of almost no scientific value. ... Basic to scientific evidence ... is the process of comparison, of recording differences, or of contrast. Any appearance of absolute knowledge, or intrinsic knowledge about singular isolated objects, is found to be illusory upon analysis. Securing scientific evidence involves making at least one comparison. For such a comparison to be useful, both sides of the comparison should be made with similar care and precision.20

A comparative experiment was therefore designed to provide edification on mathematics placement at LSSC. Perhaps some future studies of mathematics placement will also observe contrasts, rather than single-group phenomena. This experiment may suggest better ones; a description of the procedures and problems encountered may help others. At the same time, some insight into "clinical versus statistical placement" - the relative ability of human judges and actuarial methods to maximize individual success - may be obtained.

## DESIGN OF THE EXPERIMENT

Nineteen-eighty entering freshmen at LSSC were randomly assigned to one of three placement (treatment) groups.

Members of the first group were recommended for various levels of mathematics courses on the basis of their performance on the TPT, using the guidelines of Table 3. Nineteen-eighty was the cross-validating year for these standards.

Members of the second group were assigned to mathematics classes using scores on the ACTM and CAT II as criteria. Norms were based on data for 1975, 1976 and 1977 freshmen and had been cross-validated. The percentage of students in each classification who had passed the course in question was nearly identical to the corresponding ratio obtained for the TPT in 1979.

Students in the third group whose curricula required mathematics or who wished to elect a mathematics class met with members of the mathematics faculty in individual counseling sessions; 15 minutes was scheduled for each conference. Professors then recommended a specific mathematics course for each student with whom they met. Subjects in this group who did not plan to take a mathematics class attended a short talk, accompanied by a handout, on the various beginning mathematics classes at LSSC and the procedures to follow should they ever wish (or be required) to take such a class there.

The criterion variable was a dichotomous one, based on grades in fall term 1980 mathematics courses. Students achieving a grade of A, B or C were termed "successful;" those who received a grade of D, F or N, "unsuccessful." Subjects were classified according to the level of mathematics at which they were placed. The data were subjected to a "log likelihood ratio test," similar to the chi-square test, to determine if the criterion variables were independent of the mathematics placement methods. A supplementary measurement in the form of faculty opinion on the suitability of courses for each student was also analyzed.

## ORGANIZATION OF THE STUDY

The general plan of the study is to present a review of the literature on mathematics placement and related subject matter in Chapter Two. The third chapter is an account of the methodology used in collecting the data and the techniques employed in analyzing them. The results of the analysis are reported in Chapter Four. The summary, conclusions and implications for further study appear in the final chapter.[21]

# CHAPTER II

## A REVIEW OF THE LITERATURE

This chapter is composed of three parts. First, the literature of mathematics placement is surveyed; the majority of studies in this area are predictive ones.

The second part is a digest of the research relating to prediction of overall college success; this issue is similar to the problem of predicting success in mathematics.

Finally, there is a brief discussion of "clinical versus statistical prediction," motivated by experiments in forecasting college grade point average (GPA).

An overview of this material concludes the chapter.

### PLACEMENT AND PREDICTION OF SUCCESS IN MATHEMATICS

#### Reviews from the _Mental Measurements Yearbooks_

Because Buros' _Mental Measurements Yearbooks_ are widely accepted as the best source of information on published tests,[22] it is appropriate to begin with assessments of the ACTM and CAT II from that source.

The ACTM is one of four academic tests contained in the ACT Assessment (formerly the ACT Test Battery).

> The mathematics usage test consists of 40 items in arithmetic, algebra, and plane geometry with a 50-minute time limit. The emphasis on geometry varies among different forms of the test. The problems are of the traditional variety, reflecting little or no "modern mathematics."[23]

Hills says that

> The scores for the four parts seem to be a little less reliable than one might expect from a carefully developed test, roughly in the high seventies to high eighties ... However, a modest sacrifice in reliability of individual scores in order to measure four different characteristics would be an acceptable compromise, and the composite score based on all four scores has an estimated reliability of about .90. The reliabilities of these tests, then, are acceptable...
>
> Statistics for a 1975-76 form (18D) supplied to the reviewer indicate internal-consistency reliabilities of about .9 for the four separate parts.[24]

But Wallace warns that "since many students do not finish the parts of the ACT within the time limits, it may be concluded that the reliability coefficients reported are spuriously high."[25]

Hills, in the eighth and most recent Mental Measurements Yearbook, reports the results of validity studies:

> The cross-validated multiple correlations of ACT scores with college grade averages can be represented by a value in the low .40s, ranging from .20 to .56. The typical correlation between ACT scores and college grades in the appropriate subject would be

> about .4. Adding ACT scores to high school grade averages for predicting college grade averages increases the multiple correlation by about .1, on the average.26

However,

> With tests whose titles seem to suggest that the scores may be closely related to success in particular college courses, such as English and mathematics, it is important to note that there is no technically acceptable evidence that these tests are effective for placement.27

The need felt by Dr. Thesing and individuals at other institutions to develop their own examinations may have arisen because

> There has been a reprehensible neglect of new test development or experimentation associated with the ACT. Intensive attempts toward the evolvement of more varied yet pertinent measures of intellectual skills are long overdue.28

Blommers described the three Cooperative Mathematics tests: Algebra I, Algebra II and Algebra II, as "reasonably fair for students whose first or second courses in algebra were either of the traditional or modern kind."[29]

> ...the items appear to be of excellent quality, the emphasis throughout being on understanding of algebraic concepts as contrasted with manipulative skill. Speed as a possible factor is minimized by a time limit (40 minutes for 40 five-response multiple choice problems for both tests I and II) which should prove more than ample for any student who has truly attained understanding.30

The Director of Counseling and Testing at LSSC, however, reports that "most students do not finish the CAT II."[31]

This might be interpreted to the detriment of entering freshmen there, but for Travers' comment in a later review:

> Speededness data ... indicate that while over 95 percent of the students answered 30 of the 40 items, only between 68 and 71 percent of the students completed the tests. Therefore, speededness does appear to be a factor which may influence the scores.[32]

Thus the reported Kuder-Richardson formula 20 reliability coefficients of 0.84 to 0.89 are probably, like those of the ACTM, artificially inflated.[33]

Describing the CAT II specifically, Travers says that:

> Prospective users of this test will find that the items adequately sample topics dealt with in at least the first semester of second-year algebra currently being taught. But, in some cases, this test will offer inadequate coverage since topics such as trigonometry, determinants, and properties of polynomials, which are recently more common in algebra II classes, are deferred until subsequent tests in this series...[34]

and in the most recent edition of the Mental Measurements Yearbook, Forsyth comments: "These tests probably do not sample some of the important content of many current algebra courses."[35]

On validity:

> The normative data supplied for these tests were gathered in the spring of 1962. ... It hardly seems necessary to note that the use of 1962 norms in the late '70s would yield highly suspect normative information. This seems particularly true given the current concern over declining mathematics achievement.[36]

Validity coefficients are not reported.

An additional criticism of the CAT II is that "the test items .. do not often sample cognitive behavior above the comprehension or application stage."[37]

ACT Reports to Lake Superior State College

ACT provides a number of services to colleges that utilize the Assessment; among them:

> For every participating college that cooperates by providing the necessary criterion data, separate validity studies are performed. Furthermore, the predictive effectiveness of the ACT scores is analyzed by course or subject-matter groupings within schools...38

Lake Superior State College is provided with the Standard Research Report every other year and the Basic Research Report in the alternate years, in accordance with the ACT's recommendation. In the former, for 1976-77 freshmen, the ACT listed correlation coefficients of:

(a) 0.239 between ACTM and college mathematics grade point average (MGPA):

(b) 0.248 between all four academic tests and MGPA (multiple correlation);

(c) 0.259 between high school mathematics grades and MGPA;

(d) 0.241 between CAT II and MGPA;

(e) 0.238 between high school average and MGPA.[39]

The analysis was based on data submitted for 143 students enrolled in MA 100, MA 111 and MA 181 in the fall term of

1976. The three courses were pooled in order to obtain a minimum sample size required by ACT. Two difficulties are immediately apparent.

First, MA 100 is a remedial course; MA 111 and MA 181, from the college algebra group, are at a higher level. Therefore a student scoring in the 50th percentile on ACTM (or CAT II) could be expected to perform satisfactorily in MA 100; such a student would probably find the other courses rather difficult. That is, the same test score could logically be paired with either a high grade or a low one, depending on the course in which the student enrolled.

The "restricted range" problem also arises. Although the lowest-scoring students are thoroughly represented, because MA 100 was the lowest-level mathematics course available in 1976-77, the higher-scoring students who chose calculus are not. This of course is common to all situations in which screening has taken place, and its effect is to depress the value of the correlation coefficient.

There are two ways in which meaningful coefficients might be obtained. All students might be placed in a single course for this express purpose, but this is clearly not feasible. Or, coefficients could be calculated for each course level separately and the (further) restricted range considered when making interpretations. Implicit in the use of either are assumptions that the mathematics grades form a normally distributed variable and that the ways in which the tests are used for placement will not affect the coefficients. Both of these assumptions are open to question.[40]

Unfortunately, then, the information available from the ACT must be regarded as carrying very little meaning. Additionally, the only available validity data comparing the ACTM and the TPT are from 1978, for the original version of Dr. Thesing's test, rather than the revised one employed in the present experiment. For the record, these coefficients are:

(a) 0.429 between ACTM and MGPA;

(b) 0.473 between the four ACT academic tests and MGPA:

(c) 0.230 between high school mathematics grades and MGPA;

(d) 0.276 between the original Thesing test and MGPA; for 177 students in the same three courses.[41]

## Shana'a's Survey, Review and Results

As background for "A statistical analysis of the placement program in mathematics for freshmen at the University of Oklahoma,"

> An informal survey of the principal state universities of the United States was made and that investigation indicates most placement programs are combinations of two or more measurements verying in type and emphasis. Thirty-five schools

> from twenty-nine states responded to letters of inquiry sent to their mathematics departments. Of this group, eight are using the ACT program and nine, the College Entrance Examination Boards (CEEB). Twenty-seven use locally designed tests or other standardized national examinations for placement decisions, advanced standings (with or without college credit), and exemption purposes. Twelve consider the high school grade point averages while eighteen checked the specific high school courses in mathematics taken by the student. A few emphasized individual interviews for decision purposes in placement, and choice of the major field of study played a role of diversified emphasis in most programs. Some colleges partially solve the problem by stringent entrance requirements in mathematics which help assure a certain level of mathematical maturity in the student qualified for admittance to their schools. ...
>
> Thirteen schools indicated some research had been done or was in progress to validate their placement practices. Several institutions mentioned that justification for their present practices is based more on expert opinion and lack of complaints than on conclusions reached by statistical studies.[42]

Other results in the literature were reported by Shana'a, who summarized the findings.

> In the field of mathematics, the high school mathematics grade point average and mathematical achievement tests concentrating on specific concepts and skills in algebra and trigonometry appear most commonly accepted as predictive tools. There is, however, considerable variation in placement policies, as cited in the study of the major state universities, and no one variable has yet been generally acclaimed as the reliable placement guideline for mathematics.[43]

Shana'a herself correlated seven predictive variables with course grade in initial mathematics class and concluded that

> The only available predictive variables whose correlation coefficients with the course grade criterion differ from 0 at the five percent level of significance for each of the elementary mathematics courses ... are the high school grade point average, the high school mathematics grade point average, the ACTE [ACT English] and the ACTC.[44]

She pointed out, however, that "the amount of correlation between a variable and the course grade does not necessarily indicate the power of the variable in classifying students in different courses according to their abilities and skills."[45]

She then examined "the statistics on the subset of properly placed students in our fall sample to determine which variables distinguish between the given courses with a minimum overlap in the ranges of values characteristic of each course level."[46]

Students achieving a grade of B or C in their initial mathematics course were termed "properly placed;" the "subset of students receiving grades of A" was not so classified, because these individuals "may have been capable of accelerating their educational progress by enrolling in the next advanced course."[47] She found that

> The ACTM appears to be the best single variable for use as a placement guideline. This conclusion, however, is probably reflecting an inherent bias toward this variable resulting from its use as the primary guideline for the original placement of our sample in their [sic] mathematics courses.
>
> Discriminant functions dependent on the ACTM, the ACTC, the high school mathematics grade

> point average, and the number of semesters of high school mathematics are of value in distinguishing membership in different mathematics courses at the five percent level of significance. However, they do not prove significantly better than the ACTM at this level as placement tools.[48]

The sample consisted of 1,029 freshmen "enrolled in mathematics [in] the fall of 1962 at the University of Oklahoma."[49] Shana'a cautions the reader "to remain cognizant of the inherent bias in our sample as he examines the data and conclusions of this study;" because "the sharp steady increase of mean values for the course levels with respect to the ACTM and the number of semesters of high school mathematics probably is a reflection of bias resulting from the present placement policy."[50]

Attention is now turned to other studies, not included in Shana'a's dissertation, where such straightforwardness is not often found.

## Review of Other Literature

Turner studied "Prediction of success as a mathematics major at the Minnesota State Colleges;" while the results are of limited interest in this context, it is worth noting that they "indicate that those variables most useful in predicting success in mathematics at any one of the Minnesota State Colleges were not necessarily of prime importance at another..."[51]

Regression analysis was employed, with course grades (4.0 scale) as the dependent variable. Turner's subjects were

> One hundred sixty-nine ... students who graduated with a major in mathematics from the Minnesota State Colleges during the academic year 1966-67... Complete data were not available for six graduating mathematics majors; consequently, they were not included in the study.52

Corotto studied "the relative effectiveness of a test of general scholastic aptitude and a test of mathematics achievement" to predict success in mathematics.[53]

> Although evidence is somewhat contradictory, aptitude tests and achievement tests designed to measure specific content areas appear to provide greater predictive accuracy of success in mathematics than do tests of general intelligence or ability. Seigle reports a study of the prediction of success in college mathematics at Washburn University. Results indicate that at Washburn, an entrance mathematics test was the best single predictor for success in algebra courses. When Krathwohl compared success in college algebra with scores on the Iowa Mathematical Aptitude Test, he found a significant correlation. Wallace found the Iowa Foreign Language Aptitude Test and the Mathematics Placement Test (form B-2) to be significantly more effective in predicting grades for mathematics courses than the ACE [American Council on Education Psychological Examination for College Freshmen] for three levels of algebraic difficulty.54

In comparing "the relationships between 1) success in an initial mathematics course and scores on the three scales of the ACE" (the linguistic, L; the quantitative, Q; and the combined L and Q scores, T) with 2) success in an initial mathematics course and score on a "mathematics screening test developed at the University of Houston (MST)," he found that "the ACE is a better predictor for ... a purely remedial course, while the MST is a better predictor of initial success in [college algebra]."[55]

The sample consisted of 513 students "who entered the University [of Houston] during the fall semester of the 1955-56 academic year."[56] Success was defined as a grade of A, B or C; failure, as a grade of D or F. Biserial correlation coefficients were employed in the analysis.

The means by which students were placed in the beginning courses is not specified, but it was probably the MST.

Boyce used the product-moment method to determine correlations between grades in a "basic college algebra" course and "(1) Educational Testing Service SCAT Q-score, (2) American College Testing Program Mathematics Usage Score, and (3) ACE Cooperative Mathematics Pre-Test for College Students."[57] Instructors assigned a "rank order to each student in each class" so that a reliability estimate for grades (of 0.98) could be obtained.[58]

The correlation coefficients found were, respectively: (1) 0.13, not significant; (2) 0.25, significant at the 0.05 critical level; and (3) 0.41, significant at the 0.01 level.[59]

The sample consisted of "100 qualified freshmen" but Boyce does not describe their selection, nor the year in which they matriculated at an unnamed college in Alabama. "A similar study was performed at four other institutions of higher education in the state. Similar results were obtained at each."[60]

The algebra course was "required of all prospective mathematics majors," but no mention is made of the method used, if any, to place students in the course.[61]

In "A study of the factors associated with success in first-year college mathematics," Wick investigated the records of 1,692 students at "six Minnesota and Wisconsin colleges and universities ... chosen so as to be representative of the different types of public and private institutions as well as of the different types of college curricula."[62]

The five predictor variables used in the equations

> varied from course to course, but in general they may be classified as follows: a measure of high school mathematics achievement, high school rank, mathematics placement test scores, and scores on mathematical and general scholastic aptitude tests.[63]

He found that

> 1. ...there appears to be no significant difference in the quality of preparation for first-year college mathematics as between the experimental (SMSG) and the traditional programs in high school mathematics. ...

> 2. The high school mathematics record was consistently the source of the best predictors of success in first-year college mathematics. The mathematics average over grades ten to twelve gave the highest correlations with success in college algebra courses, whereas the average in grade twelve appeared to be slightly more indicative of success in beginning calculus. ...
>
> 3. The use of multiple regression techniques for predicting success represented a considerable improvement over the use of single predictor variables.[64]

Dunn used student-reported data to define "ideal" placement groups while studying "the University of Arkansas Mathematics Entrance Exam as a placement device."[65] He found that the test's "glaring difficulty seems to be that its questions do not pose sufficient variety to be an effective discriminator between various levels of mathematical sophistication" and that "the best possibility .. for improving ... seems to be to construct a new test."[66]

At Mississippi State University, where the Cooperative Mathematics Test, Algebra III (COOP) was "used for several semesters as a placement test for the first course in calculus," Kohler computed "Pearson product-moment correlations between all pairs of the four variables" COOP, ACTM, ACT Composite score and "grade point average (4.0 scale) obtained in college algebra."[67]

He found that while "GPA correlates highest with COOP ($r = .53$) the extremely small increase over the correlation between GPA and ACTM [0.52], in this instance, does not warrant a mathematics placement testing program to provide information over and above that afforded by the ACT service."[68] However, the COOP was administered "post-instruction during the last week of the semester," while "the ACT battery was administered ... prior to a student's entering Mississippi State University," so that the ACTM may have been used for mathematics placement; the method used to place students in the algebra course is not stated.[68]

The sample consisted of 161 students "enrolled in ... college algebra for the fall semester, 1972," but "not all 161 records were complete. Therefore, the Ns varied depending upon the number of complete records."[70]

Pugh found that "the combination of SAT-M [CEEB Scholastic Aptitude Test - mathematics score] and HSR [high school rank] was the most efficacious predictor of success" for a sample "of students in one of seven different beginning mathematics courses offered by Indiana University."[71] The conclusion was "based on several multiple regression analyses."[72]

Howlett used the multiple-regression technique to determine "the best possible predictor combinations" for success in MA 132, Analytic Geometry with Vectors, at Michigan Technological University.[73] Fourteen variables were utilized; the use of the "CAT" (Cooperative Algebra Test; probably CAT II, but this is not specified) and high school class rank "proved to be the best combination of two variables and at the same time proved to be a substantial improvement in correlation from the single predictor method."[74]

The subjects were "approximately 1000 entering college freshmen at Michigan Tech ... from the freshman class of fall 1966."[75] No cross-validation is reported, nor the means used to place the students in MA 132; it may be inferred that the latter was the "CAT."

Shevel and Whitney found that ACT's Mathematics Placement Examination "improves college mathematics grade prediction ... and that the increase is greater for higher-level courses. ... Prediction of grades by ACT scores and high school grades in low level courses is not much improved by the use of the Mathematics Placement Examination."[76]

The analysis used was "multiple correlation between first semester mathematics grades and six variables (ACT Mathematics, ACT Composite, and the four ACT high school grades)."[77] Placement methods used by the colleges included in the sample are not stated.

> During the 1966-67 academic year all students enrolled in engineering technology at the

> University Community and Technical College of the University of Toledo were required to take a Cooperative Mathematics Test [CAT II] for placement purposes. Students that received a raw score of fifteen or greater on this particular test were permitted to take the introductory course [college algebra]. About 50 percent of those students ... received grades of A, B, or C. The remaining students received grades of D, F, or Drop. This survival-attrition rate dictated the necessity of identifying those students that required remedial work prior to entering the basic course.
>
> The arbitrary cutoff point ... failed to provide adequate discrimination. In retrospect it was concluded that a discriminant equation would be required in order to provide the desired discrimination.78

"Success" was defined as a grade of A, B or C in the college algebra course; the "unsuccessful" group consisted of students receiving a grade of D, F or Drop.[79] The independent variables employed were "(1) score on the Cooperative Mathematics Test (CAT II); (2) the number of years of academic high school mathematics; (3) the mean grade point in high school mathematics; and (4) age, in months beyond the seventeenth birthday."[80]

When the discriminant function obtained was applied to the original sample of 50 students, it was found to be "accurate in 90 percent of the cases," which is scarcely surprising.[81] Like Wick and Howlett, Morgan found that "a significant loss ensues when any of these variables are deleted."[82]

> At Chattanooga State Technical Institute
>
> a formerly used standardized test provided no significant correlation between the test score and a student's success in mathematics ... as a result the failure rate was high for these courses, averaging about 50% of the enrollment.
>
> Hoping to better place students in their initial mathematics course, the mathematics department designed a thirty-six question multiple-choice diagnostic exam including some of the basic concepts of arithmetic, algebra, and trigonometry. ... Class changes were recommended based on rough guidelines established by the department.
>
> Based on the known statistics of the diagnostic exam and the Otis I. Q. score, the problem is to determine the appropriate level for an entering student in mathematics at C. S. T. I. so that the probability of success in his initial mathematics course will be at least 0.7. Success in a course is defined as a C or better...83
>
> A multiple regression analysis was used (1) to establish an equation to determine the lower limit of the prediction interval, (2) to determine which of the variables, I. Q. or mathematics placement score, has more of an influence on a student's grade ... and (3) to determine how much of the variance in grade can be explained by the I. Q. and mathematics placement exam...84

The sample consisted of students "who had both a diagnostic exam grade and an Otis I. Q. score during the three quarter sequence" 1970-71.[85]

> It was concluded that the diagnostic exam was more significant for predicting purposes than the I. Q. score. The linear models ... were both statistically significant as were the individual variables of diagnostic exam and I. Q. score. ... Finally, the explained variance suggested that there was considerable room for improvement in the prediction.86

Hooper therefore suggested that "additional variables such as high school grade point average, high school mathematics grade point average, number of working hours per week ... need to be examined for possible use in the prediction equation."[87]

"To add to the knowledge about remedial programs in mathematics," Edwards undertook a

> study to determine those factors from a selected domain (of ten possible predictors) which were the best predictors of success in remedial mathematics courses in the public community junior college and to develop a regression equation based upon those predictors.
>
> A sample of 359 remedial students enrolled in seven public community colleges in the fall 1970 semester provided data for the study - six colleges were in Connecticut and one in Massachusetts. In the regression analysis, the following independent variables were identified: scores from the CEEB Comparative Guidance and Placement battery..., high school average, number of class hours for which registered, number of credit hours given in remedial mathematics, an attitude toward mathematics score from the Dutton Test, work status while attending college, and sex. The criterion variable was a dichotomous variable based on grades in remedial mathematics courses - A, B, C, or satisfactory for success and D, F, W, or unsatisfactory for failure. The total sample was divided into two random samples (main and cross-validation) stratified by college.88
>
> A regression equation was determined;
>
> When a cutoff for success and failure in remedial mathematics courses was set at 0.60 (for the computed criterion or Y-value), the regression equation led to correct predictions 71 percent of the time (for the cross-validation sample..)...

> It was concluded that prediction of success in remedial mathematics courses can be made correctly 71 percent of the time using five select predictors: high school average, mathematics test score.., attitude toward mathematics score..., sentence test score.., and mathematics interest score...[89]

Reilly employed "the discriminant analysis method ... to combine a variety of test and self-report data in describing ideal groups in remedial and regular mathematics and English courses."[90]

The original subjects were students from four junior colleges enrolled in regular and remedial courses in 1967; a cross-validation sample was drawn in 1968 from the same sources.[91]

The predictors were "selected tests of achievement, ability, and interest from the Comparative Guidance and Placement Program (CGP) Battery."[92]

Data were analyzed separately for each school; the classification rules developed for the original sample were then applied to the cross-validation group.[93] The percentage of "hits" - a "predicted assignment [which] fell into either one of the actual criterion categories" - ranged from 46.7 to 77.2 for the cross-validation sample.[94]

> The overall agreement among the schools in each sample with respect to the relative weights given each predictor variable in general was not good... This lack of agreement suggests that courses labeled remedial and regular in one junior college may be quite different from their respective counterparts in another school. It also

> follows that a test found to be a highly valid placement device for one school may be useless in another program.95

Reilly concluded that "the results tend to support the use of multivariate statistical methods in making placement decisions."[96]

At Southern West Virginia Community College, where "ACT is used for placement," Nolan found a correlation coefficient of only 0.2627 between ACTM score and grade earned in College Math I.[97] As a result of this and other low correlations between "ACT Subtest scores and grades for the corresponding academic [subjects]," he recommended that use of the ACT be discontinued.[98] The sample consisted of 22 students who "had ACT results on file and who had completed the coursework with a letter grade other than W (withdrawal) or I (incomplete)..."[99]

"A study was done at MSU on the backgrounds of students taking ... [certain] math courses [in] fall term 1977..."[100] Lappan and Phillips

> concluded that the number of years of high school math together with the high [school] math grade point is a good predictor of whether a student starts with remedial math or with a pre-calculus or calculus course. A student with at least 3 years of high school math (taken at the college preparatory level) with math grades of 3.02 [sic] and overall grade point of 3.19 will probably <u>not</u> start with a remedial math course.101

These are not the criteria used to place students in mathematics courses at Michigan State University; a local algebra placement examination (there are several forms) is administered "to determine the appropriate math course for all freshmen. The placement exam roughly [sic] required one to one and one half years of H. S. algebra."[102]

As high school grades, high school mathematics grades and type of high school mathematics classes pursued appear to be important factors in predicting success in mathematics, it might be expected that the faculty advisors at LSSC, to whom both high-school transcripts and ACT scores were available, would be more capable of selecting a mathematics course for freshmen than an algorithm based on a test score or two.

Certainly Kurtz believed that "the adviser in mathematics is in a better position ... to properly make this recommendation."[103] He saw "lack of available time and records," which "could be rectified," as "the two main objections to advisement by the mathematics advisers."[104] He assumes that the advising would be superior to (or no worse than) any provided by a professional counseling bureau; no evidence is furnished for this assumption.

However, some plausibility may be lent to this position by Foshay, Schwen and Namy who, in studying instructional development at Indiana University, found that the judgment of students' full-time academic advisors accounts for variance independently of Scholastic Aptitude Test (SAT) and GPA when predicting final course grade.[105] It was also "inferred that part of the valid variance in advisor judgment was gained in face-to-face meetings."[106]

Moreover, the Supervisor of Placement and Proficiency Testing at the University of Illinois, Urbana, warns that

> Expectancy data alone are probably not sufficient information for course selection. ... Teachers vary according to their teaching style... Students vary in terms of learning rates, styles, and subject-matter interest. ...An attempt to match teacher and student variables and merge them into the student's own academic context may provide more useful information than course-grade expectancies in any form.[107]

## PREDICTION OF SUCCESS IN COLLEGE

Although the prediction of college success is a more general problem - and apparently a more difficult one[108] - than that of predicting success in a single subject, such as mathematics, there are some parallels that justify a brief review. Additionally, at least one comparative experiment very similar to the core study of this dissertatation has been done.

Five hundred and eighty studies of college achievement, made between 1949 and 1959, were located by Fishman and Pasanella.[109] More than 90 percent used intellective criteria (college grades) as the dependent variable. Three-quarters of these investigated intellective predictors, e.g., high-school grades and test scores.[110]

"The most obvious intellective predictor is the high-school record, usually expressed as total average grade or rank in class ... this measure correlated roughly .50 with comprehensive freshman-year intellective criteria."[111]

Studies employing the "usual combination" of "an aptitude test plus high school record" produced multiple correlations ranging from 0.31 to 0.82.[112]

> In general, the use of any one intellective predictor, or more than one, with the high-school record improved the forecast of freshman average in 181 studies by .00 to .38, with an average gain of .11. It seems useless, however, to employ more than two or three intellective predictors, from both the point of view of practicality and that of efficiency.[113]
>
> Few studies came to the point of combining intellective and non-intellective predictors by means of multiple-correlation techniques. Where this was done, the gain in multiple correlation attributable to the nonintellective predictor was discouragingly small. As a result, much of the literature on nonintellective predictors dealt with attempts to improve their technical and theoretical foundations.114

A later and lengthier review of the literature by Lavin substantiates these findings.[115]

Loeb provided further verification in 1972, but pointed out that test scores may prove a more reliable predictor for disadvantaged students than high school rank.[116] She also emphasized the importance of continuing "routine validity studies."[117]

Bloom and Peters pointed out that by "considering institutional variation," "there is almost as much consistency between high school grades and college grades as there is among grades within a single institution," with correlation coefficients of 0.72 to 0.75.[118] The use of academic prediction scales entailing "statistical corrections for intraschool and interschool nonequivalence of grades" to improve prediction of college achievement is recommended.[119]

This brief review reveals a number of similarities to the literature on prediction of success in an initial college mathematics course. High school records and intellective tests are the best predictors; multivariate methods can improve prediction significantly. Predictability is reduced by differences among and within institutions, although adjustments may be made for this. The majority of the investigations consist of "one-shot case studies." However, comparative experiments have also been done and some are described in the next section.

## CLINICAL VERSUS STATISTICAL PREDICTION[120]

Sarbin suggested a test of the hypothesis that "A complete case study will increase the accuracy of prediction of behavior over that obtained from the use of statistical tables based on experience with relatively few variables."[121]

He then laid down a set of rules for such a test:

> (a) the criterion to be predicted must be subject to definition and measurement; (b) experience tables or regression equations must be available beforehand in order to make statistical predictions; (c) the individuals whose behavior is to be predicted must have had at least one clinical interview; and (d),[sic] in addition to the statistically determined variables, other data which are presumably associated with the criterion must be made available to the clinician.[122]

and designed an experiment to test the more specific hypothesis that

> By virtue of the case-study method employed, clinical counselors' predictions of academic success will be more accurate than those determined from regression equations.[123]

The subjects were 162 freshmen "who matriculated in the fall of 1939 in the arts college of the University of Minnesota."[124] No mention is made of the selection process.

Five counselors participated in the experiment; four "possessed the Ph. D. degree or its equivalent. All five had considerable experience in clinical counseling work with university students."[125]

Previously derived regression equations were available

> in which academic achievement is predicted by a combination of two variables: rank in high-school graduating class and college

> aptitude test. All subjects had one interview with the clinician prior to exposure to college classes. Data available to the predictors in addition to the measurement variables were in the form of additional tests of aptitude, achievement, vocational interest, and personality; an eight-page individual record form; a preliminary interviewer's form and impressions; and, finally, the counselor's own observations.[126]

The correlation coefficients between the clinical predictions and actual first-quarter GPA were 0.35 for the 73 men and 0.69 for the 89 women in the sample; the corresponding figures for the statistical method are 0.45 and 0.70. The differences between the two methods were not statistically significant.[127] "At any rate, we can draw the conclusion that the clinical counselors in this study did not predict college achievement more accurately than did the statistical method."[128]

Sarbin makes further intriguing points; three are of especial interest. First, the counselors tended to "overestimate the contribution of the two measurement variables [high-school rank and college aptitude test]."[129] Second, "the interviewers over-estimate the college achievement of the group studied."[130] Third, when a chi-square analysis was applied to the hypothesis that counselors can better predict "success or failure" - "where success is considered to be a 'C' average or better and failure less than a 'C' average" - the clinical predictions were not significantly different from the statistical predictions "for men, while for women the chi-square value shows a doubtful difference [at the 0.05 critical level] - favoring the actuarial predictions in regard to accuracy."[131]

> One of the first areas to be investigated by clinical psychologists, as the profession grew rapidly after World War II, was the degree to which human judgment could be used in the prediction of variables such as patient response to treatment, recidivism, or academic success. What could such judgment add to prediction that could be made on a purely statistical basis by, for example, developing linear regression equations? The statistical analysis was thought to provide a floor to which the judgment of the experienced clinician could be compared.
>
> The floor turned out to be a ceiling. Meehl (1954) reviewed approximately 20 studies in which actuarial methods were pitted against the judgments of the clinician; in all cases the actuarial method won the contest or the two methods tied. Since the publication of Meehl's book, there has been a plethora of additional studies directed toward the question of whether clinical judgment is inferior to actuarial prediction, and some of these studies have been quite extensive. But Meehl was able to conclude, some 10 years after his book was published, that there was only a single example in the literature showing clinical judgment to be superior, and this conclusion was immediately disputed by Goldberg on the grounds that even that example did not show such superiority. We know of no examples after that (within the standard limitations) that have purported to show the superiority of clinical judgment.132

What are the reasons for this outcome? Meehl suggests that

> there is a major pragmatic difference between the predictive demands made upon the clinician during therapy and those made in the purely prognostic setting. All of us expect a certain

> amount of blind-alley hypothesizing to occur in the course of a therapeutic series. Therapists form transitory hypotheses of extreme tentativity and often may not follow them up by so much as a leading question unless additional support subsequently appears in the client's spontaneous productions. ... Nobody knows what the payoff rate is for these moment-to-moment guesses that come to therapists; but the over-all success frequency might be considerably less than 50 per cent and still justify the guessing, for ... the time spent on exploration of poor guesses need not greatly detract from the positive contribution of successful ones. Presumably even the unsuccessful paths are rarely pure waste, since they contribute to such diverse concurrent aims as further getting acquainted, general desensitization, and incidental support for quite unrelated constructions...
>
> When we move over into the straight prediction situation, all this is radically changed. Here, no erroneous weighting is filler in the above harmless sense, because statistical filler is error variance. ... So that in the straight prediction setting, all bad ideas tend to subtract from the power of good ones.[133]

Meehl points out that the studies he has reviewed "all involve the prediction of a somewhat heterogeneous, crude, socially defined behavior outcome," rather than the "relatively specific and concrete predictions and postdictions" that might yield an advantage to the clinician.[134]

Dawes and Corrigan report that linear models are especially effective for situations

> in which: (a) the predictor variables have conditionally monotone relationships to criteria...; (b) there is error in the dependent variable; (c) there is error in the independent variables; and (d) deviations from optimal weighting do not make much practical difference. ... Thus the situation

> demands decision-making behavior approximately like that of a linear model if the decision-making is to be appropriate...135

So very appropriate are linear models in such contexts that random linear models - models whose weights are chosen randomly except for their signs, which are taken to have a positive relationship with the criterion - can outperform human judges on such tasks as the prediction of GPA.[136]

## SUMMARY

Reviews in the Mental Measurements Yearbooks reveal that the ACTM and CAT II are reasonably reliable and valid instruments, though not without weaknesses. Many of the correlation coefficients between each of the three tests under consideration - the ACTM, CAT II and TPT - and mathematics grade at LSSC, obtained by the American College Testing Service, are not impressive; however, not much meaning can be attached to them because of methodological difficulties.

A large number of studies exist which attempt to predict success in beginning college mathematics. The prediction of success in college algebra, and the problem of distinguishing students to be placed at that level from those who should do remedial work, occur most frequently in the literature.

The prevailing criterion variable is mathematics course grade, either utilized on a 4.0 scale or dichotomized for "success" and "failure." The use of student-supplied data and of instructors' rankings of students are also reported. A "success" rate of 70 percent appears to be a reasonable goal for mathematics placement.

Linear regression, multiple regression and discriminant analysis comprise the chosen methodologies. Multivariate procedures usually, but not always, have significantly greater predictive ability than those employing a single independent variable.

The factors found to correlate most highly with success in college mathematics are high school grades and scores on mathematics placement examinations. Less highly correlated are tests of other skills, tests of general academic ability, personality factors and sex. The form in which high school grades are utilized varies; class rank, GPA and mathematics GPA all occur. Nor does any test of mathematical ability emerge as clearly superior.

Most, but not all, studies of the ACTM and CAT II show that they have good predictive validity. A number of investigations of locally-written mathematics placement examinations may be found. Usually, not invariably, good results for these tests are reported;

however, the test has probably been the placement device employed in those instances.

No experimental comparisons of such instruments with standardized tests have been found; in fact, in the field of mathematics placement, no comparative experiments of any kind have been encountered.

Many of the "one-shot case studies" are flawed by one or more of the following disadvantages:

1. Failure to cross-validate results.
2. Samples that have not been randomly drawn from the population observed. The phrase "for whom records were available" occurs frequently.
3. Omission of the methods and standards used to place students in a beginning mathematics class. In a number of studies, the apparent superiority of one testing instrument is probably due to its use as the original placement device. This is not always stated.
4. Equating beginning courses that are not equivalent (either within the same institution or among diverse colleges or both).

One mathematics professor, who is probably speaking for many, argues that members of mathematics faculties, rather than professional counselors, should place students in beginning mathematics courses. There is some evidence that advisors' judgments can contribute to the prediction of grades, above and beyond the donation of the "usual predictors."

The literature of predicting success in college bears a number of similarities to that of forecasting success in mathematics. Nearly all the studies use college GPA in some form as the criterion variable; virtually all investigate "intellective" predictors. High school grades are unquestionably the best predictor of college GPA; scholastic aptitude (or achievement)[137] tests are next. Significant increases in correlation are obtained with multivariate analysis in most cases.

Some comparative experiments have been carried out to test the hypothesis that human judges can predict college GPA more accurately than can statistical methods. The results of these and analogous experiments in other areas show "no significant differences" or favor the actuarial approach. However, the present experiment differs in design from the "clinical versus statistical" studies, so that these outcomes may not be taken as conclusive.

It may be concluded that a comparative experiment would provide insight not previously obtained into the relative effectiveness of different methods of mathematics placement.

# CHAPTER III

## DESIGN AND METHODOLOGY

This chapter consists of four major parts. First, the abstract design of the experiment is presented, together with a brief discussion of confounding variables.

The details of implementation are described next: the participants' actions in July and September, 1980; the precautions taken to avoid a "Hawthorne effect;" the way in which subjects were randomly assigned to treatment groups and the methods used for other arrangements; and the particulars of the test-norming process.

Measurement is the third topic. The independent variables are listed. Then, because no universally accepted means of measuring "effective mathematics placement" has been found, a discussion of the difficulties involved in choosing criterion variables and a justification of those chosen accompanies their description.

Finally, the statistical methods and the variables and tables chosen for analysis will be specified.

A summary concludes the chapter.

## DESIGN OF THE EXPERIMENT

The experiment is an elaboration of the "Posttest-Only Control Group Design" described in Campbell and Stanley's Experimental and Quasi-Experimental Designs for Research:

| | | |
|---|---|---|
| R | X | $O_1$ |
| R | | $O_2$ |

where X represents the exposure of a group to an experimental variable or event, O refers to some process of observation or measurement, and R indicates random assignment of subjects to treatment groups.[138]

In this case, the design can be represented as:

| | | |
|---|---|---|
| R | $X_1$ | $O_1$ |
| R | $X_2$ | $O_2$ |
| R | $X_3$ | $O_3$ |

$X_1$ is mathematics placement based on the student's score on the Thesing Placement Test; $X_2$ is placement based on the two scores: Cooperative Algebra II Test and ACT Mathematics; $X_3$ is placement based on interviews with members of the mathematics faculty at Lake Superior State College. The observations are fall term mathematics grades, course level and instructors' judgments about the suitability of courses to students.

Largely because of the random assignment of subjects to treatment groups, Campbell and Stanley's well-known sources of internal invalidity - history, maturation, testing, instrumentation, statistical regression, selection, experimental mortality and selection-maturation interaction - are controlled by this design.[139]

Some factors that may jeopardize the external validity, or representativeness of an experiment are: the reactive or interaction effect of testing produced by a pretest; the interaction of selection biases and the experimental variable; reactive effects of experimental arrangements; and multiple-treatment interference.[140] Neither the first nor the last is applicable here: there was no pretest and each subject received only one treatment.

Experimental procedures were designed with care to prevent a "Hawthorne effect," i.e., reactive arrangements, and will be described. Almost certainly, however, there are characteristics of Lake Superior State College that interact with treatment to limit the generalizability of the results and these will be considered in Chapter Five.

# EXPERIMENTAL PROCEDURES

## General Description

Almost all entering freshmen at Lake Superior State College attend one of four "pre-orientation" sessions. Three are held in July and one in September. On the second morning of the July sessions, reading and mathematics placement tests have traditionally been given. The tests are scored by Counseling Center personnel and a summary report for all students, indicating their English and mathematics placement, is complied, reproduced during the noon hour, and then distributed to major field advisors. These are usually the department heads, but other faculty may share the advising task.

After lunch, the students attend a talk on course scheduling delivered by the Registrar of the college; then they disperse to meet with their advisors and schedule courses. The September pre-orientation differs in that only testing is done and that in the evening, and students have the next several days to schedule courses in consultation with their advisors.

A day or two before testing in 1980, members of each pre-orientation group were randomly assigned to one of three treatments ($T_i$):

$T_1$: This group was placed in mathematics classes on the basis of performance on the revised version of the Thesing Placement Test (TPT). Norms to establish cutoffs for various levels of mathematics were obtained from an examination of student performance in fall term 1979 mathematics courses.

$T_2$: This group was placed on the basis of scores on the ACTM and CAT II. Norms to determine cutoffs were derived from fall term mathematics grades for 1975, 1976 and 1977 entering freshmen and showed approximately the same percentage of students passing at each level as did the TPT in 1979.

$T_3$: Members of this group who planned to take a mathematics course during the 1980-81 school year - about half the group - met one-to-one with members of the LSSC mathematics faculty, who recommended a suitable course. Two to four professors advised at each session; fifteen minutes were allocated for each consultation. Students in this group for whom mathematics was neither required nor an elective received a "placebo treatment:" a talk, accompanied by a handout, about the beginning mathematics courses at LSSC and the procedures to follow should they ever wish to take mathematics there.

The curriculum choice of each subject in $T_3$ was ascertained. If mathematics was required in that major, the subject was assigned to a mathematics professor for advising; otherwise, the student was placed in the "placebo" group.

A form was then filled out for each student, indicating the individual's name, room assigned for mathematics placement, advisor's or test administrator's name, time of appointment, and a caveat about placement satisfaction.

FIGURE 1

MATHEMATICS PLACEMENT FORM

Name ______________________________________
First Middle Last

Room Assigned for Math Placement _____(Crawford Hall)

_______________________________________________

Time: ______________________________________

If you are dissatisfied with your English or mathematics placement, be sure to discuss this with your major academic department advisor.

File material - the high school transcript, ACT profile, and front page of the LSSC application form for admission (a copy of that document is reproduced in Appendix A) - for students who had appointments with mathematics professors was duplicated and delivered to that advisor's mailbox.

The forms were arranged alphabetically and, on the day testing was scheduled, given to students as they entered Crawford Hall. Two Counseling Center staff members and the author sat under signs labeled "A-K," "L-P," "Q-Z;" students collected their forms from the appropriate person. Those who showed up without having pre-registered were sent to the auditorium where Mr. John Truckey, Director of Counseling and Testing at LSSC, asked them to inscribe their names on blank forms and return them immediately.

While the subjects took the Cooperative Reading Test in the auditorium, the author randomly assigned latecomers to treatment groups and completed the forms Mr. Truckey had obtained from them. These were returned at the conclusion of the reading test, when all students were told to proceed to mathematics placement at the time indicated on their forms. Meanwhile, the Admissions Office was alerted to the need for file material on last-minute arrivals who'd been assigned to mathematics professors. The records were immediately duplicated and delivered. Signs in the corridors directed students to the rooms used for placement; in addition to the auditorium, these were: 204, 309, 315 and the faculty offices. All are in Crawford Hall of Science.

Administration of the TPT and CAT II followed a "standard operating procedure:" students were asked to read the test's instructions and to attempt each problem. They were assured that the test was for the college's information only. When the allotted time - 40 minutes for the CAT II, 60 for the TPT - was over, tests were collected and students sent back to the auditorium to collect their college catalogs and attend a student forum on college orientation.

Implementation of $T_3$ was more complicated. An outline for advising students had been drawn up by Dr. Thesing and distributed to the mathematics faculty; this is reproduced in Appendix A. The advisors also received a supply of forms for recording course recommendations; a copy of one is in Appendix A.

Students whose placement slips bore appointments with mathematics professors simply went to the instructor's office at the appointed time. Occasionally, one or two would not read the time carefully (or at all), but these problems were easily corrected. A few students were, necessarily, scheduled for conferences at the same time as the college orientation. Professors were supplied with catalogs for these individuals, who were sent on to lunch after the appointment.

The "placebo" talk was scheduled to begin at the same time as the mathematics tests. The administrator began with a personal introduction and then asked if anyone had made a new curriculum choice or was considering a mathematics elective. These individuals were asked to remain in the room after the meeting so that placement could be arranged without delay.

A handout describing beginning mathematics courses and placement procedures was then distributed; it is reproduced in Appendix A. Each course was characterized in terms of its level and its suitability to various curricula. Mathematics placement procedures were outlined and questions invited.

Students needing placement were then named and again asked to stay; the others were dismissed. A number of students expressed interest in a mathematics elective only after the talk had been delivered, which was an unexpected outcome.

A conference with a faculty member was arranged for students needing placement. When their placement forms had been suitably modified, the administrator called the Admissions Office to arrange for transmission of file material to the mathematics advisor. The faculty members were then notified of their new appointments.

When testing was completed, the treatment administrators went on to the Counseling Center to assist with scoring. Once scored, tests were arranged in alphabetical order, as were the forms generated by the individual conferences. ACT profiles for members of the pre-orientation group had been pulled from the files, and all this information was used to assemble the placement roster for English, developmental reading and mathematics. Upon completion, the roster was duplicated; copies were delivered to each department in the college by 1:30 p.m.

## Avoiding Reactive Arrangements

Campbell and Stanley warn that

> In the usual psychological experiment, if not in educational research, a most prominent source of unrepresentativeness is the patent artificiality of the experimental setting and the student's knowledge that he is participating in an experiment. ... The play-acting, outguessing, up-for-inspection, I'm-a-guinea-pig, or whatever attitudes so generated are unrepresentative of the school setting, and seem to be qualifiers of the effect of X [treatment], seriously hampering generalization.141

While treatment occurred at a "plausible period in the curriculum calendar" and the observations were "similarly embedded as regular events" (and at a good distance in time - 10 to 19 weeks - from treatment), as suggested by Campbell and Stanley,[142] it seemed nevertheless reasonable that a reactive, or "Hawthorne effect," could occur if prospective students were told that they were subjects in an experiment on mathematics placement. For example, those who conferred with mathematics professors might feel singled out in some especial way and consequently make greater efforts in the mathematics courses they undertook. Others might resent taking a standardized test - "I never do very well on those" - and insist on seeing a faculty member also, thus upsetting the random assignment of subjects to treatment groups.

On the other hand, it appeared that there could be no unfairness involved in leaving the subjects unaware of the existence of the experiment. Had there been no experiment, the TPT would have been used to place all students in mathematics courses; clearly, then, it was reasonable to place about one-third of them that way. Because the standardized tests had been used for mathematics placement at LSSC for a long period of time, and because the profile of successful placement was very like that of the TPT in 1979, this method was also deemed equitable. Moreover, norms for the standardized tests had already been cross-validated, while the 1980 sample was to be cross-validating for the TPT. Finally, the experience and interest of the LSSC mathematics faculty members in advising students judtified that method of placement. As in past years, students who were dissatisfied with their mathematics placement, or otherwise presented a placement problem, were encouraged to seek advice from members of the mathematics teaching staff before making a final decision.

Prospective students, then, were not told that an experiment was in progress. At each pre-orientation session, after the subjects had found seats in the Crawford Hall Auditorium and before the Cooperative Reading Test was administered, Mr. Truckey introduced himself and his assistants. He then said something like, "We're a small college and so we're able to do more individualizing and small-group advising in the mathematics area. So after you finish the reading test we're going to have now, you will go for your math advising as indicated on the slip you have."

Reasoning similar to that applied in the case of students suggested that no useful purpose would be served by advertising the experiment to the college faculty, especially the major field advisors. Therefore, few faculty members were told about the research. Some obvious exceptions were the mathematics teachers and Professor C. Ernest Kemp, Head of the Department of Earth Science, Mathematics and Physics. The only other faculty members to be informed were Dr. David Behmer and Dr. Susan Ratwik, who assisted with the methodology of the experiment. Neither was a faculty advisor to students attending pre-orientation. Some college administrators were informed; cooperation from a number of them was needed and was most generously given. However, where no one needed to know, no one was apprised, and the project was carried out very quietly.

It was almost inevitable that someone would figure out that different freshmen were getting different forms of mathematics placement. A memorandum to those involved suggested that any inquiries be met in the following way:

> We do math placement in smaller groups here at LSSC, as we've found that works out better. We want as many students as possible to see math faculty members, but because of summer schedules, meetings out of town, and so on, it isn't possible for everyone to see a math professor today. However, the department head will be available this afternoon, if you want to talk with him. Also, if you're still uncertain about your math class, even if you're already scheduled for it, we'd encourage you to get advice from one of the math professors first thing this fall.

A variation for department heads was,

> We wanted to get all students in to see the math faculty this summer, but because of summer schedules, it just wasn't possible. But Ernie [Professor Kemp] will be in this afternoon, if you want to send a student over, and of course all the faculty will be here in the fall and any student is welcome to come and discuss placement then.[143]

The author was approached once or twice by students who wondered why one was seeing a professor and another wasn't, and why all times scheduled were not the same. The "formula" of the memorandum was successful and when asked, "Would you like to see a mathematics professor?" these individuals rejected the opportunity.

The format of the placement roster presented a few problems in the experimental situation. In previous years, the roster bore the students' relevant test scores and a covering memorandum laid down the guidelines to be followed by major field advisors in interpreting these scores for mathematics placement. It was essential that instructions should not reveal the variation in placement techniques. Dr. Thesing and the author had identified six useful categories for mathematics placement (see Table 3):

1. Strongly recommend Basic Mathematics;
2. Recommend Basic Mathematics;
3. Strongly recommend Intermediate Algebra;
4. Recommend Intermediate Algebra;
5. Recommend college algebra;
6. Recommend Calculus with Analytic Geometry I if the student has an adequate trigonometry preparation.

Mathematics placement coded by a number between one and six, inclusive, with the meaning described in the companion memorandum, seemed to suit the purpose admirably. But almost immediately there was a stumbling block: mathematics advisors would be recommending a specific course, rather than a general denomination. And while those recommendations could have been coded, so that an "MA 181" became a "5," and "Basic Math" transformed to "1," that would have eliminated one of the advantages of individual advising: the ability of a professor to determine a student's curriculum choice and identify the appropriate course exactly. In some cases, moreover, accuracy could be lost in translation. For example, a professor might determine that a student majoring in chemistry had sufficient background to attempt a college algebra course but lacked the ability to succeed in the rather demanding MA 121. Thus MA 181, to be followed by MA 182, might be recommended; these two courses cover the MA 121 material in a somewhat more leisurely fashion. A code of "5," however, would almost certainly cause the student's major advisor to select MA 121.

Dr. Thesing then suggested placing an asterisk (*) in the "Mathematics" heading of each sheet of the placement roster and a corresponding message at the foot of the page: "* In cases where test scores and other indicators are especially strong in their implications, a particular course has been recommended."[144]

This solved one problem, but raised another: was it possible that listing a specific course - thus relieving the major advisors of the subtleties of interpretation - could, in and of itself, improve the results of mathematics placement? Then "specific course" represented a possible source of confounding: if significant differences were found in favor of $T_3$, they might be due to this extraneous variable. If so, alternate placement methods, less expensive and easier to implement than individual advising, should be considered; e.g., employing one or two mathematics professors to examine the files of students whose test scores placed them in categories 2, 4, 5 or 6, and make explicit recommendations.

To assess the difference in placement effectiveness, if any, between "coding" and naming a course exactly, each student who scored 0-11 on the TPT (Category 1 - Basic Mathematics) or 14-22 (Category 3 - Intermediate Algebra) was randomly assigned to one of two groups. Mathematics placement was listed as "Basic Math" or "MA 100" for members of the first group, while it was coded as "1" or "3" for those in the second. A contingency table enumerating course choices for each group would be analyzed to determine whether or not these choices were independent of nomenclature.

It was then noted that some students do not submit their ACT profiles in advance and this would cause difficulties if they were in $T_2$. This was resolved by treating missing ACTM scores as if they were zero, a move which in almost all cases could lower placement no more than one level, and attaching a double asterisk (**) to the student's level number. On the memorandum accompanying the placement roster, these double asterisks were also appended to categories 1, 2, 4 and 5 (students in $T_2$ could not be assigned to levels 3 or 6 unless the ACTM was available) and a footnote explained:

> ** A double asterisk may be attached to these categories for certain students. This indicates that ACT scores have not yet been submitted. Such students should be queried about their high school background to determine placement and, if they are still uncertain about the appropriate mathematics class, should consult a member of the mathematics faculty immediately upon returning in September for a re-evaluation of placement.[145]

Doubtful of the capacity of the major advisors to absorb all these dicta just before their advisees engulfed them, Professor Kemp called on each department head well in advance of pre-orientation to explain the new rules and distribute copies of them. The rationale for a new system was the introduction of Basic Mathematics to the curriculum, which clearly rendered mathematics placement more complex. Professor Kemp did not, however, mention the single asterisk; advisors saw it for the first time on the roster.

The major field advisors' instructions for mathematics placement in 1980 are in Appendix A, together with a facsimile of a page from the placement roster.

## Random and Other Assignments

A list of about 600 random digits between one and three, inclusive, was obtained from a table of random digits[146] in the following way: a stab of the finger (with eyes closed) was used to determine a starting point; the number nine was discarded from consideration, and one plus the remainder of each digit after division by three was recorded. Thus the digits 7, 5, 3, 0, 2, 4, 9, 1, would yield the sequence 3, 3, 1, 1, 3, 2, (ignored), 2.

Postcards, provided by LSSC and indicating a student's choice of pre-orientation session, had been mailed to the Office of Admissions by most of the prospects for the sessions. While the author read from the random number list, an assistant dealt the cards for a particular session into one of three piles clearly labeled "1," "2," or "3," which identified the treatment group. Forms for mathematics placement were then completed. As numbers were used, they were, of course, discarded from consideration. This random number list was used again to designate treatments for last-minute arrivals who had not pre-registered for the session.

The treatment administrators were Mr. David

Castner, who is an LSSC counselor, Mr. Truckey and the author. Early plans had called for a random permutation of room and administrator assignments among treatment groups, with the permutations re-randomized for each pre-orientation session. It was soon clear that this would not be practicable: rooms 204, 309 and 315 are the three largest in Crawford Hall (except for the auditorium); 180 students were expected to attend the first session; and room 309 has a capacity of only 40. Additionally, the author had designed the placebo talk and wished to deliver it at the first session; Mr. Castner and Mr. Truckey were experienced in administering standardized tests and the author was not. Consequently, room 309 and the author were assigned to the $T_3$ placebo treatment for the first session; the remaining assignments were made by random permutations of the digits (1,2). Rooms were ordered numerically and administrators alphabetically for this purpose. Six random permutations of the digits (1,2,3) were obtained from the same table of random digits (this time by observing the order in which these digits appeared among the rest; a different part of the table was of course employed) and these were used to assign rooms and administrators for the remaining pre-orientation meetings. This resulted in the following arrangement:

| | $T_1$ | $T_2$ | $T_3$ (Placebo) |
|---|---|---|---|
| Session 1 | | | |
| Room | 204 | 315 | 309 |
| Administrator | Castner | Truckey | Bone |
| Session 2 | | | |
| Room | 315 | 204 | 309 |
| Administrator | Bone | Truckey | Castner |
| Session 3 | | | |
| Room | 315 | 309 | 204 |
| Administrator | Castner | Bone | Truckey |
| Session 4 | | | |
| Room | 309 | 204 | 315 |
| Administrator | Truckey | Bone | Castner[147] |

Unfortunately, further adjustments to these plans were unavoidable. The random numbers used to assign subjects to treatment groups for the first session did not fall neatly into three groups of about 60 each; instead, $T_1$ claimed the lion's share of subjects, almost half. This rendered room 204, with a capacity of 60, unsuitable for that treatment. The most reasonable resolution seemed to be stroking out the "204" on the placement slips (Figure 1) for this group and substituting "Aud" (for "Auditorium"), which was done.

A similar problem arose before the last session. It was difficult to predict how many students would be attending, because a large number had chosen other colleges by that time. As many as 200 were anticipated; about 160 actually appeared. Room 309 was too small for $T_1$ and again the auditorium was substituted. Thus there occurred a minor but systematic difference between $T_1$ and the other treatment groups: more than half of these subjects did not move from the Crawford Hall auditorium to another room for mathematics placement, while all the members of the other treatment groups did.

Random assignment of $T_3$ students to mathematics professors had also been planned and a list of random digits from one to four had been prepared for this purpose. After the first session, however, it became clear that this was unpropitious: one faculty member ended up with 11 advisees while another had only five. This meant that the latter received more last-minute advisees (who were not randomly assigned) and therefore had less time to review his students' files. Random permutations of the digits (1,2,3,4) would have solved that problem, but it would have made assignment of students attending the placebo talk and needing placement almost unbearably complicated, especially when there were other considerations to be borne in mind. Some professors served as "back-up men" for particular sessions and were entitled to go home earlier than those whose "regular" turn of duty it was. Finally, assignment of professors to pre-orientation session was not random at all, but was at the professors' convenience. Thus no inferences could be drawn about the abilities of individual advisors. Therefore, students in sessions subsequent to the first were simply "dealt out" 1-2-3 (or 1-2 or 1-2-3-4) to the faculty members advising them.

There were no difficulties in dividing students scoring 0-11 or 14-22 on the TPT into two groups, one with a specific course listed on the placement roster and the other with codes of "1" or "3." A list of random ones and twos was derived from a previously untouched (for the purposes of this experiment) portion of the random digits table by observing whether each such digit was even or odd - yielding a value of zero or one, respectively - and adding one. Courses for the first group were named and those for the second were coded.

## Norming the TPT and the ACTM-CAT II Combination[148]

Establishing norms for the TPT was not difficult. A student achieving a grade of A, B, C, "Credit," or "Credit by examination" had been defined as "successfully placed" in a mathematics course; others were not so considered. The choice of this criterion, which is neither universal nor indisputable, will be aired in the next section. As it had been chosen and accepted, however, it was natural to use it as the basis for cutoff scores.

Because the TPT had been revised just prior to its use in 1979, only fall term mathematics grades for that year could be utilized. Data were obtained by comparing the Counseling Center's list of orientation test scores with the Registrar's fall quarter 1979 grade verification lists. The data suggested that the placement guidelines of Table 2 be revised if the percentage of students passing courses at each level of mathematics was to be maximized.[149] Employing this principle, Dr. Thesing and the author obtained the criteria of Table 3. A small exception to the "Maximization of percentage of passing students" principle was the standard set for calculus: Dr. Thesing thought that a borderline category was inadvisable here and that a cutoff score of 34 was fairer to students than one of 36.

It appeared that a student's odds for success were three out of four or better when a course was definitely recommended; they were about three out of five for students in the borderline categories choosing the higher of the two levels of mathematics; and they ranged from zero to two out of five for students taking a course more difficult than the recommended one. Table 4 summarizes this information by course choice for the various levels of placement.

Data for 1977 were examined at the same time as those for 1979. The Registrar and the Director of

TABLE 4

VALIDITY DATA FOR THE THESING PLACEMENT TEST (1979)

| Placement Category | Percent Successful | Percent Unsuccessful | Total Number of Observations |
|---|---|---|---|
| | MA 100 Intermediate Algebra | | |
| 1 | 39 | 61 | 33 |
| 2 | 59 | 41 | 17 |
| 3 | 84 | 16 | 75 |
| 4 | 87.5 | 12.5 | 8 |
| 3 and 4 | 84 | 16 | 83 |
| | MA 111, MA 121, MA 181 - College Algebra | | |
| 3 or below | 28 | 72 | 18 |
| 4 | 62.5 | 37.5 | 24 |
| 5 | 77.5 | 22.5 | 40 |
| 6* | 100 | 0 | 22 |
| 5 and 6* | 85 | 15 | 62 |
| | MA 132 Calculus with Analytic Geometry I | | |
| 5 | 0 | 100 | 2 |
| 6 | 77 | 23 | 30 |

*Presumably students whose trigonometry background was insufficient for calculus.

Admissions at LSSC had agreed that there were no pronounced characteristics that made any one of the freshmen classes 1973-77 atypical, and it was natural to begin with the closest chronologically.

The CAT II appeared to predict success at the college algebra level about as well as did the TPT; the ACTM, not quite so well. Possibly because there was only one instructor for both MA 100, where one man taught both sections, and MA 132, which consisted of a single section, little emerged in the nature of a relationship between test scores and success in those classes. This was especially true of the ACTC, which was then dropped from consideration.

It appeared that the TPT would set the standard for placement: norms for the standardized tests must be found yielding nearly the same odds for success or they could not be used. The 1977 data were insufficient to establish such norms for the elementary classes or calculus. Nineteen seventy-seven and 1976 outcomes for these courses were then merged to obtain norms which were cross-validated on 1975 information; 1976 data were used to cross-validate the cutoff scores suggested by the 1977 college algebra results.

The criteria established for placement at various levels are given in Table 5. Tables 6, 7 and 8 display, by course level, the percentages of successful and

TABLE 5

GUIDELINES FOR USE OF THE ACTM AND CAT II IN 1980

| Placement Category | Test Scores ACTM | | CAT II |
|---|---|---|---|
| 1 | 22 or less | AND | 19 or less |
| 2 | 23 - 64 | OR* | 20 - 49 |
| 3 | 23 - 64 | AND | 20 - 49 |
| 4 | 65 or more | OR* | 50 or more** |
| 5 | 65 or more | AND | 50 or more*** |
| | | or | |
| | Unavailable | AND | 85 or more |
| 6 | 85 or more | AND | 85 or more |
| | | or | |
| | 70 or more | AND | 90 or more |
| | | or | |
| | 90 or more | AND | 70 or more |

*Exclusive or

**But not eligible for Category 5

***But not eligible for Category 6

unsuccessful students meeting these standards and the total number of observations. The year or years used to establish the cutoff scores are shown first; then the cross-validating year; finally, the total results for all years.

A comparison of the 1979 TPT results with the data from the combined years for the standardized tests (abbreviated ST) is given in Table 9. The evidence indicates that students taking courses at the recommended level of placement - 3 or 4 for MA 100, 5 (or 6 without trigonometry) for college algebra, 6 for calculus - stand approximately equal chances of passing, whether the TPT or the standardized instruments are used. The penalties for ignoring placement advice - taking MA 100 at level 1 or 2, college algebra at 3 or 4, or calculus at level 5 - are somewhat less severe when the standardized tests are employed. This is markedly true for calculus; however, one would expect that as more observations become available, the results will be moderated.

An obvious difference between the two prescripts does appear when the "Total Number of Observations" column is examined. Of the 133 subjects taking MA 100 in 1979, only 17, or 13 percent, were in the borderline category 2, while the corresponding figures are 109 out of 301, or 36 percent for the standardized tests.

TABLE 6

VALIDITY DATA FOR STANDARDIZED TESTS (1975-77)
MA 100 INTERMEDIATE ALGEBRA

| Placement Category | Percent Successful | Percent Unsuccessful | Total Number of Observations |
|---|---|---|---|
| | 1976 and 1977 | | |
| 1* | 48 | 52 | 101 |
| 2 | 63 | 37 | 63 |
| 3 | 83 | 17 | 18 |
| 4 | 87.5 | 12.5 | 8 |
| 3 and 4 | 85 | 15 | 26 |
| | 1975 (Cross-Validating Year) | | |
| 1** | 51 | 49 | 39 |
| 2 | 67 | 33 | 46 |
| 3 | 82 | 18 | 17 |
| 4 | 89 | 11 | 9 |
| 3 and 4 | 85 | 15 | 26 |
| | 1975, 1976 and 1977 | | |
| 1*** | 49 | 51 | 140 |
| 2 | 65 | 35 | 109 |
| 3 | 83 | 17 | 35 |
| 4 | 88 | 12 | 17 |
| 3 and 4 | 85 | 15 | 52 |

*If students for whom one test score was missing (usually the ACTM, but not always) are omitted from consideration, these percentages become: successful - 41%; unsuccessful - 59%. There were 28 such students in category 1 (MA 100, 1976 and 1977), three in category 2 and none in the others. This pattern appears consistently in the three years studied: a fairly large number of missing scores in category 1 and a substantial swing away from "success" when these observations are not included in the computations.

**Omitting subjects for whom one test score is missing: successful - 43%; unsuccessful - 57%.

***Omitting subjects for whom one test score is missing: successful - 42%; unsuccessful - 58%.

TABLE 7

VALIDITY DATA FOR STANDARDIZED TESTS (1976 AND 1977)
MA 111, MA 121, MA 181 - COLLEGE ALGEBRA

| Placement Category | Percent Successful | Percent Unsuccessful | Total Number of Observations |
|---|---|---|---|
| | | 1977 | |
| 3 | 33 | 67 | 36 |
| 4 | 66 | 34 | 35 |
| 5 | 83 | 17 | 23 |
| 6* | 93 | 7 | 15 |
| 5 and 6* | 87 | 13 | 38 |
| | | 1976 (Cross-Validating Year) | |
| 3 | 35 | 65 | 43 |
| 4 | 63 | 37 | 51 |
| 5 | 77 | 23 | 30 |
| 6* | 86 | 14 | 7 |
| 5 and 6* | 78 | 22 | 37 |
| | | 1976 and 1977 | |
| 3 | 34 | 66 | 79 |
| 4 | 64 | 36 | 86 |
| 5 | 79 | 21 | 53 |
| 6* | 91 | 9 | 22 |
| 5 and 6* | 83 | 17 | 75 |

*Presumably students whose trigonometry background was insufficient for calculus.

TABLE 8

VALIDITY DATA FOR STANDARDIZED TESTS (1975-77)
MA 132 CALCULUS WITH ANALYTIC GEOMETRY I

| Placement Category | Percent Successful | Percent Unsuccessful | Total Number of Observations |
|---|---|---|---|
| | 1976 and 1977 | | |
| 5 or below | 26 | 74 | 27 |
| 6 | 67 | 33 | 27 |
| | 1975 (Cross-Validating Year) | | |
| 5 or below | 43 | 57 | 14 |
| 6 | 94 | 6 | 17 |
| | 1975, 1976 and 1977 | | |
| 5 or below | 32 | 68 | 41 |
| 6 | 77 | 23 | 44 |

TABLE 9

COMPARISON OF VALIDITY DATA FOR THE TPT
WITH THAT OF THE ACTM-CAT II COMBINATION

| Placement Category | Percent Successful | | Percent Unsuccessful | | Total Number of Observations | |
|---|---|---|---|---|---|---|
| | TPT | ST | TPT | ST | TPT | ST |
| MA 100 Intermediate Algebra | | | | | | |
| 1 | 39 | 49* | 61 | 51* | 33 | 140 |
| 2 | 59 | 65 | 41 | 35 | 17 | 109 |
| 3 and 4 | 84 | 85 | 16 | 15 | 83 | 52 |
| MA 111, MA 121, MA 181 - College Algebra | | | | | | |
| 3 | 28 | 34 | 72 | 66 | 18 | 79 |
| 4 | 62.5 | 64 | 37.5 | 36 | 24 | 86 |
| 5 and 6** | 85 | 83 | 15 | 17 | 62 | 75 |
| MA 132 Calculus with Analytic Geometry I | | | | | | |
| 5 or below | 0 | 32 | 100 | 68 | 2 | 41 |
| 6 | 77 | 77 | 23 | 23 | 30 | 44 |

*These become: 42% successful; 58% unsuccessful; when students for whom one test score is missing are omitted from consideration.

**Students who, presumably, lack a trigonometry background.

This is similar to each of the years considered, not an effect of one particular year.

In the other borderline category, 4, we find 23 percent of the 1979 college algebra students, compared to 36 percent for 1976 and 1977, a difference less marked than at level 2 but a difference nevertheless.

One could argue cogently that more students in borderline categories is "good," because there is correspondingly greater flexibility; students can assess their own motivation - difficult to measure objectively - and determine an appropriate course based on that assessment. An equally persuasive brief could be made for more precise guidelines. Significant differences in favor of one method of placement or the other will not resolve this point; they may indicate which is to be preferred but not why. The issue remains moot, but it is an interesting one and deserves further investigation.

## MEASUREMENT

### The Independent Variables

The independent variable of cardinal interest in this project is "treatment," or method of mathematics placement. In particular, the subjects who were recommended to a specific course or level and took that course, or one at that level, will be carefully studied.

Some students, however, did not abide by the placement advice tendered. Mathematics placement at LSSC is advisory only, partly because of the difficulties of enforcing mandatory placement. Students may take a mathematics class other than the recommended one if they obtain their major field advisor's consent to do so (and some have made their way around that requirement, as is their wont). Some may feel inadequate to the recommended course, or equipped for a more difficult one. Their advisors may suggest another choice, perhaps because the placement data were not carefully noted; perhaps for other reasons, as in the case of the advisor to the medical technology students. He believes that their first year's curriculum is so difficult that MA 111 - to be succeeded by MA 112 and MA 113 in subsequent terms - is the best way to meet the National Accrediting Agency for Clinical Laboratory Sciences' requirement for 12 units of college-level mathematics, even for very capable students.[150]

To complicate the issue further, a supplemental placement device is rather informally employed in sections of MA 132. The calculus instructors give each of their students a copy of the previous year's MA 121 (pre-calculus) final examination. The students are told to treat it like a take-home test. If they can solve the problems readily, they belong in calculus; if not, they are advised to drop it and enroll in MA 121. The students who decide to make this change may not all have been recommended to calculus to begin with. Some could have been from "level 4" or "level 5;" some could have been directed to MA 121 by a mathematics faculty member. The same phenomenon can and does occur in other courses: students may be told to take a certain course, decide to take another, then discover that the original recommendation was better and enroll in that.

Measurement and analysis are of interest for those students who took the course recommended by treatment. The point of view taken here is that intermediate gyrations are not important: students who bounced around once or twice before settling in the recommended course will not be distinguished from those who did so initially.

To the author's surprise, this attitude has proved controversial in some quarters. Certainly the two kinds of student are "different" in unidentified - perhaps unidentifiable - ways. And perhaps the term "successful placement" is a misnomer in this context; what the author and the interested parties at LSSC actually want to know is this: was the original placement advice accurate? Was the student successful or unsuccessful in the course recommended by treatment? The query as to whether or not placement was "successful" in the sense that a student was initially obedient to its commands is not addressed.

There is, moreover, no reason to suppose that the students who "bounced around" are not distributed proportionately across treatment groups, and one reason - random assignment - to suppose that they are. This question cannot be answered with accuracy because there is no permanent record of course changes made during the first six days of instruction. The author did, however, retain the preliminary class list for MA 121 and was able to identify students who added the course during or after the first day of classes. In most cases, these individuals had initially signed up for calculus and then been scared out by the MA 121 final. A rough estimate by Dr. Thesing is that these 14 students account for one-third to one-half of those who made mathematics course changes in the first six days. Their distribution across treatment groups may or may not be representative. Nine of the 14 were experimental subjects; three of these nine had been placed by $T_1$, three by $T_2$ and three by $T_3$!

Although the original assignment of subjects to treatment groups was under the experimenter's control, the direction of students to classes as a result of the treatments was not. We are led to consider six, not three, levels of the principal independent variable:

1 - students who were placed in mathematics courses by the TPT and who eventually took the recommended course (or one at the recommended level);

2 - students placed by ACTM and CAT II who eventually took the recommended course or level;

3 - students placed by faculty advising, who eventually took the recommended course;

4, 5, 6 - students in $T_{i-3}$ who, ultimately, did not take the recommended course or level.

There is another independent variable, which was used to provide insight on whether the provision of a specific course on the placement roster improves the chances of a student's choosing that course. It will be called N, for nomenclature; there are two levels:

1 - student's placement was given explicitly as Basic Math or MA 100;

2 - student's placement was coded as 1 or 3.

## Difficulties

"In studies of academic achievement the traditional criterion of performance is the student's grades."[151] Especially is this true of placement studies in mathematics, as a review of the literature demonstrated.

This criterion has given rise to numerous sources of concern. What shall be said of a student who receives an F in Basic Mathematics? Clearly the student was accurately placed; no lower placement is available; yet application of the grading scale results in a decision of "unsuccessful placement." Reilly addressed this difficulty by using a "combination of self-report data and course performance."[153]

> A student was considered to be ideally placed in a remedial course if he did not find the course repeating past learning experiences. A student who failed to obtain a grade of C or higher was also considered ideally placed regardless of his opinion of the course. All other students in remedial courses, including those neutral toward the repetition question, were defined as "misplaced remedial."[154]

At the other extreme, "the majority of students receiving 'A's' may be working at their full potential; some, however, may have been capable of accelerating their educational progress by enrolling in the next advanced course."[155] Shana'a resolved this difficulty by omitting "the subset of students receiving grades of A from the sample classified as properly placed."[156] Reilly's approach:

> A student was considered ideally placed in a regular course if he was given a grade of C or better and disagreed with the statement that the course was too difficult for him. All other regular course students were considered misplaced.[157]

is puzzling unless his statement that "a regular course is generally the highest level course in which a student may be placed"[158] is interpreted to mean that only one such course was available. Otherwise, students achieving a grade of A and finding the course "too easy" (although they were not asked that question in "regular courses") would be considered as "ideally placed." This would not correspond to the notions of ideal placement held by most.

If grades are selected as the primary or exclusive standard of successful placement, we must also ask whether or not "the uncontrollable variables such as different instructors, diversified grading philosophies, and divergent teaching techniques have relatively insignificant roles in determining the final grade."[159]

There are other, profound reasons to dispute a choice of grades as the sole criterion of accomplishment. Lavin's discussion merits quoting at some length.

> In studies of academic achievement the traditional criterion of performance is the student's grades. ...
>
> The overriding use of this criterion has been a cause of concern to some persons. There are three issues here. The first involves the question of the goals and value premises that underlie the heavy emphasis on grades as an index of academic performance. A second is whether there are other socially significant values and goals in terms of which aspects of performance besides grades become important. Third is the question of the degree to which the traditional emphasis on academic performance is related to these other values. ...

> Because of .. practical considerations and perhaps because of the ease of obtaining data on grades, high grades have assumed the status of a terminal value; academic performance, particularly as represented by high levels of academic achievement, has become an end in itself.
>
> However, if we view the educational process within a broader context, other aspects of the student role may assume greater importance. Formal education is one aspect of the socialization process. When we speak of this process, we imply that we are socializing for something - primarily for performance in a variety of adult roles. ... if career success, critical-mindedness, and creativity are valuable for a variety of personal and societal goals, what is the meaning of grades (especially the implicit value position that high grades are inherently good) in the context of these other values? In short, there is a question as to why - and for what - grades are important. ...
>
> If grades will not predict future eminence, and if the early identification of outstanding talent is a task worth pursuing, there is a need to develop additional criteria of good student performance. ...
>
> Of course, grades are not unimportant. They are unquestionably an index of competence in school work, but within the context of some personal and societal goals, reliance upon grades as the only criterion of student performance is unwarranted.160

What "additional criteria of good student performance" are relevant? Shana'a suggested that

> Ideally we could speak of a student achieving success in a course if all or any of the following occurs.
>
> 1. The student experiences a growth of his rational powers.
>
> 2. He develops an interest and curiosity about the field of study and is motivated

to continue work in that area.

3. He displays mastery of the material and is ready to study more advanced concepts in the field.

4. He has been challenged to put forth maximum efforts in constructive learning activities.

5. His time has been utilized in effective, significant ways during the course.161

(It might be noted that, having set them forth, she promptly rejected them.

..these results are often nebulous and uncontrollable, sometimes unidentifiable, and if recognized and identified, difficult to measure. Moreover, when they can be measured, no absolute scale permits one to draw a line between a success and failure dependent on the amount of variation shown. Even if the foregoing five criteria represent proper outcomes from a course, they are not .. well-defined common criteria which can be used to promote or judge educational efficiency.

We need, therefore, a criterion which is readily accessible, recognizable and acceptable to use in judging an individual's success in any course. We are .. led to the academic grade a student receives in the course. ... We will assume .. that the most efficient, practical measure of a student's academic success is his final course grade.162)

Bloom, Krathwohl and others recognized the need for a "classification of the student behaviors which represent the intended outcomes of the educational process."[163] The Cognitive Domain, the first handbook of the Taxonomy of Educational Objectives, divides "the cognitive objectives into subdivisions from the simplest behavior to the most complex."[164] Very briefly, these objectives are:

<u>1.00 Knowledge</u>

1.10 Knowledge of specifics
1.20 Knowledge of ways and means of dealing with specifics
1.30 Knowledge of the universals and abstractions in a field

<u>2.00 Comprehension</u>

2.10 Translation
2.20 Interpretation
2.30 Extrapolation

<u>3.00 Application</u>

<u>4.00 Analysis</u>

4.10 Analysis of elements
4.20 Analysis of relationships
4.30 Analysis of organizational principles

<u>5.00 Synthesis</u>

5.10 Production of a unique communication
5.20 Production of a plan, or proposed set of operations
5.30 Derivation of a set of abstract relations

<u>6.00 Evaluation</u>

6.10 Judgments in terms of internal evidence
6.20 Judgments in terms of external criteria[165]

A second handbook classifies aspects of <u>The Affective Domain</u>. "The search for a continuum that would provide a means of ordering and relating the different kinds of affective behavior" was "perhaps the most difficult part of the task of building the affective domain of the Taxonomy."[166] "Internalization" provided that means.

> This word seemed an apt description of the process by which the phenomenon or value successively and pervasively become [sic] a part of the individual. When we tried this concept as an organizing principle, we found we were able to construct a

meaningful continuum.[167]

In outline form, the affective goals delineated by Krathwohl, Bloom and Masia are:

1.0 Receiving (Attending)
  1.1 Awareness
  1.2 Willingness to Receive
  1.3 Controlled or Selected Attention

2.0 Responding
  2.1 Acquiescence in Responding
  2.2 Willingness to Respond
  2.3 Satisfaction in Response

3.0 Valuing
  3.1 Acceptance of a Value
  3.2 Preference for a Value
  3.3 Commitment

4.0 Organization
  4.1 Conceptualization of a Value
  4.2 Organization of a Value System

5.0 Characterization by a Value or Value Complex
  5.1 Generalized Set
  5.2 Characterization[168]

There is no question of the value of these characterizations. Further, the handbooks provide insight into identification and measurement of the concepts involved. The organizing principles of "complexity" and "internalization" are continua along which one might "draw a line between success and failure."

There are two major disadvantages. One is the issue of practicability illumined by Lavin. The task of developing and validating such criteria before proceeding with the present experiment would have been enormous; that of obtaining final course grades, slight. The other demurrer is perhaps less tainted. Students in Basic Mathematics and Intermediate Algebra classes are expected to exhibit some knowledge and comprehension of the material studied and, to a limited extent, an ability to apply it. We also hope for reception and at least "acquiescence in responding" from them. More suitable to higher-level courses, such as MA 121 and MA 132, would be the attributes of application and analysis, acceptance and even commitment. In a word, the criteria will not be uniform across courses unless they can be transformed to a common standard.

## The Dependent Variables

The fall term 1980 mathematics grades of experimental subjects who enrolled in one of five courses: MA 100 Intermediate Algebra; MA 111, MA 121 and MA 181, college algebra; and MA 132 Calculus with Analytic Geometry I; were examined. A dichotomous variable X was called "successful" if a student's grade was A, B, C, "Credit" or "Credit by Examination;" X became "unsuccessful" when the mark was D, F or N. This was considered to be a reasonable and meaningful measure of placement effects.

A second dependent variable, Y, was almost identical to X; however, grades of "A" were excluded from consideration.

A third variable, G, took on the exact value of the grade received: that is, G = A, B, C, D, F or N. G was not converted to any numerical scale.

Also of interest was whether one placement method tended to place students at a higher level of mathematics than the others. Thus arose the variable C, whose values are "Intermediate Algebra," "College Algebra" or "Calculus with Analytic Geometry I."

A measure, I, consisting of instructors' assessments of congruity of courses to students was employed.[169] Copies of final class lists were sent to those teaching MA 100, MA 111, MA 121, MA 181, MA 132 <u>and</u> Basic Mathematics, who were asked to characterize their freshmen students in one of four ways:

1 - student is working at the right level;

2 - student should have taken a lower-level course;

3 - student should have taken a higher-level course;

4 - I don't know this student well enough to comment.

Finally, the dependent variable of interest in the investigation of "naming versus coding" is similar, but not identical to C. It will be denoted D, and its levels are:

1 - student took Basic Mathematics, if that was recommended, or MA 100 if that was the indicated course;

2 - student took a course other than the recommended one;

3 - student did not elect any mathematics course.

## Rationale

Although no ideal method of measuring successful placement has been discovered, the dependent variables chosen resolve or attenuate many of the difficulties encountered in making such a choice.

The issues Lavin raises are far-reaching. Is a "heavy emphasis on grades" suitable to the "goals and value premises" that underlie this experiment? The answer is a qualified "yes." The emphasis is on a relatively narrow aspect of the educational process: obtaining optimum placement in fall term mathematics classes for incoming freshmen.

Neither students nor faculty members would be likely to characterize placement as "successful" if it results in a grade of D, F or N. Conversely, a grade of C or better usually indicates success; there remain, however, the questions of whether these students - especially the A students - have found the course to be a repetition of material learned earlier; could have succeeded at a higher level of mathematics; have found their acuity or creativity dulled by the relatively low level of the material. This is another "socially significant value and goal" that deserves investigation, and the "traditional emphasis on academic performance" will not illuminate it. Three approaches have been taken here: first, examination of Y, the result of eliminating A grades when calculating the dichotomous variable; next, by considering each set of grades as a separate entity (the variable G); finally, by considering the instructors' evaluations of course suitability.

The emphasis on high grades that Lavin finds disturbing has been diluted by dichotomizing the variables X and Y. If analysis of these were to reveal no significant differences among treatment groups and if examination of the teachers' opinions did not indicate a large number of A students working below the level of their capability, there are other hypotheses worth testing. Does one placement method produce a larger number of A's and B's than another? Does one method result in students being placed at a higher level of mathematics than the others? G and C were introduced to clarify these issues.

On the question of whether different instructors play a significant role in determining a student's final grade, the author's experience has been very like that of Shana'a, who maintained

> that the nature of mathematics and its objective rather than subjective characteristic of testing permits grading practices more divorced from the instructor's individual feelings than possible in fields such as English, philosophy or education.[170]

> Both [objective evaluation of student achievement and objective grading practices]...can be

> carried out quite independently of most subjective factors such as essay examinations and personality differences between teachers and students which could offset the learner's comprehension of the subject.[171]

To obtain evidence in support of this viewpoint, she tested the hypothesis that

> The correlations between the given predictive variables and the common examination grade in one multi-sectioned course are significantly greater than the correlations between the same predictive variables and the final course grades as issued by the several instructors of the course.[172]

A sample of 283 freshmen, representing 14 different sections of the same course and taught by ten different instructors, was used to test this hypothesis, which was

> rejected for each pair of correlations with the eight predictive variables at the five percent level of significance. Other studies have verified that correlations of given variables with a criterion such as a course grade which is an average of many tests tend to be equal or higher than those based on a single test score. Thus the course grade is as valid a criterion as the more uniform, standardized examination which limits variances caused by different teaching philosophies, grading practices, etc.[173]

The use of instructors' judgments represents an imperfect but optomistic attempt to transform the cognitive and affective objectives of Bloom, Krathwohl, et al., to a common standard. It is assumed that teachers of MA 132 who find their calculus students applying and analyzing their knowledge, accepting and committing themselves to mathematical principles, will rate such students as "working at the proper level;" if they are merely displaying "knowledge of specifics," the instructors would not so rate them, although Basic Mathematics pupils doing so would be considered as properly placed.

Two further discussions are in order. One concerns the classification of D students as "unsuccessfully placed," an action that may seem contradictory because D is a passing grade and permits a student to enter the next higher level of mathematics at LSSC. This, however, is a formal rule of the college and does not correspond to the actual state of affairs in mathematics. Few students receiving a grade of D go on to the next course; the chances of receiving a grade higher than F for those who do are almost non-existent. Most mathematics teachers warn students that unless they have received a grade of C or higher in the prerequisite course, they are not going to pass the present one. The college does have an established policy that D is a passing grade; probably it suffices as a prerequisite in some of the courses in other disciplines. For the mathematics teachers to obtain a variance, a proposal would have to be made and accepted by various committees and administrators. Apparently because no serious problems have arisen from the present way of doing things, no waiver has been sought nor are there plans to seek one.

Finally, a word about student-reported data. Students were not asked if their placement was appropriate, if the course was a repetition of earlier work, if the material was too difficult, for a combination of reasons. The author confesses to a somewhat jaundiced view on the subject of such data: on student-faculty evaluation forms, for example, students report their class attendance as nearly perfect regardless of the actual state of affairs. Thus students unsuccessful in mathematics would - in this writer's opinion - be very apt to lay the responsibility on placement, rather than poor study habits or other personal shortcomings.

As a supplement to the other variables, however, these appraisals might have been useful and even illuminating, could they have been collected without demolishing all the efforts that had gone into avoiding a "Hawthorne effect." This was not possible. First, logistics dictated that such data be collected before the end of fall quarter 1980; otherwise some subjects would have been difficult or impossible to locate, whereas a questionnaire could be distributed easily in the fall term mathematics courses. If students were required to name themselves, however, there would be a justifiable concern as to whether or not their instructors would see the responses, and to what end. Even if this had no effect on their answers - which this author considers very unlikely - it would certainly serve to make them aware of their mathematics placement. Those who had wondered vaguely about the "small group" placement for mathematics would be set to thinking about it all over again.

If replies were collected anonymously, identification of subjects by treatment group would become infeasible. It could have been made clear to subjects at the time of the experiment that they were in one of three groups - and just which one - so that they could provide that information at term's end on a self-report questionnaire. Very definite distinctions would have had to be made, so that students would not confuse $T_1$ with $T_2$, nor assume that their contacts with major field advisors meant, necessarily, that they were in $T_3$. This of course would have nullified completely the efforts to avoid reactive arrangements, while entailing a dependence on students' memories over a period as long as 19 weeks.

Finally, much student-reported data is obtained informally by faculty members and used, after filtration, in their judgments about placement. Some of this information will therefore be reflected in the variable I.

## Criterion Contamination

The full-time mathematics faculty of Lake Superior State College was in the summer of 1980 comprised of Professors Bernard Arbic, Galen Harrison, Reino Hakala, Thomas Mickewich, Gerald Samson, Gary Thesing and Paul Wilson. All participated in the experiment by advising some of the subjects; the number of advisees was: Arbic - 17; Harrison - 18; Hakala - 14; Mickewich - 15; Samson - 24; Thesing - 17; Wilson - 71.

In August, 1980, Dr. Hakala became Dean of Arts and Science at Governor's State University in Park Forest South, Illinois, and therefore did not teach at LSSC in the fall. Each of the others taught at least one section of the relevant courses (Basic Mathematics, Intermediate Algebra, college algebra and Calculus I). The number of experimental subjects taught by each was: Arbic - 17; Harrison - 41; Mickewich - 18; Samson - 10; Thesing - 70; Wilson - 66.

There were also two part-time faculty members. One was Mr. William Gough, who taught a section of MA 111 with 11 experimental subjects in it; the other was the author, who taught MA 121, with 23 subjects. Dr. Charles Jones of the Department of Biology and Chemistry taught a section of MA 100; there were 23 experimental subjects in that class.

Professors Arbic, Harrison, Mickewich, Samson, Wilson, Gough and Jones made no effort to obtain placement data about any of their students. The author looked up placement information for each of the MA 121 students early in the term. Midway through fall quarter, Dr. Thesing received from the Counseling Center a list of all freshmen together with their mathematics placement data. In past years, he has distributed this list to the other mathematics teachers but, because it came too late to be useful for placement changes, he did not do so this time. He did examine his own students' records.

When asked if any of their advisees were also their mathematics students, Professors Arbic, Mickewich, Samson and Thesing replied, "Not as far as I know," or words to that effect. Professor Wilson said, "Just one, and she's not going to make it." Professor Harrison, in response to the same question: "One, and I made a mistake. She's too smart for 181." The exact number of students whose mathematics professor was also their $T_3$ advisor, and the grades and ratings they received, will be reported, by faculty member, in Chapter Four.

Remembering her days as one of three mathematics teachers at Mills College in Oakland, California, the author had anticipated that conversations about individual advising would abound. On the contrary, it appears that no discussions of this kind took place at all.

It is possible that instructors' awareness of which students enrolled in the course originally and which added it later (almost always within the first six days of instruction) constitutes a form of criterion contamination. Certainly some instructors can identify some students who scheduled the course by the drop/add method; there is no question about it if the student added after the six-day "grace period," because this is difficult and complicated. It is hard to believe, however, that it would make a difference in the assessment of students. The instructor, whose part in the drop/add procedure is merely to scrawl a signature on a form presented during the first busy days of school, would not always be able to identify these students, especially as the department head and department secretary may sign unless the class is full.

Dr. Thesing's facetious remark to the author: "I guess that you and I are the only ones who are contaminated," is not technically true; however, we appear to be the only two who are thoroughly contaminated.

# ANALYSIS[174]

## Contingency Tables and the Log Likelihood Test

All the variables in this study are classification variables, or categorical attributes. An exception might have been made for G; omitting grades of N and utilizing the familiar four-point system would have permitted the use of analysis of variance (ANOVA) to test the null hypothesis that mean values of mathematics grades are equal for all three methods of mathematics placement.[175]

This was not feasible. The definition of "successful placement" adopted implies a much greater difference between the grades of "C" and "D" than between any other adjacent pair, so that grades may not be regarded as continuous, as z-scores, or even as "interval data," as required by the ANOVA assumptions.[176]

"A contingency table is .. constructed for the purpose of studying the relationship between ... variables of classification"[177] and several such tables will appear in the analysis. A chi-square test is usually employed to test the hypothesis that the variables are independent; it is defined by the equation $\chi^2 = \sum \frac{\{O - E\}^2}{E}$ where O denotes the observed frequency in a cell of the contingency table; E is the expected frequency; and the sum is taken over all entries in the table.[178] This value is compared with the appropriate value from the chi-square distribution; i.e., the value corresponding to the chosen critical region and the number of degrees of freedom. Should the computed value exceed the tabled value, the null hypothesis of independence between or among the classification variables is rejected.[179]

> An alternative criterion to the usual statistic for comparing observed frequencies with those expected under a particular hypothesis is the likelihood ratio criterion, $\chi^2_L$, given by
>
> $\chi^2_L = 2 \sum \text{Observed} \log_e(\text{Observed}/\text{Expected})$
>
> which like $\chi^2$ has a chi-square distribution when the hypothesis is true. (The degrees of freedom of $\chi^2_L$ are ... the same as for $\chi^2$).
>
> ...The two statistics will take similar values for many tables. However, Ku and Kullback.., Williams.., and others show that, in general, $\chi^2_L$ is preferable to $\chi^2$ ...180

Sokal and Rohlf point out that this log likelihood test, or "G-test,"

> can be recommended on theoretical grounds and has also been shown to be more efficient for variously structured data by empirical studies. The major advantage of the G-test is in more complicated designs where it is simpler to carry out than the chi-square test and often permits certain types of detailed analyses which are impossible using $\chi^2$ ... 181

The log likelihood test has been chosen to analyze the hypotheses of interest in this experiment. It is especially attractive because the three-way contingency table whose analysis is fundamental to this study is one of the "more complicated designs" referred to by Sokal and Rohlf.

## Data Based on Grades

### Data Matrix

Data based on fall term 1980 grades will be laid out in a three by six by six matrix M, whose entries will consist of the total number of subjects falling into certain categories. The first dimension, rows, of M designates "mathematics course level." Rows will be subscripted by the letter i; i = 1 corresponds to MA 100 Intermediate Algebra; i = 2, to the college algebra courses, MA 111, MA 121 and MA 181; i = 3, to MA 132 Calculus with Analytic Geometry I.

The second dimension, columns, is course grade, subscripted by j; values of j = 1, 2, 3, 4, 5, 6, correspond to grades of A, B, C, D, F, N, respectively.

The third dimension, "layers," signifies treatment group for those who took mathematics at the level that was recommended by placement for k = 1, 2 and 3. Values of k = 4, 5 and 6 indicate the students who took a course at a level different from that prescribed by $T_{k-3}$.

## Preliminary Analysis

Denote the (i,j,k)th element of M by $m_{ijk}$. A six-element contingency table $(a_k)$ will be generated from M in the following way:

$$a_k = \sum_{j=1}^{6} \sum_{i=1}^{3} m_{ijk} \quad .$$

To each $a_k$ will then be added the number of Basic Mathematics students in $T_k$ to obtain:

FIGURE 2

TREATMENT GROUP-OUTCOME CONTINGENCY TABLE FORMAT

| | TREATMENT GROUP | | |
|---|---|---|---|
| OUTCOME | $T_1$ | $T_2$ | $T_3$ |
| Took placement advice | $a_1$ | $a_2$ | $a_3$ |
| Did not take advice | $a_4$ | $a_5$ | $a_6$ |

This table will be used to test the null hypothesis that "taking placement advice (ultimately) is independent of the method of mathematics placement." The five percent critical value of chi-squared for n = 2 degrees of freedom is 5.99; it will be compared to the G-value (the log likelihood value) for these data.

It is assumed that the test will confirm the hypothesis of independence; if not, careful interpretation of further results is mandated.

## The Fundamental Analysis

For k = 1, 2, 3, let

$$t_{11k} = \sum_{j=1}^{3} m_{1jk} \qquad t_{12k} = \sum_{j=4}^{6} m_{1jk}$$

$$t_{21k} = \sum_{j=1}^{3} \sum_{i=2}^{3} m_{2jk} \qquad t_{22k} = \sum_{j=4}^{6} \sum_{i=2}^{3} m_{2jk}$$

This creates the following two by two by three contingency table, denoted T.

FIGURE 3

TREATMENT GROUP-COURSE LEVEL-SUCCESS
CONTINGENCY TABLE FORMAT

| | TREATMENT GROUP | | | | | |
|---|---|---|---|---|---|---|
| COURSE LEVEL | $T_1$ | | $T_2$ | | $T_3$ | |
| | S | U | S | U | S | U |
| Intermediate Algebra | $t_{111}$ | $t_{121}$ | $t_{112}$ | $t_{122}$ | $t_{113}$ | $t_{123}$ |
| College Algebra and Calculus | $t_{211}$ | $t_{221}$ | $t_{212}$ | $t_{222}$ | $t_{213}$ | $t_{223}$ |

"S" stands for "successful placement;" i.e., a course grade of A, B or C, and corresponds to the sum from j = 1 to 3. "U" denotes "unsuccessful placement," or a grade of D, F or N, and corresponds to the sum from j = 4 to 6.

The entries for the college algebra group of courses and those for MA 132, Calculus I, have been pooled in each category, corresponding to the sum from i = 2 to 3, because five is the minimum cell size required for a chi-square or log likelihood analysis, and it is not anticipated that there will be so many as five students from each treatment group receiving a grade of D, F or N in MA 132. If cell size permits, these data will not be pooled, and MA 132 will produce a third row of T.

The log likelihood test will be applied to these data. The G-value will be compared to a chi-square distribution with $IJK - I - J - K + 2 = 7$ degrees of freedom (where I = number of rows, J = number of columns and K = number of layers). At the five percent critical level, this value is 14.0671. The null hypothesis: that the level of mathematics class taken, success in mathematics courses, and method of mathematics placement are jointly independent, will be rejected if the G-value is greater than 14.0671.

## Further Analysis if the Null Hypothesis Is Rejected

If the G-test is significant, three components of the overall G-value are of interest:

(1) Is course level independent of placement method? (D. f. = $IK - I - K + 1 = 2$; five percent critical value of chi-square = 5.99).

(2) Is success in mathematics independent of placement method? (D. f. = JK - J - K + 1 = 2; five percent critical value of chi-square = 5.99).

(3) Is there an interaction between course level and success? (D. f. = IJ - I - J + 1 = 1; chi-square at five percent level = 3.84).

The appropriate G-value will be calculated to test each of these hypotheses.[182] Should the null hypothesis be rejected in the first or second case, suitable one-degree-of-freedom tests will be applied to determine more precisely which treatment or treatments accounts for the difference.[183]

Further Analysis if the Null Hypothesis Is Not Rejected

If the first G-test does not lead to rejection of the null hypothesis, one further hypothesis is of interest: are mathematics grades independent of placement method? This question would not be meaningful if, for example, one method placed most students in remedial mathematics while another put most of them into calculus. Failure to reject the fundamental null hypothesis, however, indicates that this is not the case and, moreover, that students tend to be equally successful regardless of placement method. Then it may be valuable to discover which technique - if any - produces higher grades or fewer F's. To assess this, the contingency table defined by

$$c_{kj} = \sum_{i=1}^{3} m_{ijk}$$

for j = 1, 2, 3, 4, 5, 6, and k = 1, 2, 3, will be constructed and used to test the null hypothesis that grades assigned are independent of placement method.

FIGURE 4

TREATMENT GROUP-GRADES CONTINGENCY TABLE FORMAT

| | GRADE | | | | | |
|---|---|---|---|---|---|---|
| PLACEMENT METHOD | A | B | C | D | F | N |
| $T_1$ | $c_{11}$ | $c_{12}$ | $c_{13}$ | $c_{14}$ | $c_{15}$ | $c_{16}$ |
| $T_2$ | $c_{21}$ | $c_{22}$ | $c_{23}$ | $c_{24}$ | $c_{25}$ | $c_{26}$ |
| $T_3$ | $c_{31}$ | $c_{32}$ | $c_{33}$ | $c_{34}$ | $c_{35}$ | $c_{36}$ |

Again the G-value will be computed and compared with the five percent critical value of chi-square. In this case, degrees of freedom = KJ - K - J + 1 = 10 and the five percent critical value of chi-square is 18.3070. If the null hypothesis is rejected, the appropriate one-degree-of-freedom tests will be applied to determine the source of the differences.

## The Variable Y

The fundamental analysis will also be performed with grades of "A" omitted from the count of those who succeed in a recommended course. The null hypothesis is unchanged, as are the three components of interest if the null is rejected. If the null hypothesis is not rejected, no further analysis will be required, as deleting "A" grades will yield no new information from the contingency table $(c_{ij})$.

## Data Based on Instructors' Judgments

Data supplied by the variable I will be displayed in a three by four array denoted A. Rows of A correspond to placement method, with i = 1, 2, 3, indicating $T_1$, $T_2$, and $T_3$, respectively. Columns correspond to the four categories described on page 93: j = 1 means the course is appropriate for the student; j = 2, that the student should have taken a lower-level course; j = 3, that the student should have taken a higher-level course; and j = 4, that the teacher does not know the student well enough to comment.

A is a contingency table, for which the G-value will be computed and compared to 12.59, the five percent critical value of chi-square for (I-1)(J-1) = 6 degrees of freedom. This will test the null hypothesis that "successful placement, as assessed by the students' mathematics instructors, is independent of mathematics placement method." If the null hypothesis is rejected, one-degree-of-freedom tests will be used to find the source of significance.

FIGURE 5

TREATMENT GROUP-INSTRUCTORS' JUDGMENTS
CONTINGENCY TABLE FORMAT

| | JUDGMENT | | | |
|---|---|---|---|---|
| TREATMENT GROUP | 1 | 2 | 3 | 4 |
| $T_1$ | $a_{11}$ | $a_{12}$ | $a_{13}$ | $a_{14}$ |
| $T_2$ | $a_{21}$ | $a_{22}$ | $a_{23}$ | $a_{24}$ |
| $T_3$ | $a_{31}$ | $a_{32}$ | $a_{33}$ | $a_{34}$ |

## Analysis of Naming Versus Coding

A hypothesis of interest, especially if earlier analyses reveal significant differences in favor of faculty advising, is whether or not that effect might be due to the fact that a specific course is recommended, whereas major field advisors must interpret a code number when students are placed in mathematics courses by testing instruments.

The subjects chosen to test the null hypothesis that "taking the recommended course (eventually) was independent of whether the course was named or coded on the placement roster" were those placed by the TPT whose recommended placement was Basic Mathematics or Intermediate Algebra; that is, subjects in $T_1$ whose placement was coded "1" or "3," respectively. These students had been randomly assigned to one of two groups at the time each placement roster was made up; mathematics placement for the first group was listed by course name; that for the second group was coded. The data matrix, which is also a contingency table, has three rows and two columns; the rows correspond to the levels of D and the columns to the levels of N.

FIGURE 6

NOMENCLATURE-OUTCOME CONTINGENCY TABLE FORMAT

| | NOMENCLATURE | |
|---|---|---|
| OUTCOME | Named | Coded |
| Student took recommended course | $c_{11}$ | $c_{12}$ |
| Student took a different course | $c_{21}$ | $c_{22}$ |
| Student did not take mathematics | $c_{31}$ | $c_{32}$ |

The G-value will be computed for this table and compared to 5.99, the five percent critical value for chi-square with $(I-1)(J-1) = 2$ degrees of freedom.

## Confidence Intervals

An estimate of the success rate, r, for each method of placement is of interest, whatever the aftermath of the statistical tests may be. This parameter is approximated by the quotient

$$r' = \frac{\text{number of successful placements}}{\text{total number of placements}},$$

but how closely? And with what degree of certainty?

Because more than 25 observations are available in all categories of interest, the standard normal approximation to the binomial distribution may be used to obtain a 95 percent confidence interval of $r' - 1.96\ s < r < r' + 1.96\ s$, where s is an approximation to $\sigma$, the standard deviation. Let n = the total number of observations; then $s = \sqrt{r' \cdot (1 - r') \div n}$.
A 99 percent confidence interval is given by $r' - 2.58\ s < r < r' + 2.58\ s.$[184]

Ordinarily, the 95 percent interval will be used, but not always. If, for example, it is possible to obtain a 99 percent interval bounded away from a 70 percent success rate for any of the placement methods, this will be done.

Other suggestive questions are: what percentage of students selects the course recommended by placement? What proportion of students is placed at a remedial level of mathematics? And, subject to the outcome of the statistical tests, others may become inviting. For instance, if (but only if) the analysis of "naming versus coding" yields a significant result, confidence intervals for "percentage choosing recommended course" in each classification will be calculated.

## SUMMARY

A comparative experiment was designed along lines recommended by Campbell and Stanley. Subjects were randomly assigned to one of three treatment groups; each group received a different method of mathematics placement. Safeguards were taken to preclude a "Hawthorne effect;" neither students nor their major advisors (with one exception) were aware that any research was in progress.

Norms for placement by Dr. Thesing's test were obtained from an examination of 1979 data and were to be cross-validated in 1980. Norms for the standardized tests used 1975, 1976 and 1977 information; this included a cross-validating year.

Four dependent variables of primary interest were selected: "success" - a grade of A, B or C in beginning mathematics - versus "failure" - a grade of D, F or N; the same variable with grades of "A" omitted; course level; and instructors' ratings. A preliminary report on criterion contamination was included.

The major null hypotheses to be tested at the five percent critical level are:

(1) taking placement advice is independent of placement method;

(2) course level, success and placement method are jointly independent;

(3) instructors' judgments are independent of placement method.

The second of these entails analysis of a three-way contingency table. The "log likelihood test," similar to the chi-square test, is especially advantageous in complex situations and was therefore selected. Computed values will be compared with tabled values of chi-square for the suitable number of degrees of freedom in each case. Alternate analysis strategies, dependent on the outcome of these tests, were described.

Confidence intervals, obtained from the normal approximation to the binomial distribution, will be placed around parameters of interest.

# CHAPTER IV

# STATISTICAL ANALYSIS

This chapter presents the analyses of variables based on mathematics grades assigned to experimental subjects, and the instructors' ratings of these students. A comparison of the results of "coding" versus "naming" course recommendations on the placement roster is also included; this issue was discussed in detail in Chapter Three. Finally, confidence intervals are placed around parameters of interest.

## DATA BASED ON GRADES

### The Data Matrix

The data matrix M was obtained by counting grades for students at each beginning level of mathematics, except MA 090, Basic Mathematics, in each placement group (Table 10). Several points should be noted.

Students in the borderline placement category "2" who took either Basic Mathematics or Intermediate Algebra were considered to have taken the course recommended by treatment. Similarly, those in the borderline category "4" who chose Intermediate Algebra or one of the college algebra courses were so regarded. This applies only

TABLE 10

THE DATA MATRIX M

| TREATMENT GROUP | COURSE LEVEL | LAYERS 1, 2 AND 3 SUBJECTS ACCEPTED ADVICE GRADES | | | | | | LAYERS 4, 5 AND 6 SUBJECTS REJECTED ADVICE GRADES | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | A | B | C | D | F | N | A | B | C | D | F | N |
| $T_1$ | MA 100 | 3 | 12 | 12 | 4 | 2 | 3 | 0 | 0 | 4 | 3 | 2 | 3 |
| | MA 111, 121, 181 | 7 | 11 | 8 | 5 | 2 | 1 | 0 | 1 | 2 | 1 | 1 | 6 |
| | MA 132 | 0 | 3 | 2 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| $T_2$ | MA 100 | 5 | 9 | 10 | 5 | 2 | 1 | 0 | 0 | 3 | 3 | 3 | 3 |
| | MA 111, 121, 181 | 10 | 5 | 4 | 3 | 3 | 1 | 0 | 2 | 3 | 3 | 0 | 0 |
| | MA 132 | 1 | 4 | 2 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| $T_3$ | MA 100 | 1 | 3 | 7 | 2 | 4 | 4 | 0 | 2 | 0 | 0 | 0 | 0 |
| | MA 111, 121, 181 | 4 | 11 | 7 | 11 | 2 | 10 | 1 | 2 | 0 | 0 | 0 | 0 |
| | MA 132 | 1 | 0 | 2 | 0 | 0 | 2 | 0 | 0 | 0 | 1 | 0 | 0 |

to students in the first two treatment groups, placement by testing.

Students in those two groups whose placement category was "6" but who selected one of the college algebra courses were also considered to have taken the course recommended by treatment. It is assumed that these individuals lacked a background in trigonometry sufficient for success in MA 132, Calculus and Analytic Geometry I.

There are three subjects from $T_3$, the group placed by mathematics faculty members, whose status is anomalous. One selected MA 111, Freshman Mathematics I, although MA 181, Applied Mathematics I was recommended. Because both courses are at the same beginning level, this student was considered to have taken the course recommended by treatment.

The other two cases are less clear-cut. The advisors recommended "MA 181 or higher" for one and "MA 111 or MA 132" to the other. The students chose, respectively, MA 181 and MA 111, and they have also been classified as taking the course recommended by treatment. As one of the purported advantages of individual advising is the opportunity to resolve ambiguous cases and make clear-cut recommendations, this may seem a questionable decision. This is especially true in light of the fact that no borderline categories between college algebra and calculus were designed for the testing treatments.

Three students assigned to $T_3$ enrolled in mathematics courses but did not consult a faculty advisor. The most likely explanation is that they had not intended to take mathematics but changed their minds at the last minute. They have not been included in this study and it must be noted that, despite the precautions taken, $T_3$ was subject to mortality to a greater extent than were the other treatment groups.

## Preliminary Analysis

The layers of M correspond to treatment groups; each treatment group was dichotomized according to whether or not subjects took the course recommended by treatment. Elements of M were summed across the other two dimensions. Eleven students from $T_1$, one from $T_2$ and one from $T_3$ took MA 090, Basic Mathematics, as directed by placement. Two subjects from $T_3$ chose MA 090 although MA 100, Intermediate Algebra, had been recommended. Totals were then adjusted to include these 15 subjects.

TABLE 11

TREATMENT GROUP-OUTCOME CONTINGENCY TABLE

| OUTCOME | TREATMENT GROUP $T_1$ | $T_2$ | $T_3$ | Totals |
|---|---|---|---|---|
| Took placement advice | 88 | 67 | 72 | 227 |
| Did not take advice | 23 | 21 | 8 | 52 |
| Totals | 111 | 88 | 80 | 279 |

The G-value for this table is 6.362, which is significant at the 95th percentile; the tabled value of $\chi^2$ with two degrees of freedom is 5.991.[185] A Texas Instruments TI-55 was used for all statistical calculations, which are accurate to eight decimal digits. Details of each may be found in Appendix B.

A comparison of the two methods of placement by testing is shown in Table 12. A G-value of 0.281, which is not statistically significant, is found for this table.

When the two testing methods are pooled and compared to $T_3$, the resulting G-value is 6.081, which is significant at the 97.5th percentile (degrees of freedom = 1; tabled value of $\chi^2$ = 5.024). The contingency table is Table 13.

TABLE 12

TESTING METHODS-OUTCOME CONTINGENCY TABLE

| OUTCOME | TREATMENT GROUP $T_1$ | $T_2$ | Totals |
|---|---|---|---|
| Took placement advice | 88 | 67 | 155 |
| Did not take advice | 23 | 21 | 44 |
| Totals | 111 | 88 | 199 |

TABLE 13

PLACEMENT METHOD-OUTCOME CONTINGENCY TABLE

| OUTCOME | TREATMENT GROUP $T_1 + T_2$ | $T_3$ | Totals |
|---|---|---|---|
| Took placement advice | 155 | 72 | 227 |
| Did not take advice | 44 | 8 | 52 |
| Totals | 199 | 80 | 279 |

We may tentatively conclude, then, that students are more likely to choose a course recommended by a mathematics professor than one to which they are directed as a result of testing. Careful interpretation of further results in this light is mandated.

Two points should be kept in mind. If the two students from $T_3$ whose inclusion in the "took advice" group is dubious are omitted, the G-value for the resulting contingency table, Table 14, is 5.961, which is not significant at the 95th percentile (d.f. = 2; tabled $\chi^2$ = 5.991).

TABLE 14

ADJUSTED TREATMENT GROUP-OUTCOME CONTINGENCY TABLE

| OUTCOME | TREATMENT GROUP $T_1$ | $T_2$ | $T_3$ | Totals |
|---|---|---|---|---|
| Took placement advice | 88 | 67 | 70 | 225 |
| Did not take advice | 23 | 21 | 8 | 52 |
| Totals | 111 | 88 | 78 | 277 |

The other consideration is extremely interesting. When the behavior "did not take placement advice" is subdivided into "took a more difficult course than the recommended one" and "took an easier one," a striking fact emerges. Only seven students fall into the latter category and all are from $T_3$.

## The Fundamental Analysis

The contingency table T classifies subjects by treatment group, course level and success. Only those who selected the course recommended by treatment are enumerated. "Success" indicates a grade of A, B or C in the chosen course; "unsuccessful" students are those who received a grade of D, F or N.

TABLE 15

TREATMENT GROUP-COURSE LEVEL-SUCCESS CONTINGENCY TABLE

| | TREATMENT GROUP | | | | | |
|---|---|---|---|---|---|---|
| | $T_1$ | | $T_2$ | | $T_3$ | |
| COURSE LEVEL | S | U | S | U | S | U |
| Intermediate Algebra | 27 | 9 | 24 | 8 | 11 | 10 |
| College Algebra and Calculus | 31 | 10 | 26 | 8 | 25 | 25 |

Totals for this table are shown in Appendix B. The overall G-value is 19.530, which is significant at the 99th percentile (7 degrees of freedom; tabled $\chi^2$ = 18.475.) The three components set out in Chapter Three will therefore be investigated. Is course level independent of placement method? Is success independent of placement method? Is course level independent of success?

(1) Is course level independent of placement method? The G-value is 6.499, which is significant at the 5% critical level; not higher. (D.f. = 2; tabled $\chi^2$ = 5.991. Tabled $\chi^2$ at the 2½% critical level is 7.378).

A comparison of the two testing methods yields a G-value of 0.021, which is not significant.

TABLE 16

TESTING METHODS-COURSE LEVEL CONTINGENCY TABLE

| COURSE LEVEL | TREATMENT GROUP $T_1$ | $T_2$ | Totals |
|---|---|---|---|
| Intermediate Algebra | 36 | 32 | 68 |
| College Algebra and Calculus | 41 | 34 | 75 |
| Totals | 77 | 66 | 143 |

When the results of placement by testing are pooled and examined side-by-side with the data from faculty advising (Table 17), the ensuing G-value is 6.456, which is significant at the 2½% level; not higher. (D.f. = 1; tabled $\chi^2$ = 5.024; tabled $\chi^2$ at 1% level = 6.635).

It therefore seems likely that faculty advisors tend to place students at a higher level of mathematics than do the tests. Other considerations strengthen this viewpoint. When the subjects enrolled in MA 090, Basic

Mathematics, are included in the appropriate cells of the contingency table to produce Table 18, the G-value is 9.008, which is significant at the ½ of 1% critical level (d.f. = 1; tabled $\chi^2$ = 7.879).

TABLE 17

PLACEMENT METHOD-COURSE LEVEL CONTINGENCY TABLE

| COURSE LEVEL | TREATMENT GROUP $T_1 + T_2$ | $T_3$ | Totals |
|---|---|---|---|
| Intermediate Algebra | 68 | 21 | 89 |
| College Algebra and Calculus | 75 | 50 | 125 |
| Totals | 143 | 71 | 214 |

TABLE 18

COURSE LEVEL EFFECT WITH MA 090 INCLUDED

| COURSE LEVEL | TREATMENT GROUP $T_1 + T_2$ | $T_3$ | Totals |
|---|---|---|---|
| Basic Mathematics and Intermediate Algebra | 80 | 22 | 102 |
| College Algebra and Calculus | 75 | 50 | 125 |
| Totals | 155 | 72 | 227 |

When one considers that seven of the eight students from $T_3$ who did not choose the course recommended by treatment selected an easier one, while all 44 subjects from $T_1$ and $T_2$ who disregarded placement advice enrolled in a more difficult course, the conclusion is further substantiated.

(2) Is success independent of placement method? The G-value is 12.974, which is significant at the ½ of 1% critical level. (D.f. = 2; tabled $\chi^2$ = 10.597). A contrast of $T_1$ and $T_2$ (Table 19) yields a G-value of 0.004, which is not significant.

TABLE 19

TESTING METHOD-SUCCESS CONTINGENCY TABLE

| SUCCESS | TREATMENT GROUP $T_1$ | $T_2$ | Totals |
|---|---|---|---|
| Successful | 58 | 50 | 108 |
| Unsuccessful | 19 | 16 | 35 |
| Totals | 77 | 66 | 143 |

A comparison of testing with advising (Table 20) produces a G-value of 12.971. This is significant at the 99.5th percentile. (D.f. = 1; tabled $\chi^2$ = 7.879). Thus there is a strong indication that students placed by members of the mathematics faculty are less successful than those placed by testing.

TABLE 20

PLACEMENT METHOD-SUCCESS CONTINGENCY TABLE

| SUCCESS | TREATMENT GROUP $T_1 + T_2$ | $T_3$ | Totals |
|---|---|---|---|
| Successful | 108 | 36 | 144 |
| Unsuccessful | 35 | 35 | 70 |
| Totals | 143 | 71 | 214 |

(3) Is course level independent of success? The G-value is 0.391, which is not significant.

## The Variable Y

Deletion of "A" grades from the table T reduces the number of entries in the "Successful" columns; the result is Table 21. Totals are in Appendix B.

The comprehensive G-value for this table is 18.107; this is significant at the 2½ % critical level; not higher. (D.f. = 7; tabled $\chi^2$ = 16.013; tabled $\chi^2$ at 1% level = 18.475). Results for the three components of interest, identical to those of the fundamental analysis, follow. The contingency tables may be found in Appendix B.

TABLE 21

TREATMENT GROUP-COURSE LEVEL-SUCCESS CONTINGENCY TABLE
(A GRADES OMITTED)

| | TREATMENT GROUP | | | | | |
|---|---|---|---|---|---|---|
| | $T_1$ | | $T_2$ | | $T_3$ | |
| COURSE LEVEL | S | U | S | U | S | U |
| Intermediate Algebra | 24 | 9 | 19 | 8 | 10 | 10 |
| College Algebra and Calculus | 24 | 10 | 15 | 8 | 20 | 25 |

(1) Is course level independent of treatment? The G-value is 7.537, significant at the 2½% critical level. (D.f. = 2; tabled $\chi^2 = 7.378$).

A one-degree-of-freedom comparison of $T_1$ with $T_2$ yields G = 0.258; this is not significant.

When results of the two testing methods are pooled and contrasted with faculty advising, G = 7.279 is obtained. This is significant at the 99th percentile.

(2) Is success independent of treatment? G = 10.209, significant at the 99th percentile (d.f. = 2; tabled $\chi^2 = 9.210$). $T_1$ versus $T_2$ produces a G-value of 0.181; it is not significant. Inspection of the results of testing against those of counseling yields G = 10.029, significant at the ½ of 1% critical level.

(3) It appears that course level and success are mutually independent; G = 1.345, which is not significant.

The results are consistent with those obtained when grades of "A" are utilized.

## A Final Consideration

In view of the apparent superiority of testing over counseling, it is interesting to compare, for subjects in $T_1$ and $T_2$, the success of those who accepted placement advice with the success of those who rejected it and chose a more difficult course. When "A" grades are included in the count of "successful" students, the result is Table 22.

TABLE 22

ACCEPTANCE-SUCCESS CONTINGENCY TABLE

| SUCCESS | TREATMENT GROUP $T_1 + T_2$ | $T_4 + T_5$ | Totals |
|---|---|---|---|
| Successful | 108 | 15 | 123 |
| Unsuccessful | 35 | 29 | 64 |
| Totals | 143 | 44 | 187 |

The G-value is 24.677, which is extremely significant: the tabled value of $\chi^2$ with d.f. = 1 is 7.879 at the

99.5th percentile.

Exclusion of "A" students lowers the number of "successful" observations for $T_1$ and $T_2$, the "acceptors," only; nevertheless, the G-value is 17.145 for Table 23.

TABLE 23

ACCEPTANCE-SUCCESS CONTINGENCY TABLE (A GRADES OMITTED)

| SUCCESS | TREATMENT GROUP $T_1 + T_2$ | $T_4 + T_5$ | Totals |
|---|---|---|---|
| Successful | 82 | 15 | 97 |
| Unsuccessful | 35 | 29 | 64 |
| Totals | 117 | 44 | 161 |

## DATA BASED ON INSTRUCTORS' RATINGS

### The Data Matrix

During the last week of fall term, 1980, a memorandum was sent to teachers of beginning mathematics courses at Lake Superior State College.

> Here, as promised ... are copies of your final class lists in the mathematics courses: 090, 100, 111, 121, 181 and/or 132. Could you please assess the suitability of each course for each [freshman] student in one of the following ways:
>
> 1 - indicates that this is the right course (or course level) for this student;

2 - this student should have taken a lower-level course (please indicate the course you think would have been most appropriate);

3 - this student should have taken a higher-level course (again, please name the course or course level you think would have been most appropriate);

4 - I don't know this student's abilities well enough to comment. ...

> Thank you very much for your cooperation. All information will be treated with strictest confidentiality.186

A comment that accompanied Mr. Gough's ratings highlights the fact that they are not redundant and carry different connotations than do grades:

> There are a number of students who receive low grades simply because they do not put forth a reasonable effort. This is evidenced by repeated absences, late hand-in assignments, and generally shoddy work. [Four] such students have been classified as being in the right course level even though they earned [three] D's and [one] F. Taking MA 100 probably would not be an answer to their problems...187

At the other extreme of achievement, three of the five subjects who, in the opinion of their instructors, "should have taken a higher-level course," received grades of "B," not "A," from those instructors; the remaining two were "A" students.

When ratings were tallied for students who chose the course recommended by treatment, the result was the matrix A, which is shown in Table 24. Totals for those who disregarded placement advice are given in Table 25.

TABLE 24

THE DATA MATRIX A

| TREATMENT GROUP | INSTRUCTOR'S RATING 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $T_1$ | 72 | 9 | 4 | 3 |
| $T_2$ | 58 | 5 | 0 | 4 |
| $T_3$ | 46 | 18 | 1 | 7 |

TABLE 25

INSTRUCTORS' RATINGS FOR STUDENTS WHO REJECTED PLACEMENT ADVICE

| TREATMENT GROUP | INSTRUCTOR'S RATING 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $T_4$ | 9 | 10 | 0 | 4 |
| $T_5$ | 13 | 5 | 0 | 3 |
| $T_6$ | 8 | 0 | 0 | 0 |

Earlier remarks concerning borderline and anomalous cases apply here also. In addition, some MA 111 students and one MA 181 student were rated "3" with the "most appropriate course" given as MA 121. Since all begin at the same level - albeit MA 121 is the most difficult - these ratings were tallied as "1's" instead. By the same token, an MA 121 student for whom the course was too difficult was rated "1" because MA 111 was considered most appropriate.

## Statistical Analysis of A

Because a minimum cell size of five is required for the analysis of A, only the subset enumerating ratings of "1" and "2" was utilized. The overall G-value for Table 26 is 11.110; this is significant at the 99.5th percentile. (D.f. = 2; tabled $\chi^2$ = 10.597).

TABLE 26

TREATMENT GROUP-RATINGS CONTINGENCY TABLE

| TREATMENT GROUP | INSTRUCTOR'S RATING 1 | 2 | Totals |
|---|---|---|---|
| $T_1$ | 72 | 9 | 81 |
| $T_2$ | 58 | 5 | 63 |
| $T_3$ | 46 | 18 | 64 |
| Totals | 176 | 32 | 208 |

A comparison of $T_1$ with $T_2$ revealed no significant differences: G = 0.414 for Table 27.

Pooling results of testing and comparing them with the counseling method gives Table 28 with a G-value of 10.696, which is significant at the ½ of 1% critical level (d.f. = 1; tabled $\chi^2$ = 7.879).

TABLE 27

TESTING METHOD-RATINGS CONTINGENCY TABLE

| TREATMENT GROUP | INSTRUCTOR'S RATING 1 | 2 | Totals |
|---|---|---|---|
| $T_1$ | 72 | 9 | 81 |
| $T_2$ | 58 | 5 | 63 |
| Totals | 130 | 14 | 144 |

TABLE 28

PLACEMENT METHOD-RATINGS CONTINGENCY TABLE

| TREATMENT GROUP | INSTRUCTOR'S RATING 1 | 2 | Totals |
|---|---|---|---|
| $T_1 + T_2$ | 130 | 14 | 144 |
| $T_3$ | 46 | 18 | 64 |
| Totals | 176 | 32 | 208 |

Therefore it may be inferred that ratings are not independent of placement method, and that there is a tendency on the part of faculty advisors in mathematics to place students at a higher level of that subject than is suited to their capabilities.

In this light, a comparison of the ratings for students placed by testing who accepted placement advice with those who rejected it may be in order. The contingency table is Table 29; the associated G-value is 17.479, which is highly significant when compared with the tabled value of $\chi^2$ = 7.879 at the 99.5th percentile (d.f. = 1).

TABLE 29

ACCEPTANCE-RATINGS CONTINGENCY TABLE

| TREATMENT GROUP | INSTRUCTOR'S RATING 1 | 2 | Totals |
|---|---|---|---|
| $T_1 + T_2$ ("Acceptors") | 130 | 14 | 144 |
| $T_4 + T_5$ ("Rejectors") | 22 | 15 | 37 |
| Totals | 152 | 29 | 181 |

All the results obtained from instructors' ratings are consistent with those obtained from the matrix M. This is true whether marks of "A" are employed in the statistical calculations or not. Because of this consistency, and because instructors characterized only two "A" students as working below their proper level, the variable Y will no longer be utilized for measurement. When "success based on grades" is discussed, it will mean a grade of A, B or C.

## CRITERION CONTAMINATION

The subject of criterion contamination was addressed in Chapter Three. A reporting of the exact number of students whose placement advisor was also their beginning mathematics instructor was deferred to this chapter. These numbers, together with the ratings assigned and the final grades those students received, are shown in Table 30.

The advisors appear to be rather hard on themselves: out of the 14 subjects listed, only seven were considered properly placed by the same professor who did the placement.

## NAMING VERSUS CODING

There is an apparent tendency for students placed by mathematics professors to abide by the advice in greater proportions than those placed by tests. Is this due to the effect of the interview, or is it simply because a placement code need not be interpreted by the students' major advisors?

To clarify this issue, $T_1$ students whose placement code was "1" or "3" were randomly assigned to one of two groups: course names were entered on the placement roster for one group, while codes were entered for the other. The results are displayed in Table 31.

TABLE 30

RATINGS AND GRADES ASSIGNED BY STUDENTS' PLACEMENT ADVISORS

| Professor | No. of Students | Ratings Assigned | Final Grades |
|---|---|---|---|
| Arbic | 1 | 1 | B |
| Harrison | 5 | 1, 1, 1, 2, 3* | B, B, D, N, A |
| Mickewich | 3 | 1**, 2, 4 | Z***, D, Z*** |
| Samson | 1 | 1 | C |
| Thesing | 4 | 1, 1, 2, 2 | C, C, D, N |
| Wilson | 0**** | | |

*This rating was changed to "1" because Professor Harrison had recommended "MA 181 or higher," and rated the "most appropriate course" as MA 121 or MA 132.

**Dr. Mickewich recommended MA 100 to this student, who instead chose MA 090. The rating of "1" means that Dr. Mickewich considered MA 090 to be the appropriate level for this student.

***The grade of "Z" means "deferred;" the student has an indefinite amount of time to complete work in the course. All MA 090 students who did not drop the course received this grade.

****The name of the student to whom Professor Wilson referred as "not going to make it" did not appear on the final class list for the course in question.

The G-value for this table is 3.272, which is not significant at the 5% critical level. (D.f. = 2; tabled $\chi^2$ = 5.991). The one-degree-of-freedom test for the first two rows yields G = 0.677, which is also not significant at the 5% level; see Appendix B. (Tabled $\chi^2$ = 3.841).

TABLE 31

NOMENCLATURE-OUTCOME CONTINGENCY TABLE

| OUTCOME | NOMENCLATURE Named | Coded | Totals |
|---|---|---|---|
| Student took recommended course | 25 | 10 | 35 |
| Student took a different course | 13 | 10 | 23 |
| Student did not take mathematics | 29 | 26 | 55 |
| Totals | 67 | 46 | 113 |

Therefore the disposition of $T_3$ subjects to accept placement recommendations may not be attributed to the fact that these courses are named explicitly on the placement roster.

## CONFIDENCE INTERVALS

### Success Rates for Placement Methods

For measurement based on grades, i.e., A, B or C indicating "success," the 1979 data showed that 79%

of the students scoring 12 or more on the TPT (171 out of 216) were "successfully placed" by the criteria garnered from that data. Students earmarked for Basic Mathematics are not included in this figure, because the course was not available in fall term 1979, and because of the previously mentioned difficulties involved in measuring success by grades in such a course. Cross-validation in 1980 yields a success rate of 75% (58 out of 77 students).

Application of the formula given in Chapter Three to the success rate of 78.16% for all 293 observations yields a 99% confidence interval of

$71.93\% < r < 84.38\%$ for r, the true rate of success.

Eighty-eight students from $T_1$ took the course indicated by the TPT. Of these, 72 were rated by their instructor as working at the right level; nine "should have taken a lower-level course;" four "should have taken a higher-level course;" and three were not known well enough for comment. Even if it is assumed that these three were not correctly placed, a success rate of 81.82% is obtained, together with a 99% confidence interval $71.21\% < r < 92.43\%$ for the true rate.

Data based on 1975, 1976 and 1977 fall term mathematics grades show a success rate of 73% (266 out of 366 observations) for the ACTM-CAT II combination. Again, this is for students who placed in category "2" -

borderline Basic Mathematics/Intermediate Algebra - or higher. The figure for 1980 is 76%, judged by the same criterion (50 out of 66 students). In all, 316 out of 432 students were "successful" - 73.15% - and a 95% confidence interval of $68.97\% < r < 77.33\%$ is realized. (The 99% interval is $67.65\% < r < 78.65\%$).

Fifty-eight of the 67 subjects who took the course recommended by $T_2$ were rated as working at the right level, only five "should have taken a lower-level course" and there were no "mis-placements" in the other direction. If the same conservative assumption made for the TPT - that those who weren't known well enough for comment were not properly placed - is adopted, the ACTM-CAT II combination is still 86.57% successful, with a 99% confidence interval of $75.82\% < r < 97.32\%$ for the true rate r.

Only 1980 data are available to estimate the rate of success for faculty advising. At the Intermediate Algebra level or higher, 71 took the course recommended by treatment and 36, or 50.70%, were "successful" by the "A, B or C" standard. A 95% confidence interval for the actual rate is $39.07\% < r < 66.01\%$. The 99% interval is $35.40\% < r < 66.01\%$.

When the instructors' ratings are examined, we find that 18 out of 65 $T_3$ students taking the suggested course were rated as "should have taken a lower-level course." This time, the assumption that the "unknown" students were unsuccessfully placed is not made; only those rated "1," "2" or "3" have been counted. This is partly because the number, seven, is relatively large, but primarily because adoption of a testing method will be urged in Chapter Five; therefore, omission of these cases is the conservative approach here. One student "should have taken a higher-level course;" 46, or 70.77%, were rated as working at the right level. The 95% confidence interval for success rate is $59.71\% < r < 81.83\%$.

In view of the large G-value obtained when students from $T_1$ and $T_2$ - "acceptors" of testing advice - are contrasted with those in $T_4$ and $T_5$ - "rejectors" of that advice - a confidence interval for the rate of success among students who "place themselves" is in order. When grades are the criteria, we find 15 out of the 44 students in this group achieving a grade of A, B or C, while 29 received a D, F or N. An estimate for the rate is 34.09%; the 95% confidence interval is $20.08\% < r < 48.10\%$. Thus it seems likely that more than half of these students do not succeed.

When instructors' ratings are used, the picture is somewhat brighter: 22 out of 37, or 59.46%, were rated as working at the right level. The seven who weren't known well enough for comment are omitted to obtain a conservative estimate. The 95% confidence interval is $43.64\% < r < 75.38\%$. However, if we call those seven "unsuccessful," as was done for the tests, the success rate is only 50% and the 95% confidence interval is $48.89\% < r < 51.11\%$, which is bounded well below the 70% level.

## Remedial Placement Rates

Of the 279 subjects taking mathematics, 147, or 52.67%, were placed at a remedial level: Basic Mathematics or Intermediate Algebra. Tests placed 123 out of 199 students at a remedial level; this is a rate of 61.81% and a 99% confidence interval is $52.92\% < r < 70.69\%$. Mathematics advisors, on the other hand, placed only 24 out of 80 advisees in remedial mathematics. This is a 30% rate with a 99% confidence interval of $16.78\% < r < 43.22\%$.

Data for 1979 show 151 students placed at a remedial level by current standards. These are the students whose placement category would now be "1," "2," or "3," plus the category "4" students who selected Intermediate Algebra. One hundred and eighteen would be in category "5" or "6," or among those from category "4" who selected college algebra. Similar classifications for 1975, 1976 and 1977 yield 380 who would now be placed at a remedial level by the standardized tests, and 246 above that level. For the five years 1975-77 and 1979-80, then, the rate of remedial placement for testing methods is 59.78% for 654 out of 1094 students. A 99% confidence interval of $55.96\% < r < 63.61\%$ is obtained.

## Rate of Placement Acceptance

Two hundred and twenty-seven of the 279 experimental subjects, or 81.36%, chose the course recommended by treatment. The 95% confidence interval is $76.79\% < r_A < 85.93\%$ or, to look at it another way, 18.64% rejected placement advice and $14.07\% < r_R < 23.21\%$ is the 95% confidence interval.

The rejection rate for testing advice is 22.11% (44 out of 199 subjects); the 95% confidence interval is $16.34\% < r_R < 27.88\%$; while only 10% of the $T_3$ subjects rejected advice; that $3.43\% < r_R < 16.57\%$ is 95% probable. Thus, though the two rejection rates may be equal, the upper bound for counseling is very close to the lower bound for testing.

# CHAPTER V

## SUMMARY AND RECOMMENDATIONS

The purpose of this study was to determine the best available means of placing students in beginning mathematics courses at Lake Superior State College in Sault Ste. Marie, Michigan.

Three methods were compared: placement based on a locally written test; placement based on two standardized tests; and faculty advising.

The local instrument was written by Dr. Gary Thesing of the LSSC mathematics faculty. The standardized tests employed were the American College Testing Program's mathematics subtest (ACTM) and the Educational Testing Service's Cooperative Mathematics Test, Algebra II (CAT II). The experiment was cross-validating for the former test; norms for the latter two had been cross-validated previously.

All seven full-time members of the LSSC mathematics faculty participated in the advising; each had taught at the college for at least seven years.

A preliminary result, significant at the five percent critical level, suggested that students are more likely to take placement advice when it is tendered by a mathematics professor than when it is the result of testing.

Measurement was based on fall term 1980 mathematics grades. "Success" was defined as a grade of A, B or C; "unsuccessful" students were those receiving a D, F or N. The subjects placed in Intermediate Algebra, a remedial course, were contrasted with those placed in college algebra and in calculus. A "log likelihood test," similar to the chi-square test, was used to analyze the resulting three-way contingency table. "A" students, who "may be capable of accelerating their educational progress by enrolling in the next advanced course,"[188] were then excluded from the sample and a second analysis was performed.

Both analyses were statistically significant. One-degree-of-freedom tests revealed differences between testing and advising for the criterion variables "success" and "course level." Confidence intervals, obtained from the normal approximation to the binomial distribution, showed that students receiving faculty advice tended to be placed at a higher level of mathematics than those placed by tests, and that they were less likely to succeed. There were no interaction effects.

Instructors judged their fall term 1980 mathematics students as being at the right course level, too high or too low. Analysis of these ratings produced a statistically significant difference between counseling and testing at the one-half of one percent critical level. Students rated as "should have taken a lower-level course" occurred in greater proportion in the group placed by faculty members.

Five students were rated as "should have taken a higher-level course," but only two were "A" students. For this reason, and because the analysis with "A" grades omitted was entirely consistent with the results obtained when they were included, it was judged reasonable to include "A" students among the "successfully placed."

One-degree-of-freedom tests failed to reveal any statistically significant differences between the two testing methods for any of the dependent variables; however, the "success" effect was significant at the one-half of one percent level when students who accepted a testing recommendation were compared with those who rejected it. Confidence intervals confirmed that the former group was much more likely to be successful.

An analysis of "naming" versus "coding" course recommendations on the placement roster did not produce a statistically significant result.

A review of the literature indicated that a 70 percent "success" rate is an acceptable standard for mathematics placement. Although other comparative experiments have not been found, the results of predictive studies suggest that "a test found to be a highly valid placement device for one school may be useless in another program."[189]

Loeb points out that "changes in selectivity and possibly in grading standards at the college level have been shown to influence the relative efficiency of predictors of college performance."[190] It may be inferred that placement effectiveness is not necessarily constant over time, especially if the collegiate mathematical calibre changes greatly. The latter is an actuality at LSSC, formerly Sault Branch of Michigan Technological University, where the bulk of students selecting mathematics were once well-prepared for college algebra,[191] while in 1980, more than half were placed at a remedial level.

The following definition will therefore be adopted and used throughout this chapter.

<u>Definition.</u> By a "good test" (or a "good method" or a "good actuarial method," etc.) is meant a test or method for mathematics placement, together with criteria based on that test or method, employed at a particular

institution and satisfying the following:

(1) The institution has adopted a measurable definition of "successful placement" and, within the terms of this definition, the criteria have at least a 70 percent success rate.

(2) The success rate has been cross-validated there within five years of the time the test or method has been employed.

## CONCLUSIONS

1. The Thesing Placement Test is a good test for Lake Superior State College.

2. The ACTM-CAT II combination is probably a good testing method for Lake Superior State College.

3. Mathematics faculty advisors place students at a higher level of beginning college mathematics than a good actuarial method does.

4. Students placed in beginning college mathematics courses by members of that college's mathematics faculty are less likely to succeed than those placed by a good actuarial method.

5. When a good actuarial method is employed, students who accept placement recommendations are more likely to succeed than those who reject the advice.

# DISCUSSION

## The Tests

From the information that has been presented, it is obvious that the Thesing Placement Test satisfies the definition of a "good test" used here. Chances are 99 out of 100 that its success rate is greater than 71 percent for either of the chosen measurements, grades or instructors' ratings, even when conservative assumptions are made about the latter.

While the rate of success based on grades may not actually be 70 percent or more for the standardized tests, the lower bound of the 95 percent confidence interval, 68.97 percent, is so close as scarcely to be worth quarreling about. Moreover, the rate obtained in each of the four years studied was over 70 percent, and the evidence of the instructors' ratings is impressive. The ACTM-CAT II combination, then, is probably a good testing method for LSSC.

In addition to being good placement methods in the sense used here, both discriminate well at all levels of beginning mathematics and not just the Intermediate Algebra/college algebra interface. (The distribution of the pre-calculus final examination in Calculus and Analytic Geometry I classes probably enhances the effects at that level). The literature shows that this is not invariably true of placement procedures.

Neither method requires a great deal of time for administration and "we have the students for orientation anyway."[192] Neither requires additional data such as high school grade point average, number of semesters of academic high-school mathematics, age, etc. This is an important point not previously aired: utilization of such information for placement purposes would be complicated and expensive at LSSC just now. If the college had the facilities of, say, the State University of New York College at Cortland, where the data is available in computer-readable form before orientation,[193] an investigation would be in order to see if discriminant analysis, based on test scores and these other variables, could improve placement. This is not a feasible option at present.

Which testing method is to be preferred? In the opinion of this writer, the answer is, "Definitely, the Thesing Placement Test." Why?

First, because it is a "local product, based on how our students do, not on national norms. It makes more sense."[194] It has "face validity."

Next, because "the content is related to our particular program, our mathematics courses."[195] It has "content validity."

The overall success rate based on grades is slightly higher for the TPT; on the other hand, the ACTM and CAT II are more successful when instructors' ratings are employed for measurement. More important is the simplicity of the TPT's placement guidelines (Table 3), compared to the complicated rules governing use of the standardized tests (Table 5). Most important of all is the fact that a number of students - 24 of the 155 placed by $T_2$ in 1980 - have failed to submit the ACT profile before pre-orientation; this caused some to drop a category or two below the proper one for mathematics placement, while the TPT does not require the supplemental score.

The TPT's reliability coefficient, reported in Chapter One, is higher than those described for the standardized tests in Chapter Two; the latter appear to be inflated and the former is not. Sample size for the standardized tests is larger, but the sample is not nearly so germane to the objectives of LSSC.

Finally, the TPT is less expensive. At $5.50 per 100 answer sheets for the CAT II,[196] this is a minor consideration, but a pleasant one.

Choice of a placement procedure should not mean an end to corroboration. Reliability and validity data for the TPT should be verified at least every five years, and preferably every three.

Comparative experiments in mathematics placement are long overdue, especially in the field of testing. A contrast of two actuarial methods would be child's play: as long as time periods were compatible, for example, randomization could be "handled in the mixed ordering of materials for distribution,"[197] eliminating the need for the assignments and placement slips that were used at LSSC. It is hard to understand why so many "one-shot case studies" have been done, and so few comparisons, especially when the former often purport to compare two tests.

## Clinical Versus Statistical Placement

The third and fourth conclusions are founded partly on the results of the statistical computations: the high significance levels of the one-degree-of-freedom tests and the disparities among confidence intervals for some of the "success" and "remedial placement" ratios. Using the grades criterion, for example, we can be almost certain that the TPT is more than 70 percent successful, and equally sure that the rate for faculty placement is less than 70 percent. Testing methods at LSSC will presumably place more than half the mathematics students at a remedial level; advisors probably won't.

An intuitive, though imprecise, account of the

1980 outcome at LSSC is the following. Faculty members placed entering freshmen at a higher level of mathematics than the tests did; this is shown by the one-degree-of-freedom tests and the confidence intervals for remedial placement. In a number of cases, this echelon was "too high" for the student's ability or experience. Evidence is found in the instructors' ratings, where the relatively large proportion of faculty advisees who "should have taken a lower-level course" could occur by chance only once in 200 times. Lacking the qualifications for success, these students were unable to earn a grade of "C" or better; this is demonstrated by the differing rates of "successful placement" based on grades.

But "has any responsible scientist ever made up his mind about such a matter on the basis of a single experiment? ... in actual practice, a single finding seldom even tempts us to such closure of judgment."[198]

Support for a generalization of these results beyond the bounds of Lake Superior State College and the year 1980 may be found in the literature of "clinical versus statistical prediction." The Sarbin experiment in particular bears a close resemblance to the present one. The criteria were similar, moreso when Sarbin's were dichotomized for "success" and "failure." Especially striking is the fact that Sarbin's counselors "over-predicted" college success, just as the LSSC mathematics faculty "over-predicted" success in mathematics.

The ability of actuarial methods to predict at least as well as, and often better than, human prognosticators, is so thoroughly well-established[199] that a justification for using faculty advisors in this experiment is in order. The author confesses to having been unaware of these results when the project was conceived and to having thought of the Thesing Placement Test "as a floor to which the judgment of the experienced" mathematics teachers could be compared.[200] And indeed, the primary concern of everyone involved was that students placed by testing would not be unfairly treated. By the time the experiment was underway, some reference material had been gathered, but it was not perused until August, 1980; that is, after the first three groups of subjects had been processed. Even then, the material seemed to be of academic interest; it was still anticipated that the teachers would do a better job of placement than the tests, and only when the results began to take shape did the theory assume meaning in this context.

After all, "prediction" and "placement" are not synonomous. "To predict" means "to tell or declare beforehand; foretell; prophesy;" a prediction is "a predicting; also, that which is foretold; a prophecy."[201]

"Placement" is "the act or business of finding jobs, lodgings, or other positions for applicants;"[202] in the industrial setting

> The broad aim of personnel decision making is to estimate or measure as accurately as possible each person's individuality and to place him in an assignment for which his pattern of predicted job behavior is appropriate both to his own long-term goals and to the goals of his employer.[203]

Placement involves manipulation, or "interference;" prediction does not. In fact, this experiment meets neither of the "standard limitations" for "clinical versus statistical prediction" studies.

> The first of these limitations is that the comparative validity has always been evaluated by comparing the correlation between the criterion and the judges' predictions with the cross-validated correlation between the criterion and the predictions of the actuarial model. ...
>
> The second limitation is that both the clinical predictions and those of the actuarial model are made on the basis of the same codable input.[204]

The violation of the first is immaterial. The criterion variables were categorical attributes, for which correlation is an inappropriate analysis; the actuarial results were cross-validated, which is essential. Sarbin used a chi-square test in similar circumstances.

The second distinction is more fundamental. Decisions were not made on the basis of the same input (though it was all codable, saving the interview). With the exception of the ACTM, the data were unique to treatment. In addition, different subjects were - necessarily - assigned to different treatments, whereas in the straight prediction setting, the different methods make forecasts for the same individuals.

It would seem that more research into "clinical versus statistical placement" is in order. The present result, together with the testimony of the "clinical versus statistical prediction" studies, renders some experiments inadvisable, but not all. An inquiry would be especially suitable wherever placement is now done by faculty members.

A long acquaintance with the faculty advisors at LSSC, culminating in the opinion, "if they can't place students better than tests do, nobody can," is the last mainstay for the conclusions. Why didn't they? One can only speculate.

The Director of Counseling and Testing at LSSC says, "When I talk to students, they tend to be overly optimistic about their abilities, so an interview could lead you to think they will perform better than they actually can."[205] The mathematics faculty members also advanced this reason during a discussion of the results.[206]

Without stretching the point too far, it seems that Meehl's description of client-clinician communications (quoted in Chapter Two) resembles some of the exchanges that take place in the classroom and in office hours. The author is reminded of some precalculus students who had trouble with trigonometry and were not greatly aided by an explanatory handout. It was eventually revealed that they did not know what an isosceles triangle was (or an equilateral one). The handout, however, was the springboard for this discovery, and students and teacher became better acquainted in the process. Thus the "bad idea" of the trigonometry handout did not essentially subtract from the "good idea" of a more elementary approach, as it would in a straight prediction situation.

Some insightful comments were made by members of the Department of Earth Science, Mathematics and Physics at LSSC when the results of this research were presented at a recent meeting. Dr. Richard Zabelka suggested that, "Since the final objective of a mathematics course is passing a test, it might make sense to use a test for placement."[207] It would be interesting to try this experiment in fields where classroom endeavors are different.

Dr. Lewis Brown noted that the indictment of cultural bias can scarcely be laid against a mathematics test; this may make it fairer and more effective than testing in some other contexts.[208]

The statement that "almost certainly .. there are characteristics of Lake Superior State College that interact with treatment to limit the generalizability of the results" was made in Chapter Three with the expectation that these results would favor faculty advising and might not then be applicable to large universities, where professors are not so well acquainted with their students; where they may be more committed to research than to teaching, etc. It now seems likely that the actual results would generalize both to similar and larger collegiate settings. Schools whose mathematics faculty numbers but two or three might constitute an exception. Criterion contamination could produce an effect in favor of advising; a "double blind" experiment would be in order in these settings.

The fifth conclusion rests on a special case of "clinical versus statistical placement;" here, the students are the "clinicians."

Both the contingency tables and the confidence intervals show that students placed by tests who accept placement advice are more likely to succeed than those who don't. Techniques to induce acceptance should be contemplated, as they would probably be beneficial to students.

The department heads and other advisors to LSSC students should certainly be apprized of this research and its implications. A short speech to students at pre-orientation, highlighting some of the results and perhaps describing the beginning mathematics courses, as the "placebo" talk did, might aid acceptance. Reinforcement of the test's placement by members of the mathematics faculty would probably have an effect, as students appear likely to take the teachers' advice.

The placement codes seem to be easier to interpret than test scores and will be easier still in the future, as the (*) and (**) categories discussed in Chapter Three can be dispensed with; it would do no harm to continue using the codes. Naming the course recommendations should not be necessary, judging by the outcome of the "naming" versus "coding" analysis.

An area not yet explored is student satisfaction. The TPT is a good test when instructors evaluate its results, either with grades or with ratings. Is it a good test when students do the measuring? If students can be persuaded to select the course to which the TPT directs them, suitable questionnaires could be distributed in fall term classes and unidentified replies collected. However, 22 percent of the students placed by testing in 1980 chose another course, and such a large number could create an undesirable bias in the data. If this cannot be remedied, the "acceptors" and "rejectors" could be identified and different forms distributed to each group; this would preserve anonymity.

## RECOMMENDATIONS

1. The Thesing Placement Test should be used for mathematics placement at Lake Superior State College.[209]

2. Reliability and validity studies should continue, preferably every three years and at least every five. In particular, it would be interesting to determine the "goodness" of the Thesing Placement Test as measured by student-reported data.

3. Means of inducing students to accept the prognoses of the Thesing Placement Test are recommended. Reinforcement by members of the mathematics faculty would probably have an effect.

4. Genuine comparative experiments in mathematics placement should be carried out whenever possible.

5. The phenomenon of "clinical versus statistical placement" probably merits further investigation.

# APPENDIX A

# SUPPLEMENTARY MATERIAL

# APPENDIX A

## TABLE 32

### THESING PLACEMENT TEST DIFFICULTY DATA (1979)

| Difficulty Range* (n = 394) | Number of Test Items in that Range |
|---|---|
| 0 - 79 | 0 |
| 80 - 99 | 3 |
| 100 - 119 | 3 |
| 120 - 139 | 5 |
| 140 - 159 | 6 |
| 160 - 179 | 5 |
| 180 - 199 | 2 |
| 200 - 219 | 7 |
| 220 - 239 | 6 |
| 240 - 259 | 4 |
| 260 - 279 | 1 |
| 280 - 299 | 2 |
| 300 - 319 | 1 |
| 320 - 399 | 0 |
| Total: | 45 |

*The "difficulty" of an item is defined as the number of students who answered the item <u>correctly</u>.

# APPENDIX A

## TABLE 33

## THESING PLACEMENT TEST DISCRIMINATION DATA SUBJECTS RANKED BY TPT SCORE (1979)

| Range of Discrimination Index* | Number of Test Items in that Range |
|---|---|
| -100 - 0 | 0 |
| 1 - 10 | 0 |
| 11 - 20 | 1 |
| 21 - 30 | 4 |
| 31 - 40 | 8 |
| 41 - 50 | 12 |
| 51 - 60 | 11 |
| 61 - 70 | 6 |
| 71 - 80 | 3 |
| 81 - 200 | 0 |
| | Total: 45 |

*The discrimination index of an item is defined here as the percentage of students from the upper one-third who chose the correct answer minus the percentage from the lower one-third who did so.

APPENDIX A

TABLE 34

THESING PLACEMENT TEST DISCRIMINATION DATA
SUBJECTS RANKED BY ACTM SCORE
(1979)

| Range of Discrimination Index* | Number of Test Items in that Range |
|---|---|
| -100 - 0 | 0 |
| 1 - 10 | 1 |
| 11 - 20 | 2 |
| 21 - 30 | 9 |
| 31 - 40 | 18 |
| 41 - 50 | 10 |
| 51 - 60 | 5 |
| 61 - 200 | 0 |
| Total: | 45 |

*See page 163.

# APPENDIX A

## LAKE SUPERIOR STATE COLLEGE
## APPLICATION FOR ADMISSION
## FIRST PAGE

**LAKE SUPERIOR STATE COLLEGE**
OFFICE OF ADMISSIONS, SAULT SAINTE MARIE, MICHIGAN 49783
PHONE: 1-800 682-4800 (Toll free Michigan)
COLLEGE PHONE: (906) 632-6841, Ext. 231

**Application for Admission and Financial Aid**
(See enclosed yellow sheet for instructions)

**PART I** To Be Filled Out By Applicant

Date ____________ 19 ___

LAST NAME | FIRST NAME | MIDDLE NAME

DATE OF BIRTH Mo. Day Year | *female *male *(optional) | Veteran YES NO (DD214 Required) | On Active Duty YES NO

U. S. Citizen? YES NO
If not, which country? ____________

SOCIAL SECURITY NUMBER** (Non U.S. citizens leave blank)

How long have you resided in Michigan? ____ Years Months

Michigan Resident YES NO

Country of birth ____________

State (Province) of birth ____________

PERMANENT HOME ADDRESS City of birth ____________

Number Street | TELEPHONE NUMBER area code number

City | State Province or Country | Zip code

County (if in Michigan) | CURRENT MAILING ADDRESS IF DIFFERENT

Name and address of parents or guardians | Parents resident status ____ State County

Mr. and Mrs. ____ First Name Last Name Number Street City State

List your CURRICULUM SELECTION on this line (See back page) ____________

Indicate length of program: 1 year, 2 years, 3 years, 4 years, Transfer 2+1 | If undecided, indicate Liberal Arts Assoc. Degree

Intended Term of Entrance: Fall, Winter, Spring, Summer 19 ___

Have you ever applied for admission to Lake Superior State College before? If yes, When?

Do you plan to live in a residence hall? YES NO | Do you wish to apply for Financial Aid? YES NO | If you have checked the yes box you must also file a Financial Aid Form (FAF) or Family Financial Statement (FFS) prior to March 1st. (See your counselor for assistance.) Request that a copy of the needs analysis be forwarded to Lake Superior State College.

OPTIONAL
Religious Preference

List High School from which you graduated or will be graduated
NAME OF SCHOOL | ADDRESS | DATE OF H.S. GRADUATION (Use Numbers) Mo Yr

List each College or University attended starting with current enrollment, include LSSC if previously attended
COLLEGE (UNIVERSITY) | ADDRESS | Mo Yr to Mo Yr | Credits earned | Degree earned

Failure to list all institutions previously attended may result in dismissal from LSSC.

**Your Social Security number will become your permanent student number. If you do not wish to provide this number, an alternate student number will be assigned. Note: Non U.S. citizens leave this section blank.

**OFFICIAL USE ONLY**
Res. Nonres. Foreign
RESIDENT - (County)
NONRESIDENT - (State)
FOREIGN - (Country)
Curriculum Advisor
Curriculum Advisor
Res. Hall
Fin. Aid
Veteran
Religion
Ethnic
H.S.
College

**FOR OFFICIAL USE ONLY**

Date Application Received Mo. Day 19 Year | Adm. Couns. | Date of 1st action Mo. Day 19 Year
Pred. GPA ____ | GPA (college Gr 13) | Adm. Couns. | Date of 2nd action Mo. Day 19 Year
G.E.D. ____ | GPA (H.S.) | Adm. Couns. | Date of 3rd action Mo. Day 19 Year
CS/ES | Class Rank

Action: 1st 2nd 3rd
1. WA
2. REJ
3. TA
4. FA
5. on Trial
6. Cancelled
7. Part Time
Cond

___ Grad. Verif.
___ Final Grades
ACT: E M SS NS CUM

ALG ___ GEO ___ TRG ___ SrM ___
BIO ___ CHM ___ PHY ___ SSr ___
ENG ___ LAN ___ Def.

Checks indicate items needed:
H.S. Trans. ____ Coll. Transcript
Fin. Aid App. ____ Coll. Recomm.
ACT Profile ____ App. Fee
DD Form 214 ____ Other
RN/LPN Lic. ____ Ent. exams
BSN app. ____ G.E.D. Score
Mal. Ins. ____ NLN Scores

Comp. Center ___ date | Fin. Aid ___ date
O.S.S. ___ date | Housing ___ date

7/80

# APPENDIX A

## OUTLINE FOR ADVISING STUDENTS FOR FALL 1980 MATHEMATICS COURSE

1. Introduce self - obtain student's name.

2. Obtain and examine student's folder. Inquire about student's intended major.

3. Using High School Transcript (math courses and grades, class rank, etc.) and ACT Scores (math and composite) and inquiry (considering borderline cases) decide the appropriate course for fall:

| | | | | | | |
|---|---|---|---|---|---|---|
| Basic Math | or | MA 100 | or | MA 111<br>MA 181<br>MA 121 | or | MA 132<br>(MA 183?) |

Score guidelines:

Basic Math if no algebra in High School or if grades poor or if taken long ago

MA 100 if no 2nd year algebra in High School

MA 111, 181, or MA 121 if apparently ready

MA 132 if excellent High School grades, has trigonometry, ACT scores high (183?) plus motivation

4. When appropriate course has been decided, tell student, "The recommended course for you is __________. I shall provide the advisor you see this afternoon with the recommendation so that you can enroll in __________."

5. Record recommended course on sheet provided.

6. Make brief notes on why recommended course seems most appropriate.

# APPENDIX A

## COURSE RECOMMENDATION FORM

Student's Name______________________________________________
Last                    First                    Middle or Initial

Math Class Recommended for Fall Term 1980____________________

Why does this course seem most appropriate? (If possible, list factors briefly in order of importance, or indicate the order of importance).

______________________________________
Mathematics Advisor's Signature or Initials

# APPENDIX A

## BEGINNING MATHEMATICS COURSES AND MATHEMATICS PLACEMENT PROCEDURES

### LAKE SUPERIOR STATE COLLEGE

Mathematics Placement 1980 Entering Freshmen

The following courses are those in which most students begin a mathematics sequence at LSSC. A complete list of mathematics offerings can be found on pages 266 to 270 of the 1979-80 college catalog.

Basic Mathematics (not yet in catalog)
Arithmetic operations, exponentials, proportions, percentages; graphing; introduction to algebra.

MA 100 Intermediate Algebra
Elementary operations, first degree equations, one unknown, products and factoring, algebraic fractions, exponents and radicals, quadratic equations, functions and systems of equations.

MA 111 Freshman Mathematics I
Sets, analytic geometry topics, functions and graphs, logarithms, probability, with problem solving applications of these concepts.

MA 121 Pre-Calculus Mathematics
Basic theory of functions including polynomial, exponential, logarithmic and trigonometric functions. Inequalities, topics from analytic geometry and plane trigonometry. Provides the essential background for calculus and subsequent upper level mathematics.

MA 132 Calculus and Analytic Geometry I
Inequalities, analytic geometry, limits, continuity, differentiation and applications of derivatives.

MA 181 Applied Mathematics I
A course in algebra and trigonometry covering elementary operations, functions and graphs, fractions, systems of linear equations, quadratic equations, exponents and radicals, trigonometric functions, graphs of trigonometric functions, vectors and logarithms.

# APPENDIX A

## BEGINNING MATHEMATICS COURSES AND MATHEMATICS PLACEMENT PROCEDURES

MA 207 Principles of Statistical Methods
Descriptive statistics, probability distributions, techniques of statistical inference including tests of hypotheses, selected non-parametric tests.

In addition, the Department of Business and Economics offers

SC 104 Business Mathematics
A comprehensive review of whole numbers, fractions and decimals. Problem solving with percentages, discounts, and payments, payrolls, insurance, taxes, finance charges, depreciation and inventory.

Although these are the courses in which a student is most likely to begin the study of mathematics at LSSC, they are by no means the only possibilities. It is important that a suitable choice of mathematics class be made, one that is neither so elementary that you are bored by the material nor so difficult that you cannot comprehend it. For this reason, we urge you to take the following steps if you should ever wish to take a mathematics class here, or if your curriculum requires that you take one.

### Counseling Center

You may arrange with the receptionist at the Counseling Center for placement by means of a mathematics placement test. There is no fee for this test. The time required is approximately one hour. The counseling center staff will notify your major field advisor of the test results and you and your advisor can make an appropriate selection in consultation.

### Department of Earth Science, Mathematics and Physics

One of the advantages of a relatively small college like LSSC is the opportunities it affords for personal acquaintance with members of the faculty. If you have met a member of the mathematics faculty, or heard about one you think you would like to talk to, or just simply prefer to obtain placement advice from a mathematics professor, you are welcome to do so. Arrange for an appointment with the secretary of the department. Be sure to explain that the conference is for the purpose of mathematics placement, so that there will be time to obtain copies of your high school transcript and ACT profile before the consultation.

# APPENDIX A

MAJOR ADVISORS' INSTRUCTIONS FOR MATHEMATICS PLACEMENT (1980)

Students have been assigned to the following categories.

Category (numbered 1 through 6)

**1. Strongly recommend Basic Mathematics, pre-requisite for MA 100.

**2. Recommend Basic Mathematics; however, students who, in their own judgment, are adequately prepared for MA 100 may pursue that course at their own risk.

3. Strongly recommend MA 100 Intermediate Algebra.

**4. Recommend MA 100 Intermediate Algebra. Students who, in their own judgment, are adequately prepared, may pursue one of the courses: MA 111, MA 121, MA 181 at their own risk.

   MA 111 Freshman Mathematics I - for biology, business, social science, etc.

   MA 121 Pre-Calculus Mathematics - for mathematics, chemistry, engineering, and students wanting a rigorous calculus course.

   MA 181 Applied Mathematics I - for technology.

**5. Recommend MA 111, MA 121, or MA 181 depending on the student's major field.

6. This student should be well prepared to enter directly into MA 132 Calculus and Analytic Geometry if the student has an adequate trigonometry background and ambitious goals. Students in a technology program are encouraged to take MA 132. However, they may wish to consult the mathematics department about testing out of MA 181 and MA 182 so they can enter directly into MA 183 Introduction to Calculus if they have a strong trigonometry background.

**A double asterisk may be attached to these categories for certain students. This indicates that ACT scores have not yet been submitted. Such students should be queried about their high school background to determine placement and, if they are still uncertain about the appropriate mathematics course, should consult a member of the mathematics faculty immediately upon returning in September for a re-evaluation of placement.

APPENDIX A

FACSIMILE OF PAGE FROM PLACEMENT ROSTER

| NAME | EN 101 Regular or EN 184 | EN 101 Honors | EN 101 Remedial | EN 105 Developmental Reading | Mathematics* |
|---|---|---|---|---|---|
| Student #1 | | √ | | | MA 132 |
| Student #2 | √ | | | | MA 132 |
| Student #3 | √ | | | √ | 1** |
| Student #4 | √ | | | | 5 |
| Student #5 | √ | | | | MA 181 |
| Student #6 | | | √ | √ | MA 100 |
| Student #7 | √ | | | | Basic Math |
| Student #8 | √ | | | | 2 |
| Student #9 | √ | | | | 3 |
| Student #10 | √ | | | | MA 100 |
| Student #11 | √ | | | | Basic Math |
| Student #12 | | √ | | | 4 |
| Student #13 | | | √ | √ | MA 181 |
| Student #14 | √ | | | √ | Basic Math |
| Student #15 | √ | | | √ | MA 100 |
| Student #16 | √ | | | | 2** |
| Student #17 | | √ | | | |

*In cases where test scores and other indicators are especially strong in their implications, a particular course has been recommended.

# APPENDIX B

# STATISTICAL CALCULATIONS

# APPENDIX B

## STATISTICAL CALCULATIONS

The statistical calculations for the log likelihood tests of Chapter Four will be found in this Appendix. Computations for the confidence intervals given in that chapter are not presented; they are easily obtained from the formulae presented in the last section of Chapter Three.

Two by two and two by three contingency tables will be called "small tables" in this Appendix. General rules for calculating the G-values of small tables are given. The rules are then applied to the small tables of Chapter Four to show the derivation of their G-values. (A complete list of intermediate steps is shown for the first two tables only). The one-degree-of-freedom test for "naming" versus "coding" course recommendations on the placement roster is also computed and the related small table is displayed.

The fundamental analysis will then be given in detail, followed by the same analysis with "A" grades omitted. Small tables associated with the latter test were not given in Chapter Four. They are shown here, and their G-values calculated.

## G-VALUES FOR SMALL TABLES

Figure 7 represents an arbitrary two by two contingency table. The G-value for Figure 7 is given by the formula

$G = 2 \cdot \{\sum_{i=1}^{4} a_i \ln a_i - \sum_{i=1}^{4} T_i \ln T_i + S \ln S\}$, where ln denotes "logarithm base e," the natural logarithm function, and e is "Euler's number;" its approximate value is 2.7182818.[210]

FIGURE 7

TWO BY TWO CONTINGENCY TABLE

| | | | Totals |
|---|---|---|---|
| | $a_1$ | $a_2$ | $T_1$ |
| | $a_3$ | $a_4$ | $T_2$ |
| Totals | $T_3$ | $T_4$ | S |

An arbitrary two by three contingency table is shown in Figure 8. Its G-value is $G = 2 \cdot \{\sum_{i=1}^{6} a_i \ln a_i - \sum_{i=1}^{5} T_i \ln T_i + S \ln S\}$.[211]

FIGURE 8

TWO BY THREE CONTINGENCY TABLE

| | | | | Totals |
|---|---|---|---|---|
| | $a_1$ | $a_2$ | $a_3$ | $T_1$ |
| | $a_4$ | $a_5$ | $a_6$ | $T_2$ |
| Totals | $T_3$ | $T_4$ | $T_5$ | S |

G-VALUES FOR THE SMALL TABLES OF CHAPTER FOUR

Table 11: G = 2 • {88 ln 88 + 67 ln 67 + 72 ln 72 + 23 ln 23 + 21 ln 21 + 8 ln 8 - 227 ln 277 - 52 ln 52 - 111 ln 111 - 88 ln 88 - 80 ln 80 + 279 ln 279} = 2 • {394.00564 + 281.71441 + 307.92996 + 63.934971 + 16.635532 - 1231.4637 - 205.46467 - 522.75785 - 394.00564 - 350.56213 + 1571.1081} = 2 • 3.1810132 = 6.3620265.

Table 12: G = 2 • {88 ln 88 + 67 ln 67 + 23 ln 23 + 21 ln 21 - 155 ln 155 - 44 ln 44 - 111 ln 111 - 88 ln 88 + 199 ln 199} = 2 • {394.00564 + 281.71441 + 72.116367 + 63.934971 - 781.73089 - 166.50434 - 522.75785 - 394.00564 + 1053.3677} = 2 • 0.14031876 = 0.28063752.

All the original calculations were made with eight decimal digits of accuracy, as shown above; however, for the purpose of clarity, further results for small tables will be rounded to three decimal places.

Table 13: G = 2 • {155 ln 155 + 72 ln 72 + ... - 227 ln 227 - 52 ln 52 - ... + 279 ln 279} = 2 • {781.73089 + ... + 1571.1081} = 6.081.

Table 14: G = 2 • {88 ln 88 + ... - 225 ln 225 - ... + 277 ln 277} = 5.961.

Table 16: G = 2 • {36 ln 36 + ... - 68 ln 68 - ... + 143 ln 143} = 0.043.

Table 17: G = 2 • {68 ln 68 + ... - 89 ln 89 - ... + 214 ln 214} = 6.456.

Table 18: $G = 2 \cdot \{80 \ln 80 + \ldots - 102 \ln 102 - \ldots + 227 \ln 227\} = 9.008$.

Table 19: $G = 2 \cdot \{58 \ln 58 + \ldots - 108 \ln 108 - \ldots + 143 \ln 143\} = 0.004$.

Table 20: $G = 2 \cdot \{108 \ln 108 + \ldots - 144 \ln 144 - \ldots + 214 \ln 214\} = 12.971$.

Table 22: $G = 2 \cdot \{108 \ln 108 + \ldots - 143 \ln 143 - \ldots + 187 \ln 187\} = 24.677$.

Table 23: $G = 2 \cdot \{82 \ln 82 + \ldots - 97 \ln 97 - \ldots + 161 \ln 161\} = 17.145$.

Table 26: $G = 2 \cdot \{72 \ln 72 + \ldots - 81 \ln 81 - \ldots + 208 \ln 208\} = 11.110$.

Table 27: $G = 2 \cdot \{72 \ln 72 + \ldots - 81 \ln 81 - \ldots + 144 \ln 144\} = 0.414$.

Table 28: $G = 2 \cdot \{130 \ln 130 + \ldots - 144 \ln 144 - \ldots + 208 \ln 208\} = 10.696$.

Table 29: $G = 2 \cdot \{130 \ln 130 + \ldots - 144 \ln 144 - \ldots + 181 \ln 181\} = 17.479$.

Table 31: $G = 2 \cdot \{25 \ln 25 + \ldots - 35 \ln 35 - \ldots + 113 \ln 113\} = 3.272$.

## Naming Versus Coding

The reduced nomenclature-outcome contingency table is shown as Table 35.

TABLE 35

REDUCED NOMENCLATURE-OUTCOME CONTINGENCY TABLE

| OUTCOME | NOMENCLATURE Named | Coded | Totals |
|---|---|---|---|
| Student took recommended course | 25 | 10 | 35 |
| Student took a different course | 13 | 10 | 23 |
| Totals | 38 | 20 | 58 |

Table 35: G = 2 • {25 ln 25 + ... - 35 ln 35 - + 58 ln 58} = 0.677.

## THE FUNDAMENTAL ANALYSIS

Tables 36 and 37 together represent the contingency table for the fundamental analysis and all relevant totals. The data has been rearranged and notation altered so that computations could be patterned after an example of Sokol and Rohlf's.[212] Treatment, or placement method, is represented by the letter P; success by S (U indicates unsuccessful subjects); and mathematics course level by M. Let A, B and C denote, respectively, the number of levels of each variable. Then A = 3, B = 2 and C = 2.

a. P x S x M = $\sum$ f ln f = 27 ln 27 + 31 ln 31 + 20 ln 10 + 9 ln 9 + 24 ln 24 + 26 ln 26 + 16 ln 8 + 11 ln 11 + 50 ln 25 = 642.84343.

b. P x S = $\sum$ f ln f = 58 ln 58 + 19 ln 19 + 50 ln 50 + 16 ln 16 + 36 ln 36 + 35 ln 35 = 784.85647.

TABLE 36

CONTINGENCY TABLE FOR THE FUNDAMENTAL ANALYSIS

| PLACEMENT METHOD | SUCCESS | MATHEMATICS COURSE LEVEL: Intermediate Algebra | College Algebra and Calculus | Totals |
|---|---|---|---|---|
| $T_1$ | S | 27 | 31 | 58 |
| | U | 9 | 10 | 19 |
| | Totals | 36 | 41 | 77 |
| $T_2$ | S | 24 | 26 | 50 |
| | U | 8 | 8 | 16 |
| | Totals | 32 | 34 | 66 |
| $T_3$ | S | 11 | 25 | 36 |
| | U | 10 | 25 | 35 |
| | Totals | 21 | 50 | 71 |
| Grand Totals | | 89 | 125 | 214 |

TABLE 37

SUCCESS-COURSE LEVEL CONTINGENCY TABLE

| SUCCESS | MATHEMATICS COURSE LEVEL: Intermediate Algebra | College Algebra and Calculus | Totals |
|---|---|---|---|
| S | 62 | 82 | 144 |
| U | 27 | 43 | 70 |
| Totals | 89 | 125 | 214 |

c. $P \times M = \sum f \ln f = 36 \ln 36 + 41 \ln 41 + 32 \ln 32 + 34 \ln 34 + 21 \ln 21 + 50 \ln 50 = 771.59907.$

d. $M \times S = 62 \ln 62 + 82 \ln 82 + 27 \ln 27 + 43 \ln 43 = 867.95251.$

Marginal Totals

e. $P = 77 \ln 77 + 66 \ln 66 + 71 \ln 71 = 9.3.6405.$

f. $M = 89 \ln 89 + 125 \ln 125 = 1003.0279$

g. $S = 144 \ln 144 + 70 \ln 70 = 1013.0478.$

Grand Total

h. Grand total $= 214 \ln 214 = 1148.3189$

The overall G-value is $2 \cdot \{a - e - f - g + 2h\} = 2 \cdot \{9.76503\} = 19.53006$. Degrees of freedom $= ABC - A - B - C + 2 = 3\ 2\ 2 - 3 - 2 - 2 + 2 = 7$.

(1) Test of the hypothesis that Mathematics Course Level is independent of Placement Method. $G = 2 \cdot \{c - e - f + h\} = 2 \cdot \{3.37068\} = 6.49914$. D.f. $= AC - A - C + 1 = 2$.

(2) Test of the hypothesis that Success is independent of Placement Method. $G = 2 \cdot \{b - e - g + h\} = 2 \cdot \{6.48707\} = 12.97414$. D.f. $= AB - A - B + 1 = 2$.

(3) Test of the hypothesis that Mathematics Course Level and Success are mutually independent. $G = 2 \cdot \{d - f - g + h\} = 2 \cdot \{0.19571\} = 0.39142$. D.f. $= BC - B - C + 1 = 1$.

## THE VARIABLE Y

Tables 38 and 39 are similar to Tables 36 and 37, and the notation is identical. However, students achieving a grade of "A" have been omitted from the count of "successful" students.

a. P x S x M = 48 ln 24 + 30 ln 10 + 9 ln 9 + 19 ln 19 + 15 ln 15 + 16 ln 8 + 20 ln 20 + 25 ln 25 = 511.62186.

b. P x S = 48 ln 48 + 19 ln 19 + 34 ln 34 + 16 ln 16 + 30 ln 30 + 35 ln 35 = 632.49277.

c. P x M = 34 ln 34 + 33 ln 33 + 27 ln 27 + 23 ln 23 + 20 ln 20 + 45 ln 45 = 627.59943.

d. M x S = 53 ln 53 + 59 ln 59 + 28 ln 28 + 43 ln 43 = 701.71938.

Marginal Totals

e. P = 67 ln 67 + 50 ln 50 + 65 ln 65 = 748.65073.

f. M = 80 ln 80 + 102 ln 102 = 822.30936.

g. S = 112 ln 112 + 70 ln 70 = 825.86654.

Grand Total

h. Grand total = 182 ln 182 = 947.12922.

The overall G-value is 2 • {a - e - f - g + 2h}= 18.10734. D.f. = 7; dimensions have not been altered.

(1) Test of the hypothesis that Mathematics Course Level is independent of Placement Method. G = 2 • {c - e - f + h} = 7.53712. Contingency tables for the one-degree-of-freedom tests are Tables 40 and 41.

TABLE 38

CONTINGENCY TABLE FOR THE FUNDAMENTAL ANALYSIS (A GRADES OMITTED)

| PLACEMENT METHOD | SUCCESS | MATHEMATICS COURSE LEVEL: Intermediate Algebra | College Algebra and Calculus | Totals |
|---|---|---|---|---|
| $T_1$ | S | 24 | 24 | 48 |
| | U | 9 | 10 | 19 |
| | Totals | 33 | 34 | 67 |
| $T_2$ | S | 19 | 15 | 34 |
| | U | 8 | 8 | 16 |
| | Totals | 27 | 23 | 50 |
| $T_3$ | S | 10 | 20 | 30 |
| | U | 10 | 25 | 35 |
| | Totals | 20 | 45 | 65 |
| Grand Totals | | 80 | 102 | 182 |

TABLE 39

SUCCESS-COURSE LEVEL CONTINGENCY TABLE (A GRADES OMITTED)

| SUCCESS | MATHEMATICS COURSE LEVEL: Intermediate Algebra | College Algebra and Calculus | Totals |
|---|---|---|---|
| S | 53 | 59 | 112 |
| U | 27 | 43 | 70 |
| Totals | 80 | 102 | 182 |

TABLE 40

TESTING METHOD-COURSE LEVEL CONTINGENCY TABLE
(A GRADES OMITTED)

| COURSE LEVEL | TREATMENT GROUP $T_1$ | $T_2$ | Totals |
|---|---|---|---|
| Intermediate Algebra | 33 | 27 | 60 |
| College Algebra and Calculus | 34 | 23 | 57 |
| Totals | 67 | 50 | 117 |

TABLE 41

PLACEMENT METHOD-COURSE LEVEL CONTINGENCY TABLE
(A GRADES OMITTED)

| COURSE LEVEL | TREATMENT GROUP $T_1 + T_2$ | $T_3$ | Totals |
|---|---|---|---|
| Intermediate Algebra | 60 | 20 | 80 |
| College Algebra and Calculus | 57 | 45 | 102 |
| Totals | 117 | 65 | 182 |

Table 40: $G = 2 \cdot \{33 \ln 33 + \ldots - 60 \ln 60 - \ldots + 117 \ln 117\} = 0.258.$

Table 41: $G = 2 \cdot \{60 \ln 60 + \ldots - 80 \ln 80 - \ldots +$

182 ln 182} = 7.279.

(2) Test of the hypothesis that Success is independent of Placement Method. G = 2 • {b − e − g + h} = 10.20944. Tables 42 and 43 are associated with the one-degree-of-freedom tests.

TABLE 42

TESTING METHOD-SUCCESS CONTINGENCY TABLE
(A GRADES OMITTED)

| SUCCESS | TREATMENT GROUP $T_1$ | $T_2$ | Totals |
|---|---|---|---|
| Successful | 48 | 34 | 82 |
| Unsuccessful | 19 | 16 | 35 |
| Totals | 67 | 50 | 117 |

TABLE 43

PLACEMENT METHOD-SUCCESS CONTINGENCY TABLE
(A GRADES OMITTED)

| SUCCESS | TREATMENT GROUP $T_1 + T_2$ | $T_3$ | Totals |
|---|---|---|---|
| Successful | 82 | 30 | 112 |
| Unsuccessful | 35 | 35 | 70 |
| Totals | 117 | 65 | 182 |

Table 42: $G = 2 \cdot \{48 \ln 48 + \ldots - 82 \ln 82 - \ldots + 117 \ln 117\}$ $= 0.181$.

Table 43: $G = 2 \cdot \{82 \ln 82 + \ldots - 112 \ln 112 - \quad + 182 \ln 182\}$ $= 10.029$.

(3) Test of the hypothesis that Mathematics Course Level and Success are mutually independent. $G = 2 \cdot \{d - f - g + h\} = 1.3454$.

BIBLIOGRAPHY

Aleamoni, Lawrence M., "Can Grade Point Average Be More Accurately Predicted?" Journal of Educational Psychology, 69(3), 227-230, 1977.

The ACT Standard Research Service Report, Lake Superior State College, Summary Analysis, Standard Report, 1976-77 Freshmen, The American College Testing Program, Iowa City, Iowa, 1977.

The ACT Standard Research Service Report, Lake Superior State College, Summary Analysis, Standard Report, 1978-79 Freshmen, The American College Testing Program, Iowa City, Iowa, 1979.

American Heritage Dictionary of the English Language, American Heritage Publishing Co., New York, 1970.

Anastasi, Anne, Psychological Testing (fourth edition), McMillan, New York, 1976.

Bloom, Benjamin S. (Ed.) et al., Taxonomy of Educational Objectives Handbook I: Cognitive Domain, Longman, Inc., New York, 1977.

Bloom, Benjamin S. and Peters, Frank R., The Use of Academic Prediction Scales for Counseling and Selecting College Entrants, The Free Press of Glencoe, Inc., New York, 1961.

Boyce, Richard W., "The Prediction of Achievement in College Algebra," Educational and Psychological Measurement, 24(2), 419-420, 1964.

Buros, Oscar K. (Ed.), The Sixth Mental Measurements Yearbook, The Gryphon Press, Highland Park, New Jersey, 1965.

Buros, Oscar K. (Ed.), The Seventh Mental Measurements Yearbook, The Gryphon Press, Highland Park, New Jersey, 1972.

Buros, Oscar K. (Ed.), The Eighth Mental Measurements Yearbook, The Gryphon Press, Highland Park, New Jersey, 1978.

# BIBLIOGRAPHY

Campbell, Donald T. and Stanley, Julian C., Experimental and Quasi-Experimental Designs for Research, Rand McNally College Publishing Co., Chicago, 1966.

Cook, Desmond L., "The Hawthorne Effect in Educational Research," Phi Delta Kappan, 44, 116-22, 1962.

Corotto, Loren V., "The Prediction of Success in Initial College Mathematics Courses," The Journal of Educational Research, 56(5), 268-71, January, 1963.

Dawes, R. M. and Corrigan, Bernard, "Linear Models in Decision Making," Psychological Bulletin, 81(2), 95-106, 1974.

Dunn, James E., "A Study of the University of Arkansas Mathematics Entrance Exam as a Placement Device," The Journal of Experimental Education, 34(3), 62-8, 1966.

Dunnette, Marvin D., Personnel Selection and Placement, Brooks/Cole Publishing Co., Monterey, California, 1966.

Edwards, Ronald R., "The Prediction of Success in Remedial Mathematics Courses in the Public Community Junior College," The Journal of Educational Research, 66(4), 157-60, 1972.

Everitt, B. S., The Analysis of Contingency Tables, Chapman and Hall, London, 1979.

Fishman, Joshua A. and Pasanella, Ann K., "College Admission-Selection Studies," Review of Educational Research, 30(4), 298-310, 1960.

Foshay, W. R., et al., "A Multivariate Analysis of Academic Advising," ERIC Document ED 155 194, April, 1977.

Frisbie, David A., "Some Limitations of Expectancy Data Used with College-Level Course Placement," Measurement and Evaluation in Guidance, 9(4), 166-71, 1977.

Glass, Gene V. and Stanley, Julian C., Statistical Methods in Education and Psychology, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1970.

Goldberg, Lewis R., "Diagnosticians vs. Diagnostic Signs: the Diagnosis of Psychosis vs. Neurosis from the MMPI," Psychological Monographs, 79 (9, Whole No. 602), 1965.

Goldberg, Lewis R., "Man Versus Model of Man: a Rationale Plus Some Evidence for a Method of Improving Clinical Methods," Psychological Bulletin, 73(6), 422-432.

Goldman, Roy D. and Slaughter, Robert E., "Why College Grade Point Average is Difficult to Predict," Journal of Educational Psychology, 68(1), 9-14, 1976.

Hays, William L., Statistics, Holt, Rinehart and Winston, New York, 1963.

Hoel, Paul G., Elementary Statistics, John Wiley and Sons, New York, 1960.

Hooper, Herbert L., Jr., "The Development of a Prediction Model to Determine the Appropriate Level of Entering Students in Mathematics," ERIC Document ED 065 336, September, 1971.

Howlett, John L., "A Study of Placement Methods for Entering College Freshmen in the Proper Mathematics Sequence at Michigan Technological University," The Mathematics Teacher, December, 1969, 651-659.

Humphreys, Lloyd G., "Prediction of Grades Is Fleeting; a Comment," Journal of Educational Psychology, 68(5), 521, 1976.

Johnson, Van C., "An Assessment of the Motivation Factor in the Estimation of Academic Achievement of Eleventh Grade Indian Students and the Factored Dimensions of the M-Scales an Exploratory Study," Ph. D. Thesis, Michigan State University, 1963.

Kohler, Emmett T., "The Relationship Between the Cooperative Mathematics Test, Algebra III, ACT Mathematics Usage Test, ACT Composite and Grade Point Average in College Algebra," Educational and Psychological Measurement, 33, 929-31, 1973.

Krathwohl, David R., et al., Taxonomy of Educational Objectives Handbook II: Affective Domain, David McKay Company, New York, 1974.

Krathwohl, William C., "Predictions of Average Class Achievement by Means of Aptitude Tests," Journal of Engineering Education, 37, 234-42, 1946.

Kurtz, Ray, "Who Should Place College Freshmen in Mathematics?" The Mathematics Teacher, November, 1969, 557-9.

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

Li, Jerome C. R., Statistical Inference, Edwards Brothers, Ann Arbor, Michigan, 1964.

Loeb, Jane, "High School Performance as Predictive of College Performance," NASSP Bulletin, September, 1972, 19-26.

MCTM Monograph on the state of mathematical preparedness of entering college students (mimeographed, untitled, undated).

Mauger, Paul A., "Is the Prediction of Grades Fleeting Only in Illinois?" Journal of Educational Psychology, 68(5), 520, 1976.

Meehl, Paul E., Clinical Versus Statistical Prediction, University of Minnesota Press, Minneapolis, 1954.

Meehl, Paul E., "Seer Over Sign: the First Good Example," Journal of Experimental Research in Personality, 1, 27-32, 1965.

Mehrens, William A., Readings in Measurement and Evaluation in Education and Psychology, Holt, Rinehart and Winston, New York, 1976.

Mehrens, William and Lehmann, Irvin J., Measurement and Evaluation in Education and Psychology (Second Edition), Holt, Rinehart and Winston, New York, 1978.

Morgan, William P., "Prediction of Success in Junior College Mathematics," The Mathematics Teacher, March, 1970, 260-3.

Nolan, Edwin J., "The Relationship Between ACT Sub-Test Scores and Grades Earned: a Correlational Study," ERIC Document ED 131 902, 1977.

Pugh, Richard, et al., "An Analysis of Achievement Behavior in Selected Mathematics Courses," ERIC Document ED 087 634, August, 1969.

Reilly, Richard R., "Student Placement by a Maximum Likelihood Procedure," Measurement and Evaluation in Guidance, 7(1), 8-15, 1974.

Rozeboom, William W., "The Fallacy of the Null-Hypothesis Significance Test," Psychological Bulletin, 57, 416-28, 1960.

Sarbin, Theodore R., "A Contribution to the Study of Actuarial and Individual Methods of Prediction," American Journal of Sociology, 48, 593-602, 1943.

Sawyer, Jack, "Measurement and Prediction, Clinical and Statistical," Psychological Bulletin, 66(3), 178-200, 1966.

Shana'a, Joyce Adrian, "A Statistical Analysis of the Placement Program in Mathematics for Freshmen at the University of Oklahoma," Ph. D. Thesis, University of Oklahoma, Norman, 1966.

Shevel, Linda R., and Whitney, Douglas R., "Predictive Validity of the Mathematics Placement Examination," Educational and Psychological Measurement, 29, 895-901, 1969.

Siegle, William F., "Prediction of Success in College Mathematics at Washburn University," Journal of Educational Research, 47, 577-588, 1954.

Sokal, Robert R. and F. James Rohlf, Biometry, W. H. Freeman and Company, San Francisco, 1969.

Turner, V. Dean, "Prediction of Success as a Mathematics Major at the Minnesota State Colleges," Ph. D. Thesis, University of Oklahoma, Norman, 1968.

Webster's New Collegiate Dictionary, G. & C. Merriam Co., Springfield, Massachusetts, 1951.

Wick, Marshall E., "A Study of the Factors Associated with Success in First-Year College Mathematics," The Mathematics Teacher, November, 1965, 642-8.

Wheeler, Robert G., "Computer Predicts Student Success as They Register," College Management, April, 1971, 36-38.

## FOOTNOTES

1. Shana'a, pp. 2-3.

2. Lake Superior State College 1980-81 Catalog, p. 270.

3. Ibid., pp. 270-1; curriculum information is from an undated memorandum provided by the LSSC Counseling Center.

4. From an undated memorandum provided by the LSSC Counseling Center.

5. Conversations with Dr. Gary Thesing, 1978-79.

6. Paraphrased from an undated memorandum provided by the LSSC Counseling Center.

7. Conversations with Dr. Gary Thesing, 1978-79.

8. A computer program written in the FORTRAN language by this author was used for the analysis.

9. Allan Lange, et al., "Using Item Analysis to Improve Tests," in Mehrens (1976), pp. 120-3.

10. Mehrens and Lehmann (1978), pp. 273-303.

11. From an undated memorandum provided by the LSSC Counseling Center.

12. From an undated memorandum written by Dr. Reino Hakala for the Department of Earth Science, Mathematics and Physics at LSSC.

13. The norms were determined by Dr. Thesing and this author; further description is in Chapter Three.

14. Kurtz, p. 559.

15. Meehl (1954), p. 127.

16. Shana'a, p. 1 and p. 85.

17. Morgan, pages 260, 261 and 263.

18. Conversation with Professor Douglas Hall, 1979.

19. Campbell and Stanley, p. 6.

20. Ibid.

21. Johnson, p. 11.

22. Mehrens and Lehmann (1978), pp. 394-5; Anastasi, p. 21.

23. Wimburn L. Wallace, in Buros (1972), p. 613.

24. John R. Hills, in Buros (1978), p. 623.

25. Wallace, op. cit., p. 614.

26. Hills, op. cit., p. 623.

27. Ibid., p. 624.

28. Wallace, op. cit., p. 615.

29. Paul Blommers, in Buros (1965), p. 887.

30. Ibid.

31. Conversation with Mr. John Truckey, 1980.

32. Kenneth J. Travers, in Buros (1972), p. 895.

33. Ibid.

34. Ibid.

35. Robert A. Forsyth, in Buros (1978), p. 445.

36. Ibid.

37. Travers, op. cit., 895. This is a reference to Bloom (1977); see Chapter Three.

38. Wallace, op. cit., p. 614.

39. ACT (1977), tables T-2.1, T-2.2 and T-3.2.

40. Hays, Chapter 15; Shana'a, p. 47; Frisbie, pp. 168-9.

41. ACT (1979), tables T-2.1, T-2.2 and T-3.2.

42. Shana'a, pp. 16-17.

43. Ibid., p. 21.

44. Ibid., p. 84.

45. Ibid., p. 47.

46. Ibid.

47. Ibid., p. 33.

48. Ibid., p. 85.

49. Ibid., p. 45.

50. Ibid., p. 45 and p. 47.

51. Turner, p. 67.

52. Ibid., pp. 15-16.

53. Corotto, p. 268.

54. Ibid.

55. Ibid., p. 271 and 269.

56. Ibid., p. 268.

57. Boyce, p. 419.

58. Ibid.

59. Ibid., p. 420.

60. Ibid.

61. Ibid., p. 419.

62. Wick, p. 643.

63. Ibid., pp. 646-7.

64. Ibid., pp. 647-8.

65. Dunn, pp. 62-8.

66. Ibid., p. 65.

67. Kohler, p. 929 and p. 930.

68. Ibid., pp. 921-2.

69. Ibid., pp. 929-30.

70. Ibid.

71. Pugh, p. 1.

72. Ibid.

73. Howlett, p. 654.

74. Ibid., p. 655.

75. Ibid., p. 652.

76. Shevel and Whitney, p. 900.

77. Ibid., p. 896.

78. Morgan, p. 260.

79. Ibid., p. 261.

80. Ibid.

81. Ibid., p. 262.

82. Ibid., p. 263

83. Hooper, pp. 6-7.

84. Ibid., p. 4.

85. Ibid., p. 8 and p. 7.

86. Ibid., p. 4.

87. Ibid., p. 16.

88. Edwards, p. 157.

89. Ibid., pp. 159-60.

90. Reilly, p. 8.

91. Ibid., p. 9.

92. Ibid., p. 10.

93. Ibid.

94. Ibid., pp. 13-14.

95. Ibid., pp. 12-13.

96. Ibid., p. 13.

97. Nolan, p. 4.

98. Ibid., p. 4 and p. 6.

99. Ibid., p. 4.

100. MCTM Monograph, p. 27.

101. Ibid., p. 28.

102. Ibid., p. 27.

103. Kurtz, p. 557.

104. Ibid., p. 559.

105. Foshay, et al., p. 1.

106. Ibid.

107. Frisbie, p. 171.

108. Cf. Aleamoni, Humphreys, Mauger.

109. Fishman and Pasanella, p. 298.

110. Ibid., pp. 298-9.

111. Ibid., p. 300.

112. Ibid., p. 301.

113. Ibid.

114. Ibid., p. 303.

115. Lavin.

116. Loeb, pp. 19-23.

117. Ibid., pp. 24-5.

118. Bloom and Peters, p. 110 and p. 109.

119. Fishman and Pasanella, p. 302; Bloom and Peters, Chapters 2, 3, 4, 5.

120. The author is indebted to Dr. Susan Ratwik for suggesting this line of investigation.

121. Sarbin, p. 593.

122. Ibid.

123. Ibid., p. 594.

124. Ibid.

125. Ibid., p. 593.

126. Ibid., p. 594.

127. Ibid.

128. Ibid.

129. Ibid., p. 596.

130. Ibid.

131. Ibid., p. 597.

132. Dawes and Corrigan, pp. 96-7; some references have been omitted.

133. Meehl (1954), pp. 120-22; emphasis in the original.

134. Ibid., p. 122 and p. 123.

135. Dawes and Corrigan, p. 105.

136. Ibid., p. 95 and pp. 102-3.

137. They are very similar; see, e.g., Anastasi, p. 17.

138. Campbell and Stanley, p. 25 and p. 6.

139. Ibid., p. 5 and p. 8.

140. Ibid., pp. 5-6.

141. Ibid., p. 20.

142. Ibid., p. 21.

143. Memorandum to Mr. John Truckey from the author, July 1, 1980.

144. Placement Roster for 1980 entering freshmen at LSSC.

145. From an undated memorandum to LSSC faculty advisors from the LSSC Counseling Center (cover memorandum to 1980 Placement Rosters).

146. Glass and Stanley, pp. 510-12.

147. Memorandum to Mr. John Truckey from the author, July 6, 1980.

148. The author is indebted to Dr. Susan H. Ratwik for her advice and assistance on obtaining these norms.

149. Biserial correlation coefficients were obtained for the relationship between "success" (0 or 1) and TPT score at each of three levels of mathematics. For Intermediate Algebra, the college algebra group and Calculus I these were, respectively, $r = 0.44$; $r = 0.56$; and $r = 0.34$. They were not, however, used in obtaining the cutoff scores.

150. Conversation with Dr. Arthur Duwe, 1980.

151. Lavin, p. 14.

152. Howlett, for example, used A=1, B=2, C=3, D=4, and F=5. This was "not consistent with the normal grading system at the university, but because of the design of the computer program it was the most efficient means of operation with the system used... Results were .. translated into the normal university four-point grading system of A=4." (Howlett, p. 652).

153. Reilly, p. 8.

154. Ibid., p. 9.

155. Shana'a, p. 30 and p. 33.

156. Ibid., p. 33.

157. Reilly, p. 9. The words "regular" and "remedial" have been reversed, apparently by mistake, in the summary statements of criteria that appear on this page.

158. Ibid.

159. Shana'a, p. 26.

160. Lavin, pp. 14-16; emphasis in the original. Footnotes have been omitted.

161. Shana'a, p. 6; footnote omitted.

162. Ibid., pp. 6-7.

163. Bloom et al., (1977), p. 12.

164. Ibid., p. 15.

165. Ibid., from the table of contents (pages unnumbered).

166. David R. Krathwohl et al., p. 24.

167. Ibid., p. 28.

168. Ibid., from the table of contents, pp. xii-xiv.

169. This measurement was suggested by Dr. William Mehrens.

170. Shana'a, p. 8.

171. Ibid., p. 27.

172. Ibid.

173. Ibid., pp. 27-8. The reference is to Fishman and Pasanella.

174. The author is indebted to Dr. David Behmer, Dr. Larry Lezotte and Dr. Susan Ratwik for assistance with this section.

175. Glass and Stanley, Chapters 15 and 17; Hays, Chapter 12.

176. Glass and Stanley, p. 340; Hays, p. 364.

177. Hoel, p. 162.

178. Ibid., p. 163.

179. Hoel, pp. 157 ff.; Hays, pp. 589 ff.; Glass and Stanley, pp. 329 ff.

180. Everitt, p. 79.

181. Sokal and Rohlf, pp. 560-1.

182. Ibid., pp. 602-6.

183. Ibid.; Li, pp. 499-500.

184. Hoel, pp. 62-68, 95-98 and 240.

185. Tabled values of $\chi^2$ are taken from Hoel, p. 245.

186. Memorandum to instructors of beginning mathematics classes at LSSC, fall term 1980, from the author; November [sic], 1980.

187. Personal communication from Mr. William Gough, November, 1980.

188. Shana'a, p. 30 and p. 33.

189. Reilly, p. 13.

190. Loeb, p. 25.

191. Conversations with Professor Emeritus Franklin Otis, 1980, and Dean Emeritus C. Ernest Kemp, 1981.

192. Conversation with Mr. John Truckey, 1981.

193. Wheeler, pp. 36-7.

194. Conversation with Mr. John Truckey, 1981.

195. Ibid.

196. Buros (1978), p. 444.

197. Campbell and Stanley, p. 26.

198. Rozeboom, p. 420.

199. Meehl (1954); Dawes and Corrigan; Dunnette, p. 171.

200. Dawes and Corrigan, p. 97.

201. _Webster's New Collegiate Dictionary_, p. 665.

202. _American Heritage Dictionary of the English Language_, p. 539.

203. Dunnette, p. 183; emphasis added.

204. Dawes and Corrigan, p. 97; emphasis in the original.

205. Conversation with Mr. John Truckey, 1981.

206. January 14, 1981, meeting of the Department of Earth Science, Mathematics and Physics at LSSC.

207. Ibid.

208. Ibid.

209. Inquiries about the Thesing Placement Test may be directed to Dr. Gary Thesing, Associate Professor of Mathematics, Department of Earth Science, Mathematics and Physics, Lake Superior State College, Sault Ste. Marie, MI 49783.

210. This formula was derived from the instructions on p. 599, Sokol and Rohlf.

211. Ibid.

212. Ibid., pp. 601-607; the example on pp. 602-6 has been emulated.