1971

# The impact of method of evaluation upon achievement in elementary foreign language courses

Barbara Von Wittich
*Iowa State University*

# The impact of method of evaluation upon achievement in elementary foreign language courses

by

Barbara von Wittich

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of
The Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Major Subject: Education

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Area

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Of Science and Technology
Ames, Iowa

1971

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

# INTRODUCTION

One reason for the declining interest in foreign languages in American colleges and universities may very well be the failure of the foreign language profession to provide the majority of students with a successful learning experience. Since academic success and failure traditionally are expressed in terms of grades, one must agree that evaluation policies are "a serious matter" (61) provoking "profound educational questions" (30). Their psychological and educational impact can enhance or hamper the learning process.

## Purpose of Investigation

The present investigation is concerned with the effect of grading methods upon achievement in the area of foreign language learning. Its primary goal is

(1) to find concrete evidence as to the effect of certain grading methods upon student achievement in elementary foreign language courses;

(2) to contribute to the clarification of the pass-fail issue.

## Theoretical Background of Investigation and Review of Related Research

The theory behind this investigation is that student evaluation in elementary foreign language courses should be based on the assumption that language learning is a function of the following factors:

(1) The academic environment

(2) The cumulative aspect[1] of language learning

(3) The time element in language learning

(4) The power of grades.

Each of these factors will be discussed briefly in connection with recent research done in these areas.

The academic environment

Language study at a college or university takes place in a complex educational setting where each course makes its demand upon the student's time. The situation is quite different in an institution with humanistic trends and in a university with main emphasis on the sciences and technology, like Iowa State University, where students enroll in foreign language classes mainly to fulfill graduation requirements, while carrying average course loads of 16 to 17 credit hours per quarter.

In this type of institution, foreign language instruction as well as evaluation procedures at the elementary level should be approached in such a way as to facilitate the learning process, to motivate and encourage the learner, and to reduce the number of dropouts by assuring a maximum of success with a minimum of time required to master the learning tasks. Any other approach is unrealistic and opposed to Ortega's (72) "principle of economy in education" which takes into consideration the student's limited capacity to learn. An application of this principle is the establishment of clear-cut "behavioral objectives"[2] in language learning. "What made the difference" Steiner (102) said, after an ex-

[1] The student must have mastered one step before being able to take the next one.

[2] Clearly defined learning tasks.

periment with "performance objectives" in a French course, "was that the students knew what to expect" and that "they could concentrate on learning the material rather than spending their time guessing what would be on the test." According to this method, the student is judged in terms of "learning outcomes" (39), in terms of "what he can actually do rather than in terms of the amount of time he has spent studying."

The cumulative aspect of language learning

Elementary language learning is a cumulative process, essentially, the learning of a skill. Therefore, frequent evaluation of student achievement is necessary to decide if the learner is ready to take the next step.

Because of this cumulative nature of language learning, the emphasis upon mastery of clearly defined learning tasks is of greatest importance. Politzer (82) said that "every experienced language teacher knows the chances that the pupil who does not have the foundations will catch up and at the same time learn the new material are practically nil."

Also Pimsleur et al. (78), in their study of underachievement, emphasized the cumulative aspect of second language learning. Of the students who earn an "A" the first year, more than half will get a lower grade the second year. Unless the student really learns, unless he masters the material presented in the first year, rather than merely "covering" it, he will be unable to succeed in the second course. They stated:

> "In the average King City foreign language class, six students out of 30 are under-achievers. The number of under-achievers would be higher still, except that half the students drop out of the language course after the first year; by the third year nine-tenth of the students have dropped. Yet, no provision is made to identify under-achievers, diagnose the reasons for their foreign language difficulty, and help them to overcome it."

This is quite an indictment of the pedagogical inertia of foreign language teachers, and the conclusion reached by Pimsleur et al. (78) in a research project conducted under the auspices of the Ohio State University Foundation involving a sample of 10,000 high school students. This kind of fatalistic attitude toward underachievement is symptomatic not only of the high school level but even more so of the colleges and universities where adverse conditions concerning foreign languages are stoically accepted as inevitable. A third-quarter elementary foreign language course, thus, often becomes a torture for both students and instructor because basic subject matter had been covered but not mastered in the preceding courses.

The "mastery" approach, though quite naturally used in the acquisition of skills such as typing, swimming, piano playing, etc., up to now has been applied to academic subject matter predominantly in experimental and laboratory settings. One of the still rare exceptions is a vocational teacher education program at the University of Massachusetts where "the meeting of performance criteria is pretty much an all-or-nothing proposition," and where "conventional grading has become obsolete. Students meet the performance criteria or they don't. It is a pass-fail system" (49). This pass-fail system, however, is quite different from the pass-fail system at Iowa State University which will be discussed later.

The history of mastery learning which goes back to the 1920's in the United States has been described by Block (11) who also gathered a bibliography accompanied by abstracts which summarize the most relevant research in this area. Kim et al. (53) found that mastery learning strategies were most effective for the slower student.

The time factor in foreign language learning

Basic to the mastery learning concept is the assumption that, given sufficient time, every student can attain a mastery level of initially established objectives. This, according to Valette (110) and others, means, for instance, that "all students can learn a foreign language, although some will take more time." The "special gift myth," she says, is being used by teachers, students, and administrators to justify the great number of failures in foreign language classes. The traditional set of expectations whereby the teacher assumes that one-third will fail or hardly get by, and others will learn some but not enough, according to Bloom (14) "is the most wasteful and destructive aspect of the present educational system."

Carroll (20) equates aptitude with the time needed to master a task. In this sense, students with a high aptitude are those who learn more rapidly. Also Washburne (114) stated that failure is an unnecessary phenomenon in foreign language learning if time becomes the variable, and success the constant. Strangely enough, in American education, at least as far as foreign languages are concerned, time has been magically equated with competence. Thus, for instance, college foreign language requirements are defined, not in terms of competence levels, but "in terms of the

number of hours (measured in 'years') spent sitting in the language classroom" (111).

The power of grades

Much has been written in recent years about grades and grading policies in general. Most of it is criticism based on theory and personal opinion rather than on experimental evidence. This fact was recognized by a Conference on College Grading Systems (25) held in Pennsylvania in 1963 with 52 colleges participating, among them Amherst, Antioch, Brandeis, Haverford, Sarah Lawrence, and Wesleyan. Future research in the following areas was recommended:

(1) The effects of grading and grade emphasis on students. Does grade emphasis really detract from true learning? Is the "lash of grades" necessary as an incentive to continuous application to studies?

(2) Experimentation with "external examinations," either oral with outside examiners, or in cooperation with E.T.S.[1]

(3) Careful study of what goes into grades.

(4) The relation between aspects of the academic environment; for example, course load, freedom of election, etc., and grades.

(5) Study of the admission procedures of professional and other graduate schools.

---

[1]Educational Testing Service.

Another nationwide discussion of grading policies and student evaluation took place during the Ninth Annual Meeting of the Council of Graduate Schools in the United States (28) in December, 1969, in Washington, D.C. The reports of these two conferences provide a comprehensive survey of grading practices and problems in this country. There seems to be common agreement that present grading policies are inadequate, even dangerous. Under main attack is the "deceptive appearance of objectivity and precise evaluation" of letter grades which "in reality conceal a host of assumptions, variables and methods" by which they are determined. According to Professor Simon (95) of the University of Massachusetts, "there is literally not a shred of research evidence which supports the present grading system." One could counter that the "normal curve" is based on considerable statistical evidence. Its use, however, for the evaluation of student achievement has been seriously questioned by Bloom (14) on the grounds that education is a purposeful activity, not a random process. Therefore, the achievement distribution should be very different from the normal curve if our instruction is effective. "In fact, our educational efforts may be said to be unsuccessful to the extent that student achievement is normally distributed."

The effect of traditional evaluation strategies upon the mental health and self-concept of students was the object of recent investigations by Bloom (13), Modu (64), Thorshen (107), and others. Under the regime of the normal curve, students at the lower end of the distribution are exposed to the traumatic experience of continuous failure, while changes in evaluation methods may "prevent severe emotional disturbances among college students." This is especially important in foreign

language learning with its cumulative psychological and educational effects.

## Organization of Study

This study was designed to investigate the impact of the following methods of evaluation upon three different groups of Iowa State University students:

Group I: Spring Quarter, 1970.
The impact of the pass-fail system upon achievement in elementary foreign language courses.

Group II: Fall Quarter, 1970.
Grading procedures based on a cumulative point system versus evaluation on the basis of the final examination only.

Group III: Winter Quarter, 1971
Student evaluation based on self-pacing as an attempt at "flexibility within a fixed time-block system" (82).

A brief introduction to these three approaches and their theoretical background will be given in subsequent chapters.

# METHOD OF PROCEDURE

## Sources of Data

The data for this investigation were gathered from the following sources:

(1) Class lists on which instructors had recorded final grades.

(2) Cumulative student records filed in the Office of Admissions and Records, Iowa State University.

(3) Achievement and aptitude tests as described in subsequent chapters.

(4) Student questionnaire.

(5) Faculty interviews concerning the pass-fail system.

## Organization of Study

This study was designed to investigate the impact of the following methods of evaluation upon achievement in elementary foreign language courses:

(1) the impact of the pass-fail system;

(2) evaluation methods based on a cumulative point system versus evaluation on the basis of the final examination only; and

(3) student evaluation based on self-pacing within a fixed time-block system versus traditional lock-step evaluation.

The following three student populations were involved:

Group I: The total student population enrolled in elementary foreign language courses during the Spring Quarter, 1970. N = 895.

Group II: 72 students enrolled in 4 sections of German 131 during the Fall Quarter, 1970.

Group III: 173 students enrolled in German 132, second elementary course, during the Winter Quarter, 1971.

Eliminated were students who had had two or more years of high school instruction in the target language, and students who did not complete the course.

Detailed descriptions and discussions of these groups are contained in subsequent chapters.

## Criteria of Achievement

Each of the three student populations involved was used to investigate the impact of specific methods of evaluation upon achievement in elementary foreign language courses, as measured by the following criteria:

Group I: Final grade in elementary foreign language courses as recorded by instructors on the class lists filed in the Office of Admissions and Records at Iowa State University. These grades were not the result of uniform evaluation methods.

The French grades were based on a departmental final examination counting 25%, on language laboratory tests counting 25%, and on an instructor grade counting 50%. The German and Spanish grades were based completely on evaluation by individual instructors who also graded the final departmental

test according to their own methods. The Russian grades were based mainly on homework, weekly quizzes, and a final examination counting approximately 25%.

The adequacy of the assumption of uniform grading standards was tested by analysis of variance by section and instructor, and will be discussed in connection with the statistical analysis of Group I. Over 50% of the 25 instructors involved were native speakers. The majority had the rank of instructor, and only 28% ranked from assistant to full professor.

Group II: (1) Raw score on the objective part of the final examination in German 131, based on the text by Schulz-Griesbach, DEUTSCH FÜR AMERIKANER (91). This test consists of 95 multiple-choice questions emphasizing knowledge of vocabulary as well as of idiomatic and grammatical structure. It includes a 30-item listening-comprehension test. The KR-20 reliability estimates ranged from .77 to .94 for the 13 sections of German 131.

(2) Letter grade received on the objective part of the final examination. These grades were distributed as follows: total score possible = 95 points, A = 85, B = 76, C = 67, D = 58.

(3) Instructor grade. This grade does not represent uniform evaluation procedures. The 13 sections of German 131 were taught by 5 different instructors who used 5 different methods of assigning grades. In the main, however, these grades were based on the results of a number of quizzes, a midterm and the departmental final examination, which besides an objective multiple choice test, also included an individually scored written part.

(4) Total sum of unit-quiz scores. The total possible score was 420; A = 370, B = 330, C = 290, D = 250.

(5) Total final test score including both the objective and the written parts of the final examination in German 131. The total possible score was 310 points.

(6) Course grade based on the total final test score, where A = 270, B = 240, C = 210, D = 180.

(7) Total sum of scores comprising the points accumulated on all examinations given during the Fall Quarter, 1970, in German 131. The total possible score was 730 points.

(8) Course grade based on the total sum of scores. These grades were distributed as follows: A = 650, B = 580, C = 510, D = 430.

Group III: (9) Raw score on the objective part of the final examination based on the same text as in German 131. This test consists of 150 multiple choice questions including a 25-item listening comprehension test. Emphasized are vocabulary, idiomatic and grammatical structure. The KR-20 reliability estimates for this test which was administered to 208 students, ranged from .85 to .95 for the 12 sections of German 132.

(10) Final grade for German 132 based on the score received on the objective part of the final examination. These grades were distributed as follows: Total possible score = 150, A = 135, B = 120, C = 105, D = 90.

(11) Instructor grade for German 132 assigned on an individual basis by each of the 5 instructors teaching the 12 sections.

Variables Affecting Achievement

To account for the variability of the dependent variables, the following independent variables were investigated:

(1) Cumulative grade-point average.

(2) Course load.

(3) Method of evaluation.

Group I: Pass-fail versus letter-grade evaluation. The instructors recorded letter grades for all students because they did not know which students were enrolled under the pass-fail system. These letter grades were converted to pass-fail grades by the Office of the Registrar.

Group II: A.M. versus P.M; i.e., evaluation on the basis of cumulative points for the forenoon group versus evaluation based on the final test score only for the afternoon group.

Group III: Traditional lock-step evaluation for the control group versus evaluation based on self-pacing for the experimental group.

(4) Sex.

(5) Cumulative credit hours, end of Spring Quarter, 1970 (Group I).

(6) Year in college (Group III).

(7) College: Sciences and Humanities, Engineering, Agriculture, Education, Home Economics, Agricultural Engineering.

(8) Language studied in Group I: French, German, Russian, and Spanish.

(9) Course, section, and instructor.

(10) ACT composite score available for 750 students in Group I. The American College Testing Program is administered to high school students to compare them with other college-bound students. These academic tests cover four subject areas:

English usage, mathematics usage, social studies reading, and natural sciences reading. The ACT tests deal with intellectual skills and abilities, not with specific and detailed subject matter content. They are supposed to measure aptitude required of college students, and thus are indicators of academic success. According to a study based on 4,283 freshmen completing their first quarter at Iowa State University in the Fall of 1968, ACT ranked third as a predictor of academic success after High School Rank and the Minnesota Scholastic Aptitude Test for students in Sciences and Humanities (46).

(11) Language aptitude, as measured by the total score on the Pimsleur Language Aptitude Battery (76) intended for use in grades 7 through 12, is defined by the author in terms of three factors:

a. verbal intelligence; that is, the knowledge of words and the ability to reason analytically in using verbal materials;

b. motivation, an expression of interest in studying a modern foreign language; and

c. auditory ability, the ability to receive and process information through the ear.

The test consists of 6 parts covering an academic aptitude measure, an interest estimate, vocabulary knowledge, language analysis, sound discrimination, and sound-symbol association. The total score was used with Group II. The composite raw

score of parts 3, 4, 5, and 6 was used with Group III to estimate language aptitude. Part I of this battery, i.e., grade-point average in academic areas other than foreign languages, based on the students' recording the grades last received in English, Mathematics, Social Studies, and Science, was excluded in favor of their cumulative grade-point average considered to be a more reliable estimate of academic achievement and motivation.

(12) Interest (Group III) was estimated on the basis of Part 2 of the Pimsleur Language Aptitude Battery where the student is asked to consider his interest in terms of "how useful a foreign language is to him, how much he will enjoy it, and how interested he is in foreign languages as compared with other subjects." The estimate is based on five possible choices ranging from "rather uninterested" to "strongly interested."

(13) Motivation for learning German was estimated separately for Group III on the basis of the following four choices:
I am taking this course

1. to fulfill requirements;
2. to fulfill requirements, and also because I think German may be useful to me in the future;
3. to fulfill requirements, and also because I am interested in the German people and their way of life;
4. voluntarily, because _____ (please complete).

These choices are a modified application of Gardner and Lambert's "Orientation Index" (38) which is supposed to classify students into "integratively oriented" and "instrumentally oriented." In Gardner and Lambert's investigation, students were presented with four reasons for studying French and asked to rank them as to their personal relevance. The possible reasons for studying French were that it would

a. be useful in obtaining a job;
b. helpful in understanding the French-Canadian people and their way of life;
c. permit meeting and conversing with more and varied people; and
d. make one a better-educated person.

"Integrative" motivation was supposed to reflect interest in the foreign culture and its people. "Instrumental" motivation indicated a more utilitarian kind of motivation for learning the foreign language.

In the present study, "integrative" motivation, as expressed by choices 3 and 4, was weighted heavier than "instrumental" motivation under the assumption that it correlates more highly with success in foreign language learning according to Gardner and Lambert's conclusion "that the student's attitude toward language study and toward the speakers of the language he is studying can have profound influence over and above those of aptitude."

(14) Course sequence (Group I).

Hypotheses Tested and Statistical Methods Employed

The statistical analysis of the various groups included in this investigation is divided into three major parts, each concerned with a different approach to evaluation of achievement:

Part 1: Pass-fail system versus letter-grade system.

Part 2: Cumulative point system versus evaluation based on the final examination.

Part 3: Traditional lock-step evaluation versus self-pacing.

The specific objectives connected with each part are stated in hypothesis form. These hypotheses are tested by using linear, additive, multiple regression models. Each independent variable is investigated in turn to test the null hypothesis that $\beta = 0$, against the alternative hypothesis that $\beta \neq 0$ by considering the regression models in which the given variable is included. The significance of the contribution of these variables to the variability in the dependent variable was estimated on the basis of calculated t-values which were compared to the tabular t-value with the degrees of freedom for residual. These values, if significant at the .05 level, were marked with an asterisk. The corresponding data are included in the Appendix. Note, that the stated level of significance is not the actual level because more than one t-test was performed using the same data. However, it is unknown.

Preliminary investigations of differences between groups due to method of evaluation or to instructor used special cases of multiple

regression, namely, t-tests, analysis of variance and the technique of chi-square where other variables are ignored.

# INVESTIGATION OF THE IMPACT OF THE PASS-FAIL SYSTEM UPON ACHIEVEMENT IN ELEMENTARY FOREIGN LANGUAGE COURSES

## Introduction

When interviewed, thirteen out of seventeen foreign language instructors at Iowa State University felt that pass-fail students "try to get by with a minimum of effort," thus rendering elementary foreign language sequences meaningless; four instructors were of the opinion that there is no difference in the performance of students who study under the pass-fail system and of those working for a letter grade. The passing level under the pass-fail system at Iowa State University is a D-grade, defined in the Iowa State Student Handbook (45) as "passing but unsatisfactory."

The pass-fail controversy does not only exist in the area of foreign languages. Since its evolvement, this grading system has been under continuous investigation at Iowa State University and in other places. Some recent discussions in academic and professional circles are proof that its merits have not been sufficiently tested or universally accepted on the American scene. Thus, the American Bar Association (57) endorsed in October 1970 a statement by the Law School Admissions Council on the Pass-Fail Grading Systems questioning the practicability of such systems for the purpose of admissions decisions to law school:

> "The adoption by an increasing number of colleges and universities of pass-fail or similar grading systems for some or all their students' work has implications for the law school admissions process. When a student with a transcript bearing such grades seeks to enter law school, law school admissions committees will be deprived of data that has served them well in the past in making the admissions decision. In the belief that college and university faculties and administrations who are considering conversion of a conventional grading system to a pass-fail or some variant system may be interested in the

> possible effect of such grading systems upon their graduates who seek admission to law school, the Law School Admission Test Council issues this statement. ... Where an applicant for admission to law school submits a transcript in which all or virtually all of his grades are on a pass-fail basis, and submits no other indication of his level of achievement in college, the admissions committee can make little specific use of his college work in predicting his law school grades. This means that this prediction must be based upon the LSAT (Law School Admission Test) score, even though the committee would much prefer not to place sole reliance on the test scores in making this prediction. ... Where the applicant for admission to law school submits a transcript containing some conventional grades and some pass-fail grades, the admissions committee can develop a grade-point average for the portion of the student's college work bearing the conventional grades. However, many admissions officers will not feel justified in assigning to that average the conventional weight. They may well assume that the student chose to receive a conventional grade in those courses in which he gauged his probabilities for a premium grade to be good. ... Furthermore, the committee may reasonably assume that the applicant did not make the same effort in the courses graded on a pass-fail basis as he did in those graded on the conventional basis."

A variety of reactions was elicited by the Student Curriculum Council of Wilson High School in Portland, Oregon (71) which polled 213 colleges and universities for their opinion about an experimental grading system supposed to relieve the pressure of grades. Answers ranged from

> "Grades represent diligence, time budgeting, self-discipline, effort, and perseverance. What are you afraid of?"

from the University of South Carolina, to

> "Go to it. We're seriously studying a similar proposal for ourselves."

from Colorado College.

Quite a few institutions pointed to the difficulty of dealing with descriptive records and the necessity of relying in such cases too heavily on the results of College Board test scores.

The educational considerations in accepting the pass-fail system by colleges and universities were summarized in a report of the Pass-Fail Committee of Phi Beta Kappa (75) in December, 1969:

> "The inquiries of the Committee have included a review of many detailed studies concerning pass-fail options. The studies show that proponents of the pass-fail option are generally agreed on these assumptions:
>
> (1) The pass-fail option permits the student to study or learn without pressure or emotional strain.
>
> (2) Under the option the student does not feel repressed or inhibited by a grading system.
>
> (3) Students have an opportunity to pursue courses in 'academically unfamiliar' areas without fear of a poor grade.
>
> (4) Students following pass-fail options should display greater motivation and intellectual curiosity than those under traditional programs."

The various patterns of the pass-fail options practiced in the United States are described in several reports (48), (75), (84), which all agree that

(1) The pass-fail option is seldom permitted in the area of a student's specialization on the assumption that letter grades are necessary to predict success in professional or graduate schools. Thus, the pass-fail option is reserved for "disparate fields of study" without endangering the grade-point average.

(2) Students are allowed to take one pass-fail course per term in the large majority of cases. There is a limit in the number of pass-fail courses which count toward graduation, the maximum encountered being 50 percent.

(3) Pass-fail courses are usually limited to juniors and seniors with a grade-point average of 2.5 or above.

(4) In some institutions the letter grades reported by the instructors are converted to pass-fail in the registrar's office to prevent dual grading standards. If necessary, the passing grades can be quantified again.

The reports also agree as to certain tendencies connected with this kind of grading systems:

(1) Students taking advantage of the pass-fail system are in a minority.

(2) Contrary to the purpose of pass-fail grades, students tend to stay close to their major.

(3) Academic achievement under pass-fail tends to be lower, students often aiming at the minimum in order to have more time to study for courses in the major area.

(4) In the majority of institutions, the "pass" level seems to be the D-grade.

The members of the Pass-Fail Study Committee of Phi Beta Kappa concluded in December, 1969, after a detailed study of 121 chapters that the pass-fail options do not, at this time, constitute a serious problem for most of the Phi Beta Kappa chapters in their selection of members. They stated, however, that the trend toward more pass-fail options may present difficult problems in the future. A year later, in December, 1970, the society recognized that the above study is "fast becoming outdated, and it will soon have to face the pass-fail system head on."

At Iowa State University, the Pass/Not Pass system (45) formerly called pass-fail system, follows basically the same general pattern. With the permission of their college and department, students who have earned at least 60 credit hours are eligible to take a maximum of two courses per quarter under this grading system. These courses should not be in the student's major subjects nor directly supporting them. Except for restrictions on its own undergraduate majors, a department may not deny the availability of any of its course offerings on a Pass/Not Pass basis. There are no restrictions as to student grade-point average. A maximum of six courses under this system may be used to meet degree requirements. Courses offered only on a Satisfactory/Fail basis do not count in the six-course limitation unless the student elects to take the course on a Pass/Not Pass basis. Pass/Not Pass courses, unless an F-grade was received, cannot be repeated. The student chooses to take a given course under the Pass/Not Pass system at the time of pre-classification and may change his classification or registration from Pass/Not Pass to a graded basis or reverse within a period ending 14 calendar days after classes begin. The names of students classified on a Pass/Not Pass basis are not identified on the class lists. The letter grade given by the instructor is recorded by the Office of the Registrar as "P" or "NP." Neither a "P" nor a "NP" affect the student's quality-point average. A "Pass" credit on the transcript can only be interpreted as "D" or better. Criteria for the operation of this system are being evaluated annually.

The statistical data for the Spring Quarter 1970 as indicated in Table 1 show that 13.5% of the sophomores, 26.6% of the juniors, and 41.4% of

Table 1. Quality-point averages of students enrolled in pass/not pass courses at Iowa State University, Spring quarter 1970[a]

| | Cumulative average | Total university enrollment | Percent | Students enrolled in pass-fail | Percent |
|---|---|---|---|---|---|
| | 3.75 | 62 | 2 | 9 | 2 |
| | 3.50 | 122 | 4 | 16 | 4 |
| | 3.25 | 178 | 5 | 35 | 8 |
| | 3.00 | 286 | 9 | 53 | 12 |
| | 2.75 | 409 | 13 | 68 | 15 |
| Sophomore | 2.50 | 547 | 17 | 70 | 16 |
| | 2.25 | 636 | 20 | 87 | 20 |
| | 2.00 | 564 | 17 | 63 | 14 |
| | 1.75 | 299 | 9 | 32 | 7 |
| | Less | 152 | 5 | 7 | 2 |
| | Totals | 3255 | | 440 | |
| | 3.75 | 89 | 3 | 27 | 3 |
| | 3.50 | 145 | 4 | 46 | 5 |
| | 3.25 | 220 | 7 | 69 | 8 |
| | 3.00 | 311 | 9 | 102 | 12 |
| | 2.75 | 416 | 13 | 126 | 14 |
| Junior | 2.50 | 563 | 17 | 149 | 17 |
| | 2.25 | 647 | 20 | 195 | 22 |
| | 2.00 | 548 | 17 | 128 | 15 |
| | 1.75 | 283 | 9 | 32 | 4 |
| | Less | 89 | 3 | 7 | 1 |
| | Totals | 3311 | | 881 | |
| | 3.75 | 96 | 3 | 46 | 3 |
| | 3.50 | 190 | 5 | 91 | 6 |
| | 3.25 | 330 | 9 | 162 | 10 |
| | 3.00 | 422 | 11 | 191 | 12 |
| | 2.75 | 556 | 15 | 238 | 15 |
| Senior | 2.50 | 763 | 20 | 320 | 20 |
| | 2.25 | 666 | 18 | 265 | 17 |
| | 2.00 | 633 | 17 | 233 | 15 |
| | 1.75 | 118 | 3 | 20 | 1 |
| | Less | 7 | 0 | 1 | 0 |
| | Totals | 3781 | | 1567 | |

[a]Data used with permission of Registrar, Iowa State University.

the seniors were enrolled in Pass/Not Pass courses at Iowa State University, the largest groups of the sophomores and juniors having a grade-point average of 2.25, and the seniors of 2.50. Only 10% of the students enrolled during the Spring Quarter took their first Pass/Not Pass course the first quarter they were eligible. The majority of students enrolled in Pass/Not Pass courses as first quarter juniors.

As is evident from Table 2, the participation in these courses is growing in volume. The enrollment in Pass/Not Pass courses increased from 577 students in the Fall 1967 to 3,148 during the Spring 1970. The greatest involvement in Pass/Not Pass courses seems to take place in the Spring Quarter of each year. For this reason, the Spring Quarter 1970 with the greatest enrollment to date was selected for the investigation of the impact of the pass-fail system upon achievement in elementary foreign language courses.

## Description of Population

The investigation of Group I includes the total student population enrolled during the Spring Quarter 1970 in elementary courses in French, German, Russian, and Spanish. The analysis of this population was undertaken after the end of the Spring Quarter. The group had not been subjected to any experimental treatment except for a difference in the evaluation methods according to which 305 out of 895 students were evaluated on a pass-fail basis; the remaining 590 received letter grades on their cumulative records.

Table 2. Number of students enrolled under the pass/not pass system at Iowa State University between Fall 1967 and Winter 1971[a]

| | Students | Courses |
|---|---|---|
| 1967-68 | | |
| Fall 1967 | 577 | 711 |
| Winter 1968 | 943 | 1165 |
| Spring 1968 | 1351 | 1698 |
| 1968-69 | | |
| 1st SS 1968 | 198 | 206 |
| 2nd SS 1968 | 163 | 175 |
| Fall 1968 | 861 | 1063 |
| Winter 1969 | 1386 | 1974 |
| Spring 1969 | 1979 | 2436 |
| 1969-70 | | |
| 1st SS 1969 | 308 | 371 |
| 2nd SS 1969 | 254 | 276 |
| Fall 1969 | 1783 | 2291 |
| Winter 1970 | 2506 | 3073 |
| Spring 1970 | 3148 | 4162 |
| 1st SS 1970 | 373 | 429 |
| 2nd SS 1970 | 269 | 301 |
| Fall 1970 | 2219 | 2683 |
| Winter 1971 | 2866 | 3431 |

[a]Data used with permission of Registrar, Iowa State University.

The distribution of students in Group I according to language, course, section, grade, sex, college, and year in college, is presented in Tables 3, 4, 5 and 6.

Table 3. Distribution of Iowa State University students enrolled in elementary foreign language courses during the Spring quarter 1970

| Language | Course No. | Number of Sections | Total enrollment | Grade distribution | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | A | B | C | D | F | P[a] | I[b] | D[c] |
| French | 102 | 6 | 148 | 20 | 25 | 33 | 17 | 7 | 31 | - | 15 |
| French | 103 | 10 | 260 | 20 | 41 | 60 | 38 | 1 | 72 | 8 | 20 |
| Subtotal | | 16 | 408 | 40 | 66 | 93 | 55 | 8 | 103 | 8 | 35 |
| German | 132 | 4 | 88 | 10 | 14 | 18 | 7 | 4 | 9 | 2 | 24 |
| German | 133 | 9 | 210 | 29 | 40 | 39 | 21 | 8 | 38 | 3 | 32 |
| German | 133X | 1 | 15 | 5 | 3 | 3 | - | - | 1 | - | 3 |
| Subtotal | | 14 | 313 | 44 | 57 | 60 | 28 | 12 | 48 | 5 | 59 |
| Russian | 123 | 3 | 71 | 18 | 16 | 22 | 4 | 1 | 2 | 2 | 6 |
| Spanish | 152 | 5 | 141 | 28 | 30 | 21 | 4 | 8 | 17 | 11 | 22 |
| Spanish | 153 | 14 | 398 | 60 | 100 | 99 | 15 | 11 | 87 | 4 | 22 |
| Subtotal | | 19 | 539 | 88 | 130 | 120 | 19 | 19 | 104 | 15 | 44 |
| Total | | 52 | 1,331 | 190 | 269 | 295 | 106 | 40 | 257 | 30 | 144 |

[a]Departmental "Pass" for students who have had 2 or more years of high school instruction in the language.

[b]Incomplete.

[c]Drop.

From the total enrollment of 1,331, 257 cases were eliminated because of previous language study of two or more years at the high school level; 179 students had not completed the elementary courses. The remaining 895 cases were divided into two groups:

Table 4. Group I. Distribution of students by language studied and by sex; N = 895

| | Pass - Fail | | | | Letter grade | | | |
|---|---|---|---|---|---|---|---|---|
| Language | Male | Female | Total | Percent | Male | Female | Total | Percent |
| French | 61 | 28 | 89 | 29 | 91 | 81 | 172 | 29 |
| German | 52 | 12 | 64 | 21 | 97 | 44 | 141 | 24 |
| Russian | 13 | 3 | 16 | 5 | 32 | 14 | 46 | 8 |
| Spanish | 112 | 24 | 136 | 45 | 134 | 97 | 231 | 39 |
| Total | 238 | 67 | 305 | 100 | 354 | 236 | 590 | 100 |
| Percent | 78 | 22 | | | 60 | 40 | | |

Table 5. Group I. Distribution of students by language[a] studied and by college; N = 895

| | Pass - Fail | | | | | | Letter grade | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| College | F | G | R | S | Total | % | F | G | R | S | Total | % |
| Sciences & Humanities | 77 | 53 | 13 | 122 | 265 | 87 | 147 | 119 | 42 | 190 | 498 | 84.4 |
| Engineering | 5 | 7 | 3 | 2 | 17 | 6 | 6 | 11 | 3 | 6 | 26 | 4.4 |
| Agriculture | 1 | 3 | 0 | 2 | 6 | 2 | 4 | 7 | 0 | 15 | 26 | 4.4 |
| Education | 1 | 1 | 0 | 2 | 4 | 1 | 6 | 0 | 1 | 4 | 11 | 1.9 |
| Home Economics | 5 | 0 | 0 | 8 | 13 | 4 | 9 | 4 | 0 | 16 | 29 | 4.9 |
| Total | 89 | 64 | 16 | 136 | 305 | 100 | 172 | 141 | 46 | 231 | 590 | 100 |

[a] F = French; G = German; R = Russian; S = Spanish.

Table 6. Group I. Distribution of students by language[a] studied and by year in college; N = 895

| Year in college | Pass-Fail | | | | | | Letter grade | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | F | G | R | S | Total | % | F | G | R | S | Total | % |
| Freshman | 0 | 1 | 1 | 2 | 4 | 1.3 | 47 | 28 | 10 | 46 | 131 | 22 |
| Sophomore | 27 | 10 | 3 | 39 | 79 | 25.9 | 57 | 30 | 10 | 64 | 161 | 27 |
| Junior | 33 | 29 | 2 | 50 | 114 | 37.4 | 39 | 33 | 21 | 79 | 183 | 31 |
| Senior | 29 | 24 | 10 | 45 | 108 | 35.4 | 23 | 34 | 5 | 37 | 99 | 17 |
| Graduate | 0 | 0 | 0 | 0 | 0 | 0 | 6 | 5 | 0 | 5 | 16 | 3 |
| Total | 89 | 64 | 16 | 136 | 305 | | 172 | 141 | 46 | 231 | 590 | |

[a] F = French; G = German; R = Russian; S = Spanish.

(1) 305 students enrolled under the pass-fail system (34.2%);

(2) 590 students enrolled under the traditional letter-grade system.

Of the 895 students involved in this investigation, 29% studied French, 23% German, 7% Russian, and 41% Spanish.

The statistical tabulations for Group I indicate that 78% of the pass-fail foreign language students, and 60% of the letter-grade students were males. Over 80% in both groups were enrolled in the College of Sciences and Humanities. The Spanish students are the largest group followed by the French and German students. The Russian students are the smallest group.

In the pass-fail group only 1.3% were freshmen, as compared to 22.2% in the letter-grade group. The percentage of seniors enrolled under the pass-fail system was twice that of seniors enrolled under the letter-grade system.

## Hypotheses Tested and Statistical Methods Employed

The investigation of Group I involves four hypotheses, each discussed in turn.

Hypothesis 1:

After adjusting for other variables, there is no difference in achievement in elementary foreign language courses between students enrolled under the pass-fail system and those enrolled under the letter-grade system.

For the investigation of null hypothesis 1, as a preliminary step, t-tests based on separate group means, with groups of unequal size, were performed to investigate if the pass-fail and the letter-grade groups differ as to grade-point average, ACT score, or course load. Then the statistical technique of multiple regression was applied. The final course grade, $X_1$, considered the dependent variable, was regressed on the independent variables as listed in the introduction; namely,

| | |
|---|---|
| $X_2$ | grading system: pass-fail versus letter grade |
| $X_3$ | sex |
| $X_4$ | cumulative grade-point average by end of Spring Quarter 1970 |
| $X_5$ | ACT composite score (available for 750 students) |

| | |
|---|---|
| $X_6$ | course load during the Spring Quarter 1970 |
| $X_7$ | cumulative credit hours by end of Spring 1970 |
| $X_{8-11}$ | college[1] |
| $X_{12-14}$ | language[1] |
| $X_{15}$ | course sequence. |

The complete model in raw score form with 15 variables is:

$$X_1 = \beta_0 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4 + \beta_5X_5 + \beta_6X_6 + \beta_7X_7 + \beta_8X_8 + \beta_9X_9 + \beta_{10}X_{10} + \beta_{11}X_{11} + \beta_{12}X_{12} + \beta_{13}X_{13} + \beta_{14}X_{14} + \beta_{15}X_{15} + \epsilon$$

where the unknown regression coefficients, $\beta$, are estimated from the multiple regression program which also computed correlation matrices, F-ratios, t-values for each combination of variables, and where $\epsilon$ is the unexplained variation in the dependent variable, $X_1$.

The following assumptions were made for this analysis:

1. The students enrolled in elementary foreign language courses during the Spring Quarter 1970 are representative of foreign language students at Iowa State University as to aptitude, motivation, grade-point average, sex ratio, course load, and selection of courses on the basis of pass-fail or letter grade evaluation. In other words, on the average, this group, though including the total student population enrolled in elementary foreign language courses during the Spring Quarter 1970, is assumed to be a random sample from the population of all students over time who study foreign languages at Iowa State University.

2. The grades received are a satisfactory measure of achievement in elementary foreign language courses at Iowa State University.

---

[1]For computer analysis, each college and each language is treated as a separate variable.

3. The mean final grade is a linear, additive function of the independent variables.

4. The residuals, i.e., the deviations of final grade from the regression plane of the independent variables, are independently distributed.

Note: The assumption of independence does not hold exactly since students within sections tend to interact and thus produce correlated responses.

5. The residuals are normally distributed.

6. The residuals have common variance.

7. The independent variables are considered fixed, so that the conclusions hold for groups of students who have the same distribution of values for the independent variables as this group.

Hypothesis 2:

There is no difference among the average grades for students taught by different instructors. That is, there is no significant variation in the mean achievement of students taught by different instructors, as measured by final grade based on individual teaching and evaluating methods.

Hypothesis 3:

There is no difference among the average grades for sections within an instructor. That is, there is no significant variation in the mean achievement of sections taught by the same instructors.

Hypothesis 4:

There is no difference among the average grades for students studying different languages.

The effect of 25 instructors upon achievement in 52 sections of elementary foreign language courses taught during the Spring Quarter 1970 was investigated by analysis of variance rather than by multiple regression because of the large number of instructor variables and of sections which could not be handled by the computer regression program.

The statistical model is a completely randomized design with sub-sampling and unequal numbers. Instructors are considered to be treatments, sections are considered to be experimental units, and students to be sampling units:

$$Y_{ijk} = \mu + A_i + B_{ij} + \epsilon_{ijk}$$

$Y_{ijk}$ = final course grade for the $k^{th}$ students in the $j^{th}$ section of the $i^{th}$ instructor

$\mu$ = the overall average grade for the population

$A_i$ = the effect of the $i^{th}$ instructor

$B_{ij}$ = the effect of the $j^{th}$ section within the $i^{th}$ instructor

$\epsilon_{ijk}$ = the random deviation associated with the $ijk^{th}$ student; i.e., the unexplained variation

$i = 1, 2. \ldots. 25$ instructors

$j = 1, 2, \ldots. b_i$ sections for the $i^{th}$ instructor

$k = 1, 2, \ldots. c_{ij}$ students for the $j^{th}$ section within the $i^{th}$ instructor

$c_{11} = 24$; $c_{12} = 18$; $c_{13} = 16$;

$c_{21} = 18$; $c_{22} = 8$; $c_{23} = 9$;

$c_{31} = 18$;

$c_{41} = 26$; $c_{42} = 21$;

$c_{51} = 6$;

$c_{61} = 13$; $c_{62} = 16$;

$c_{71} = 15$;

$c_{81} = 18$; $c_{82} = 10$;

$c_{91} = 26$;

$c_{10,1} = 23$; $c_{10,2} = 24$;

$c_{11,1} = 11$; $c_{11,2} = 14$; $c_{11,3} = 5$;

$c_{12,1} = 14$; $c_{12,2} = 15$; $c_{12,3} = 7$;

$c_{13,1} = 9$; $c_{13,2} = 13$; $c_{13,3} = 15$;

$c_{14,1} = 26$; $c_{14,2} = 14$;

$c_{15,1} = 11$;

$c_{16,1} = 14$;

$c_{17,1} = 23$; $c_{17,2} = 24$;

$c_{18,1} = 25$; $c_{18,2} = 25$; $c_{18,3} = 24$;

$c_{19,1} = 16$; $c_{19,2} = 16$; $c_{19,3} = 14$;

$c_{20,1} = 9$;

$c_{21,1} = 31$;

$c_{22,1} = 9$; $c_{22,2} = 17$;

$c_{23,1} = 14$; $c_{23,2} = 13$; $c_{23,3} = 10$;

$c_{24,1} = 38$; $c_{24,2} = 37$; $c_{24,3} = 33$;

$c_{25,1} = 19$; $c_{25,2} = 10$; $c_{25,3} = 16$.

The sums of squares for instructors and sections within instructors were calculated using unweighted section means; that is, no adjustments for different section sizes were made. The sum of squares for students within sections within instructors was calcuated as usual and divided by the harmonic mean of the number of students within sections; that is, by the quotient of the number of sections and of the sum of reciprocals of the number of students in each section to put it on the same basis as the other two sums of squares.

The assumptions used for this analysis can be summarized as follows:

1. The students enrolled in elementary foreign language courses are a random sample of the total student population at Iowa State University; that is, it is assumed that they are representative of Iowa State University foreign language students as far as language aptitude, motivation, sex ratio, and academic ability are concerned.

2. The final grade assigned by individual instructors is a satisfactory measure of achievement in elementary foreign language courses.

3. The analysis of variance model is linear; i.e., a first degree polynomial.

4. The analysis of variance model is additive.

5. The unexplained variation is distributed normally.

6. The unexplained variation is independently distributed. It is realized that this assumption is not completely satisfied in this situation because the grades of students within a section will tend to correlate with each other. It is hoped that through random assignment of students to sections the correlation between errors will not favor any particular section. Ostle (73) states that "in general the conse-

quences are not serious when the assumptions made in connection with analysis of variance are not strictly satisfied."

8. It is assumed that instructors and sections are picked at random from a larger population over time. That is, if the analysis were to be repeated, different instructors and different numbers of sections taught by an instructor would be used.

## Findings

### Preliminary analysis

The contrasts between the grade distributions for pass-fail and for letter grade students in elementary foreign language courses and for Iowa State University students in general are presented in Tables 7 through 11.

Table 7. Group I. Distribution of grades in elementary foreign language courses, Spring quarter 1970 - pass-fail versus letter-grade students

| Grade | Pass-fail students | | Letter-grade students | |
|---|---|---|---|---|
| | Number | Percent | Number | Percent |
| A | 18 | 5.9 | 172 | 29.2 |
| B | 63 | 20.7 | 207 | 35.1 |
| C | 153 | 50.1 | 141 | 23.9 |
| D | 69 | 22.6 | 37 | 6.3 |
| F | 2 | 0.7 | 33 | 5.5 |
| Total | 305 | 100.0 | 590 | 100.0 |

Table 8 compares the letter grades received (but not recorded) by students in elementary pass-fail foreign language courses with the letter grades received (and recorded) by the same students in courses taken under letter-grade system during the Spring Quarter 1970.

Table 8. Group I. Distribution of grades received by students in elementary foreign languages courses under the pass-fail system compared to grades received by the same students in other subjects taken under the letter-grade system during the Spring quarter 1970

| Grade | Pass-fail | | Letter grade | |
|---|---|---|---|---|
| | Number | Percent | Number | Percent |
| A | 18 | 5.9 | 361 | 31 |
| B | 63 | 20.7 | 473 | 40 |
| C | 153 | 50.1 | 292 | 25 |
| D | 69 | 22.6 | 47 | 4 |
| F | 2 | 0.7 | 8 | 1 |
| Total | 305 | 100.0 | 1,181 | 100.0 |

Table 9 presents the data for all students in pass-fail courses enrolled at Iowa State University during the Spring Quarter 1970.

Table 9. Grade distributions of the same Iowa State University students in pass-fail courses and in non-pass-fail courses Spring 1970

| Grade | Pass-fail | | Letter grade | |
|---|---|---|---|---|
| | Number | Percent | Number | Percent |
| A | 232 | 8 | 3,614 | 32 |
| B | 874 | 29 | 4,445 | 40 |
| C | 1,340 | 44 | 2,629 | 23 |
| D | 493 | 16 | 445 | 4 |
| F | 112 | 4 | 139 | 1 |
| Total | 3,051 | 100 | 11,272 | 100 |

Table 10. Group I. Grade distribution of elementary foreign language students by grade-point average, Spring quarter 1970

| Student's | QPA[a] | A | Percent | B | Percent | C | Percent | D | Percent | F | Percent | Total All letter grades |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3.75-4.00 | P-F[b] | 2 | 29 | 3 | 43 | 1 | 14 | 1 | 14 | 0 | 0 | 7 (100%) |
| | L[c] | 19 | 83 | 4 | 17 | 0 | 0 | 0 | 0 | 0 | 0 | 23 |
| 3.50-3.74 | P-F | 4 | 27 | 5 | 33 | 4 | 27 | 2 | 13 | 0 | 0 | 15 |
| | L | 40 | 69 | 15 | 26 | 3 | 5 | 0 | 0 | 0 | 0 | 58 |
| 3.25-3.49 | P-F | 3 | 15 | 4 | 20 | 9 | 45 | 4 | 20 | 0 | 0 | 20 |
| | L | 37 | 54 | 27 | 39 | 5 | 7 | 0 | 0 | 0 | 0 | 69 |
| 3.00-3.24 | P-F | 1 | 2 | 13 | 35 | 20 | 51 | 5 | 12 | 0 | 0 | 39 |
| | L | 76 | 46 | 63 | 38 | 26 | 16 | 0 | 0 | 1 | 0.6 | 166 |
| 2.75-2.99 | P-F | 2 | 5 | 9 | 24 | 20 | 52 | 7 | 19 | 0 | 0 | 38 |
| | L | 59 | 39 | 65 | 43 | 28 | 18 | 0 | 0 | 0 | 0 | 152 |
| 2.50-2.74 | P-F | 2 | 3 | 15 | 25 | 27 | 47 | 15 | 25 | 0 | 0 | 59 |
| | L | 66 | 27 | 105 | 44 | 60 | 25 | 9 | 4 | 0 | 0 | 240 |
| 2.25-2.49 | P-F | 1 | 2 | 10 | 16 | 40 | 62 | 12 | 18 | 1 | 2 | 64 |
| | L | 43 | 17 | 121 | 48 | 84 | 32 | 7 | 2 | 2 | 1 | 257 |
| 2.00-2.24 | P-F | 3 | 6 | 5 | 9 | 28 | 53 | 16 | 30 | 1 | 2 | 53 |
| | L | 17 | 10 | 65 | 37 | 67 | 38 | 25 | 14 | 2 | 1 | 176 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1.75-1.99 | P-F | 0 | 0 | 0 | 0 | 5 | 50 | 5 | 50 | 0 | 0 | 10 |
| | L | 4 | 11 | 8 | 22 | 19 | 51 | 3 | 8 | 3 | 8 | 37 |
| 1.74 below | P-F | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | L | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 100 | 0 | 0 | 3 |
| Total | P-F | 18 | 5 | 64 | 20 | 154 | 51 | 67 | 21 | 2 | 1 | 305 |
| | L | 361 | 31 | 473 | 40 | 292 | 25 | 47 | 4 | 8 | 1 | 1181 |

---

[a]QPA = Quality Point Average of students taking elementary foreign language courses at Iowa State University during the Spring Quarter 1970.

[b]P-F = Number of letter grades received by students taking elementary foreign language courses under the Pass-Fail system during the Spring Quarter 1970.

[c]L = Number of letter grades received by the same students in other than pass-fail courses during the Spring Quarter 1970.

Table 11. Grade distribution by quality point average of all Iowa State University students enrolled in pass-fail courses during the Spring quarter 1970[a]

| Students' QPA[b] | A | Percent | B | Percent | C | Percent | D | Percent | F | Percent | Total all letter grades |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 3.75-4.00[c,d] | 30 | 35 | 36 | 42 | 18 | 21 | 1 | 1 | 0 | 0 | 85 (100%) |
| | 299 | 87 | 45 | 13 | 1 | 0 | 0 | 0 | 0 | 0 | 345 (100%) |
| 3.50-3.74 | 45 | 25 | 79 | 44 | 48 | 27 | 7 | 4 | 1 | 1 | 180 |
| | 443 | 71 | 172 | 27 | 14 | 2 | 1 | 0 | 0 | 0 | 630 |
| 3.25-3.49 | 45 | 17 | 109 | 40 | 96 | 35 | 18 | 7 | 3 | 1 | 271 |
| | 569 | 55 | 372 | 36 | 88 | 8 | 4 | 1 | 1 | 0 | 1,034 |
| 3.00-3.24 | 29 | 8 | 142 | 37 | 167 | 44 | 37 | 10 | 6 | 2 | 381 |
| | 635 | 43 | 634 | 43 | 176 | 12 | 7 | 1 | 6 | 1 | 1,458 |
| 2.75-2.99 | 35 | 7 | 177 | 38 | 198 | 42 | 48 | 10 | 9 | 2 | 467 |
| | 540 | 33 | 807 | 48 | 299 | 18 | 20 | 1 | 6 | 0 | 1,672 |
| 2.50-2.74 | 28 | 5 | 162 | 28 | 259 | 45 | 109 | 19 | 20 | 3 | 578 |
| | 547 | 25 | 1,003 | 46 | 538 | 25 | 67 | 3 | 23 | 1 | 2,178 |
| 2.25-2.49 | 12 | 2 | 101 | 17 | 307 | 53 | 131 | 23 | 28 | 5 | 579 |
| | 386 | 19 | 805 | 39 | 721 | 35 | 111 | 6 | 25 | 1 | 2,048 |
| 2.00-2.24 | 8 | 2 | 62 | 14 | 221 | 50 | 120 | 27 | 31 | 7 | 442 |
| | 177 | 11 | 542 | 34 | 654 | 41 | 173 | 11 | 46 | 3 | 1,592 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1.75- | 0 | 0 | 6 | 10 | 25 | 41 | 20 | 33 | 10 | 16 | 61 |
| 1.99 | 17 | 6 | 62 | 22 | 121 | 43 | 58 | 20 | 26 | 9 | 284 |
| 1.74- | 0 | 0 | 0 | 0 | 1 | 14 | 2 | 29 | 4 | 57 | 7 |
| Below | 1 | 3 | 3 | 10 | 17 | 55 | 4 | 13 | 19 | 6 | 31 |
| Totals | 232 | 8 | 874 | 29 | 1,340 | 44 | 493 | 16 | 112 | 4 | 3,051 |
| | 3,614 | 32 | 4,445 | 40 | 2,629 | 23 | 445 | 4 | 139 | 1 | 11,272 |

---

[a]Reproduced by permission of Registrar, Iowa State University.

[b]Quality-point average as of end of Spring Quarter 1970.

[c]First line: letter grades received (but not recorded) by students in pass-fail courses, Spring 1970.

[d]Second line: total letter grades received (and recorded) in courses other than pass-fail by above pass-fail enrollees.

For the 305 students enrolled under the pass-fail system, in addition, the number of grades received during the Spring Quarter 1970 in other than pass-fail courses was recorded according to general grade-point average in order to compare the grade distributions of the language students with the grade distributions of the entire university student population enrolled in all pass-fail courses. These grade distributions are seen in Tables 10 and 11.

Preliminary t-tests, ignoring other variables, indicate that the letter grade group and the pass-fail group did not differ significantly as to grade-point average (GPA), ACT composite score, or course load. The data for these t-tests which were based on separate group means with groups of unequal size, assumed to be random samples from two populations, are summarized in Table 12.

Table 12. Group I. Preliminary t-tests - summary of data

| | Mean | | Variance | | Calculated |
|---|---|---|---|---|---|
| | L-G[a] | P-F[b] | L-G | P-F | t[c] |
| GPA | 2.71 | 2.68 | .32 | .22 | .81 |
| ACT | 26.18 | 26.00 | 11.55 | 10.91 | .75 |
| Study load | 15.95 | 16.18 | 9.03 | 8.43 | .01 |

[a] L-G = letter-grade group (N = 484).

[b] P-F = pass-fail group (N = 266).

[c] Tabular t = 1.96 (.05).

As evident from Table 12, in both groups the average grade-point was 2.7, the average ACT score = 26, and the average course load, 16 credit hours. There was, however, a pronounced difference between the grade distributions of the two groups, as shown in Table 10.

Multiple regression analysis

The purpose of this analysis was to control for those independent variables which are assumed to contribute significantly to the variation in the dependent variable, $X_1$, course grade.

As evident from the correlation matrix in Table 13, computed for N = 895, and for a subsample of N = 750 including only students for whom an ACT score was available, the highest correlation between the dependent variable, $X_1$, course grade, and the independent variables is that between $X_1$ and $X_4$, cumulative grade-point average ($r_{14} = .49$). Next in size are the negative correlations of $X_1$ with $X_2$, grading system ($r_{12} = -.29$), and with $X_7$, cumulative credit hours ($r_{17} = -.27$).

The highest correlations between the independent variables occurred between $X_2$, grading system and $X_7$, cumulative credit hours ($r_{27} = .95$) and between $X_4$, grade-point average and $X_5$, composite ACT score ($r_{45} = .45$). With sample sizes N = 750 and 895, these correlations are significant at the .01 level.

A summary of data for the variables used in this analysis is given in Table 14.

The remaining variables, i.e., $X_2$, $X_3$, $X_{8-11}$, $X_{12-14}$, and $X_{15}$, are codes for the grading system (0 = P-F; 1 = L-G), sex (1 = male; 2 = female), college, language, and course sequence. In the computer regression program, every college and every language are treated as separate

Table 13. Product-moment correlation coefficient matrix for Group I

| Variables | | $X_1$ | $X_2$ | $X_3$ | $X_4$ | $X_5$ | $X_6$ | $X_7$ | $X_{15}$ |
|---|---|---|---|---|---|---|---|---|---|
| $X_1$ | Final grade | 1.000[a]<br>1.000 | | | | | | | |
| $X_2$ | Grading system | -.289<br>-.297 | 1.000<br>1.000 | | | | | | |
| $X_3$ | Sex[b] | .204<br>.183 | -.195<br>-.181 | 1.000<br>1.000 | | | | | |
| $X_4$ | Grade-point average | .493<br>.511 | -.022<br>-.047 | .154<br>.144 | 1.000<br>1.000 | | | | |
| $X_5$ | Act | -.188 | -.026 | -.005 | .450 | 1.000 | | | |
| $X_6$ | Course load | .173<br>.126 | .036<br>.056 | .105<br>.118 | .282<br>.206 | .182 | 1.000<br>1.000 | | |
| $X_7$ | Cum. credit hours | .273<br>.273 | .945<br>.941 | -.214<br>-.194 | -.016<br>-.034 | -.029 | .056<br>.075 | 1.000<br>1.000 | |
| $X_{15}$ | Course sequence | .012<br>.009 | .115<br>.096 | -.084<br>-.061 | .077<br>.068 | .123 | .024<br>.021 | .101<br>.084 | 1.000<br>1.000 |

[a]The upper correlations are based on N = 750; the lower correlations are based on N = 895.

[b]Sex is a discrete variable; males were coded 1, females 2.

Table 14. Group I. Summary of data - multiple regression analysis

| Variable | Mean | | Standard deviation | |
|---|---|---|---|---|
| | N = 750 | N = 895 | N = 750 | N = 895 |
| $X_1$ = course grade | 2.51 | 2.54 | 1.06 | 1.07 |
| $X_4$ = cumulative grade-point average | 2.70 | 2.72 | .53 | .54 |
| $X_5$ = ACT composite score[a] | 26.12 | - | 3.36 | - |
| $X_6$ = course load | 16.00 | 15.87 | 2.97 | 3.04 |
| $X_7$ = cumulative credit hours as per Spring Quarter 1970 | 50.60 | 48.66 | 72.32 | 71.96 |

[a]Only available for subsample size N = 750.

variables; in the regression analysis, as presented in this investigation, these variables are treated as two groups.

To test the null hypothesis of no difference in student achievement in elementary foreign language courses under the pass-fail system and under the letter grade system, each independent variable used in the complete model was investigated in turn in various combinations to investigate if it contributed significantly to the variation in course grade, $X_1$.

Grading system, $X_2$

Table 15 lists those models in which the grading system, $X_2$, is included as an independent variable. The null hypothesis that $\beta_2 = 0$ is tested against the alternative hypothesis that $\beta_2 \neq 0$ for each model,

Table 15. Group I. Summary of the calculated t-values for various multiple regression models used to investigate the effect of grading system, $X_2$, on the dependent variable, course grade, $X_1$

| Model | Independent variables included in regression model | Degrees of freedom | Calculated "t" |
|---|---|---|---|
| 1-IA[a] | $X_2X_3X_4X_5X_6X_7X_{8\text{-}11}X_{12\text{-}14}X_{15}$ | 735 | -3.26** |
| 2-IA | $X_2X_3X_4X_5X_6X_7X_{12\text{-}14}X_{15}$ | 739 | -3.29** |
| 3-IA | $X_2X_3X_4X_5X_6X_7X_{8\text{-}11}$ | 739 | -2.80** |
| 4-IA | $X_2X_3X_4X_6X_7X_{8\text{-}11}X_{12\text{-}14}X_{15}$ | 736 | -3.26** |
| 5-IA | $X_2X_4X_5X_6X_7X_{15}$ | 743 | -2.71** |
| 6-IA | $X_2X_4X_5X_6X_7$ | 744 | -2.69** |
| 7-IA | $X_2X_4X_5X_6X_{15}$ | 744 | -9.30** |
| 8-IA | $X_2X_4X_5X_7X_{15}$ | 744 | -2.79** |
| 9-IA | $X_2X_4X_5X_6$ | 745 | -9.33** |
| 10-IA | $X_2X_4X_5$ | 746 | -9.25** |
| 11-IA | $X_2X_5X_7$ | 746 | -2.80** |
| 12-IA | $X_2X_5X_{15}$ | 746 | -8.28** |
| 13-IA | $X_2X_5X_6$ | 746 | -8.55** |
| 20-IA | $X_2X_5$ | 747 | -8.27** |
| - | $X_2$ | 748 | -8.29** |
| 1-IB[b] | $X_2X_3X_4X_6X_7X_{8\text{-}11}X_{12\text{-}14}X_{15}$ | 881 | -3.60** |
| 2-IB | $X_2X_4X_6X_7X_{15}$ | 889 | -3.42** |
| 3-IB | $X_2X_4X_6$ | 891 | -10.07** |

[a] 1A= ACT-group (N = 750).

[b] 1B= Combined group (N = 895).

** Indicates $p < .01$.

Table 15 (continued)

| Model | Independent variables included in regression model | Degrees of freedom | Calculated "t" |
|---|---|---|---|
| 4-IB | $X_2X_4X_7$ | 891 | -3.47** |
| 5-IB | $X_2X_4X_{15}$ | 891 | -9.94** |
| 6-IB | $X_2X_6X_{15}$ | 891 | -9.69** |
| 7-IB | $X_2X_7X_{15}$ | 891 | -3.72** |
| 12-IB | $X_2X_4$ | 892 | -9.99** |
| 13-IB | $X_2X_6$ | 892 | -9.62** |
| 14-IB | $X_2X_7$ | 892 | -3.67** |
| 15-IB | $X_2X_{15}$ | 892 | -9.36** |
| - | $X_2$ | 893 | -9.28** |
| | Tabular "t" = 1.96 (.05)<br>Tabular "t" = 2.58 (.01) | | |

and the value of the calculated test statistic "t" is compared to a tabular "t" using the degrees of freedom for residual, and the .05 level of significance. The calculated "t" which exceeds the corresponding tabular "t" is marked with an asterisk.

Notice that the stated level of significance is not the actual level because more than one "t" test is performed using the same data. However, it is unknown. These comments hold for all the independent variables discussed in turn.

From Table 15 it is evident that grading system, $X_2$, explains a significant portion of the variation in course grade, $X_1$, when the effect of the independent variables $X_3$, sex, $X_4$, grade-point average, $X_5$, ACT, $X_6$, course load, $X_7$, cumulative credit hours, $X_{8-11}$, college, $X_{12-14}$, language, and $X_{15}$, course sequence, is accounted for when testing at the .05 level of significance. Whenever $X_7$, cumulative credit hours, is included in a model, the calculated "t" which tests the hypothesis that grading system, $X_2$, has no effect, becomes smaller; that is, less significant. This implies that inclusion of $X_7$, cumulative credit hours, in the model reduces the amount of variation left for $X_2$, grading system, to explain. One would expect, then, that $X_2$ and $X_7$ are correlated and, indeed, the correlation coefficient is quite large, $r_{27} = .94$, for the combined data (N = 895). Despite this high correlation between grading system, $X_2$, and cumulative credit hours, $X_7$, there is still a significant portion of the variation in course grades, $X_1$, explained by the grading system, $X_2$, after adjustment for cumulative credit hours, $X_7$.

The high correlation between grading system, $X_2$, and cumulative credit hours, $X_7$, can be attributed to the fact that fewer freshmen than sophomores, juniors and seniors are enrolled in pass-fail courses because of the eligibility prerequisite of 60 credit hours.

The negative t-values connected with the various models including the grading system as an independent variable provided evidence that the pass-fail group (coded "1" for the computer regression program; the letter grade group was coded "0") received lower grades than the letter grade group.

Sex, $X_3$[1]

Table 16 lists those models in which sex, $X_3$, is included as an independent variable. The null hypothesis that $\beta_3 = 0$ is tested against the alternative hypothesis that $\beta_3 \neq 0$, and the calculated "t" is compared with the tabular "t" with the degrees of freedom for the residual mean square for each model which includes $X_3$, and the level of significance is set at .05.

Table 16. Group I. Summary of calculated t-values, for various multiple regression models, used to investigate the effect of sex, $X_3$, on the dependent variable, final grade, $X_1$

| Model | Independent variables | Degrees of freedom | Calculated "t" |
|---|---|---|---|
| 1-IA | $X_2X_3X_4X_5X_6X_7X_{8\text{-}11}X_{12\text{-}14}X_{15}$ | 735 | 1.76 |
| 2-IA | $X_2X_3X_4X_5X_6X_7X_{12\text{-}14}X_{15}$ | 739 | 2.60** |
| 3-IA | $X_2X_3X_4X_5X_6X_7X_{8\text{-}11}$ | 739 | 1.32 |
| 4-IA | $X_2X_3X_4X_6X_7X_{8\text{-}11}X_{12\text{-}14}X_{15}$ | 736 | 1.75 |
| 1-1B | $X_2X_3X_4X_6X_7X_{8\text{-}11}X_{12\text{-}14}X_{15}$ | 881 | 1.34 |
| - | $X_3$ | 893 | 5.59** |

**Indicates $p < .01$.

[1]Code: 1 = male; 2 = female.

Inspection of Table 16 indicates that $X_3$, sex, explains a significant portion of the variation in $X_1$, final grade, when the effect of $X_{8-11}$, college, is removed from the complete regression model which adjusts for $X_2$, grading system, $X_4$, grade-point average, $X_5$, ACT score, $X_6$, course load, $X_7$, cumulative credit hours, $X_{12-14}$, language, and $X_{15}$, course sequence.

The t-value for $X_3$, sex, is also significant when $X_3$ is used alone (calculated t = 5.59 with 893 degrees of freedom; tabular t = 2.58 (.01)).

Since $X_3$, sex, is correlated significantly with both $X_1$, final grade ($r_{13}$ = .20) and $X_{8-11}$, college, it explains a larger portion of the variation in $X_1$, final grade, when $X_{8-11}$, college, is omitted from the regression model.

Grade-point average, $X_4$

Table 17 lists those models in which grade-point average, $X_4$, is included as an independent variable. The null hypothesis that $\beta_4 = 0$ is tested against the alternative hypothesis that $\beta_4 \neq 0$, and the calculated "t" is compared with the tabular "t" with the degrees of freedom for the residual mean square for each model which includes $X_4$, and the level of significance is set at .05.

Table 17 provides evidence to reject the null hypothesis that grade-point average, $X_4$, does not explain a significant portion of the variation in the dependent variable, $X_2$, course grade, after adjusting for the effect of the independent variables when testing at the .05 level of significance. Thus, there is evidence that grade-point average does explain a significant portion of the variance in final grade, $X_1$, after

Table 17. Group I. Summary of calculated t-values for various multiple regression models used to investigate the effect of grade-point average, $X_4$, on the dependent variable, course grade, $X_1$

| Model | Independent variables included in regression model | Degrees of freedom | Calculated "t" |
|---|---|---|---|
| 1-IA | $X_2X_3X_4X_5X_6X_7X_{8-11}X_{12-14}X_{15}$ | 735 | 14.28** |
| 2-IA | $X_2X_3X_4X_5X_6X_7X_{12-14}X_{15}$ | 739 | 14.33** |
| 3-IA | $X_2X_3X_4X_5X_6X_7X_{8-11}$ | 739 | 13.73** |
| 4-IA | $X_2X_3X_4X_6X_7X_{8-11}X_{12-14}X_{15}$ | 736 | 15.66** |
| 5-IA | $X_2X_4X_5X_6X_7X_{15}$ | 743 | 14.28** |
| 6-IA | $X_2X_4X_5X_6X_7$ | 744 | 14.30** |
| 7-IA | $X_2X_4X_5X_6X_{15}$ | 744 | 14.28** |
| 8-IA | $X_2X_4X_5X_7X_{15}$ | 744 | 15.04** |
| 9-IA | $X_2X_4X_5X_6$ | 745 | 14.30** |
| 10-IA | $X_2X_4X_5$ | 746 | 15.07** |
| 14-IA | $X_4X_5X_6$ | 746 | 13.74** |
| 15-IA | $X_4X_5X_7$ | 746 | 15.08** |
| 16-IA | $X_4X_5X_{15}$ | 746 | 14.39** |
| 21-IA | $X_4X_5$ | 747 | 14.38** |
| 1-IB | $X_2X_3X_4X_6X_7X_{8-11}X_{12-14}X_{15}$ | 881 | 17.63** |
| 2-IB | $X_2X_4X_6X_7X_{15}$ | 889 | 17.44** |
| 3-IB | $X_2X_4X_6$ | 891 | 17.51** |
| 4-IB | $X_2X_4X_7$ | 891 | 18.19** |
| 5-IB | $X_2X_4X_{15}$ | 891 | 18.14** |

**Indicates that $p < .01$.

Table 17 (continued)

| Model | Independent variables included in regression model | Degrees of freedom | Calculated "t" |
|---|---|---|---|
| 8-IB | $X_4X_6X_7$ | 891 | 17.53** |
| 9-IB | $X_4X_6X_{15}$ | 891 | 17.23** |
| 10-IB | $X_4X_7X_{15}$ | 891 | 18.20** |
| 12-IB | $X_2X_4$ | 892 | 18.21** |
| 16-IB | $X_4X_6$ | 892 | 17.21** |
| 17-IB | $X_4X_7$ | 892 | 18.25** |
| 18-IB | $X_4X_{15}$ | 892 | 17.76** |
| - | $X_4$ | 893 | 17.74** |

adjustment for all or some of the independent variables. This is expected from the size of the simple correlation coefficient, $r_{14}$, which for the combined data (N = 895) is .51, significant at the .01 level with 894 degrees of freedom. The highly significant t-values indicate that, on the average, the higher the grade-point average the better the final grade obtained in elementary foreign language courses.

ACT, $X_5$

Table 18 lists those models in which ACT, $X_5$, is included as an independent variable. The null hypothesis that $\beta_5 = 0$ is tested against the alternative hypothesis that $\beta_5 \neq 0$, and the calculated "t" is compared with the tabular "t" with the degrees of freedom for the residual

Table 18. Group I. Summary of calculated t-values for various multiple regression models used to investigate the effect of ACT, $X_5$, on the dependent variable, course grade, $X_1$

| Model | Independent variables included in regression model | Degrees of freedom | Calculated t |
|---|---|---|---|
| 1-IA | $X_2X_3X_4X_5X_6X_7X_{8-11}X_{12-14}X_{15}$ | 735 | 0.21 |
| 2-IA | $X_2X_3X_4X_5X_6X_7X_{12-14}X_{15}$ | 739 | 0.10 |
| 3-IA | $X_2X_3X_4X_5X_6X_7X_{8-11}$ | 739 | -1.02 |
| 5-IA | $X_2X_4X_5X_6X_7X_{15}$ | 743 | -1.57 |
| 6-IA | $X_2X_4X_5X_6X_7$ | 744 | -1.54 |
| 7-IA | $X_2X_4X_5X_6X_{15}$ | 744 | -1.56 |
| 8-IA | $X_2X_4X_5X_7X_{15}$ | 744 | -1.46 |
| 9-IA | $X_2X_4X_5X_6$ | 745 | -1.53 |
| 10-IA | $X_2X_4X_5$ | 746 | -1.43 |
| 11-IA | $X_2X_5X_7$ | 746 | 5.25** |
| 12-IA | $X_2X_5X_{15}$ | 746 | 5.12** |
| 13-IA | $X_2X_5X_6$ | 746 | 4.40** |
| 14-IA | $X_4X_5X_6$ | 746 | -1.27 |
| 15-IA | $X_4X_5X_7$ | 746 | -1.47 |
| 16-IA | $X_4X_5X_{15}$ | 746 | -1.12 |
| 17-IA | $X_5X_6X_7$ | 746 | 4.33** |
| 18-IA | $X_5X_6X_{15}$ | 746 | 4.48** |
| 19-IA | $X_5X_7X_{15}$ | 746 | 5.10** |
| 20-IA | $X_2X_5$ | 747 | 5.25** |

**Indicates that $p < .01$.

Table 18 (continued)

| Model | Independent variables included in regression model | Degrees of freedom | Calculated t |
|---|---|---|---|
| 21-IA | $X_4X_5$ | 747 | -1.19 |
| 22-IA | $X_5X_6$ | 747 | 4.48** |
| 23-IA | $X_5X_7$ | 747 | 5.20** |
| 24-IA | $X_5X_{15}$ | 747 | 5.23** |
| - | $X_5$ | 748 | 5.25** |

mean square for each model which includes $X_5$, and the level of significance is set at .05.

Table 18 provides evidence that ACT, $X_5$, explains a significant portion of the variation in final grade, $X_1$, if grade-point average, $X_4$, is removed from any of the regression models. The explanation for this is that ACT, $X_5$, is significantly correlated with achievement in elementary foreign language courses, as measured by final grade, $X_1$, ($r_{15}$ = .188, N = 750, level of significance = .01). However, ACT, $X_5$, is also highly correlated with grade-point average, $X_4$ ($r_{45}$ = .45), N = 750) and, therefore, does not reduce significantly the unexplained variance in final grade, $X_1$, when grade-point average, $X_4$, is included in the model.

Course load, $X_6$

Table 19 lists those models in which course load is included as an independent variable. The null hypothesis that $\beta_6 = 0$ is tested against the alternative hypothesis that $\beta_6 \neq 0$, and the calculated "t" is compared to the tabular "t" with the degrees of freedom for the residual mean square for each model which includes $X_6$ and the level of significance set at .05.

From Table 19 it is evident that the influence of course load, $X_6$, depends on which other independent variables are included in the regression model and on which sample is investigated, the larger group (N = 895) which includes students for whom ACT scores were not available, or the smaller group (N = 750) whose members had ACT scores.

Thus, course load, $X_6$, explains a significant portion of the variation in final grade, $X_1$, after adjustment for the other independent variables for the sample of N = 750 but not for the sample of size N = 895.

Course load, $X_6$, becomes insignificant when language, $X_{12-14}$, is excluded from the complete model when N = 750, but remains significant when college, $X_{8-11}$, is excluded. For both samples when college, $X_{8-11}$, language, $X_{12-14}$, and grade-point average, $X_4$, are excluded simultaneously from the model leaving grading system, $X_2$, sex, $X_3$, ACT, $X_5$ (for sample size N = 750), cumulative credit hours, $X_7$, and course sequence, $X_{15}$, or any combination of them, course load, $X_6$, becomes significant at the .01 level.

Table 19. Group I. Summary of calculated t-values for various multiple regression models used to investigate the effect of course load, $X_6$, on the dependent variable, course grade, $X_1$

| Model | Independent variables included in regression model | Degrees of freedom | Calculated t |
|---|---|---|---|
| 1-IA | $X_2X_3X_4X_5X_6X_7X_{8\text{-}11}X_{12\text{-}14}X_{15}$ | 735 | 2.08* |
| 1a-IA | $X_2X_4X_5X_6X_7X_{8\text{-}11}X_{12\text{-}14}X_{15}$ | 736 | 2.21* |
| 2-IA | $X_2X_3X_4X_5X_6X_7X_{8\text{-}11}X_{15}$ | 738 | 1.42 |
| 2a-IA | $X_2X_3X_4X_5X_6X_7X_{12\text{-}14}X_{15}$ | 739 | 2.16* |
| 3-IA | $X_2X_3X_4X_5X_6X_7X_{8\text{-}11}$ | 739 | 1.43 |
| 4-IA | $X_2X_3X_4X_6X_7X_{8\text{-}11}X_{12\text{-}14}X_{15}$ | 736 | 2.10* |
| 5-IA | $X_2X_4X_5X_6X_7X_{15}$ | 743 | 1.69 |
| 6-IA | $X_2X_4X_5X_6X_7$ | 744 | 1.69 |
| 7-IA | $X_2X_4X_5X_6X_{15}$ | 744 | 1.67 |
| 9-IA | $X_2X_4X_5X_6$ | 745 | 1.67 |
| 13-IA | $X_2X_5X_6$ | 746 | 4.50** |
| 14-IA | $X_4X_5X_6$ | 746 | 1.18 |
| 17-IA | $X_5X_6X_7$ | 746 | 4.62** |
| 18-IA | $X_5X_6X_{15}$ | 746 | 3.96** |
| 22-IA | $X_5X_6$ | 747 | 3.96** |
| 1-IB | $X_2X_3X_4X_6X_7X_{8\text{-}11}X_{12\text{-}14}X_{15}$ | 881 | 1.75 |
| 2-IB | $X_2X_4X_6X_7X_{15}$ | 889 | 1.45 |
| 3-IB | $X_2X_4X_6$ | 891 | 1.46 |
| 6-IB | $X_2X_6X_{15}$ | 891 | 4.50** |

*Indicates $p < .05$

**Indicates $p < .01$.

Table 19 (continued)

| Model | Independent variables included in regression model | Degrees of freedom | Calculated t |
|---|---|---|---|
| 8-IB | $X_4X_6X_7$ | 891 | 1.56 |
| 9-IB | $X_4X_6X_{15}$ | 891 | 0.75 |
| 11-IB | $X_6X_7X_{15}$ | 891 | 4.60** |
| 13-IB | $X_2X_6$ | 892 | 4.53** |
| 16-IB | $X_4X_6$ | 892 | 0.74 |
| 19-IB | $X_6X_7$ | 892 | 4.62** |
| 20-IB | $X_6X_{15}$ | 892 | 3.79** |
| - | $X_6$ | 893 | 3.72** |

Course load, $X_6$, when used alone, that is, without adjusting for other variables, explains a significant portion of the variation in final grade, $X_1$.

Cumulative credit hours, $X_7$

Table 20 lists those models in which cumulative credit hours, $X_7$, is included as an independent variable. The null hypothesis that $\beta_7 = 0$ is tested against the alternative hypothesis that $\beta_7 \neq 0$, and the calculated "t" is compared to the tabular "t" with the degrees of freedom for the residual mean square for each model which includes $X_7$ and the level of significance set at .05.

Table 20. Group I. Summary of calculated t-values for various multiple regression models used to investigate the effect of cumulative credit hours, $X_7$, on the dependent variable, course grade, $X_1$

| Model | Independent variables included in regression model | Degrees of freedom | Calculated t |
|---|---|---|---|
| 1-IA | $X_2X_3X_4X_5X_6X_7X_{8-11}X_{12-14}X_{15}$ | 735 | 0.22 |
| 2-IA | $X_2X_3X_4X_5X_6X_7X_{12-14}X_{15}$ | 739 | 0.21 |
| 3-IA | $X_2X_3X_4X_5X_6X_7X_{8-11}$ | 739 | 0.06 |
| 4-IA | $X_2X_3X_4X_6X_7X_{8-11}X_{12-14}X_{15}$ | 736 | 0.21 |
| 5-IA | $X_2X_4X_5X_6X_7X_{15}$ | 743 | -0.38 |
| 6-IA | $X_2X_4X_5X_6X_7$ | 744 | -0.39 |
| 8-IA | $X_2X_4X_5X_7X_{15}$ | 744 | -0.27 |
| 11-IA | $X_2X_5X_7$ | 746 | 0.10 |
| 15-IA | $X_4X_5X_7$ | 746 | -8.79** |
| 17-IA | $X_5X_6X_7$ | 746 | -8.11** |
| 19-IA | $X_5X_7X_{15}$ | 746 | -7.74** |
| 23-IA | $X_5X_7$ | 747 | -7.74** |
| 1-IB | $X_2X_3X_4X_6X_7X_{8-11}X_{12-14}X_{15}$ | 881 | 0.10 |
| 2-IB | $X_2X_4X_6X_7X_{15}$ | 889 | 0.01 |
| 4-IB | $X_2X_4X_7$ | 891 | 0.10 |
| 7-IB | $X_2X_7X_{15}$ | 891 | 0.58 |
| 8-IB | $X_4X_6X_7$ | 891 | -9.41** |
| 10-IB | $X_4X_7X_{15}$ | 891 | -9.25** |
| 11-IB | $X_6X_7X_{15}$ | 891 | -8.95** |

**Indicates that $p < .01$.

Table 20 (continued)

| Model | Independent variables included in regression model | Degrees of freedom | Calculated t |
|---|---|---|---|
| 14-IB | $X_2X_7$ | 892 | 0.56 |
| 17-IB | $X_4X_7$ | 892 | -9.30** |
| 19-IB | $X_6X_7$ | 892 | -8.90** |
| 21-IB | $X_7X_{15}$ | 892 | -8.54** |

From Table 20 it is evident that cumulative credit hours, $X_7$, explains a significant portion of the variation in final grade, $X_1$, only when grading system, $X_2$, is omitted from the model, regardless of what other variables are included.

This can be explained by the high correlation of cumulative credit hours, $X_7$, with both grading system, $X_2$ ($r_{27}$ = .94 when N = 895) and final grade, $X_1$ ($r_{12}$ = .27 when N = 895). When grading system, $X_2$, is omitted from a model, the portion of variation in final grade which had been explained by grading system is now explained by cumulative credit hours.

College, $X_{8\text{-}11}$

The four degrees of freedom associated with this group represent the contrasts between the Colleges of Sciences and Humanities, Engineering, Agriculture, Education and Home Economics. For the computer program, these variables were individually coded as follows:

| | |
|---|---|
| Sciences and Humanities | 1000 |
| Engineering | 0100 |
| Education | 0001 |
| Agriculture | 0010 |
| Home Economics | 0000 |

In the regression analysis, these variables were eliminated as a group to investigate their contribution in explaining the variability in final grade, $X_1$.

The following analysis of regression Table 21 tests the null hypothesis that $\beta_8 = \beta_9 = \beta_{10} = \beta_{11} = 0$ against the alternative hypothesis that at least one of $\beta_8$, $\beta_9$, $\beta_{10}$, $\beta_{11} \neq 0$.

Table 21. Group I. Analysis of regression: effect of the elimination of College, $X_{8-11}$

| Source of variation | D.F.[a] | Sum of squares | Mean square |
|---|---|---|---|
| Regression on $X_2X_3X_4X_5X_6X_7X_{8-11}X_{12-14}X_{15}$ | 14 | 334.471 | |
| Regression on $X_2X_3X_4X_5X_6X_7X_{12-14}X_{15}$ | 10 | 328.890 | |
| Difference due to elimination of $X_{8-11}$ | 4 | 5.581 | 1.395 |
| Residual | 735 | 515.008 | 0.701 |

Calculated F = 1.9913
Tabular $F_{4,735}$ = 2.38 (.05)

[a]Degrees of freedom.

The calculated F-ratio is not significant when testing at the .05 level. This means that the elimination of College, $X_{8-11}$, from the regression model would be justified since it did not explain a significant portion of the variability of the dependent variable, $X_1$, final grade, in elementary foreign language courses during the Spring Quarter 1970, after adjusting for the other independent variables contained in the complete model (N = 750).

Language, $X_{12-14}$

The three degrees of freedom associated with this group represent the contrasts between French, German, Russian, and Spanish. For the computer program, these variables were individually coded as follows:

| | |
|---|---|
| French | 100 |
| German | 010 |
| Russian | 001 |
| Spanish | 000 |

In the regression analysis, these variables were eliminated as a group to investigate their contribution in explaining the variability in final grade, $X_1$.

The following analysis of regression Table 22 tests the null hypothesis that $\beta_{12} = \beta_{13} = \beta_{14} = 0$ against the alternative hypothesis that at least one of $\beta_{12}$, $\beta_{13}$, $\beta_{14} \neq 0$.

The calculated F-ratio is significant at the .01 level which means that language explained a significant portion of the variation in final grade, $X_1$. Thus, there is sufficient evidence to reject hypothesis 4 of no difference in average final grades for students studying different languages.

Table 22. Group I. Analysis of regression: effect of the elimination of language, $X_{12-14}$

| Source of variation | D.F.[a] | Sum of squares | Mean square |
|---|---|---|---|
| Regression on $X_2X_3X_4X_5X_6X_7X_{8-11}X_{12-14}X_{15}$ | 14 | 334.471 | |
| Regression on $X_2X_3X_4X_5X_6X_7X_{8-11}X_{15}$ | 11 | 287.645 | |
| Difference due to elimination of $X_{12-14}$ | 3 | 46.826 | 15.61 |
| Residual | 735 | 515.008 | 0.70 |

Calculated F = 22.27
Tabular $F_{3,735}$ = 2.62 (.05)
= 3.83 (.01)

[a]Degrees of freedom.

Course sequence, $X_{15}$

Table 23 lists the regression models in which course sequence, $X_{15}$, is included. The null hypothesis that $\beta_{15} = 0$ is tested against the alternative hypothesis that $\beta_{15} \neq 0$, and the calculated "t" is compared with the tabular "t" with the degrees of freedom for the residual mean square for each model which includes $X_{15}$. The level of significance is set at .05.

From Table 23 it is evident that there is no evidence to reject the null hypothesis that course sequence, $X_{15}$, does not explain a significant portion of the variation in final grade, $X_1$, after adjusting for any or all of the independent variables when testing at the .05 level of significance. The tabular "t" with the degrees of freedom for the residual mean square is 1.96.

Table 23. Group I. Summary of the calculated t-values for various multiple regression models used to investigate the effect of course sequence, $X_{15}$, on the dependent variable, course grade, $X_1$

| Model | Independent variables included in regression model | Degrees of freedom | Calculated t |
|---|---|---|---|
| 1-IA | $X_2X_3X_4X_5X_6X_7X_{8-11}X_{12-14}X_{15}$ | 735 | 0.10 |
| 2-IA | $X_2X_3X_4X_5X_6X_7X_{12-14}X_{15}$ | 739 | -0.47 |
| 4-IA | $X_2X_3X_4X_6X_7X_{8-11}X_{12-14}X_{15}$ | 736 | -0.09 |
| 5-IA | $X_2X_4X_5X_6X_7X_{15}$ | 743 | 0.36 |
| 7-IA | $X_2X_4X_5X_6X_{15}$ | 744 | 0.37 |
| 8-IA | $X_2X_4X_5X_7X_{15}$ | 744 | 0.34 |
| 12-IA | $X_2X_5X_{15}$ | 746 | 0.65 |
| 16-IA | $X_4X_5X_{15}$ | 746 | -0.72 |
| 18-IA | $X_5X_6X_{15}$ | 746 | -0.33 |
| 19-IA | $X_5X_7X_{15}$ | 746 | 0.48 |
| 24-IA | $X_5X_{15}$ | 747 | -0.33 |
| 1-IB | $X_2X_3X_4X_6X_7X_{8-11}X_{12-14}X_{15}$ | 881 | -0.08 |
| 2-IB | $X_2X_4X_6X_7X_{15}$ | 889 | 0.07 |
| 5-IB | $X_2X_4X_{15}$ | 891 | 0.07 |
| 6-IB | $X_2X_6X_{15}$ | 891 | 1.11 |
| 7-IB | $X_2X_7X_{15}$ | 891 | 1.21 |
| 9-IB | $X_4X_6X_{15}$ | 891 | -0.89 |
| 10-IB | $X_4X_7X_{15}$ | 891 | -0.12 |
| 11-IB | $X_6X_7X_{15}$ | 891 | 0.93 |

**Indicates that $p < .01$.

Table 23 (continued)

| Model | Independent variables included in regression model | Degrees of freedom | Calculated t |
|---|---|---|---|
| 15-IB | $X_2X_{15}$ | 892 | 1.19 |
| 18-IB | $X_4X_{15}$ | 892 | -0.88 |
| 20-IB | $X_6X_{15}$ | 892 | 0.18 |
| 21-IB | $X_7X_{15}$ | 892 | 1.01 |

Instructors

The effect of 25 instructors upon achievement in 52 sections of elementary foreign language courses taught during the Spring Quarter 1970 is tested by the analysis of variance Table 24, which involves the following two hypotheses:

Hypothesis 2:

There is no difference among the average grades for students taught by different instructors. That is, there is no significant variation in the mean achievement of students taught by different instructors as measured by the final grade based on individual teaching and evaluation methods.

Hypothesis 3:

There is no difference among the mean grades for sections within an instructor. That is, there is no significant variation in the mean achievement of sections taught by the same instructor as measured by the final grade.

Table 24. Analysis of variance of final grades in elementary foreign language courses among instructors, and among sections within instructors[a]

| Source of variation | Degrees of freedom | Sum of squares | Mean square | F-ratio |
|---|---|---|---|---|
| Instructors | 24 | 5.48 | .2283 | 2.5479 |
| Sections within instructors | 27 | 2.42 | .0896 | .9813 |
| Students within sections within instructors | 848 | 77.44 | .0913 | |
| Total | 899 | | | |

[a] Unweighted means analysis.

There is evidence to reject hypothesis 2 at the .05 level of significance. The calculated F = 2.55, which exceeds the tabular F = 1.93 with 24 and 27 degrees of freedom at the .05 level of significance.

It is, therefore, concluded with a risk of 5 in 100 of being wrong that there is a difference in the mean achievement of students taught by different instructors. That is, on the average, students achieve differently depending on their instructor.

There is insufficient evidence to reject hypothesis 3. The calculated F = .9813 is smaller than the tabular F = 1.67 with 27 and 848 degrees of freedom at the .05 level of significance.

It is, therefore, concluded that there is no difference in the mean achievement of students taught by the same instructor.

## Discussion

Since the focal point of interest in this part of the investigation was a confrontation of two methods of evaluation, namely, of the pass-fail and letter-grade systems, the first independent variable to be discussed is the grading system and its impact upon achievement in elementary foreign language courses.

### Grading system, $X_2$

Preliminary t-tests, ignoring other variables, indicated that the letter-grade students and the pass-fail students did not differ significantly as to grade-point average, ACT composite score or course load. Thus, differences in achievement between the two groups, if not attributable to other variables investigated by analysis of multiple regression, were assumed to be due to the impact of method of evaluation.

Analysis of multiple regression substantiated this assumption. The grading system explained a highly significant portion of the variation in final grades in elementary foreign language courses after adjusting for all, or some, of the other independent variables involved.

The negative t-values connected with the various models including the grading system as an independent variable provided evidence that the pass-fail group (coded "1" for the computer regression program; the

control group was coded "0") received lower grades than the letter grade group.

The analysis of Group I, that is, of the students enrolled in elementary foreign language courses during the Spring Quarter 1970, indicates that there is a striking difference in the performance of students in foreign language courses taken under the pass-fail grading system and under the letter-grade system of evaluation. These contrasts are illustrated by Figure 1, where almost identical grade distributions were obtained for pass-fail students in foreign languages, and in other subjects taken under the pass-fail system. These distributions are characterized by oppositely skewed trends. This inverse relationship is also reflected in the negative correlation between grading system, $X_2$, and final grade, $X_1$ ($r_{12} = -.30$). The distributions of grades according to grade-point average, as shown in Tables 10 and 11, clearly indicate that even students with a high grade-point average differ drastically in their performance in pass-fail courses, foreign language or other, from their performance in non-pass-fail courses. Thus, for instance, Iowa State University students with grade-point averages ranging from 3.75 to 4.00 received 35% A's in pass-fail courses as compared to 87% in courses taken under the letter-grade system.

The logical conclusion, considering these findings, is that elementary foreign language courses or any subjects involving the learning of a skill should not be offered under the pass-fail grading system if results and adequate progress are expected.

This does not mean that advanced foreign language courses involving literature and other aspects of the foreign culture, for which elementary

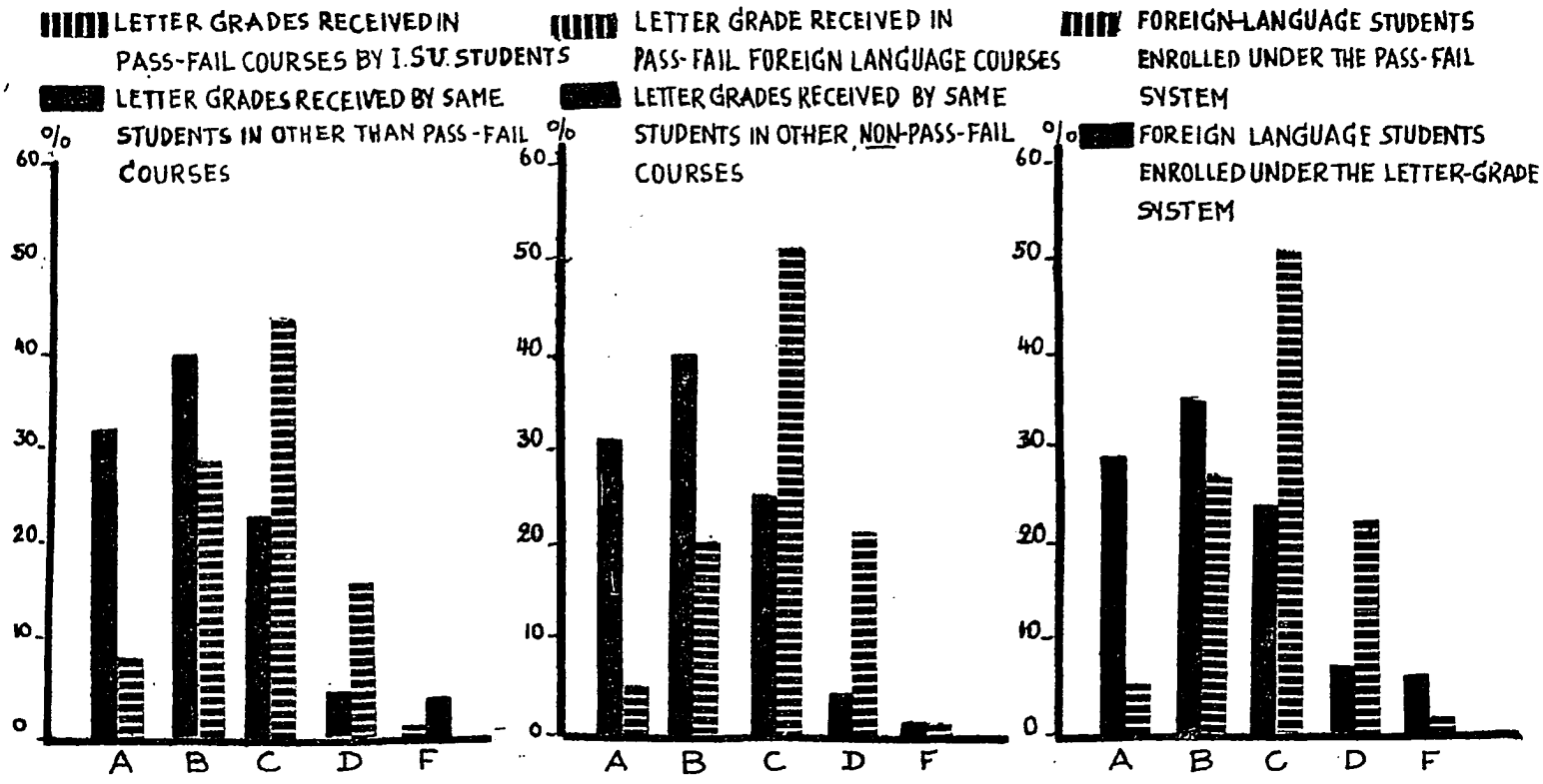



FIGURE 1. COMPARISON OF LETTER-GRADE AND PASS-FAIL GRADE DISTRIBUTIONS - SPRING QUARTER 1970
IOWA STATE UNIVERSITY

courses are a prerequisite, could not be taken under the pass-fail system if only a widening of horizons is the goal. However, foreign language majors, like majors in other subjects, should not be allowed to fulfill the requirements by taking language courses under the pass-fail system as it is organized now in most institutions, namely, with a D passing level.

Instructors

The instructor variable and its manifestations in student achievement, as measured by the final grades for the Spring Quarter 1970, was investigated by analysis of variance. Analysis of multiple regression could not be applied because the computer was not able to handle hypothesis 1 which tested differences in achievement among 52 elementary foreign language sections taught by 25 instructors.

For Group I, since no objective measures of achievement had been employed, final grades would have presented a rather elusive quantity if not for the normalizing effect of sample size (N = 895). For the pass-fail sample, the grade pattern had become so consistent that differences in achievement between pass-fail and letter-grade students had to be attributed to the effect of the grading systems rather than to the instructor variable.

However, analysis of variance provided statistical evidence that there is a difference in mean achievement of students taught by different instructors. The differences within sections taught by the same instructor were insignificant.

To facilitate research in foreign language teaching and to improve the learning situation for the students in elementary foreign language

courses, it is recommended that instructors of elementary course sequences agree as to common, precisely defined learning tasks, so-called behavioral goals, as discussed in the introduction to this investigation and use uniform methods of evaluation while given freedom of how to achieve these goals.

Grade-point average, $X_4$

Grade-point average has been proven to be the best predictor of success in foreign language study by previous research (19, 112). These findings were confirmed by the present investigation. Grade-point average explained a highly significant portion of the variability in achievement in elementary foreign language courses, as measured by final grade.

Course load, $X_6$

Course load turned out to be one of the most puzzling variables in this investigation. According to the results of research in areas other than foreign languages, course load had not been a factor interfering with achievement (37, 50); that is, increased course loads did not result in lower grades. These findings were corroborated, with one exception, by the results of the present investigation.

For Group I (N = 895 and the ACT subsample, N = 750), which involved all the students enrolled in elementary foreign language courses during the Spring Quarter 1970, course load was significantly correlated with grade-point average ( r = .28 when N = 750; r = .21 when N = 895), which means that students with a higher grade-point average tended to carry higher course loads.

Table 25 confirms these findings. Thus, 82 percent of the students with grade-point averages ranging from 2.50 to 4.00, i.e., A and B students, carried between 21 and 25 credit hours as compared to only 18 percent of the students with grade-point average 2.25 and below. At the other extreme, 44 percent of the A and B students carried course-loads of 10 credit hours and less, as compared to 56 percent of the C, D, and F students. The relatively small difference between students with a high grade-point average and those with lower grade-point averages at the low end of the course load ( 1 - 10 credit hours) was caused by 8 graduate students (38% of the high grade-point range) who had grade-point averages between 3.19 and 3.88, and who carried only 7 to 10 credit hours during the Spring Quarter 1970.

Inspection of Figure 2 indicates that there were no significant differences between pass-fail and letter grade students as far as course load is concerned.

For the ACT subsample (N=750), course load explained a significant portion of the variation in final grade after adjusting for all the other independent variables involved. It became, however, insignificant when the language variable was removed from the complete model. This is due to the fact that language as an independent variable explained a significant portion of the variability in final grade. Thus, its exclusion from the regression model resulted in a considerable increase in error, that is, in unexplained variation, which weakened the significance of the other independent variables included in the reduced model.

The elimination of the college variable from the complete model did not have any effect upon the significance of course load in explaining

Table 25. Group I. Distribution of course load by grade-point average

| GPA | Credit hours 1-10 Number | 1-10 Percent[a] | 11-15 Number | 11-15 Percent | 16-20 Number | 16-20 Percent |
|---|---|---|---|---|---|---|
| 1. 3.75-4.00 | 2 | 5.7 | 10 | 28.6 | 19 | 54.3 |
| 2. 3.50-3.74 | 3 | 5.0 | 8 | 13.3 | 46 | 76.7 |
| 3. 3.25-3.49 | 3 | 4.2 | 18 | 25.0 | 48 | 66.6 |
| 4. 3.00-3.24 | 3 | 3.0 | 29 | 29.0 | 61 | 61.0 |
| 5. 2.75-2.99 | 5 | 4.1 | 40 | 33.1 | 72 | 59.5 |
| 6. 2.50-2.74 | 5 | 3.3 | 48 | 31.6 | 88 | 57.9 |
| 7. 2.25-2.49 | 9 | 5.7 | 51 | 32.5 | 91 | 58.0 |
| 8. 2.00-2.24 | 9 | 6.8 | 55 | 41.3 | 69 | 51.9 |
| 9. 1.75-1.99 | 8 | 13.3 | 31 | 51.7 | 21 | 35.0 |
| 10. 1.74-below | 1 | 20.0 | 2 | 40.0 | 2 | 40.0 |
| Total | 48 | | 292 | | 517 | |
| | 5.4% | | 32.6% | | 57.8% | |

[a] Percentages were calculated for course load within grade-point range, i.e., for rows.

| Credit hours | | | | | |
|---|---|---|---|---|---|
| 21-25 | | 26-above | | Total | |
| Number | Percent | Number | Percent | Number | Percent |
| 2 | 5.7 | 2 | 5.7 | 35 | 3.9 |
| 1 | 1.7 | 2 | 3.3 | 60 | 6.7 |
| 3 | 4.2 | 0 | 0 | 72 | 8.0 |
| 7 | 7.0 | 0 | 0 | 100 | 11.2 |
| 4 | 3.3 | 0 | 0 | 121 | 13.5 |
| 10 | 6.6 | 1 | 0.6 | 152 | 17.0 |
| 6 | 3.8 | 0 | 0 | 157 | 17.5 |
| 0 | 0 | 0 | 0 | 133 | 14.9 |
| 0 | 0 | 0 | 0 | 60 | 6.7 |
| 0 | 0 | 0 | 0 | 5 | 0.6 |
| 33 | | 5 | | 895 | |
| 3.7% | | 0.5% | | | 100.0% |

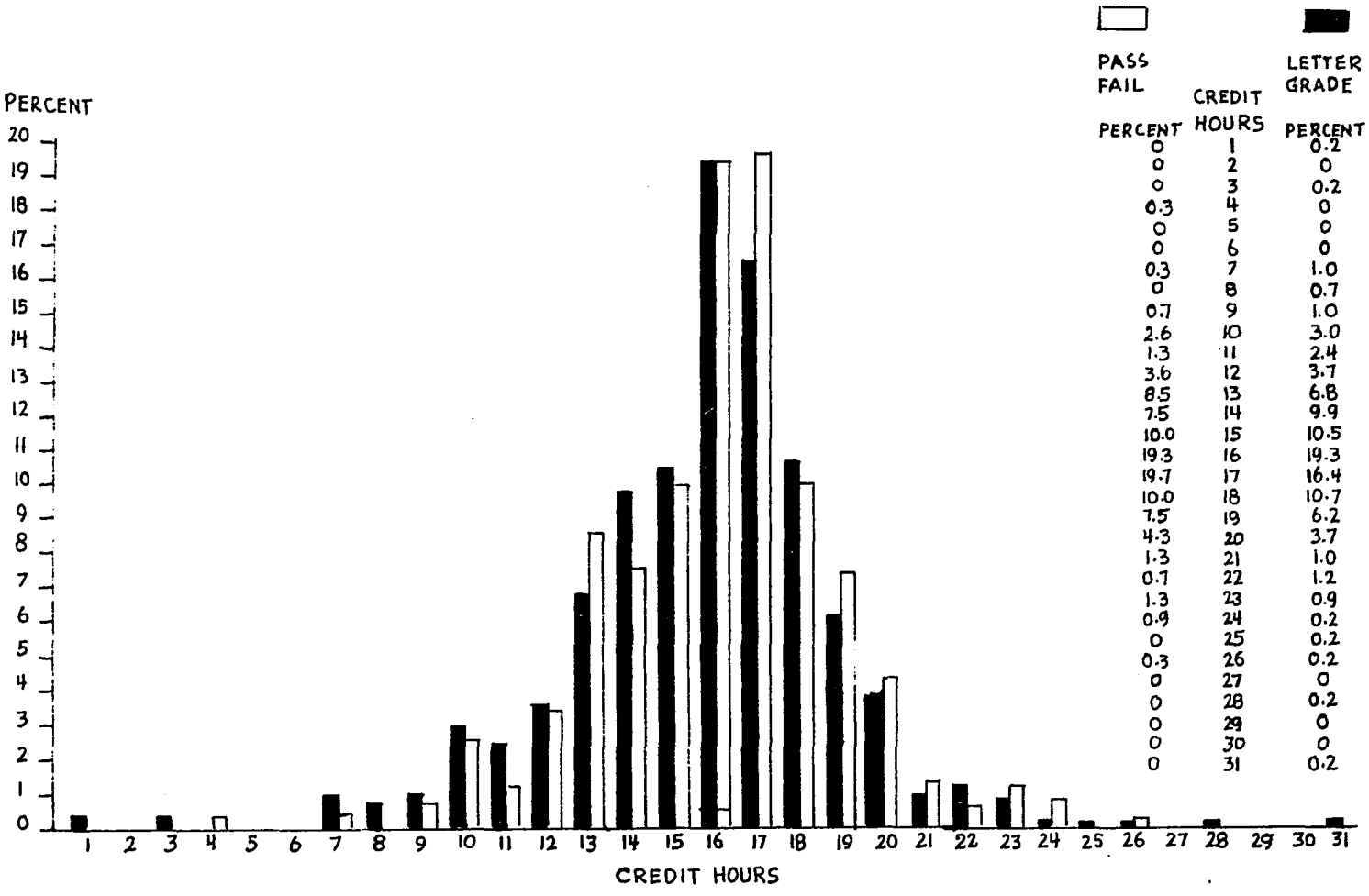



FIGURE 2. SPRING QUARTER 1970: DISTRIBUTION OF PASS-FAIL AND LETTER GRADE STUDENTS ENROLLED IN ELEMENTARY FOREIGN LANGUAGE COURSES BY COURSE LOAD

the variation in final grade for the ACT subsample (N = 750). This can be attributed to the fact that the college variable did not affect achievement in elementary foreign language courses, as will be seen later.

After adjusting for other variables, course load did not explain a significant portion of the variation in final grade for the combined group (N = 895). This discrepancy in results could be ascribed to random sampling variation, or it may be that students who do not have ACT scores differ from those who do with regard to their response to course load. Also, the mean course loads differed slightly for the two groups. The ACT group had an average course load of 16.03 with standard deviation 2.97; the combined group, which included also the students without ACT scores, had a mean course-load of 15.87 with standard deviation 3.04.

For both groups, course load explained a significant portion of the variation in final grade whenever the college variable, language, and grade-point average were excluded simultaneously. The increase in significance of course load when grade-point average was eliminated from the reduced models, according to statistical theory, can be explained by the correlation of grade-point average with both final grade and course load. If two independent variables correlate highly with each other, and one correlates highly with the dependent variable, then if this independent variable is omitted from the model, the residual variation is changed very little, and the independent variable remaining in the model explains the portion of the variation in the dependent variable which had been explained by both independent variables.

Cumulative credit hours, $X_7$

Cumulative credit hours, as an independent variable, explained a significant portion of final grade for Group I only when the grading system was eliminated from the regression model regardless of what other variables were included. This was explained by the high correlation of cumulative credit hours with both grading system ($r = .94$ when $N = 895$) and final grade ($r = .27$ when $N = 895$). When grading system is omitted from a model, the portion of variation in final grade which had been explained by grading system is now explained by cumulative credit hours. Vice versa, as evident from Table 15, the t-values for grading system decrease when cumulative credit hours is included in the regression model.

The high correlation between grading system and cumulative credit hours could be ascribed to the fact that out of 375 freshmen and sophomores, only 83, that is, 23 percent, were enrolled in pass-fail elementary foreign language courses during the Spring Quarter 1970 as compared to 44 percent of the juniors and seniors. One of the reasons for this may be that a student must have 60 cumulative credit hours before he can take pass-fail courses.

The negative t-values indicate that, on the average, as the accumulation of credit hours increased, achievement in elementary foreign language courses, as measured by final grade, decreased.

Sex, $X_3$

The sex ratio at Iowa State University is about 30 percent females and 70 percent males. This proportion held also for the various samples of foreign language students. Sex as an independent variable was in-

vestigated in connection with Groups I and III.

In Group I, which included 303 coeds, i.e., 33 percent of the total population enrolled in elementary foreign language courses during the Spring Quarter 1970, sex, after adjusting for other variables, explained a significant portion of the variation in final grade when the effect of the college variable was removed from the complete model. The explanation for this is that sex is significantly correlated with both final grade and the college variable. Therefore, when college is omitted from the regression model, sex becomes more significant in explaining the variation in final grade. In this group, 87 percent of the pass-fail students and 84.4 percent of the letter grade students were enrolled in the College of Sciences and Humanities. The positive t-values suggest that the coeds received better grades in elementary foreign language courses than the male students (computer code: 1 = male; 2 = female). These findings confirm research at the secondary school level (22) that girls "have higher scores on tests and tend to get higher marks in language courses in school, particularly in the upper grades (Grades 11-13)" (21); however, insufficient evidence until now was available to compare adult males and females.

ACT composite score, $X_5$

This variable was investigated only in connection with Group I where a sample of 750 students having ACT scores was available.

The ACT composite score explained a significant portion of the variance in final grade in elementary foreign language courses during the Spring Quarter 1970, if grade-point average was removed from the regression model. The explanation for this is that ACT is significantly

correlated with final grade but also with grade-point average ( r = .45 when N = 750), and, therefore, does not reduce significantly the unexplained variation in final grade when grade-point average is included in the model.

Notice that ACT, when used independently of other variables is significant, but less significant than grade-point average when used alone. Thus, the t-value for grade-point average when used as a single variable is 15.38 with 798 degrees of freedom, while that of ACT is only 5.25 but still significant at the .01 level with the same number of degrees of freedom.

Thus, when grade-point average is not available, the ACT score is useful. According to the expectancy data for the total freshman student body (46), ACT ranks fourth (r = .35) as a predictor of first-quarter grade-point average at Iowa State University, after high school rank as best predictor (r = .44), the Minnesota Scholastic Aptitude Test (MSAT), second (r = .41), and the English Placement Test, third (r = .37).

Language, $X_{12-14}$

Language explained a significant portion of the variation in final grade in elementary foreign language courses taught during the Spring Quarter 1970; in other words, final grade varied according to language. On the basis of the corresponding t-values, it could be said that grades tended to be highest in Spanish and lowest in German. Whether this result is attributable to differences in language difficulty, to the quality of students electing certain languages, to differences in instructors, or to other factors, needs further investigation.

College, $X_{8-11}$

The college variable did not explain a significant portion of the variation in achievement in elementary foreign language courses.

The great majority of students, more than 80 percent, were enrolled in the College of Sciences and Humanities. The number of students from the remaining colleges was too small to warrant conclusions as to the influence, if any, of the college variable upon achievement in elementary foreign language courses.

Course sequence, $X_{15}$

During the Spring Quarter 1970 only the second and third courses of the elementary foreign language sequences were taught. Since, according to previous research which emphasizes the cumulative aspect of language learning (78), the achievement level has a tendency to decline with each subsequent course, the failure of course sequence to explain the variation in final grade could be attributed to the possibility that the weaker students had been gradually eliminated. Further investigation of course sequence, by taking into account the reasons for attrition, is indicated.

## EVALUATION BASED ON A CUMULATIVE POINT SYSTEM VERSUS EVALUATION BASED ON THE RESULTS OF THE FINAL EXAMINATION

### Introduction

This part of the investigation constitutes a transition from the lock-step[1] evaluation methods typical for the traditional elementary language courses as taught during the Spring Quarter 1970 to evaluation of achievement based on self-pacing[2] as discussed in connection with Group III.

The approach to Group II is based on the assumption that traditional grading policies with their cumulative effect can be detrimental to the satisfactory progress in elementary foreign language courses to certain kinds of students. To this category belong:

(1) the slower learner;

(2) the student who is merely aiming at a passing grade; and

(3) the student with a heavy study load.

The motivated but slow learner can be penalized by a cumulative evaluation process for obvious reasons. If his final achievement is satisfactory, it should not be weighted down by a low or unsatisfactory mid-term grade, or by other low grades received during the academic term. Knowing that he has an opportunity to catch up removes some of the pressure and frustration sometimes connected with foreign language study. This is specifically true if the subject is entirely new to the

[1] Uniform test schedule.

[2] Within the academic quarter system, students were allowed to proceed at their own pace, taking the tests when they were ready.

student who first must learn how to approach it. In other words, the psychological effect of initial failure can be reduced by giving the student enough time to familiarize himself with the new situation. Creamer (27) investigated "mid-semester deficiency grading practices" and arrived at the conclusion that mid-semester deficiency reports "discouraged the marginal student from doing his best and deprived him of motivation to achieve a higher grade." Also Borland (15), in an investigation of the effects of midterm grades on the academic performance of college freshmen in Ohio, stated that further study is needed "into the factors relevant to grading systems such as the manner in which these grades are used by students, instructors, and student personnel workers."

The second category of students whose achievement might be affected by cumulative grading policies are those who merely work for a passing grade which in the pass-fail system at most universities is a "D." This bookkeeper-type of student adjusts his effort to the absolute minimum required. He works perhaps during the first weeks of the academic term when things are relatively easy and begins to relax as soon as the passing minimum has been secured. In subsequent elementary courses, he becomes a burden and a negative influence upon the rest of his class because he is not able to perform at a higher level. For this kind of student, evaluation on the basis of the final examination can certainly be a motivating factor.

The student with a heavy study load is penalized by lock-step evaluation because he does not always have the time to prepare himself for current quizzes due to conflict with other requirements. This kind of

student should also have the opportunity to receive a good final grade, even if he cannot always be up-to-date during the course. As long as the university system encourages high study loads, the student should be helped to cope with them.

## Description of Population

Group II, involving the total student population (N = 264) enrolled in 13 sections of German 131 during the Fall Quarter 1970, was used for two purposes:

1. to collect data for the investigation of Group III;
2. to test two different methods of evaluation on a sub-sample of 72 students enrolled in 4 sections of German 131 which were taught by the same instructor.

Table 26 presents the distributions of students in Group II where the sub-sample of 173 students represents the group which subsequently took German 132, and which was used for comparative purposes for the statistical analysis of Group III.

German 131, the first elementary course, was taught by 5 different instructors. More than 80% of the students taking the course during the Fall Quarter 1970 were enrolled in the College of Sciences and Humanities; approximately 70% were sophomores and juniors. The sex ratio of 70% males and 30% females was typical of Iowa State University.

## Method of Procedure

### Experimental design

Group II, consisting of 72 subjects enrolled in 4 sections of German 131 during the Fall Quarter 1970, taught by the same instructor, was

Table 26. Group II. Distribution of elementary German students by sex, college, and year in college, Fall 1970

| | N = 264 | | N = 173 | | N = 72 | |
|---|---|---|---|---|---|---|
| | Number | Percent | Number | Percent | Number | Percent |
| Sex: | | | | | | |
| Male | 187 | 71.0 | 120 | 69.0 | 51 | 71.0 |
| Female | 77 | 29.0 | 53 | 31.0 | 21 | 29.0 |
| College: | | | | | | |
| Sciences and Humanities | 214 | 81.1 | 145 | 84.0 | 62 | 86.0 |
| Engineering | 22 | 8.3 | 14 | 8.0 | 1 | 1.4 |
| Agriculture | 7 | 2.7 | 4 | 2.0 | 5 | 7.0 |
| Education | 3 | 1.1 | 1 | 0.5 | 1 | 1.4 |
| Home Economics | 14 | 5.3 | 8 | 5.0 | 2 | 2.8 |
| Agricultural Engineering[a] | 4 | 1.5 | 1 | 0.5 | 1 | 1.4 |
| Year in College: | | | | | | |
| Freshman | 47 | 18.0 | 29 | 17.0 | 11 | 15.0 |
| Sophomore | 94 | 36.0 | 66 | 38.0 | 24 | 33.0 |
| Junior | 88 | 33.0 | 52 | 30.0 | 27 | 38.0 |
| Senior | 28 | 10.0 | 20 | 12.0 | 7 | 10.0 |
| Graduate | 7 | 3.0 | 6 | 3.0 | 3 | 4.0 |

[a]Belongs to both the College of Agriculture and the College of Engineering.

subjected to the following treatments:

(1) The two morning classes, called the A.M. group, consisting of 42 subjects, were informed that their grades would be based upon achievement on a series of unit tests plus the grade received on the final examination.

(2) The two afternoon classes, called the P.M. group, consisting of 29 subjects (9 students had to be eliminated because of two or more years

of high school German), were informed that any tests or quizzes administered during the academic quarter were only designed to teach, not to evaluate. The final grade would be based on the score obtained on the final examination. In accordance with university policy, this group of students had to be informed at midterm about D's or F's received which, however, were said to be totally inconclusive as far as final grade was concerned.

Both groups were instructed how to study a foreign language effectively by a discussion of Moulton's (65) "Study Hints for Language Students." All four experimental sections were administered the Pimsleur Language Aptitude Battery (76). The total score on this diagnostic test is supposed to measure academic achievement, motivation, verbal and auditory ability.

Variables

The following criteria of achievement were included in the statistical analysis:

$X_2$ = sum of unit quiz scores

$X_3$ = score on the objective part of the final examination

$X_4$ = total final test score, including both the objective and the written part of the final examination

$X_5$ = course grade based on $X_4$, the total final test score

$X_6$ = total sum of scores, including unit and final test scores

$X_7$ = course grade based on $X_6$, total sum of scores.

The following independent variables, assumed to affect achievement, were selected:

$X_1$ = method of evaluation: A.M. versus P.M.

$X_8$ = cumulative grade-point average as at end of Fall Quarter 1970

$X_9$ = course load during the Fall Quarter 1970

$X_{10}$ = language aptitude as measured by the total score on the Pimsleur Language Aptitude Battery.

Hypotheses tested

Objectives were stated in hypothesis form,and each hypothesis was discussed in turn.

Hypothesis 5:

There is no difference in the total unit quiz scores, $X_2$, between the A.M. and the P.M. groups after adjusting for grade-point average, $X_8$, course load, $X_9$, and language aptitude, $X_{10}$.

Hypothesis 6:

There is no difference in the objective final test score, $X_3$, between the A.M. and the P.M. groups after adjusting for the independent variables $X_8$, $X_9$, and $X_{10}$.

Hypothesis 7:

There is no difference in total final test score, $X_4$, between the A.M. and the P.M. groups after adjusting for the independent variables $X_8$, $X_9$, and $X_{10}$.

Hypothesis 8:

There is no difference in the total sum of scores, $X_6$, based on the sum of $X_2$ and $X_4$, between the A.M. and the P.M. groups after adjusting for the independent variables $X_8$, $X_9$, and $X_{10}$.

Hypothesis 9:

There is no difference in the distribution of grades based on $X_4$, final test score, and that based on $X_6$, total sum of scores, or, there is no difference between the grade distributions of $X_5$ and $X_7$.

Statistical methods employed

A preliminary analysis using tables of grade distributions preceded the analysis of multiple regression which tested hypotheses 5 through 9.

Multiple regression models involving $X_1$, method of evaluation, $X_8$, grade-point average, $X_9$, course load, and $X_{10}$, language aptitude were fit for each of the dependent variables; that is, the unknown parameters $\beta_0$, $\beta_1$, $\beta_8$, $\beta_9$ and $\beta_{10}$ were estimated. The question whether an independent variable contributed significantly to explaining the variation in the dependent variable was investigated by a t-test. This is equivalent to fitting a reduced model, one which omits the independent variable under consideration, and investigating the reduction in regression sum of squares to see if it is significant.

Null hypothesis 9 that the distribution of grades based on $X_4$, final test score, is independent of the distribution of grades based on $X_6$, total sum of scores, was investigated by three chi-square tests of independence. The first had the alternative hypothesis of dependence; the second had the alternative hypothesis of dependence due to agreement, and the third had the alternative hypothesis of dependence not due to agreement. In the first test, expected values are calculated by multiplying a row total by a column total and dividing by the overall total. If the calcula-

ted chi-square exceeds the tabular chi-square, the null hypothesis is rejected. The second chi-square test involves a comparison of the observed and expected numbers for the diagonal cells, where the two methods agree, with the off-diagonal cells where the methods disagree on their assignment of grades. When the methods agree on their assignment of grades, most of the observations will lie in the diagonal cells. When the methods disagree on their assignment of grades, most of the observations will lie in the off-diagonal cells. These observations are compared to the same grouping of expected values based on the null hypothesis that the two distributions are independent. If significant, the distributions are assumed not to be independent because they are in agreement, a special case of dependency. The probability of finding a significant chi-square value by chance, using this method, is half the probability listed in the chi-square table, because if the deviations between observations and expectations are summed for the diagonal cells, and also for the off-diagonal cells, differences in direction of the deviations will cancel each other. The third chi-square which has the alternative hypothesis that the two distributions are dependent because they do not agree in their assignment of grades is calculated by subtracting the second chi-square from the first.

## Findings

A summary of raw score and grade distributions for Group II, by section, on the objective final test is presented in Table 27.

Table 27. Group II. Raw score and grade distribution by section on objective final test

| Section number | N | Sum of raw scores | Mean score | Distribution of grades on objective final test | | | | | Grade points received on objective final test | | | | | Total grade points | Mean grade point |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | A | B | C | D | F | A | B | C | D | F | | |
| 1 | 19 | 1358 | 71.47 | 2 | 6 | 7 | 2 | 2 | 8 | 18 | 14 | 2 | 0 | 42 | 2.21 |
| 3 | 24 | 1651 | 68.79 | 2 | 5 | 9 | 4 | 4 | 8 | 15 | 18 | 4 | 0 | 45 | 1.88 |
| Subtotal | 43 | 3009 | 69.98 | 4 | 11 | 16 | 6 | 6 | 16 | 33 | 32 | 6 | 0 | 87 | 2.02 |
| | | | Percent: | 9% | 26% | 37% | 14% | 14% | | | | | | | |
| 12 | 15 | 1048 | 69.86 | 2 | 3 | 6 | 2 | 2 | 8 | 9 | 12 | 2 | 0 | 31 | 2.07 |
| 13 | 14 | 980 | 70.00 | 1 | 4 | 5 | 2 | 2 | 4 | 12 | 10 | 2 | 0 | 28 | 2.00 |
| Subtotal | 29 | 2028 | 69.93 | 3 | 7 | 11 | 4 | 4 | 12 | 21 | 22 | 4 | 0 | 59 | 2.03 |
| | | | Percent: | 10% | 24% | 38% | 14% | 14% | | | | | | | |
| Total | 72 | 5037 | 69.96 | 7 | 18 | 27 | 10 | 10 | 28 | 54 | 54 | 10 | 0 | 146 | 2.02 |

Analysis of multiple regression

Null hypotheses 5 through 9 of no difference in achievement between the A.M. and the P.M. groups were tested by using four different linear regression models, each including a measure of achievement in German 131 as a dependent variable, and the following independent variables, chosen on the basis of previous research because they had been shown to influence achievement in foreign languages:

$X_1$ = method of evaluation

$X_8$ = grade-point average

$X_9$ = course load

$X_{10}$ = language aptitude.

A summary of mean scores and standard deviations for the variables used with Group II is presented in Table 28. The number of observations was N = 72.

The data in Table 28 indicate that the average sum of unit quiz scores was 248 points, equivalent to a D-grade; the average total sum of scores was 477 points, also a "D"; the mean total final test score was 232, equivalent to a C+.

The average score on the Pimsleur Language Aptitude Battery (PLAB) was 92 which is somewhat above the average for a sample of 529 Iowa State University German and Spanish students (mean = 90) who had taken this test during the Fall Quarter 1970 and during the Winter Quarter 1971.

Table 28. Group II. Analysis of regression - summary of means and standard deviations; N = 72

| | Variables | Mean | Standard deviation |
|---|---|---|---|
| $X_2$ = | sum of quiz scores | 248.46 | 83.72 |
| $X_3$ = | objective score on final examination | 69.89 | 11.30 |
| $X_4$ = | total final test score (objective and written) | 232.29 | 53.15 |
| $X_5$ = | course grade based on $X_4$ | 2.42 | 1.35 |
| $X_6$ = | total sum of scores | 477.31 | 127.31 |
| $X_7$ = | course grade based on $X_6$ | 1.57 | 1.31 |
| $X_8$ = | cumulative grade-point average | 2.76 | 0.58 |
| $X_9$ = | course load - Fall Quarter 1970 | 15.57 | 2.76 |
| $X_{10}$ = | language aptitude score on PLAB | 91.85 | 10.45 |
| $X_1$ = | method of evaluation was coded as follows: | | |
| | A.M. group = 1; P.M. group = 2 | | |

The distributions of evaluation measures used in German 131 are recorded in Table 29.

A summary of F-ratios testing the significance of regression, multiple correlation coefficients, and standard errors is presented in Table 30.

As evident from Table 30, the ratios of the regression mean squares and the residual mean squares in the six models involving $X_1$, method of evaluation, $X_8$, cumulative grade-point average, $X_9$, course load, and $X_{10}$, language aptitude as independent variables, are highly significant at the .01 level with 4 and 67 degrees of freedom.

Table 29. Group II. Distribution of evaluation measures in German 131

| $X_2$ Sum of unit quiz scores | $X_3$ Objective score | $X_5$ Total final course grade based on $X_4$, total final test score | $X_7$ Course grade based on $X_6$, total sum of scores |
|---|---|---|---|
| T[a] = 420 | T = 95 | T = 310 | T = 730 |
| A = 370 | A = 85 | A = 270 | A = 650 |
| B = 330 | B = 76 | B = 240 | B = 580 |
| C = 290 | C = 67 | C = 210 | C = 510 |
| D = 250 | D = 58 | D = 180 | D = 430 |

[a]T = total score possible.

Table 30. Group II. Summary of F-ratios, multiple $R^2$, and standard errors; N = 72

| Model[a] | Dependent variables | F-ratio 4.67 d.f. | Multiple $R^2$ | Standard error |
|---|---|---|---|---|
| 1-II | $X_2$ = sum of unit quiz scores | 22.67** | .58 | 56.18 |
| 2-II | $X_3$ = objective score on final test | 16.41** | .49 | 8.26 |
| 3-II | $X_4$ = total final test score | 12.25** | .42 | 41.58 |
| 4-II | $X_5$ = course grade based on $X_4$ | 15.44** | .48 | 1.00 |
| 5-II | $X_6$ = total sum of scores | 25.46** | .60 | 82.56 |
| 6-II | $X_7$ = course grade based on $X_6$ | 21.01** | .56 | .09 |

[a]See Appendix. In each of the 6 models, the following independent variables were used; $X_1$, method of evaluation; $X_8$, grade-point average; $X_9$, course load, $X_{10}$, language aptitude.

**Exceeds tabular F (.01) with 4 and 67 degrees of freedom = 3.62.

Coefficients of correlation were computed between the variables used in the regression analysis of Group II as indicated in Table 31.

Inspection of the correlation matrix indicates that method of evaluation $X_1$, is correlated negatively with all the variables involved in the analysis of Group II. This means that the A.M. group had higher values for the independent variables than the P.M. group. Except for the correlation with $X_2$, sum of unit quiz scores, and $X_6$, total sum of scores, these correlations are not significantly different from zero with sample size N = 72.

There is a substantial correlation between $X_2$, sum of unit quiz scores and both $X_6$, total sum of scores, and $X_7$, course grade based on $X_6$ ( $r_{26} = .94$; $r_{27} = .89$).

All measures of achievement are significantly correlated with grade-point average, $X_8$.

Language aptitude, $X_{10}$, has the highest correlation with $X_3$, the objective score on the final test ( $r_{3,\ 10} = .429$).

The lowest correlations occurred for course load.

A summary of the results of analyses of multiple regression is presented in Table 32. For each independent variable the null hypothesis was tested that it does not contribute significantly to explaining the variation in the independent variable when the other independent variables are included in the model. Tables containing individual analyses for each model are included in the Appendix.

As evident from Table 32, $X_1$, method of evaluation, explains a significant portion of the variability in the dependent variables, $X_2$, sum of unit quiz scores and in $X_6$, total sum of scores, but not in $X_7$,

Table 31. Group II. Product moment correlation matrix - German 131; N = 72

| | $X_1$[a] | $X_2$[b] | $X_3$[c] | $X_4$[d] | $X_5$[e] | $X_6$[f] | $X_7$[g] | $X_8$[h] | $X_9$[i] | $X_{10}$[j] |
|---|---|---|---|---|---|---|---|---|---|---|
| $X_1$ | 1.000 | | | | | | | | | |
| $X_2$ | -.350 | 1.000 | | | | | | | | |
| $X_3$ | -.007 | .662 | 1.000 | | | | | | | |
| $X_4$ | -.045 | .599 | .795 | 1.000 | | | | | | |
| $X_5$ | -.065 | .664 | .815 | .898 | 1.000 | | | | | |
| $X_6$ | -.256 | .944 | .796 | .807 | .844 | 1.000 | | | | |
| $X_7$ | -.186 | .894 | .769 | .741 | .828 | .930 | 1.000 | | | |
| $X_8$ | -.086 | .690 | .656 | .594 | .672 | .734 | .717 | 1.000 | | |
| $X_9$ | -.212 | .279 | .244 | .367 | .306 | .332 | .252 | .228 | 1.000 | |
| $X_{10}$ | -.034 | .325 | .429 | .271 | .276 | .336 | .389 | .361 | -.047 | 1.000 |

[a]Method of evaluation. This is a discrete variable, coded: AM-group = 1; PM-group = 2.

[b]Sum of unit quiz scores.

[c]Objective score on final test.

[d]Total final test score.

[e]Course grade based on $X_4$.

[f]Total sum of scores.

[g]Course grade based on $X_6$.

[h]Cumulative grade-point average.

[i]Course load.

[j]Language aptitude score (PLAB).

Table 32. Group II. Summary of calculated t-values for German 131

| Model | Dependent variables | Independent variables $X_1$ Method of evaulation | $X_8$ GPA | $X_9$ Course load | $X_{10}$ Language aptitude |
|---|---|---|---|---|---|
| 1-II | $X_2$ - Total unit quiz score | -3.39** | 6.91** | 1.01 | 1.15 |
| 2-II | $X_3$ - Objective final test score | .89 | 5.59** | 1.63 | 2.60* |
| 3-II | $X_4$ - Total final test score | .61 | 4.84** | 2.75* | 1.05 |
| 4-II | $X_5$ - Course grade based on $X_4$ | .29 | 6.23** | 1.88 | .68 |
| 5-II | $X_6$ - Total sum of scores | -2.08* | 7.56** | 1.89 | 1.25 |
| 6-II | $X_7$ - Course grade based on $X_6$ | -1.28 | 6.93** | 1.10 | 1.86 |

*Exceeds tabular t (.05) = 1.99 in absolute value.

**Exceeds tabular t (.01) = 2.65 in absolute value.

course grade based on $X_6$.

$X_8$, cumulative grade-point average as of Fall 1970, explains a significant portion of the variability in all the dependent variables involved in the analysis of Group II.

$X_9$, course load, after adjusting for method of evaluation, $X_1$, grade-point average, $X_8$, and $X_{10}$, language aptitude, explains a significant portion of the variability in the dependent variable $X_4$, total

final test score. Course load does not explain a significant portion of the variability in any of the other dependent variables.

$X_{10}$, language aptitude, explains a significant portion of the variability of $X_3$, objective final test score, after adjusting for the other independent variables involved. In connection with the other dependent variables, it is insignificant.

Each of these variables is now discussed in turn.

Method of evaluation, $X_1$

After adjusting for grade-point average, $X_8$, course load, $X_9$, and language aptitude, $X_{10}$, method of evaluation accounted for differences in achievement between the two treatment groups in total unit quiz score, $X_2$, and in the total sum of scores, $X_6$, as shown in Table 33. Method of

Table 33. Group II. Grade distribution by method of evaluation, German 131

| | Method of evaluation: | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | $X_2$ - Sum of unit quiz scores | | | | $X_6$ - Total sum of scores | | | |
| | A.M. | | P.M. | | | A.M. | | P.M. | |
| Grade | N | Percent | N | Percent | Grade | N | Percent | N | Percent |
| A | 4 | 9 | 2 | 7.1 | A | 4 | 9 | 2 | 7 |
| B | 7 | 16 | 3 | 10.3 | B | 9 | 21 | 4 | 14 |
| C | 11 | 26 | 3 | 10.3 | C | 12 | 28 | 6 | 21 |
| D | 9 | 21 | 5 | 17.2 | D | 8 | 19 | 5 | 17 |
| F | 12 | 28 | 16 | 55.1 | F | 10 | 23 | 12 | 41 |
| Total | 43 | 100% | 29 | 100% | | 43 | 100% | 29 | 100% |

evaluation did not explain a significant portion of the variability in course grade based on $X_6$, total sum of scores.

On the basis of the corresponding t-values (see Table 32) it can be said that the A.M. group scored significantly higher on $X_2$, total unit quiz score, and on $X_6$, total sum of scores in German 131.

As evident from Table 32, there is sufficient evidence to reject the null hypotheses 5 and 8 that there is no significant difference in total unit quiz score and in the total sum of scores, $X_2$ and $X_6$, between the A.M. and the P.M. groups.

After adjusting for other variables, method of evaluation, $X_1$, did not affect significantly the variation in the scores on the objective part of the final examination nor the total score on the final examination. Table 34 shows the grade distributions based on the objective score obtained on the final test as compared to the total final test score.

Table 34. Group II. Grade distributions: final examination in German 131

| | Method of evaluation: | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | $X_3$ - Objective score on final test | | | | $X_4$ - total final test score | | | |
| | A.M. | | P.M. | | | A.M. | | P.M. | |
| Grade | N | Percent | N | Percent | Grade | N | Percent | N | Percent |
| A | 7 | 16 | 3 | 10 | | 12 | 28 | 7 | 24 |
| B | 9 | 21 | 7 | 24 | | 12 | 28 | 9 | 31 |
| C | 15 | 35 | 11 | 38 | | 10 | 23 | 4 | 14 |
| D | 6 | 14 | 4 | 14 | | 4 | 9 | 4 | 14 |
| F | 6 | 14 | 4 | 14 | | 5 | 12 | 5 | 17 |
| Total | 43 | 100% | 29 | 100% | | 43 | 100% | 29 | 100% |

Method of evaluation also influenced final achievement as measured by course grade. This is illustrated by Table 35 which shows that the P.M. students whose grades were based on the final test score, $X_4$, on the average received better grades than the A.M. students whose grades were based on the total sum of scores accumulated during the academic term.

Table 35. Group II. Distribution of course grades for the two treatment groups

| A.M. $X_7$ - Course grade based on total sum of scores | | | P.M. $X_5$ - Course grade based on total final test score | | |
|---|---|---|---|---|---|
| Grade | Number | Percent | Grade | Number | Percent |
| A | 4 | 9 | A | 7 | 24 |
| B | 9 | 21 | B | 9 | 31 |
| C | 12 | 28 | C | 4 | 14 |
| D | 8 | 19 | D | 4 | 14 |
| F | 10 | 23 | F | 5 | 17 |
| Total | 43 | 100% | | 29 | 100% |

Hypothesis 9 was further investigated by three chi-square tests of independence. The first chi-square tested the alternative hypothesis of dependence; the second, the alternative hypothesis of dependence due to agreement; and the third, the alternative hypothesis of dependence due to disagreement.

Chi-square test of independence:

1. Null hypothesis 9 is that the distribution of grades based on $X_4$, total final test score, is independent of the distribution of grades based on $X_6$, total sum of scores in German 131. The alternative hypothesis is that these two distributions are dependent. Expected values were calculated by multiplying a row total by a column total and dividing by the overall total of the distribution in Table 36.

Table 36. Group II. Observed number of students; contingency table for variables $X_4$ and $X_6$ ; N = 72

| | | $X_6$ - total sum of scores | | | | | |
|---|---|---|---|---|---|---|---|
| | | A | B | C | D | F | Total |
| $X_4$ - total final test score | A | 6 | 9 | 3 | | 1 | 19 |
| | B | | 4 | 13 | 4 | | 21 |
| | C | | | 2 | 6 | 6 | 14 |
| | D | | | | 3 | 4 | 7 |
| | F | | | | 1 | 10 | 11 |
| Total | | 6 | 13 | 18 | 14 | 21 | 72 |

The result is Table 37.

Table 37. Group II. Expected number of students; variables $X_4$ and $X_6$

| | | $X_6$ - total sum of scores | | | | | |
|---|---|---|---|---|---|---|---|
| | | A | B | C | D | F | Total |
| $X_4$, total final test score | A | 1.58 | 3.43 | 4.75 | 3.69 | 5.55 | 19 |
| | B | 1.75 | 3.79 | 5.25 | 4.08 | 6.13 | 21 |
| | C | 1.16 | 2.52 | 3.54 | 2.72 | 4.06 | 14 |
| | D | .58 | 1.26 | 1.75 | 1.36 | 2.05 | 7 |
| | F | .93 | 2.00 | 2.71 | 2.15 | 3.21 | 11 |
| Total | | 6.00 | 13.00 | 18.00 | 14.00 | 21.00 | 72 |

Since many of the expected values were smaller than 5, the classes were grouped so that most of the expected values were 5 or larger, which resulted in the distribution as recorded in Table 38.

Table 38. Group II. Observed and expected number of students; variables $X_4$ and $X_6$; N = 72

| | | Number of students | $X_6$ - total sum of scores | | | |
|---|---|---|---|---|---|---|
| | | | A + B | C | D + F | Total |
| $X_4$, total final test score | A + B | Expected: | 10.55 | 10.00 | 19.45 | 40 |
| | | Observed: | 19 | 16 | 5 | 40 |
| | C | Expected: | 3.68 | 3.54 | 6.78 | 14 |
| | | Observed: | 0 | 2 | 12 | 14 |
| | D + F | Expected: | 4.77 | 4.46 | 8.77 | 18 |
| | | Observed: | 0 | 0 | 18 | 18 |
| | Total | Expected: | 19 | 18 | 35 | 72 |
| | | Observed: | 19 | 18 | 35 | 72 |

The calculated chi-square to test the hypothesis of independence versus the alternative hypothesis of dependence is

$$\Sigma (O - E)^2 / E = 48.40$$

This chi-square has 4 degrees of freedom and exceeds the tabular chi-square with 4 degrees of freedom at the .05 level of significance, namely, 9.488. Therefore, the null hypothesis of independence is rejected. There is evidence that the two distributions of grades are dependent; i.e., there is some relationship between the grades based on $X_6$, total sum of scores, and those based on $X_4$, total score on the final examination.

2. Chi-square test of independence based on the alternative hypothesis that the distribution of grades based on $X_4$ is dependent of the distribution of grades based on $X_6$ due to agreement in their assignment of grades.

The data for this test as presented in Table 39 were taken from Table 38.

Table 39. Group II. Observed and expected number of students; chi-square test

| | Number of students | | | | |
|---|---|---|---|---|---|
| | O<br>Observed | E<br>Expected | (O - E) | $(O - E)^2$ | $\frac{(O - E)^2}{E}$ |
| Diagonal cells | 39 | 22.86 | $(16.14)^2$ | 260.49 | 11.39 |
| Off-diagonal cells | 33 | 49.14 | $(-16.14)^2$ | 260.49 | 5.30 |
| | | | | $\chi^2$ | = 16.69 |

The calculated chi-square 16.69 with 1 degree of freedom exceeds the tabular chi-square with 1 degree of freedom at the .05 level of significance, 2.706. There is evidence to reject the null hypothesis of independence because the two methods of evaluation are in agreement.

3. Chi-square test of independence based on the alternative hypothesis that the two distributions are dependent because they do not agree in their assignment of grades.

This chi-square is calculated by subtracting the second chi-square from the first:

| | Calculated chi-square | d.f. | Tabular chi-square |
|---|---|---|---|
| Chi-square test of independence No. 1 | 48.40** | 4 | 9.488 |
| Chi-square test of independence No. 2 | 16.69* | 1 | 2.706 |
| Chi-square test of independence No. 3 | 31.71* | 3 | 7.815 |

The calculated chi-square 31.71 with 3 degrees of freedom exceeds the tabular chi-square with 3 degrees of freedom at the .05 level of significance, 7.815. There is evidence to reject the null hypothesis of independence because the two methods of evaluation disagree in their assignment of grades. In other words, these two grading methods are not independent. They agree in the sense that grades assigned using $X_4$, total final test score, are almost never lower than grades assigned using $X_6$, total sum of scores. This is shown by the empty lower left off-diagonal cells in Table 36. In addition, they disagree in the sense that grades assigned using $X_6$, total sum of scores, are lower than grades

assigned using $X_4$, total final test score. This is shown by the occupied upper right off-diagonal cells in Table 36.

Grade-point average, $X_8$

As evident from Table 32, grade-point average, $X_8$, is highly correlated with all measures of achievement employed in connection with Group II and explains a significant portion of their variation. Its highest t-values occur in conjunction with $X_6$, total sum of scores ($t = 7.56^{**}$), $X_7$, course grade based on the total sum of scores ($t = 6.93^{**}$), and with $X_2$, total sum of unit quiz scores ($t = 6.91^{**}$). Relatively lower values are associated with $X_4$, final test score ($t = 4.84^{**}$), and $X_3$, score on the objective part of the final examination ($t = 5.59^{**}$).

Course load, $X_9$

For Group II, as evident from Table 32, the t-value for course load is significant at the .05 level only in conjunction with total final test score, $X_4$, as a dependent variable. The correlation between these two variables ($r_{49} = .37$) is significant at the .01 level with sample size N = 72. The t-values connected with course load are smallest for total unit quiz score, $X_2$, and for the course grade based on the total sum of scores in German 131, $X_6$. An explanation for this could be that students with higher course loads, while not studying regularly for the unit quizzes which constituted 58 percent of the total sum of scores, made a concentrated effort to do well on the final examination.

Language aptitude, $X_{10}$

As evident from Table 32, language aptitude, $X_{10}$, did not explain a significant portion of the variation in total unit quiz score, $X_2$, in the total score on the final examination, $X_4$, in course grade based on

$X_4$, in the total sum of scores, $X_6$, and in course grade based on $X_6$. Language aptitude explained a significant portion of the variation in the dependent variable, if this variable was the score on the objective part of the final examination, $X_3$. This variable and language aptitude are significantly correlated ($r_{3,10} = .43$ when $N = 72$).

## Discussion

### Method of evaluation, $X_1$

The students belonging to Group II were evaluated according to their membership in the forenoon (A.M.) or afternoon (P.M.) sections of German 131 during the Fall Quarter 1970. The A.M. treatment involved evaluation on the basis of a cumulative point system, the P.M. treatment based the course grade upon the results of the final examination for reasons discussed in the Introduction.

The following graphical presentations illustrate the achievement of the two groups on the various criteria used in the experimental group of German 131.

Inspection of Figure 3 indicates for the combined group (N = 72) that grades based on the P.M. method of evaluation (mean grade = 2.42 = C+) were higher than those based on the A.M. method (mean grade = 1.57 = C-).

Chi-square tests provided evidence to reject null hypothesis 9 that the two methods of evaluation are independent.

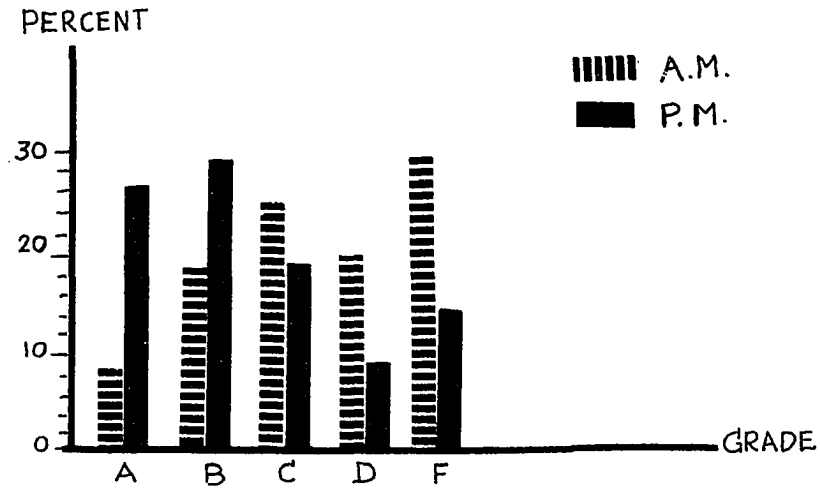



Figure 3. Group II. Comparison of course grade distributions based on two different methods of evaluation

These two methods are in agreement in their assignment of grades and at the same time, they are different in their assignment of grades. This can be explained by looking at Table 40.

Table 40. Group II. Contingency table for variables $X_5$ and $X_7$ - combined group; N = 72

| Method of evaluation | $X_7$ - Course grade based on total sum of scores | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Grade | A | B | C | D | F | Total | Percent |
| $X_5$ - Course grade based on total final test score | A | 6 | 9 | 3 | | 1 | 19 | 26.4 |
| | B | | 4 | 13 | 4 | | 21 | 29.2 |
| | C | | | 2 | 6 | 6 | 14 | 19.4 |
| | D | | | | 3 | 4 | 7 | 9.7 |
| | F | | | | 1 | 10 | 11 | 15.3 |
| | Total | 6 | 13 | 18 | 14 | 21 | 72 | |
| | Percent | 8.3% | 18.1% | 25% | 19.4% | 29.2% | | 100% |

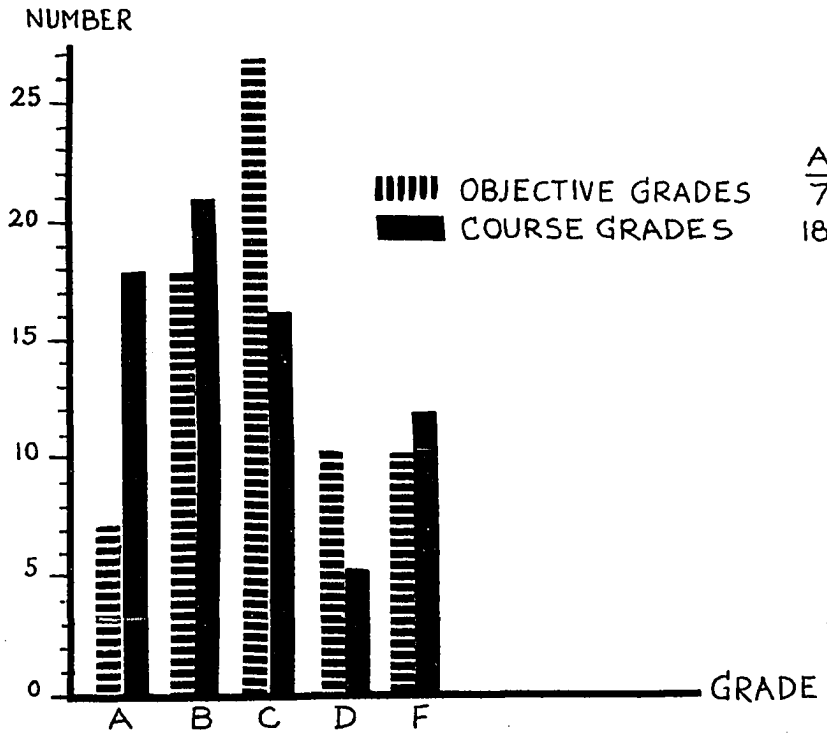
The method $X_7$ assigns more lower grades than method $X_5$. Except for 1 person out of 72 involved, the grade assigned by $X_7$ was not higher than the grade assigned by $X_5$. In other words, when grades are assigned on the basis of total sum of scores they are likely to be lower than those assigned on the basis of final test score. This can be explained by the fact that the grade based on evaluation method $X_7$ includes both the total sum of unit quiz scores and the total score on the final examination. Thus, to get a good grade by method $X_7$ requires sustained motivation throughout the academic term. A student evaluated by method $X_7$ is penalized for low scores on unit quizzes even if he should achieve a perfect score on the final examination. On the other hand, method of evaluation, $X_5$, does not take into account failure during the quarter and allows for a good grade if the student is successful on the final examination.

The two methods are in agreement in that method $X_7$ will not assign good grades to students where method $X_5$ does not. Looking at these grade distributions, there is only 1 case out of 72 where method $X_5$ assigned an "F" and method $X_7$ assigned a "D."

Figure 4 compares, for the combined group (N = 72), the distribution of grades on the objective part of the final examination and of course grades.

Inspection of Figure 4 indicates that the objective part of the final examination is a more rigorous measure of achievement than course grade as assigned by both methods of evaluation, i.e., either on the basis of points accumulated during the academic term (A.M.) or on the basis of the complete final test which consisted of an objective and of

a written part (P.M.).




| | A | B | C | D | F | TOTAL |
|---|---|---|---|---|---|---|
| OBJECTIVE GRADES | 7 | 18 | 27 | 10 | 10 | 72 |
| COURSE GRADES | 18 | 21 | 16 | 5 | 12 | 72 |

Figure 4. Group II. Course grade versus grade on objective final test

Inspection of Figure 5 indicates that the grade distributions on the objective part of the final examination more closely approximate the normal curve than the distribution of grades on the complete final test. However, the differences in achievement between the A.M. and the P.M. groups are minimal, both having a mean raw score of 70 points on the objective part of the final examination.

Figure 6 presents the contrast in performance between the two treatment groups on the unit quizzes and in total sum of scores. It is evident that the A.M. group is superior on both measures of achievement.

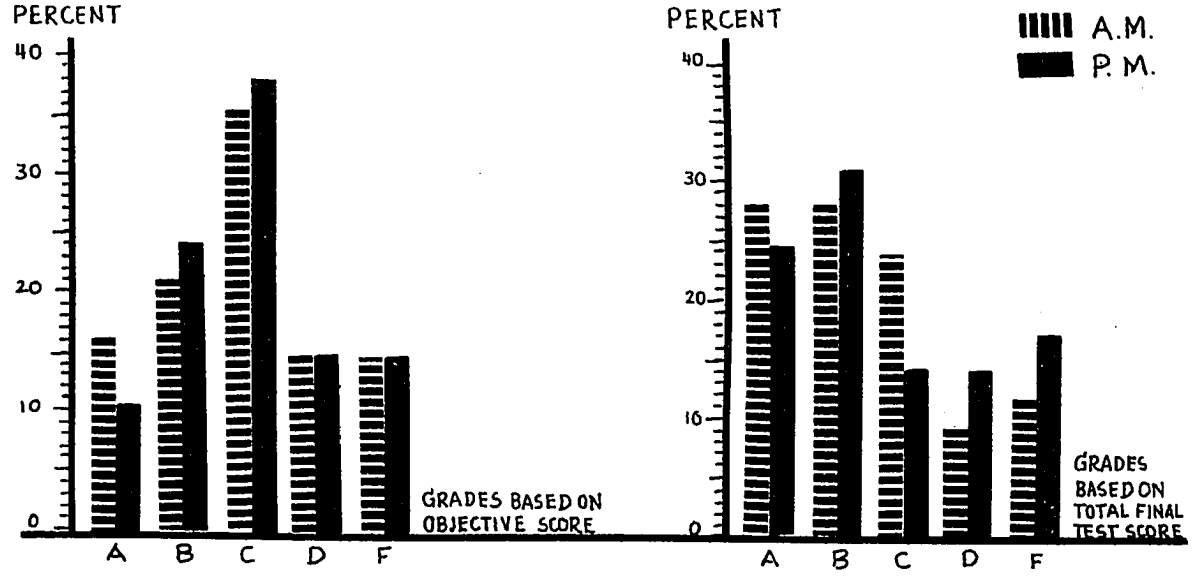



Figure 5. Group II. Grade distributions of the treatment groups on the final examination in German 131

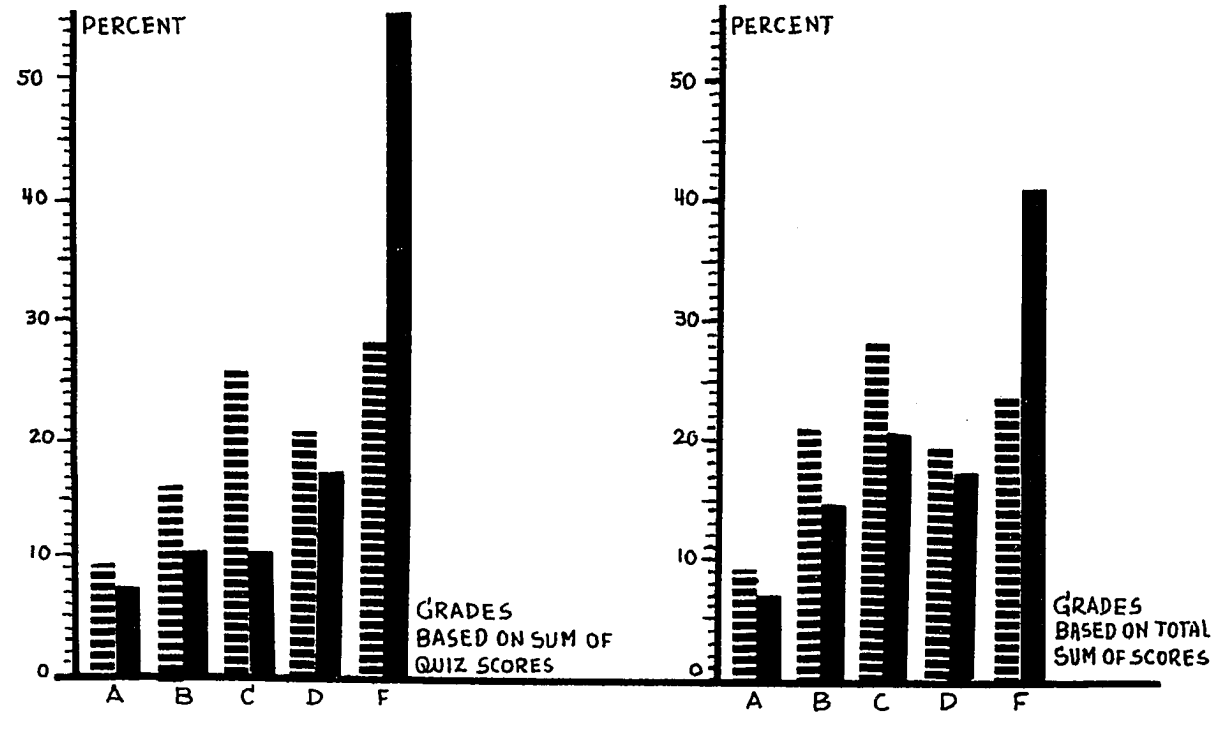



Figure 6. Group II. Distribution of grades on the unit quizzes and on the total sum of scores for the two treatment groups

These results, consistent with what was expected on the basis of theoretical considerations, were confirmed by analysis of multiple regression. Thus, as evident from Table 32, method of evaluation, after adjusting for grade-point average, course load, and language aptitude, explained a significant portion of the variation in total unit quiz score and in the total sum of scores when testing at the .05 level of significance. This evidence justified the rejection of null hypotheses 5 and 8 of no difference among the two treatment groups on these criteria. The negative t-values indicate that the A.M. group (coded: 1) had higher scores than the P.M. group.

These findings confirmed the assumption that the A.M. group who were told that their final grade would be based on the total sum of scores accumulated during the academic quarter were more motivated to perform well on the unit quizzes which constituted 58 percent of the final grade for this group than the P.M. group who knew that their final grade would be the result of their performance on the final examination.

While method of evaluation explained a significant portion of the variation in the total sum of scores, it did not explain the variation in the course grade based on the total sum of scores. This could be explained by the fact that grades based on a 70-point interval scale do not reflect achievement as accurately as raw scores. Thus, for instance, a student with a total score of 649 on this criterion would receive the same grade as a student with a score of 580, namely, a "B."

That method of evaluation, after adjusting for other variables, would not explain a significant portion of the variation in the scores on the objective part of the final examination, in total final test score, and

in the grades based on it, is consistent with what was expected from the effect of the two treatments. As already mentioned, Figure 5 indicates that the A.M. and the P.M. groups have almost identical distributions on these measures of achievement due to the fact that they did not differ in motivation to perform well on these criteria.

Considering these results, the question arises which group achieved better in German 131, the A.M. group who made a sustained effort, or the P.M. group who did not perform evenly, as shown by the lower scores on the unit quizzes and their lower total sum of scores.

A look at Figure 5 indicates that the grade distributions of the two treatment groups on the objective part of the final examination are almost identical. Figure 4 indicates that the objective part of the final examination is a more rigorous measure of achievement than any of the other criteria. It requires more aptitude and more ability to think in terms of the foreign language than, for instance, the written part of this examination which requires mainly memorization. It also is easier to perform well on unit tests than on a final examination involving more subject matter.

Since there is no difference in achievement on the objective part of the final examination where both groups obtained a mean score of 70 points, it can be assumed that there is no difference in achievement between the two groups. However, further investigation is indicated to gather evidence which approach to studying a foreign language leads to better ultimate results, sporadic but intensive efforts, or sustained effort.

In favor of the method of evaluation based on the final test is the fact that it does not penalize students who, for some reason or another, cannot achieve evenly during the academic term. This method enables them to continue, instead of dropping out, when other requirements interfere with their even level of performance in foreign language study.

Grade-point average, $X_8$

In Groups II, where six different measures of achievement were used, the highest t-value for grade-point average occurred in connection with the total sum of scores which reflects a combination of motivation and aptitude, two important factors inherent in grade-point average. The sum of scores obtained in German 131 is the result of both sustained effort, as reflected by achievement on the unit tests, and of reasoning ability as measured by the objective part of the final examination. The lower t-value of grade-point average connected with the total score on the final examination reflects, perhaps, a lesser degree of required ability since two-thirds of this score was based on the results of memorization, and only one-third, the objective part, on the ability to reason in terms of the syntactical and grammatical structures of the foreign language.

## COMPARISON OF LOCK-STEP EVALUATION OF STUDENT ACHIEVEMENT WITH EVALUATION BASED ON SELF-PACING

### Introduction

The specific purpose of this part of the study was to investigate the possibility of reducing underachievement and failure by controlling the time factor in elementary foreign language learning through evaluation policies based on the following concepts and procedures:

(1) Self-pacing[1] combined with tutoring of the slower students.

(2) Precise definition of behavioral objectives.

(3) Mastery learning.

(4) Economy of effort.

The theoretical basis for these concepts, which have already been discussed in the general introduction, is the assumption that "all students can learn a foreign language, although some will take much more time ..." (69).

### Description of Population

Group III involved 173 students enrolled in 12 sections of German 132 which were taught by 5 instructors during the Winter Quarter 1971. The whole group had taken German 131, the first elementary course, during the Fall 1970. Eliminated were students who had not been enrolled in German 131 during the Fall Quarter 1970, and those who had not taken the departmental final examinations for either one or both quarters.

The experimental group consisted of 56 students enrolled in 4 sections taught by the same instructor. The control group consisted of

[1] Within the academic quarter system, students were allowed to proceed at their own pace, taking the tests when ready.

117 students enrolled in 8 sections taught by four different instructors.

Table 41 shows the distribution of the two groups by sex, college, and year in college:

Table 41. Group III. Distribution of elementary German students by sex, college, and year in college, Winter 1971

| | Experimental group | | Control group | |
|---|---|---|---|---|
| | Number | Percent | Number | Percent |
| Sex: | | | | |
| Males | 39 | 70 | 81 | 69 |
| Females | 17 | 30 | 36 | 31 |
| College: | | | | |
| Sciences and Humanities | 50 | 89 | 95 | 81 |
| Engineering | 5 | 9 | 9 | 8 |
| Agriculture | 1 | 2 | 3 | 2 |
| Education | - | - | 1 | 1 |
| Home Economics | - | - | 8 | 7 |
| Agricultural Engineering[a] | - | - | 1 | 1 |
| Year in college: | | | | |
| Freshman | 14 | 25 | 15 | 13 |
| Sophomore | 22 | 39 | 44 | 38 |
| Junior | 15 | 27 | 37 | 31 |
| Senior | 4 | 7 | 16 | 14 |
| Graduate students | 1 | 2 | 5 | 4 |

[a]Belongs to both the College of Agriculture and the College of Engineering.

The sex ratio was the same in both the experimental and the control groups, i.e., approximately 70% male and 30% female, corresponding to the sex ratio for the entire student population at Iowa State University during the Fall and Winter Quarters 1970-71:

| Iowa State University: | Fall 1970 | | Winter 1971 | |
|---|---|---|---|---|
| | Number | Percent | Number | Percent |
| Males: | 13,451 | 68.56 | 12,912 | 69.32 |
| Females: | 6,169 | 31.44 | 5,715 | 30.68 |

## Method of Procedure

The analysis of Group III was concerned with the investigation of differences in achievement in German 132 between the experimental and the control groups, due to differences in methods of evaluation.

### Experimental design

The control group of 117 students was taught and evaluated by four different instructors using traditional methods which expect students to proceed at the same pace and to submit to uniform test schedules throughout the academic term.

The experimental group of 56 students was subjected to the following treatment:

(1) At the beginning of the course, these students received a handout describing the behavioral objectives for German 132 as follows:

a. The student understands both printed and spoken German based on the vocabulary and grammar as presented in lessons 8 through 13 of the Schulz-Griesbach-von Hofe text, Deutsche Sprachlehre für Amerikaner (91).

b. The student acquires new vocabulary and learns to speak and to pronounce correctly by memorizing, with accuracy and fluency, selected idomatic expressions and sentences taken verbatim from the text. To facilitate the learning process, these sentences for each chapter

of the text are presented to the students on handouts with an English translation and recorded on tape for language laboratory practice. The method of teaching vocabulary and structure within the context of sentences, is based on the following practical considerations:

Vocabulary learned in context is meaningful and remembered longer.

In the beginning stages of foreign language learning, the memorization of correct model sentences is preferable to incorrect improvisation by students and allows for subsequent variation and application in different situations.

Eliciting responses by means of English equivalents furthers the acquisition of an active vocabulary which is not easily achieved by using German questions and answers alone, as expected in the textbook. English equivalents also prevent misunderstanding or ignorance of meaning.

These basic sentences, constituting a minimum requirement for passing the course, are assumed to be a time-saving element for those students who because of a heavy course load would otherwise not achieve to capacity, or fail. These sentences allow them to concentrate on a precisely defined learning task. This method is not supposed to prevent the instructor and the students from engaging in more sophisticated activities, if so desired.

Elementary foreign language learning is the learning of a skill, and as such cannot be taught efficiently on a trial-and-error basis, i.e., by leaving responses completely to the student. Errors reinforce incorrect habits.

The knowledge of a basic core of sentences facilitates the progress of students in subsequent courses, under different instructors, and makes the conducting of classes in the foreign language possible.

c. The student is able to form and answer both orally and in writing simple questions in German by using the interrogatives: wer, was, wie, welcher, welche, welches, was für ein, wo, wohin, woher, wann, wieviel, wie viele, womit, etc.

d. The student understands the logic of German grammar and is able to make a syntactical analysis using German. For this purpose, he learns commonly used German grammatical terminology and knows how to apply the basic grammatical rules as they appear in the text.

Detailed definitions of behavioral objectives were issued for each chapter covered during the Winter Quarter 1971.

(2) In another handout, the experimental group was informed about the following evaluation policies used in this course:

a. To pass German 132, all students, no matter whether enrolled under the pass-fail system or under the letter grade system, had to master 80 percent of the sentences listed for each chapter; i.e., they had to be able to recite them fluently and with accuracy, analyze them grammatically, and handle them correctly with appropriate changes in tense and parts of speech during previously announced oral quizzes for each chapter on a "pass" or "fail" basis.

b. Those students who failed at the first attempt were given as many opportunities as necessary to meet the chapter requirements.

(3) Language laboratory attendance was not enforced but fluency and correctness of oral recitation were emphasized.

(4) The final examination consisted of three parts:

a. a 25-item listening comprehension test;

b. a 125-item multiple choice test based on the grammar and vocabulary of six textbook units;

c. 20 sentences to be translated from English into German. These sentences had been practiced and analyzed grammatically during the academic term. Each sentence counted 10 points with 5-point deductions for grammatical errors, missing words, wrong word order, and two-point deductions for misspellings.

d. The following passing standards were established: 100 points on the objective part, and 160 points on the written part were necessary to pass and to receive a "B;" 120 points on the objective part, and 160 points on the written part, to receive an "A."

Students were warned that those who did not appear for the final examination would receive an "F" for the course. As a result, nobody was missing.

(5) Those students who did not meet these standards on one or both parts of the final test, or who had not removed the chapter deficiencies, were given an "incomplete" grade with the indication that it could be removed by taking another test during the first week of the Spring Quarter 1971, and by removing the chapter deficiencies during the first two weeks.

(6) When the students reported for the second test, they were informed about three alternatives concerning their grade for German 132:

a. Pass the second examination with a different set of sentences to be translated, and the passing score on the objective part of the final examination set 20 points higher to achieve a "B."

b. Accept the traditional letter grade on the basis of their previous achievement.

c. Keep the "incomplete" grade until ready to remove it within the time limit of one year. After the expiration of one year "incomplete" grades will be changed to an "F."

By the end of the first week of the Spring Quarter 1971, 35 out of 47 incomplete grades had been removed as follows:

24 students took the second examination grade with gains ranging from 0.5 to 2.0 points;

11 students accepted the traditional grade;

13 students kept the incomplete grade.

As evident from Table 42, the self-pacing experiment involved 552 test interviews lasting from a few minutes to almost an hour. To handle this volume, the instructor used an average of 15 office hours per week.

Unsuccessful attempts at the unit tests by individual students ranged from 1 to 13. One single successful attempt at each unit was made only by seven students.

Table 42. Group III. Experiment in self-pacing in elementary German; distribution of oral unit tests

| Section | Number of students | Successful attempts | Unsuccessful attempts | No attempt | Total attempts |
|---|---|---|---|---|---|
| 1 | 23 | 119 | 57 | 7 | 176 |
| 2 | 15 | 64 | 50 | 16 | 114 |
| 9 | 20 | 95 | 50 | 15 | 145 |
| 10 | 14 | 69 | 40 | 6 | 117 |
| | 72[a] | 347 | 205 | 44 | 552 |

[a]This number includes all students enrolled in the experimental group without restrictions as specified for the statistical analysis.

## Statistical analysis

### Hypothesis 10:

That there is no difference in achievement in German 132 between students evaluated on a self-pacing basis and those subjected to a uniform test schedule throughout the academic term was tested using several approaches.

First, ignoring other variables, a preliminary analysis of trends in achievement within the experimental and control groups as well as of differences in achievement between them was conducted using both raw scores and grades received on the objective part of the final examinations in German 131 and 132.

A chi-square test of independence based on the frequency distribution of students by letter grade on the objective part of the final examination for the experimental and the control groups tested a version of hypothesis 10; namely, the null hypothesis that the distribution of grades on the objective part of the final examination is the same for both the experimental and the control groups, i.e., independent of the classification into experimental and control. The alternative hypothesis was that the distribution of grades is different, that is, not independent for the experimental and the control groups.

The second test was a chi-square test of the same location. For this test, the obtained chi-square was partitioned into two parts. The first part was used to investigate whether the two grade distributions differ significantly in terms of location, i.e., to see if one distribution tended to have higher scores than the other. To do this, the A's, B's, and C's were treated as one class and the D's and F's as another. This classification was chosen because it divided the total number of students in each group into approximately equally sized classes. The expectations for these two classes for both the experimental group and for the control group were obtained by addition of the expectations used for the chi-square test of independence. The hypothesis being tested was:

$H_o$: Distribution of grades on the objective part of the final examination is the same for both the experimental and the control groups, i.e., independent of the classification into experimental and control.

$H_A$: The distribution of grades for the experimental group has a different location from the distribution of grades for the control group.

The third chi-square test investigated the differences due to reasons other than location, that is, whether the two grade distributions differed significantly if the effect of location was removed. This chi-square value was obtained by subtraction. The hypothesis tested is:

$H_o$: The distribution of grades is the same for both the experimental and the control groups.

$H_A$: The distribution of grades for the experimental group is different from the distribution of grades for the control group for reasons other than differences in location.

To investigate whether method of evaluation affected the variances of the two groups, the null hypothesis of no difference was tested using an F-ratio involving raw scores on the objective part of the final examination.

The assumption of equal variance justified the use of preliminary t-tests to investigate further hypothesis 10 of no difference in achievement between the experimental and the control groups. These tests of

equal means were based on pooled variance for related samples, and on repeated measurement in the case where the experimental and the control groups served as their own controls when comparing achievement in elementary German 131 and 132.

The next analysis used the technique of multiple regression to control for independent variables which had been ignored in the preceding investigation, and which were assumed to contribute to the variability in achievement in elementary foreign language courses, in this case, $X_2$, raw score on the objective part of the final examination in German 132. The following independent variables were included in the multiple regression analysis:

$X_1$ = method of evaluation: self-pacing versus lock step

$X_5$ = grade-point average, end of Winter Quarter 1971

$X_6$ = course load during the Winter Quarter 1971

$X_{7-11}$ = instructors*

$X_{12}$ = sex

$X_{13-16}$ = year in college*

$X_{17-21}$ = college*

$X_{22}$ = language aptitude as measured by the composite raw score on parts 2, 4, 5, and 6 of the Pimsleur Language Aptitude Battery (76)

$X_{23}$ = interest, as estimated by the interest score on the Pimsleur Language Aptitude Battery (PLAB)

$X_{24}$ = motivation, as estimated from a 4-choice scale.

Note that each variable marked with an asterisk actually represents the contrasts among individual variables. For example, $X_{7-11}$, instructors,

represents four contrasts among the 5 instructors involved.

To predict achievement in German 132 from achievement in German 131 and to investigate the differences in results due to the use of the following dependent variables: (1) objective score on the final examination; (2) objective final grade; or (3) instructor grade, the following variables were used:

$X_1$ = score on the objective part of the final examination in German 131

$X_2$ = grade based on $X_1$, score obtained on the objective part of the final examination in German 131

$X_3$ = instructor grade for German 131

$X_4$ = cumulative grade-point average, end of Fall Quarter 1970

$X_5$ = course load during the Fall Quarter 1970

$X_6$ = method of evaluation in German 132; i.e., traditional lock step for the control group, and self-pacing for the experimental group

$X_7$ = score obtained on the objective part of the final examination in German 132

$X_8$ = grade based on $X_7$

$X_9$ = instructor grade for German 132.

The following hypotheses were tested:

Hypothesis 11:

There is no relationship between achievement in German 132 and German 131 as measured by the score obtained on the objective part of the final examination.

Hypothesis 12:

There is no relationship between achievement in German 132 and German 131 as measured by the grade received on the basis of the objective part of the final examination.

Hypothesis 13:

There is no relationship between achievement in German 132 and German 131 as measured by the instructor grade for German 132.

Hypothesis 14:

There is no relationship between cumulative grade-point average at the end of the Fall Quarter 1970 and achievement in German 132.

Hypothesis 15:

There is no relationship between course load during the Fall Quarter 1970 and achievement in German 132.

Hypothesis 16:

There is no relationship between method of evaluation in German 132 and achievement in German 132.

The technique of analysis of multiple regression was used to test the null hypotheses. The regression models for each of the dependent variables were:

$$X_7 = \beta_o + \beta_1 X_1 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \varepsilon$$

$$X_8 = \beta_o + \beta_1 X_1 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \varepsilon$$

$$X_9 = \beta_o + \beta_1 X_1 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \varepsilon$$

$$X_7 = \beta_o + \beta_2 X_2 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \varepsilon$$

$$X_8 = \beta_o + \beta_2 X_2 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \varepsilon$$

$$X_9 = \beta_o + \beta_2 X_2 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \epsilon$$

$$X_7 = \beta_o + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \epsilon$$

$$X_8 = \beta_o + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \epsilon$$

$$X_9 = \beta_o + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \epsilon$$

where the unknown regression coefficients, $\beta$, were estimated from the multiple regression program which also computed correlation matrices, F-ratios, and t-values for each combination of variables, and where $\epsilon$ is the unexplained variation in the dependent variables.

Each of these regression models included as an independent variable cumulative grade-point average, $X_4$, course load, $X_5$, and method of evaluation, $X_6$, because these variables had been found to influence achievement in the elementary foreign language courses investigated in connection with the pass-fail system.

In addition, these models included as an independent variable some measure of achievement in German 131, either the objective final score, $X_1$, or the objective final grade, $X_2$, or the instructor grade, $X_3$. Each model was analyzed in turn.

Assumptions

For the purposes of this investigation, the following assumptions were made:

1. The students enrolled in German 131 and 132 during the Fall Quarter 1970 and during the Winter Quarter 1971 are representative of German students at Iowa State University as to aptitude, motivation, grade-point average, sex ratio, and course load.

2. The grades received are a satisfactory measure of achievement in elementary German courses at Iowa State University.

3. The mean final score and the mean final grade on the objective part of the final examination are a linear additive function of the independent variables.

4. The residuals, i.e., the deviations of final score or grade from the regression plane of the independent variables are independently distributed.

5. The residuals have common variance and are normally distributed.

6. The independent variables are considered fixed so that the conclusions hold for groups of students who have the same distributions of values for the independent variables in this group.

## Findings

### Preliminary analysis

In Table 42a, the grade distribution of raw scores on the objective part of the final examination is shown for both the experimental and the control groups in German 132.

As evident from Table 42a, the experimental group had a higher percentage of A's, B's and C's, and fewer D's and F's than the control group, but a wider range in raw scores at the low end of the distribution.

These findings were verified by the following tests as presented in Tables 43 through 46.

Consulting the table of chi-square, it was found that with 4 degrees of freedom, a computed value of 10.96 (Table 43) is significant at the .05 level. Therefore, the null hypothesis of equal grade distributions in the experimental and the control groups was rejected.

Table 42a. Group III. Frequency distribution of students by raw score and grade on objective part of the final examination for the experimental and the control groups

| Grade | Raw score | Number of students | | Total | | | |
|---|---|---|---|---|---|---|---|
| | | | | Experimental | | Control | |
| | | Experimental | Control | Number | Percent | Number | Percent |
| | 145 - 149 | 1 | 0 | | | | |
| A | 140 - 144 | 1 | 2 | | | | |
| | 135 - 139 | 2 | 3 | | | | |
| | | | | 4 | 7 | 5 | 4.3 |
| | 130 - 134 | 4 | 3 | | | | |
| B | 125 - 129 | 5 | 5 | | | | |
| | 120 - 124 | 8 | 7 | | | | |
| | | | | 17 | 30 | 15 | 13.0 |
| | 115 - 119 | 2 | 14 | | | | |
| C | 110 - 114 | 6 | 8 | | | | |
| | 105 - 109 | 8 | 10 | | | | |
| | | | | 16 | 29 | 32 | 27.3 |
| | 100 - 104 | 5 | 14 | | | | |
| D | 95 - 99 | 2 | 15 | | | | |
| | 90 - 94 | 3 | 11 | | | | |
| | | | | 10 | 18 | 40 | 34.0 |
| | 85 - 89 | 1 | 9 | | | | |
| | 80 - 84 | 0 | 11 | | | | |
| | 75 - 79 | 2 | 2 | | | | |
| | 70 - 74 | 2 | 2 | | | | |
| | 65 - 69 | 2 | 1 | | | | |
| | 60 - 64 | 0 | 0 | | | | |
| | 55 - 59 | 2 | 0 | | | | |
| | 50 - 54 | 0 | 0 | 9 | 16 | 25 | 21.4 |
| | Total | 56 | 117 | 56 | | 117 | |

Table 43. Group III. Chi-square test of independence; N = 173

| Grade | Observed number of students: Experimental | Control | Total | Expected number of students: Experimental | Control | Total |
|---|---|---|---|---|---|---|
| A | 4 | 5 | 9 | $E_{11}$ 2.91 | $E_{12}$ 6.09 | 9 |
| B | 17 | 15 | 32 | $E_{21}$ 10.35 | $E_{22}$ 21.65 | 32 |
| C | 16 | 32 | 48 | $E_{31}$ 15.53 | $E_{32}$ 32.47 | 48 |
| D | 10 | 40 | 50 | $E_{41}$ 16.18 | $E_{42}$ 33.82 | 50 |
| F | 9 | 25 | 34 | $E_{51}$ 11.03 | $X_{52}$ 22.97 | 34 |
| Total | 56 | 117 | 173 | 56.00 | 117.00 | 173 |

$\chi^2 = 10.96$

Tabular $\chi^2$ 4, (.05) = 9.49

Table 44. Group III. Chi-square test of same location; N = 173

| Grade | Observed number of students: Experimental | Control | Total | Expected number of students: Experimental | Control | Total |
|---|---|---|---|---|---|---|
| A,B,C | 37 | 52 | 89 | 28.79 | 60.21 | 89 |
| D,F | 19 | 65 | 84 | 27.21 | 56.79 | 84 |
| Total | 56 | 117 | 173 | 56.00 | 117.00 | 173 |

The test statistic is chi-square $=\Sigma(O-E)^2/E$ with (2-1)(2-1) = 1 d.f.

$\chi^2 = 7.13$ Tabular $\chi^2$ 1 d.f. (.01) = 6.63

Table 45. Group III. Chi-square test of residual differences; N = 173

| Name of test | Value | Degrees of freedom |
|---|---|---|
| Chi-square of independence | 10.96 | 4 |
| Chi-square of location | -7.13 | -1 |
| Chi-square residual | 3.83 | 3 |

Tabular $\chi^2$ 3 d.f. (.05) = 7.81

Table 46. Group III. Data for F-test of equal variance; N = 173

| Group | Degrees of freedom | Sum of squares | Mean square | Standard deviation |
|---|---|---|---|---|
| Experimental | 55 | 25,540 | 464.36 | 21.55 |
| Control | 116 | 31,100 | 268.10 | 16.37 |

The statistic is $F = \frac{\text{larger } s^2}{\text{smaller } s^2} = \frac{21.55}{16.37} = 1.31$

Tabular $F_{55,116} = 1.40$

Consulting the table of chi-square, it was found that with 1 degree of freedom, a value of 7.13 (Table 44) is significant at the .01 level. Therefore, the null hypothesis of no difference in location between the two groups was rejected at the .01 level of significance.

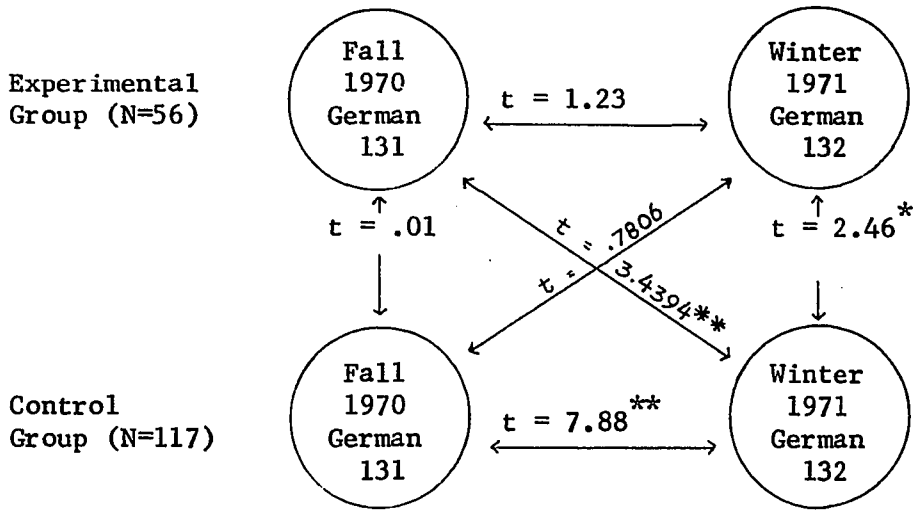
Consulting the table of chi-square, it was found that with 4 degrees of freedom, a value of 3.83 (Table 45) is not significant at the .05 level.

When testing at the .05 level of significance, the chi-square test of independence suggests that the distributions of grades for the objective part of the final examination were different for the control and the experimental groups.

Partition of this chi-square for independence indicates that the principal reason for the difference between the grade distributions was one of location. Inspection of Table 44 (chi-square test of same location) suggests that this difference in location was due to a larger percentage of A's, B's, and C's in the experimental group, relative to the number of D's and F's, than in the control group.

Consulting the F-distribution table, it was found that with 55 and 116 degrees of freedom, a value of 1.31 (Table 46) is not significant at a .05 level. Therefore, the null hypothesis of equal variance of raw scores was not rejected.

T-tests of equal means, as presented in Figure 7, were based on pooled variance for related samples and on repeated measurement. They led to a rejection of the null hypothesis of no difference in mean grade between the experimental and the control groups as far as achievement on the objective part of the final examination is concerned.




*Indicates that $p < .05$.

**Indicates that $p < .01$.

Figure 7. Group III. Preliminary t-tests: comparison of experimental and control groups

The results of the six t-tests performed are as follows:

t-test 1 indicates that there was no initial statistically significant difference between the control group and the experimental group in their performance on the objective final examination in German 131.

t-test 2 indicates a statistically highly significant difference between the experimental group and the control group in their performance on the objective final examination in German 132.

t-test 3 indicates that there was no statistically significant decline in performance of the experimental group in German 132 in comparison to German 131.

t-test 4 reveals a highly significant decline in performance of the control group in German 132 as compared with German 131.

t-test 5 indicates no statistically significant difference in the mean performance between the control group in German 131 and the experimental group in German 132.

t-test 6 indicates a highly significant difference between the performance of the experimental group in German 131 and the control group in German 132.

Multiple regression analysis

The purpose of this analysis was to test null hypothesis 10 of no difference in achievement between the experimental and the control groups after adjusting for those independent variables which were assumed to contribute significantly to the variation in the dependent variables, $X_2$, score on the objective part of the final examination, and in $X_3$, grade based on this score.

The means and standard deviations of the variables investigated in this analysis are recorded in Table 47.

Method of evaluation, $X_1$, is identical with the instructor for the experimental group. The code for the computer regression program was: 1 = experimental; 2 = control. $X_{7-11}$ identify instructors. $X_{12}$ = sex, coded 1 = males; 2 = females. $X_{13-16}$ identify year in college, and $X_{17-21}$ identify colleges.

Inspection of Table 48 indicates that, on the average, the experimental group received higher scores and grades on the objective part of the final examination than the control group. The experimental group also had a somewhat higher cumulative grade-point average. The substantial

Table 47. Group III. Means and standard deviations of dependent and independent variables for the experimental group, the control group, and the combined group

| Variables | Experimental group N = 56 Mean | S.D. | Control group N = 117 Mean | S.D. | Combined group N = 173 Mean | S.D. |
|---|---|---|---|---|---|---|
| Dependent variables: | | | | | | |
| $X_2$ | 108.16 | 21.55 | 103.92 | 16.37 | 105.29 | 18.26 |
| $X_3$ | 2.14 | 1.35 | 1.53 | 1.13 | 1.73 | 1.23 |
| $X_4$ | 2.20 | 1.49 | 2.26 | 1.00 | 2.24 | 1.18 |
| Independent variables: | | | | | | |
| $X_5$ | 2.93 | 0.52 | 2.86 | .57 | 2.88 | .55 |
| $X_6$ | 16.54 | 2.87 | 16.61 | 2.76 | 16.58 | 2.79 |
| $X_{22}^a$ | 77.09 | 9.42 | - | - | - | - |
| $X_{23}^a$ | 4.54 | 1.99 | - | - | - | - |
| $X_{24}^a$ | 2.27 | 1.15 | - | - | - | - |

where:

$X_2$ = raw score on objective part of final test

$X_3$ = grade based on raw score on objective final test

$X_4$ = instructor grade

$X_5$ = grade-point average

$X_6$ = course load

$X_{22}$ = language aptitude

$X_{23}$ = interest

$X_{24}$ = motivation

[a] Variables tested only in connection with the experimental group.

Table 48. Group III. Product-moment correlation coefficient matrix of dependent and independent variables for the experimental group; N = 56

| Variables | $X_2$ | $X_3$ | $X_4$ | $X_5$ | $X_6$ | $X_{12}$ | $X_{22}$ | $X_{23}$ | $X_{24}$ |
|---|---|---|---|---|---|---|---|---|---|
| $X_2$ | 1.000 | | | | | | | | |
| $X_3$ | .941* | 1.000 | | | | | | | |
| $X_4$ | .840* | .884* | 1.000 | | | | | | |
| $X_5$ | .576* | .580* | .550* | 1.000 | | | | | |
| $X_6$ | .223 | .195 | .179 | .028 | 1.000 | | | | |
| $X_{12}$ | .226 | .306 | .306 | .431* | .067 | 1.000 | | | |
| $X_{22}$ | .384* | .374* | .222 | .329 | .135 | .123 | 1.000 | | |
| $X_{23}$ | .177 | .213 | .281 | .110 | -.159 | .331 | .333 | 1.000 | |
| $X_{24}$ | .250 | .255 | .265 | .191 | -.088 | .185 | .284 | .631* | 1.000 |

*Significant at the .01 level with N = 56.

relationship between grade-point average and achievement as measured by raw score and the grade based on it, is reflected in the highly significant correlation coefficients as listed in Table 48, where $r_{25} = .576$ and $r_{35} = .580$. This correlation matrix also shows that language aptitude, $X_{22}$, was significantly correlated with $X_2$, raw score and $X_3$, grade based on raw score ($r_{2,22} = 384$; $r_{3,22} = .374$). The negative correlations of course load with interest, $X_{23}$, and motivation, $X_{24}$, suggest that as course load increased, interest and motivation decreased in the experimental group.

Method of evaluation, $X_1$

Table 49 lists those regression models in which method of evaluation, $X_1$, is included as an independent variable. The null hypothesis that $\beta_1 = 0$ is tested against the alternative hypothesis that $\beta_1 \neq 0$, and the calculated "t" is compared with the tabular "t" with the degrees of freedom for the residual mean square for each model which included $X_1$, and the level of significance is set at .05.

Table 49. Group III. Summary of calculated t-values for various multiple regression models used to investigate the effect of method of evaluation, $X_1$, on the dependent variables, $X_2$, raw score on objective part of the final test in German 132 and on $X_3$, grade based on $X_2$ - combined group; N = 173

| Model | Dependent variable | Independent variables included in regression model | Degrees of freedom | Calculated "t" |
|---|---|---|---|---|
| 1-IIIC | $X_2$ | $X_1X_5X_6X_7X_9X_{10}X_{12}X_{13\text{-}16}X_{17\text{-}21}$ | 156 | -0.0220 |
| 2-IIIC | $X_2$ | $X_1X_5X_6X_{12}X_{13\text{-}16}X_{17\text{-}21}$ | 159 | -0.7806 |
| 3-IIIC | $X_2$ | $X_1X_5X_6X_7X_9X_{10}X_{12}X_{17\text{-}21}$ | 160 | 0.0538 |
| 4-IIIC | $X_2$ | $X_1X_5X_6X_7X_9X_{10}X_{12}X_{13\text{-}16}$ | 161 | -0.0286 |
| 5-IIIC | $X_2$ | $X_1$ | 171 | -1.4329 |
| 6-IIIC | $X_3$ | $X_1X_5X_6X_7X_9X_{10}X_{12}X_{13\text{-}16}X_{17\text{-}21}$ | 156 | -1.1319 |
| 7-IIIC | $X_3$ | $X_1X_5X_6X_{12}X_{13\text{-}16}X_{17\text{-}21}$ | 159 | -2.8566* |
| 8-IIIC | $X_3$ | $X_1X_5X_6X_7X_9X_{10}X_{12}X_{13\text{-}16}$ | 161 | -1.1845 |
| 9-IIIC | $X_3$ | $X_1$ | 171 | -3.1322* |

*Indicates $p < .05$.

Table 49 provides evidence that method of evaluation, $X_1$, explains a significant portion of the variation in the dependent variable, $X_3$, grade received on the objective part of the final examination, for the combined group, after adjusting for $X_5$, grade-point average, $X_6$, course load, $X_{12}$, sex, $X_{13-16}$, year in college, and $X_{17-21}$, college. It remains significant when all the other variables are eliminated from the regression model. It is not significant when instructors, $X_7X_9X_{10}$, are included in the model.

Method of evaluation, $X_1$, does not explain a significant portion of the variation in the dependent variable when this variable is $X_2$, raw score on the objective part of the final examination in German 132.

Grade-point average, $X_5$

Table 50 lists those regression models in which grade-point average, $X_5$, is included as an independent variable. The null hypothesis that $\beta_5 = 0$ is tested against the alternative hypothesis that $\beta_5 \neq 0$, and the calculated "t" is compared with the tabular "t" with the degrees of freedom for the residual mean square for each model which includes $X_5$, and the level of significance is set at .05.

Table 50 provides evidence that grade-point average, $X_5$, explains a significant portion of the variation in the dependent variable, $X_2$, raw score on the objective part of the final examination in German 132, for the experimental group, after adjusting for all the other independent variables included in the complete model and for all possible combinations of them. Grade-point average is also highly significant when used independently of other variables ($t = 5.1810$ with 54 degrees of freedom).

Table 50. Group III. Summary of calculated t-values for various multiple regression models used to investigate the effect of grade-point average, $X_5$, on the dependent variable, $X_2$, raw score on the objective part of the final test in German 132 - experimental group; N = 56

| Model | Independent variables included in regression model | Degrees of freedom | Calculated "t" |
|---|---|---|---|
| 1-IIIA | $X_5X_6X_{12}X_{13\text{-}16}X_{17\text{-}19}X_{22}X_{23}X_{24}$ | 42 | 3.5202** |
| 2-IIIA | $X_5X_6X_{12}X_{13\text{-}16}X_{17\text{-}19}$ | 45 | 4.7501** |
| 3-IIIA | $X_5X_6X_{12}X_{13\text{-}16}X_{22}X_{23}X_{24}$ | 45 | 4.0374** |
| 4-IIIA | $X_5X_6X_{12}X_{17\text{-}19}$ | 49 | 4.6345** |
| 5-IIIA | $X_5X_6X_{12}X_{17\text{-}19}X_{22}X_{23}X_{24}$ | 46 | 3.7089** |
| 6-IIIA | $X_5X_6X_{12}X_{13\text{-}16}$ | 48 | 5.0095** |
| 7-IIIA | $X_5X_6X_{22}X_{23}X_{24}$ | 50 | 4.1909** |
| 8-IIIA | $X_5X_6X_{22}X_{23}$ | 51 | 4.3452** |
| 9-IIIA | $X_5X_6X_{22}X_{24}$ | 51 | 4.2288** |
| 10-IIIA | $X_5X_6X_{23}X_{24}$ | 51 | 4.8908** |
| 11-IIIA | $X_5X_{22}X_{23}X_{24}$ | 51 | 4.1977** |
| 13-IIIA | $X_5X_6X_{22}$ | 52 | 4.3700** |
| 14-IIIA | $X_5X_6X_{23}$ | 52 | 5.0989** |
| 15-IIIA | $X_5X_6X_{24}$ | 52 | 4.9188** |
| 16-IIIA | $X_5X_{22}X_{23}$ | 52 | 4.3539** |
| 17-IIIA | $X_5X_{22}X_{24}$ | 52 | 4.2572** |
| 18-IIIA | $X_5X_{23}X_{24}$ | 52 | 4.8319** |
| 23-IIIA | $X_5X_6$ | 53 | 5.2502** |
| 24-IIIA | $X_5X_{22}$ | 53 | 4.3866** |

**Indicates $p < .01$.

Table 50 (continued)

| Model | Independent variables included in regression model | Degrees of freedom | Calculated "t" |
|---|---|---|---|
| 25-IIIA | $X_5X_{23}$ | 53 | 5.0389** |
| 26-IIIA | $X_5X_{24}$ | 53 | 4.8710** |
| 34-IIIA | $X_5$ | 54 | 5.1810** |

Grade-point average explains the variation in both $X_2$, raw score on the objective part of the final examination and in $X_3$, the grade based on the raw score, for the combined groups (t = 9.58 with 156 degrees of freedom; tabular t = 2.58 at the .01 level of significance).

Course load, $X_6$

Table 51 lists those regression models in which course load, $X_6$, is included as an independent variable. The null hypothesis that $\beta_6 = 0$ is tested against the alternative hypothesis that $\beta_6 \neq 0$, and the calculated "t" is compared with the tabular "t" with the degrees of freedom for the residual mean square for each model which included $X_6$, and the level of significance is set at .05.

In contrast to grade-point average, $X_5$, the independent variable course load, $X_6$, does not explain independently a significant portion of the variation in the dependent variable, $X_2$. It is, however, significant in combination with other variables. In the experimental group,

Table 51. Group III. Summary of calculated t-values for various multiple regression models used to investigate the effect of course load, $X_6$, on the dependent variable, $X_2$, raw score on the objective part of the final test in German 132 - experimental group; N = 56

| Model | Independent variables included in regression model | Degrees of freedom | Calculated "t" |
|---|---|---|---|
| 1-IIIA | $X_5X_6X_{12}X_{13\text{-}16}X_{17\text{-}19}X_{22}X_{23}X_{24}$ | 42 | 1.7094 |
| 2-IIIA | $X_5X_6X_{12}X_{13\text{-}16}X_{17\text{-}19}$ | 45 | 0.9484 |
| 3-IIIA | $X_5X_6X_{12}X_{13\text{-}16}X_{22}X_{23}X_{24}$ | 45 | 1.5032 |
| 4-IIIA | $X_5X_6X_{12}X_{17\text{-}19}$ | 49 | 1.8559 |
| 5-IIIA | $X_5X_6X_{12}X_{17\text{-}19}X_{22}X_{23}X_{24}$ | 46 | 2.5067* |
| 6-IIIA | $X_5X_6X_{12}X_{13\text{-}16}$ | 48 | 0.9842 |
| 7-IIIA | $X_5X_6X_{22}X_{23}X_{24}$ | 50 | 2.3561* |
| 8-IIIA | $X_5X_6X_{22}X_{23}$ | 51 | 2.3776* |
| 9-IIIA | $X_5X_6X_{22}X_{24}$ | 51 | 2.1087* |
| 10-IIIA | $X_5X_6X_{23}X_{24}$ | 51 | 2.1087* |
| 12-IIIA | $X_6X_{22}X_{23}X_{24}$ | 51 | 2.3282* |
| 13-IIIA | $X_5X_6X_{22}$ | 52 | 2.3084* |
| 14-IIIA | $X_5X_6X_{23}$ | 52 | 2.1221* |
| 15-IIIA | $X_5X_6X_{24}$ | 52 | 2.0645* |
| 19-IIIA | $X_6X_{22}X_{23}$ | 52 | 2.3506* |
| 20-IIIA | $X_6X_{22}X_{24}$ | 52 | 2.3771* |
| 21-IIIA | $X_6X_{23}X_{24}$ | 52 | 1.9399 |
| 23-IIIA | $X_5X_6$ | 53 | 1.9044 |

*Indicates $p < .05$.

Table 51 (continued)

| Model | Independent variables included in regression model | Degrees of freedom | Calculated "t" |
|---|---|---|---|
| 27-IIIA | $X_6X_{22}$ | 53 | 2.2917* |
| 28-IIIA | $X_6X_{23}$ | 53 | 1.9462 |
| 29-IIIA | $X_6X_{24}$ | 53 | 1.9119 |
| 35-IIIA | $X_6$ | 54 | 1.6800 |

course load is most effective in combination with $X_5$, grade-point average, $X_{12}$, sex, $X_{17-19}$, college, $X_{22}$, language aptitude, $X_{23}$, interest, and $X_{24}$, motiviation, i.e., when all independent variables investigated in connection with Group III are included in the regression model, except for $X_{13-16}$, year in college. When $X_{13-16}$ is included in the model, course load becomes insignificant. The next highest t-value for course load ($t = 2.3776$ with 51 degrees of freedom) occurs when it is combined with grade-point average, $X_5$, language aptitude, $X_{22}$, and interest, $X_{23}$. It remains significant at the .05 level if grade-point average is eliminated from this combination; that is, if only language aptitude, $X_{22}$, and interest, $X_{23}$, are controlled. Course load is least significant ($t = .9484$ with 45 degrees of freedom) in the experimental group when language aptitude, $X_{22}$, interest, $X_{23}$, and motivation, $X_{24}$, are removed simultaneously from the regression model. Since interest, $X_{23}$, and motivation, $X_{24}$, are significantly correlated with language aptitude, $X_{22}$,

course load and language aptitude in combination with either interest or motivation explain a significant portion of the variability in the dependent variable, $X_2$, raw score on the objective part of the final examination. It seems, however, that either language aptitude or grade-point average must be included in the regression model for course load to be significant. Course load, $X_6$, is not significant in connection with the control or with the combined groups. In the combined groups (N = 173) the t-values for course load are negative when the dependent variable is $X_3$, grade obtained on the objective part of the final examination. This means that as course load increases, grades decrease. On the other hand, if the dependent variable is $X_2$, raw score, the t-values are positive, which means that scores become higher as course load increases. These results may be due to random variation.

Instructors, $X_7X_9X_{10}$

During the Winter Quarter 1971, German 132 was taught by 5 different instructors. Variables $X_7$, $X_9$, and $X_{10}$, represent the instructors of the control group. Instructor $X_8$ taught the 4 sections of the experimental group and, therefore, being identical with method of evaluation, $X_1$, instructor $X_8$ has been excluded from this analysis.

The analysis of variance Table 52 tests the null hypothesis that $\beta_7 = \beta_9 = \beta_{10} = 0$ against the alternative hypothesis that at least one of $\beta_7$, $\beta_9$, $\beta_{10} \neq 0$ for the models investigated for the combined group.

Table 52 provides evidence that instructors $X_7X_9X_{10}$ did not explain a significant portion of the variation in raw scores on the objective part of the final examination, $X_2$, after adjusting for the effect of method of evaluation, $X_1$, grade-point average, $X_5$, course load, $X_6$, sex, $X_{12}$, year

Table 52. Group III. Analysis of regression - effect of the elimination of $X_7X_9X_{10}$, instructors

| Source of variation | Degrees of freedom | Sum of squares | Mean square |
|---|---|---|---|
| Regression on $X_1X_5X_6X_7X_9X_{10}X_{12}X_{13\text{-}16}X_{17\text{-}21}$ | 16 | 25,313.567 | |
| Regression on $X_1X_5X_6X_{12}X_{13\text{-}16}X_{17\text{-}21}$ | 13 | 23,900.916 | |
| Difference due to elimination of $X_7X_9X_{10}$ | 3 | 1,412.651 | 470.884 |
| Residual | 156 | 32,006.398 | 205.169 |

Calculated F = 2.295

Tabular $F_{3,156}$ = 2.67 (.05)

in college, $X_{13\text{-}16}$, and college, $X_{17\text{-}21}$ when testing at the .05 level of significance.

Sex, $X_{12}$

Sex, $X_{12}$, does not explain a significant portion of the variation in raw scores on the objective part of the final examination, $X_2$, when testing at the .05 level of significance. Evidence is provided by Table 53 which summarizes the results of testing the null hypothesis that $\beta_{12} = 0$ against the alternative hypothesis that $\beta_{12} \neq 0$ for the various models investigated.

Table 53. Group III. Summary of calculated t-values for various multiple regression models used to investigate the effect of sex, $X_{12}$, on the dependent variable, $X_2$, raw score on the objective part of the final test in German 132 - experimental group; N = 56

| Model | Independent variables included in regression model | Degrees of freedom | Calculated "t" |
|---|---|---|---|
| 1-IIIA | $X_5X_6X_{12}X_{13\text{-}16}X_{17\text{-}19}X_{22}X_{23}X_{24}$ | 42 | -0.7491 |
| 2-IIIA | $X_5X_6X_{12}X_{13\text{-}16}X_{17\text{-}19}$ | 45 | -0.4403 |
| 3-IIIA | $X_5X_6X_{12}X_{13\text{-}16}X_{22}X_{23}X_{24}$ | 45 | -0.7512 |
| 4-IIIA | $X_5X_6X_{12}X_{17\text{-}19}$ | 49 | -0.2219 |
| 5-IIIA | $X_5X_6X_{12}X_{17\text{-}19}X_{22}X_{23}X_{24}$ | 46 | -0.5508 |
| 6-IIIA | $X_5X_6X_{12}X_{13\text{-}16}$ | 48 | -0.5305 |

Year in college, $X_{13\text{-}16}$

These variables represent the contrasts between the freshman, sophomore, junior, senior, and graduate status of students and were coded as follows for the computer multiple regression program:

| | |
|---|---|
| Freshmen | 1000 |
| Sophomores | 0100 |
| Juniors | 0010 |
| Seniors | 0001 |
| Graduate students | 0000 |

The analysis of regression Table 54 tests the null hypothesis that $\beta_{13} = \beta_{14} = \beta_{15} = \beta_{16} = 0$ against the alternative hypothesis that at least one of $\beta_{13}$, $\beta_{14}$, $\beta_{15}$, $\beta_{16} \neq 0$.

Table 54. Group III. Analysis of regression - effect of the elimination of year in college, $X_{13-16}$, combined group; N = 172

| Source of variation | Degrees of freedom | Sum of squares | Mean square |
|---|---|---|---|
| Regression on $X_1X_5X_6X_7X_9X_{10}X_{12}X_{13-16}X_{17-21}$ | 16 | 25,313.567 | |
| Regression on $X_1X_5X_6X_7X_9X_{10}X_{12}X_{17-21}$ | 12 | 22,595.381 | |
| Difference due to elimination of $X_{13-16}$ | 4 | 2,718.186 | 679.547 |
| Residual | 156 | 32,006.398 | 205.169 |

Calculated F = 3.3121

Tabular $F_{4,156}$ = 2.43 (.05 level of significance)
3.44 (.01 level of significance)

The calculated F-ratio is significant at the .05 level. This result indicates that for the combined group, $X_2$, raw score on the objective part of the final examination in German 132, varied according to year in college.

The analysis of regression Table 55 provides evidence that for the experimental group, year in college, $X_{13-16}$, does not explain a significant portion of the variation in the dependent variable, $X_2$, after adjusting for the effects of grade-point average, $X_5$, course load, $X_6$,

sex, $X_{12}$, college, $X_{17-19}$, language aptitude, $X_{22}$, interest, $X_{23}$, and motivation, $X_{24}$, when testing at the .05 level of significance.

Table 55. Group III. Analysis of regression - effect of the elimination of year in college, $X_{13-16}$ - experimental group; N = 56

| Source of variation | Degrees of freedom | Sum of squares | Mean square |
|---|---|---|---|
| Regression on $X_5X_6X_{12}X_{13-16}X_{17-19}X_{22}X_{23}X_{24}$ | 13 | 13,872.811 | |
| Regression on $X_5X_6X_{12}X_{17-19}X_{22}X_{23}X_{24}$ | 9 | 12,316.028 | |
| Difference due to elimination of $X_{13-16}$ | 4 | 1,556.783 | 389.196 |
| Residual (complete model) | 42 | 11,666.742 | 277.779 |

Calculated F = 1.40

Tabular $F_{4,42}$ = 2.59 at the .05 level of significance

College, $X_{17-21}$

The analysis of variance Table 56 tests the null hypothesis that $\beta_{17} = \beta_{18} = \beta_{19} = \beta_{20} = \beta_{21} = 0$ against the alternative hypothesis that at least one of $\beta_{17}$, $\beta_{18}$, $\beta_{19}$, $\beta_{20}$, $\beta_{21} \neq 0$ for the models investigated for the combined group.

Tables 56 and 57 provide evidence that College, as independent variable, does not explain a significant portion of the variation in raw score on the objective part of the final examination, $X_2$, when testing at the .05 level of aignificance.

Table 56. Combined Group III. Analysis of regression - effect of the elimination of $X_{17-21}$, college

| Source of variation | Degrees of freedom | Sum of squares | Mean square |
|---|---|---|---|
| Regression on $X_1X_5X_6X_7X_9X_{10}X_{12}X_{13-16}X_{17-21}$ | 16 | 25,313.567 | |
| Regression on $X_1X_5X_6X_7X_9X_{10}X_{12}X_{13-16}$ | 11 | 24,246.888 | |
| Difference due to elimination of $X_{17-21}$ | 5 | 1,066.679 | 213.336 |
| Residual | 156 | 32,006.398 | 205.169 |
| Calculated F = 1.04 | | | |
| Tabular $F_{5,156}$ = 2.19 (.05) | | | |

Table 57. Experimental Group III. Analysis of regression - effect of the elimination of $X_{17-19}$, college

| Source of variation | Degrees of freedom | Sum of squares | Mean square |
|---|---|---|---|
| Regression on $X_5X_6X_{12}X_{13-16}X_{17-19}X_{22}X_{23}X_{24}$ | 13 | 13,872.811 | |
| Regression on $X_5X_6X_{12}X_{13-16}X_{22}X_{23}X_{24}$ | 10 | 13,276.207 | |
| Difference due to elimination of $X_{17}X_{18}X_{19}$ | 3 | 596.604 | 198.87 |
| Residual | 42 | 11,666.742 | 277.78 |
| Calculated F = .7159 | | | |
| Tabular $F_{3,42}$ = 2.83 (.05) | | | |

Language aptitude, $X_{22}$

Table 58 lists those models in which language aptitude, $X_{22}$, is included as independent variable. The null hypothesis that $\beta_{22} = 0$ is tested against the alternative hypothesis that $\beta_{22} \neq 0$ for each model, and the value of the test statistic "t" calculated by the computer is compared to a tabular "t" using the degrees of freedom for residual, and the .05 level of significance.

From Table 58 it is evident that language aptitude explains a significant portion of variation in raw score on the objective part of the final examination, $X_2$, for the experimental group when testing at the .05 level. It is most significant if used as a sole predictor or in combination with course load, or both course load and interest. Language aptitude does not explain a significant portion of the variability in the dependent variable, $X_2$, if course load, interest or motivation are excluded from the regression model. The t-value for language aptitude is higher if it is combined with interest than when it is combined with motivation, $X_{24}$. Information on the effect of language aptitude, as measured by the Pimsleur Language Aptitude Battery, was not available for the control group.

Interest, $X_{23}$

Interest, $X_{23}$, does not explain a significant portion of the variation in raw scores on the objective part of the final examination, $X_2$, for the experimental group when testing at the .05 level of significance.

Evidence is provided by Table 59 which summarizes the results of testing the null hypothesis that $\beta_{23} = 0$ against the alternative hypothesis that $\beta_{23} \neq 0$ for the various models investigated.

Table 58. Group III. Summary of calculated t-values for various multiple regression models used to investigate the effect of language aptitude, $X_{22}$, on the dependent variable $X_2$, raw score on the objective part of the final test in German 132 - experimental group; N = 56

| Model | Independent variables included in regression model | Degrees of freedom | Calculated "t" |
|---|---|---|---|
| 1-IIIA | $X_5X_6X_{12}X_{13\text{-}16}X_{17\text{-}19}X_{22}X_{23}X_{24}$ | 42 | 2.2371* |
| 3-IIIA | $X_5X_6X_{12}X_{13\text{-}16}X_{22}X_{23}X_{24}$ | 45 | 2.0184* |
| 5-IIIA | $X_5X_6X_{12}X_{17\text{-}19}X_{22}X_{23}X_{24}$ | 46 | 2.1203* |
| 7-IIIA | $X_5X_6X_{22}X_{23}X_{24}$ | 50 | 1.9006 |
| 8-IIIA | $X_5X_6X_{22}X_{23}$ | 51 | 1.9546 |
| 9-IIIA | $X_5X_6X_{22}X_{24}$ | 51 | 1.9994 |
| 11-IIIA | $X_5X_{22}X_{23}X_{24}$ | 51 | 1.5848 |
| 12-IIIA | $X_6X_{22}X_{23}X_{24}$ | 51 | 2.9068** |
| 13-IIIA | $X_5X_6X_{22}$ | 52 | 2.3009* |
| 16-IIIA | $X_5X_{22}X_{23}$ | 52 | 1.6357 |
| 17-IIIA | $X_5X_{22}X_{24}$ | 52 | 1.6248 |
| 19-IIIA | $X_6X_{22}X_{23}$ | 52 | 3.0312** |
| 20-IIIA | $X_6X_{22}X_{24}$ | 52 | 2.9724** |
| 22-IIIA | $X_{22}X_{23}X_{24}$ | 52 | 2.5994* |
| 24-IIIA | $X_5X_{22}$ | 53 | 1.8954 |
| 27-IIIA | $X_6X_{22}$ | 53 | 3.4531** |
| 30-IIIA | $X_{22}X_{23}$ | 53 | 2.7210** |
| 31-IIIA | $X_{22}X_{24}$ | 53 | 2.6048* |
| 36-IIIA | $X_{22}$ | 54 | 3.0546* |

*Indicates $p > .05$.

**Indicates $p > .01$.

Table 59. Group III. Summary of calculated t-values for various multiple regression models used to investigate the effect of interest, $X_{23}$, on the dependent variable $X_2$, raw score on the objective part of the final test in German 132 - experimental group; N = 56

| Model | Independent variables included in regression model | Degrees of freedom | Calculated "t" |
|---|---|---|---|
| 1-IIIA | $X_5X_6X_{12}X_{13\text{-}16}X_{17\text{-}19}X_{22}X_{23}X_{24}$ | 42 | 0.9296 |
| 3-IIIA | $X_5X_6X_{12}X_{13\text{-}16}X_{22}X_{23}X_{24}$ | 45 | 0.7280 |
| 5-IIIA | $X_5X_6X_{12}X_{17\text{-}19}X_{22}X_{23}X_{24}$ | 46 | 0.6794 |
| 7-IIIA | $X_5X_6X_{22}X_{23}X_{24}$ | 50 | 0.1805 |
| 8-IIIA | $X_5X_6X_{22}X_{23}$ | 51 | 0.7786 |
| 10-IIIA | $X_5X_6X_{23}X_{24}$ | 51 | 0.5705 |
| 11-IIIA | $X_5X_{22}X_{23}X_{24}$ | 51 | -0.0688 |
| 12-IIIA | $X_6X_{22}X_{23}X_{24}$ | 51 | -0.0981 |
| 14-IIIA | $X_5X_6X_{23}$ | 52 | 1.3911 |
| 16-IIIA | $X_5X_{22}X_{23}$ | 52 | 0.4710 |
| 18-IIIA | $X_5X_{23}X_{24}$ | 52 | 0.8232 |
| 19-IIIA | $X_6X_{22}X_{23}$ | 52 | 0.7104 |
| 21-IIIA | $X_6X_{23}X_{24}$ | 52 | 1.3479 |
| 22-IIIA | $X_{22}X_{23}X_{24}$ | 52 | -0.3485 |
| 25-IIIA | $X_5X_{23}$ | 53 | 1.026 |
| 28-IIIA | $X_6X_{23}$ | 53 | 1.6461 |
| 30-IIIA | $X_{22}X_{23}$ | 53 | 0.4101 |
| 32-IIIA | $X_{23}X_{24}$ | 53 | 1.3437 |
| 37-IIIA | $X_{23}$ | 54 | 1.3205 |

Tabular "t" = 2.018 (.05)

Motivation, $X_{24}$

Motivation, $X_{24}$, does not explain a significant portion of the variation in the raw scores on the objective part of the final examination, $X_2$, for the experimental group when testing at the .05 level of significance, except when the model is a combination of $X_6$, course load, and $X_{24}$, motivation.

Evidence is provided by Table 60 which summarizes the results of testing the null hypothesis that $\beta_{24} = 0$ against the alternative hypothesis that $\beta_{24} \neq 0$ for the various models investigated.

Prediction of achievement in German 132 from achievement in German 131

The mean scores, mean grades, and standard deviations for the dependent and independent variables are presented in Table 61.

Product moment correlation coefficients between the dependent variables, $X_7$, score on the objective final test in German 132, $X_8$, grade on the objective final test in German 132, $X_3$, instructor grade in German 132 and the independent variables are recorded in Table 62.

The entries in Table 62 show that measures of achievement in German 132 are substantially correlated with measures of achievement in German 131 with the exception of course load, $X_5$, and method of evaluation in German 132, $X_6$.

In Table 63 the coefficients of correlation are arranged according to size.

A summary of the results of analysis of multiple regression is presented in Table 64.

Table 60. Group III. Summary of calculated t-values for various regression models used to investigate the effect of motivation, $X_{24}$, on the dependent variable, $X_2$, raw score on the objective part of the final test in German 132 - experimental group; N = 56

| Model | Independent variables included in regression model | Degrees of freedom | Calculated "t" |
|---|---|---|---|
| 1-IIIA | $X_5X_6X_{12}X_{13\text{-}16}X_{17\text{-}19}X_{22}X_{23}X_{24}$ | 42 | 0.2036 |
| 3-IIIA | $X_5X_6X_{12}X_{13\text{-}16}X_{22}X_{23}X_{24}$ | 45 | 0.3699 |
| 5-IIIA | $X_5X_6X_{12}X_{17\text{-}19}X_{22}X_{23}X_{24}$ | 46 | 0.4169 |
| 7-IIIA | $X_5X_6X_{22}X_{23}X_{24}$ | 50 | 0.7440 |
| 9-IIIA | $X_5X_6X_{22}X_{24}$ | 51 | 1.0694 |
| 10-IIIA | $X_5X_6X_{23}X_{24}$ | 51 | 0.8323 |
| 11-IIIA | $X_5X_{22}X_{23}X_{24}$ | 51 | 0.7458 |
| 12-IIIA | $X_6X_{22}X_{23}X_{24}$ | 51 | 1.1347 |
| 15-IIIA | $X_5X_6X_{24}$ | 52 | 1.5235 |
| 17-IIIA | $X_5X_{22}X_{24}$ | 52 | 0.8869 |
| 18-IIIA | $X_5X_{23}X_{24}$ | 52 | 0.8232 |
| 20-IIIA | $X_6X_{22}X_{24}$ | 52 | 1.3490 |
| 21-IIIA | $X_6X_{23}X_{24}$ | 52 | 1.3479 |
| 22-IIIA | $X_{22}X_{23}X_{24}$ | 52 | 1.1450 |
| 26-IIIA | $X_5X_{24}$ | 53 | 1.2934 |
| 29-IIIA | $X_6X_{24}$ | 53 | 2.1063* |
| 31-IIIA | $X_{22}X_{24}$ | 53 | 1.1754 |
| 32-IIIA | $X_{23}X_{24}$ | 53 | 1.3437 |
| 38-IIIA | $X_{24}$ | 54 | 1.8990 |

Tabular "t" = 2.01 approximately with 53 degrees of freedom (.05)

Table 61. Group III. Prediction of achievement in German 132 - mean scores, mean grades and standard deviations; N = 173

| Academic quarter | Variable | Mean | Standard deviation |
|---|---|---|---|
| Fall 1970 | $X_1$ = Score on objective final test, German 131 | 71.17 | 10.25 |
| | $X_2$ = Grade on objective final test, German 131 | 2.17 | 1.07 |
| | $X_3$ = Instructor grade, German 131 | 2.87 | .93 |
| | $X_4$ = Cumulative grade-point average, end of Fall Quarter 1970 | 2.89 | .57 |
| | $X_5$ = Course load, Fall Quarter 1970 | 16.08 | 2.65 |
| Winter 1971 | $X_7$ = Score on objective final test, German 132 | 105.29 | 18.26 |
| | $X_8$ = Grade on objective final test, German 132 | 1.73 | 1.23 |
| | $X_9$ = Instructor grade, German 132 | 2.28 | 1.23 |

Note: $X_6$ = method of evaluation in German 132 was coded as follows for the computer regression program: 1 = experimental group; 2 = control group

Table 62. Group III. Product-moment correlation coefficients between measures of achievement in German 132 and the prediction variables; N = 173

| Variables | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $X_1$ | 1.000 | | | | | | | | |
| $X_2$ | .926 | 1.000 | | | | | | | |
| $X_3$ | .681 | .715 | 1.000 | | | | | | |
| $X_4$ | .496 | .477 | .585 | 1.000 | | | | | |
| $X_5$ | .013 | .016 | .004 | .041 | 1.000 | | | | |
| $X_6$ | -.012 | -.019 | .021 | -.040 | .108 | 1.000 | | | |
| $X_7$ | .614 | .629 | .578 | .534 | .081 | -.109 | 1.000 | | |
| $X_8$ | .604 | .634 | .583 | .516 | .035 | -.233 | .946 | 1.000 | |
| $X_9$ | .586 | .631 | .613 | .461 | .110 | -.011 | .817 | .804 | 1.000 |

where:

$X_1$ = score on objective final test, German 131

$X_2$ = grade on objective final test, German 131

$X_3$ = instructor grade, German 131

$X_4$ = cumulative grade-point average, end of Fall Quarter 1970

$X_5$ = course load, Fall Quarter 1970

$X_6$ = method of evaluation, German 132

$X_7$ = score on objective final test, German 132

$X_8$ = grade on objective final test, German 132

$X_9$ = instructor grade, German 132

Table 63. Group III. Product-moment correlation coefficients for predictor and dependent variables arranged according to size; N = 173

| Variables | | r |
|---|---|---|
| $X_7X_8$ | Objective score, 132 - objective grade, 132 | .946 |
| $X_1X_2$ | Objective score, 131 - objective grade, 131 | .926 |
| $X_7X_9$ | Objective score, 132 - instructor grade, 132 | .817 |
| $X_8X_9$ | Objective grade, 132 - instructor grade, 132 | .804 |
| $X_2X_3$ | Objective grade, 131 - instructor grade, 131 | .715 |
| $X_1X_3$ | Objective score, 131 - instructor grade, 131 | .681 |
| $X_2X_8$ | Objective grade, 131 - objective grade, 132 | .634 |
| $X_2X_9$ | Objective grade, 131 - instructor grade, 132 | .631 |
| $X_2X_7$ | Objective grade, 131 - objective score, 132 | .629 |
| $X_1X_7$ | Objective score, 131 - objective score, 132 | .614 |
| $X_3X_9$ | Instructor grade, 131 - instructor grade, 132 | .613 |
| $X_1X_8$ | Objective score, 131 - objective grade, 132 | .604 |
| $X_1X_9$ | Objective score, 131 - instructor grade, 132 | .586 |
| $X_3X_4$ | Instructor grade, 131 - cumulative grade-point average, 131 | .585 |
| $X_3X_7$ | Instructor grade, 131 - objective score, 132 | .578 |
| $X_4X_7$ | Cumulative grade-point average, 131 - objective score, 132 | .534 |

Table 63 (continued)

| Variables | | r |
|---|---|---|
| $X_4X_8$ | Cumulative grade-point average, 131 - objective grade, 132 | .516 |
| $X_1X_4$ | Objective score, 131 - cumulative grade-point average, 131 | .496 |
| $X_2X_4$ | Objective grade, 131 - cumulative grade-point average, 131 | .477 |
| $X_4X_9$ | Cumulative grade-point average, 131 - instructor grade, 132 | .461 |
| $X_6X_8$ | Method of evaluation, 132 - objective grade, 132 | -.233 |
| $X_5X_9$ | Course load, 131 - instructor grade, 132 | .110 |
| $X_6X_7$ | Method of evaluation, 132 - objective score, 132 | -.109 |
| $X_5X_6$ | Course load, Fall 1970 - method of evaluation, 132 | .108 |
| $X_5X_7$ | Course load, Fall 1970 - objective score, 132 | .081 |
| $X_4X_5$ | Cumulative grade-point average, 131 - course load, 131 | .041 |
| $X_4X_6$ | Cumulative grade-point average, 131 - method of evaluation, 132 | -.040 |
| $X_5X_8$ | Course load, 131 - objective grade, 132 | .035 |
| $X_3X_6$ | Instructor grade, 131 - method of evaluation, 132 | .021 |
| $X_2X_6$ | Objective grade, 131 - method of evaluation, 132 | -.019 |
| $X_2X_5$ | Objective grade, 131 - course load, 131 | .016 |
| $X_1X_5$ | Objective score, 131 - course load, 131 | .013 |
| $X_1X_6$ | Objective score, 131 - method of evaluation, 132 | -.012 |
| $X_6X_9$ | Method of evaluation, 132 - instructor grade, 132 | -.011 |
| $X_3X_5$ | Instructor grade, 131 - course load, 131 | .004 |

Table 64. Group III. Prediction of achievement in German 132 - summary of analyses of multiple regression; N = 173

| Model | Dependent variable | Independent variable | b-coefficient | Calculated t .01 | Calculated t .05 | Calculated t p > .05 | F-ratio (4 and 168 d.f.) | Multiple $R^2$ |
|---|---|---|---|---|---|---|---|---|
| 1 | $X_7$ | $X_1$ | 0.829 | 7.131 | | | | |
| | | $X_4$ | 9.559 | 4.537 | | | | |
| | | $X_5$ | 0.509 | | | 1.297 | | |
| | | $X_6$ | -3.858 | | | -1.739 | 35.948** | 0.461 |
| 2 | $X_8$ | $X_1$ | 0.560 | 7.219 | | | | |
| | | $X_4$ | 0.598 | 4.253 | | | | |
| | | $X_5$ | 0.019 | | | 0.735 | | |
| | | $X_6$ | -0.580 | -3.919 | | | 38.023** | 0.475 |
| 3 | $X_9$ | $X_1$ | 0.057 | 6.844 | | | | |
| | | $X_4$ | 0.482 | 3.189 | | | | |
| | | $X_5$ | 0.045 | | | 1.592 | | |
| | | $X_6$ | -0.019 | | | -0.116 | 26.977** | 0.391 |
| 4 | $X_7$ | $X_2$ | 8.295 | 7.666 | | | | |
| | | $X_4$ | 9.551 | 4.673 | | | | |
| | | $X_5$ | 0.493 | | | 1.279 | | |
| | | $X_6$ | -3.720 | | | -1.707 | 38.766** | 0.4799 |
| 5 | $X_8$ | $X_2$ | 0.582 | 8.173 | | | | |
| | | $X_4$ | 0.579 | 4.303 | | | | |
| | | $X_5$ | 0.018 | | | 0.718 | | |
| | | $X_6$ | -0.571 | -3.983 | | | 43.360** | 0.508 |
| 6 | $X_9$ | $X_2$ | 0.614 | 8.106 | | | | |
| | | $X_4$ | 0.443 | 3.095 | | | | |
| | | $X_5$ | 0.044 | | | 1.622 | | |
| | | $X_6$ | -0.009 | | | -0.060 | 33.033** | 0.440 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 7 | $X_7$ | $X_3$ | 8.148 | 5.659 | | | |
| | | $X_4$ | 9.173 | 3.883 | | | |
| | | $X_5$ | 0.553 | | 1.345 | | |
| | | $X_6$ | -4.463 | | 1.920 | 29.241** | 0.411 |
| 8 | $X_8$ | $X_3$ | 0.592 | 6.247 | | | |
| | | $X_4$ | 0.532 | 3.425 | | | |
| | | $X_5$ | 0.023 | | 0.836 | | |
| | | $X_6$ | -0.625 | -4.087 | | 33.264** | 0.442 |
| 9 | $X_9$ | $X_3$ | 0.699 | 7.158 | | | |
| | | $X_4$ | 0.322 | 2.008 | | | |
| | | $X_5$ | 0.049 | | 1.759 | | |
| | | $X_6$ | -0.073 | | -0.464 | 28.389** | 0.403 |

Dependent variables:

$X_7$ = objective final score, 132
$X_8$ = objective final grade, 132
$X_9$ = instructor grade, 132

Independent variables:

$X_1$ = objective final score, 131
$X_2$ = objective final grade, 131
$X_3$ = instructor grade, 131
$X_4$ = grade point average, 131
$X_5$ = course load, 131
$X_6$ = method of evaluation, 132

---

**Tabulated F with 4 and 168 degrees of freedom = 3.91 (.01)

As evident from Table 64, the F-ratios for all 9 regression models are significant at the .01 level with 4 and 168 degrees of freedom. This means that each model could be used to predict achievement in German 132 from achievement measures in German 131, after adjusting for grade-point average, $X_4$, course load, $X_5$, and method of evaluation, $X_6$.

Further inspection of Table 64 indicates that the t-values for course load, $X_5$, are not significant; that is, they do not explain a significant portion of the variability in either of the achievement measures in German, i.e., in objective final score, $X_7$, in objective final grade, $X_8$, or in instructor grade, $X_9$. Therefore, course load can be eliminated from the prediction scheme without significant loss of predicting ability.

Method of evaluation, $X_6$, explains a significant portion of the variability in the dependent variable in models 2 and 8, where this variable is the grade on the objective final test in German 132, $X_8$, and where the combination of independent variables includes the score on the objective final test in German 131, $X_2$, or instructor grade in German 131, $X_3$.

On the basis of the calculated t-values, it can be concluded that

(1) There is sufficient evidence to reject null hypothesis 11 that achievement in German 131 as measured by the score obtained on the objective part of the final examination, $X_1$, is not related to achievement in German 132.

(2) There is evidence to reject the null hypothesis 12 that achievement in German 131 as measured by the grade obtained on the objective part of the final examination, $X_2$, is not related to the objective grade earned in German 132.

(3) There is evidence to reject the null hypothesis 13 that the instructor grade received in German 131, $X_3$, is not related to the instructor grade received in German 132.

(4) There is evidence to reject the null hypothesis 14 that cumulative grade-point average, $X_4$, as per Fall Quarter 1970, is not related to achievement in German 132.

(5) There is no evidence to reject the null hypothesis 15 that course load, $X_5$, in the Fall Quarter 1970 is not related to achievement in German 132.

(6) There is no evidence to reject the null hypothesis 16 that method of evaluation, $X_6$, is not related to achievement in German 132 if the measure of achievement is raw score on the objective part of the final examination, $X_7$, or instructor grade, $X_9$.

(7) There is evidence to reject the null hypothesis 16 that method of evaluation, $X_6$, is not related to achievement in German 132 if the measure of achievement is objective final grade, $X_8$.

Following are the regression equations for each of the dependent variables:

(1) $\hat{X}_7 = .8287X_1 + 9.5587X_4 + .5089X_5 - 3.8582X_6 + 16.9709$

(2) $\hat{X}_8 = .0560X_1 + .5982X_4 + .0193X_5 - .5804X_6 - 3.3236$

(3) $\hat{X}_9 = .0571X_1 + .4821X_4 + .0448X_5 - .0185X_6 - 3.8612$

(4) $\hat{X}_7 = 8.2948X_2 + 9.5508X_4 + .4930X_5 - 3.7208X_6 + 58.0221$

(5) $\hat{X}_8 = .5819X_2 + .5785X_4 + .0182X_5 - .5711X_6 - .5406$

(6) $\hat{X}_9 = .6142X_2 + .4429X_4 + .0438X_5 - .0092X_6 - 1.0168$

(7) $\hat{X}_7 = 8.1482X_3 + 9.1732X_4 + .5526X_5 - 4.4634X_6 + 54.0200$

(8) $\hat{X}_8 = .5919X_3 + .5324X_4 + .0226X_5 - .6251X_6 - .8228$

(9) $\hat{X}_9 = .6999X_3 + .3220X_4 + .0491X_5 - .0732X_6 - 1.3201$

On the basis of the data in Table 64, it can be assumed that the best combination of variables to predict the grade on the objective part of the final examination in German 132 is the linear multiple regression model No. 5 which includes the objective grade on the final test, $X_2$, grade-point average, $X_4$, course load, $X_5$, and method of evaluation, $X_6$. Since the t-value for course load is not significant, this variable could be omitted from the model without any significant loss in prediction ability.

## Discussion

### Method of evaluation

Evaluation of achievement based on self-pacing was applied to four sections of German 132 (N = 56) during the Winter Quarter 1971. The control group (N = 117) was subjected to a uniform test schedule. Only the final examination was administered to both groups at the same time. The criteria of achievement in this investigation were raw score and grade on the objective part of the final examination in German 132.

Visual inspection of Figures 8 and 9 indicates that the experimental group had relatively more A's, B's and C's and fewer D's and F's than the control group, but a wider range in raw score within the low grades.

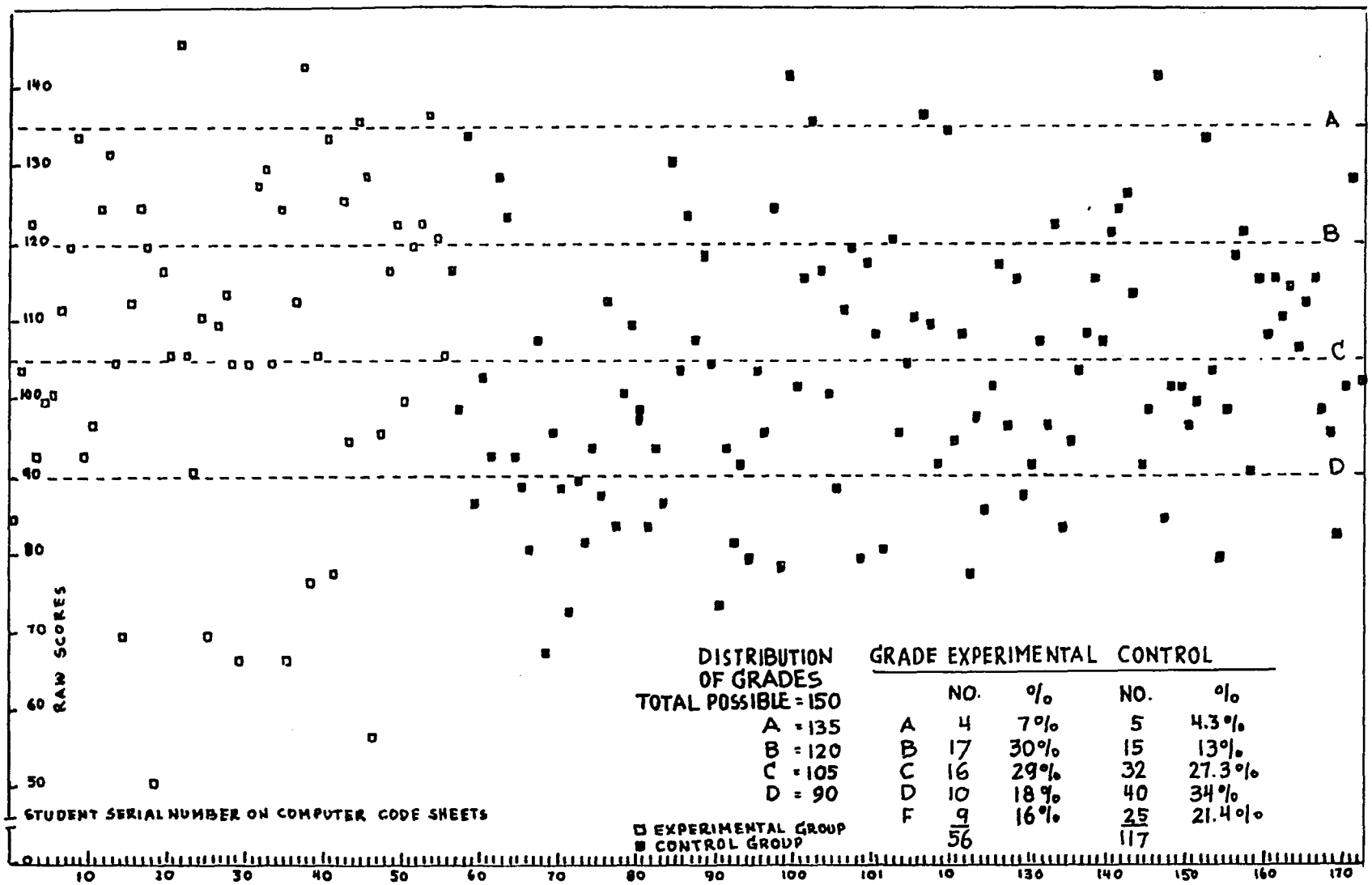



FIGURE 8. GROUP III. SCATTER DIAGRAM OF RAW SCORES ON OBJECTIVE FINAL TEST - GERMAN 132

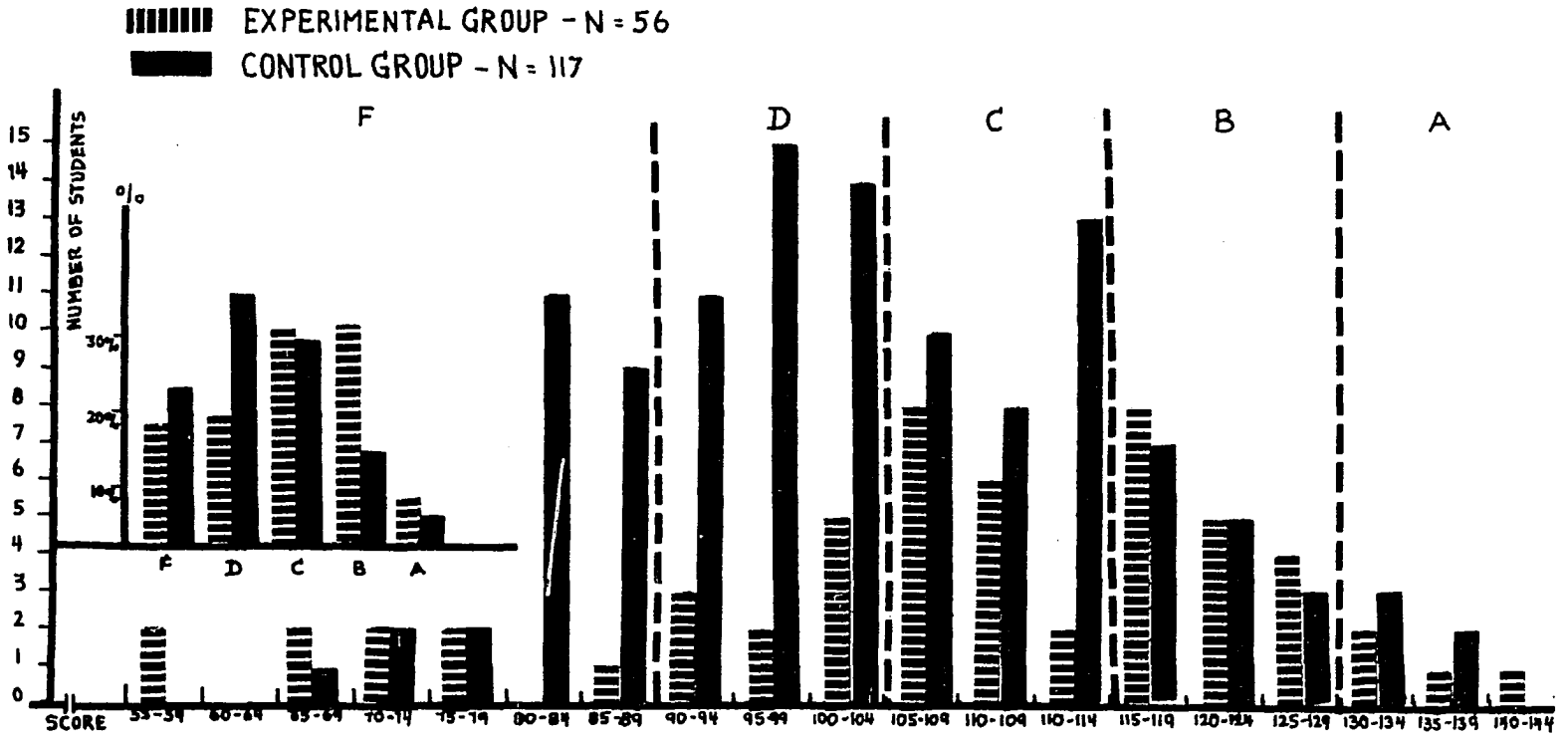



FIGURE 9. GROUP III: FREQUENCY DISTRIBUTION - NUMBER OF STUDENTS BY RAW SCORE ON OBJECTIVE PART OF THE FINAL EXAMINATION IN GERMAN 132

An explanation for the cluster of extremely low scores, as seen in Figure 8, is that these students did not study at all, taking advantage of the experimental situation which excluded F-grades, except for those students who did not take the final examination in German 132. Under ordinary circumstances, these students, probably, would have dropped out of the course or made a greater effort.

These findings were further investigated by chi-square tests which confirmed the difference in grade distribution between the experimental and the control group. The principal difference was one of location, i.e., due to a larger percentage of A's, B's and C's in the experimental group relative to the number of D's and F's. Thus, it can be said, on the basis of the chi-square test of location, that the method of evaluation based on self-pacing, which in fact was a method of teaching, attained its goal, namely, a more successful foreign language learning experience for a larger number of students.

These findings seem to confirm the results of 40 major studies cited by Block (11) that strategies of mastery learning and self-pacing "enable about three-fourth of students to learn to the same performance standards as the top fourth of students learning under conventional, group-based instructional approaches." He concludes that "individual differences need not condition student learning," and that "individual differences have largely been used as a scapegoat for ineffective instruction."

The question whether these results were really due to method of evaluation, or perhaps, to other factors was investigated by an F-test of equal variance, and by analysis of multiple regression. The F-test was not significant, and analysis of multiple regression provided evi-

dence that method of evaluation explained a significant portion of the variation in the grades on the objective part of the final examination, but not in the raw scores obtained on this test.

These, seemingly, contradictory results could be explained by the fact that $X_2$, raw score on the objective part of the final examination, was more variable than $X_3$, the grade based on this raw score. Thus, the influence of $X_1$, method of evaluation, was obscured by large residual variation in $X_2$.

The negative t-values for $X_1$, method of evaluation (Table 49) indicate that the experimental group received both higher scores and grades than the control group, even if the difference in the mean raw score was statistically not significant.

Since in this investigation the four experimental sections of elementary German were taught by the same instructor, the experiment with mastery learning and evaluation based on self-pacing should be followed by further research involving different instructors. Also a survey of student opinion about this approach to language learning, as compared to the traditional lock step system, is indicated.

It can be concluded that this experiment, even within the limits of the academic quarter system, resulted in better study habits and in a more relaxed and optimistic attitude toward language learning. After initial reluctance, the great majority of students adjusted quickly to the new pass-fail system, where passing was equivalent to the mastery of precisely defined learning tasks.

Instructors

In German 131 (Group II) and in German 132 (Group III), the investigation of differences in achievement between the experimental and the control groups was based on objective measures, that is, on the raw score obtained on the objective part of the final examination taken by all the students enrolled in these two elementary courses. The question whether it was the instructor or the method which caused differences in achievement between the treatment and the control groups is a merely rhetorical one, since only one instructor was conducting experiments in German 131 and 132 during the Fall Quarter 1970 and during the Winter Quarter 1971. To find out whether the experimental method of evaluation would lead to the same results, independently of instructor, will require further research involving several instructors working with this method.

Language aptitude

Language aptitude, as defined in the Pimsleur Language Aptitude Battery (PLAB, 76), explained a significant portion of the variation in raw scores on the objective part of the final examination in both German 131 and 132. For German 131 the total score was used, for German 132 only the composite raw score on parts 3, 4, 5, and 6. Language aptitude did not explain differences in achievement on the unit quizzes or in the total sum of scores, which seem to be more a matter of motivation than of language aptitude. The raw score on the objective part of the final examinations in German 131 and 132, as already mentioned, reflected the ability to reason in terms of the grammatical and idiomatic structure of the foreign language.

Since achievement of the experimental group on the objective part of the final examination was significantly correlated with language aptitude as measured by the PLAB ($r = .38$ when $N = 56$), this test, though intended for grades 7 through 12, is appropriate for use with college students to predict success in foreign language study.

Interest and motivation

Interest was measured by a partial score on the Pimsleur Language Aptitude Battery, where this variable is defined as an estimate of how interested the student is in studying a foreign language, how much he will enjoy it, and how interested he is in foreign language as compared with other subjects.

Interest as an independent variable, independently, or after controlling for other variables, did not explain a significant portion of the variation in raw score on the objective part of the final examination when testing at the .05 level of significance. This seems to confirm findings by Carroll (20) that "motivational differences will not make much difference in achievement."

However, motivation, estimated from a four-point scale, described under Methods and Procedures, explained a significant portion of the variation in raw score, but only when combined with course load. The negative correlations of both interest and motivation with course load, though not significant, indicate that as course load increased, interest and motivation decreased in the experimental group.

Course load

Contrary to Group II, where course load was correlated significantly with both dependent and independent variables, except for method of evaluation and language aptitude, course load in connection with Group III did not show significant correlations with any of the variables investigated for the experimental or the control groups. For the combined group, the correlations of course load with the various criteria of achievement, i.e., raw score, grade based on raw score, and instructor grade, were even negative. This means that for this group higher course loads were associated with lower scores and grades. This result is opposed to the findings for Group I where higher course loads were associated with higher grades. Since these correlations were not significant, the findings are probably due to sampling variation. Also, contrary to the findings for Group I, the correlation between grade-point average and course load was not significantly different from zero ($r = .028$ when $N = 56$; $r = .007$ when $N = 117$; $r = .013$ when $N = 173$). Thus, the presence of grade-point average in the model did not reduce the t-values for course load like in Group I. On the contrary, when both grade-point average and language aptitude were eliminated, the t-values for course load became insignificant.

These findings, though seemingly contradictory, probably can be explained by the fact that in the experimental group, course load, due to the experimental method, i.e., evaluation of achievement on the basis of self-pacing, did not interfere with achievement; in other words, achievement had become independent of course load. Grade-point average and course load were not correlated in the experimental situation of

Group III because grade-point average independently of other variables explained a significant portion of the variation in the dependent variable while course load did not. However, in connection with Group I, course load, when used independently of other variables, still was significant at the .01 level, as evident from Table 19.

Course load explained a significant portion of the variation in raw score obtained on the objective part of the final examination in German 132 only if year in college was removed from any of the regression models and if language aptitude was included (Table 51). Year in college did not influence achievement in the experimental group, but it is correlated significantly with course load. Year in college was significant at the .05 level only for the combined group (N = 173), as evident from Tables 54 and 55.

In the reduced models, that is, in those models from which both year in college and college were omitted, course load is insignificant only if the model does not contain language aptitude, or when both language aptitude and grade-point average are eliminated. An explanation for this is the substantial correlation of both language aptitude ($r = .384$) and grade-point average ($r = .576$) with the dependent variable, i.e., raw score on the objective part of the final examination in German 132. Logically, it can be assumed that language aptitude is a more pertinent factor in language learning than grade-point average. Therefore, language aptitude in combination with course load still explains a significant portion of the variation in raw score while, as evident from Table 51, the combination of grade-point average and course load does not.

Year in college

Year in college, the less precise equivalent of cumulative credit hours, investigated as an independent variable in connection with Group III, explained a significant portion of the variation in raw scores on the objective part of the final examination in German 132 for the combined group (N = 173), but not for the experimental group.

Possible reasons for the difference in findings may be due to differences in models. The experimental model included language aptitude, interest, and motivation as independent variables which the combined group did not. The combined group included the instructor variable, method of evaluation, and two additional colleges. Also different numbers of observations were involved. The larger group would detect a smaller difference as significant on the average. Another explanation may be that because of the experimental method of evaluation, individual differences in achievement associated with the status as freshmen, sophomores, juniors, or seniors, were reduced.

Sex

In connection with Group III, sex does not explain a significant portion of the variation in raw score obtained on the objective part of the final examination in German 132. The t-values were not significant and negative, which would imply that the coeds received lower scores than the male students. However, the number of females involved (N = 17) in the experimental group was too small to lead to valid conclusions.

In view of the significant findings in connection with Group I, where the sample size was sufficient to provide valid evidence, the results of the analysis of Group III must be considered inconclusive be-

cause of the small numbers involved and, therefore, the possibility of sampling error.

Prediction of achievement in German 132

Inspection of Table 61 indicated that for the 173 students involved there was a decline in achievement, as measured by the grade on the objective part of the final examination, as well as in instructor grade from a B-average to a C-average, accompanied by increased variance among students as indicated by the larger standard deviations for German 132. These findings confirmed the results of the preliminary t-tests for the control group; for the experimental group, however, the decline in achievement proved to be insignificant.

In spite of these differences in performance between German 131 and German 132, achievement in both groups is significantly correlated, as evident from the product-moment correlation matrix in Table 62.

The highest correlation between dependent and independent variables occurred between objective final grades in German 131 and German 132 ($r = .634$). With sample size $N = 173$, the probability that a correlation coefficient as large as this would occur by chance is less than one in a hundred (36).

The grade on the objective part of the final test in German 131 is also significantly correlated with the raw score on the objective part of the final test, and with the instructor grade in German 132 ($r = 629$; $r = .631$).

Thus, there was sufficient evidence to reject the null hypothesis that achievement in German 132 is not related to achievement in German 131.

The best combination of variables to predict the grade on the objective part of the final test in German 132 included the objective grade in German 131, grade-point average, method of evaluation in German 132, and course load. Because of its insignificant t-value in any combination, course load could be omitted. The contradictory findings concerning method of evaluation which explained a significant portion of the variability in the objective grade, but not in raw score, had been explained by a chi-square test of location which established that the difference between the two distributions is one of location; that is, in the experimental group, the proportion of higher grades was larger in relation to the proportion of low grades. This fact also was confirmed by the negative t-values of method of evaluation (the control group was coded: 2).

Grade-point average, as usual, was highly significant in explaining the variation in all measures of achievement in German 132. It was least significant in connection with instructor grade, which means that some other factors contributed more to that grade.

It can be concluded that the objective part of the final examination in German 131 was a reliable predictor of achievement in German 132. The Kuder-Richardson reliability estimates for this test ranged from .77 to .94, with an average of .87 for the 13 sections of German 131.

## SUMMARY AND RECOMMENDATIONS

The purpose of this investigation was to test the hypothesis that grading policies, here defined as methods of evaluation, are a factor influencing student achievement and attitude in elementary foreign language courses.

This hypothesis was based on the theory that language learning is a function of several factors, namely, of the academic environment, of cumulative mastery, of time, and of attitude.

It was postulated that, since the academic environment presents the student with an overwhelming amount of competing subject matter to be absorbed within a relatively short time, teaching and evaluation methods should be designed to facilitate the learning process. It was assumed that this could be accomplished by facing the student with precisely defined learning tasks which can be mastered within the time available, while allowing for individual difference in the amount of time necessary to achieve the specific objectives within the limits of an academic term.

In connection with this theory, the impact of the following methods of evaluation upon student achievement in elementary foreign language courses was investigated:

(1) The pass-fail system versus the letter-grade system.

(2) Method of evaluation based on a cumulative point system which takes into account all the tests given during the academic term versus evaluation based on the results of the final examination only.

(3) Traditional lock step, i.e., evaluation according to a uniform test schedule for all the students, versus evalua-

tion based on self-pacing within the academic quarter system.

The investigation included three groups of Iowa State University students enrolled in elementary foreign language courses during the Spring and Fall Quarters 1970 and during the Winter Quarter 1971.

Group I involved 305 pass-fail students and 590 letter grade students enrolled in 52 sections of elementary French, German, Russian, and Spanish, taught by 25 instructors.

Group II consisted of 264 students enrolled in 13 sections of German 131, the first course of the elementary sequence, taught by 5 instructors. The experimental group consisted of 4 sections of German 131 with a total of 72 students taught by the same instructor.

Group III involved 173 students enrolled in 12 sections of German 132, the second course in the elementary sequence, taught by 5 instructors. For this group, the objective score obtained on the final examination in German 131 was available to predict achievement in German 132. The experimental group consisted of 4 sections with a total of 56 students taught by the same instructor.

The statistical analysis of the various groups was divided into three major parts, each concerned with a different approach to evaluation of achievement. The specific objectives connected with each part were stated in hypothesis form. These hypotheses were tested using linear additive multiple regression models.

The following variables, assumed to affect achievement in elementary foreign language courses, were investigated: 1) cumulative grade-point average (Groups I, II, III); 2) course load (Groups I, II,

III); 3) method of evaluation (Groups I, II, III); 4) sex (Groups I, II); 5) cumulative credit hours (Group I); 6) year in college (Group III); 7) college (Groups I, III); 8) language (Group I); 9) instructor (Groups I, III); 10) ACT composite score (Group I); 11) language aptitude (Groups II, III); 12) interest (Group III); 13) motivation (Group III); 14) course sequence (Group I).

Achievement in elementary foreign language courses was measured by the following criteria:

(1) Instructor grades based on a variety of evaluation methods (Groups I, II, and III).

(2) Grades based on the raw score obtained on the objective part of the final examination in German 131 and 132 (Groups II and III).

(3) Grades based on the total sum of unit quiz scores, on the total score obtained on the final examination, and on the total sum of scores comprising the points accumulated on all examinations taken during the Fall Quarter 1970 (Group II).

Preliminary investigations of differences between the various groups due to method of evaluation or to instructor used special cases of multiple regression, namely, t-tests and analysis of variance, and the technique of chi-square where other variables are ignored.

The statistical evidence justified the following conclusions:

(1) There was a definite difference in achievement between students enrolled under the pass-fail system and those enrolled under the letter-grade system. The distributions of grades according to grade-point

average indicate that students with grade-point averages ranging from 3.75 to 4.00 received 35 percent A's in pass-fail courses as compared to 87 percent in courses taken under the letter grade system.

(2) The method of evaluation emphasizing final achievement did not penalize students for initial failure or the unsatisfactory results of a mid-term test. In other words, students whose grades were based on the results of the final examination only, received better grades than students whose grades were calculated by taking into consideration all the tests and quizzes given during the academic term.

(3) Evaluation of achievement based on student self-pacing combined with mastery learning provided a more successful language learning experience to a larger number of students than the application of a uniform test schedule. These results confirmed previous research (11).

(4) There was sufficient evidence to reject the null hypothesis of no difference in achievement in elementary foreign language courses among students evaluated by different methods.

(5) There was a difference in achievement among students taught by different instructors, but no difference within the sections taught by the same instructor.

(6) Grade-point average, again, proved to be the best single predictor of success in foreign language study. Its correlations with the various measures of achievement ranged from $r = .49$ to $r = .71$.

(7) Less stable, but still significant factors in achievement, were the ACT composite score, language aptitude as measured by the Pimsleur Language Aptitude Battery, the language studied, sex, cumulative credit hours, and year in college. According to these results, the ACT

score is a useful predictor if grade-point average is not available. Language aptitude was significantly correlated with the results on the objective part of the final examinations in German 131 ($r = .43$) and in German 132 ($r = .38$) which measured the ability to reason in terms of the grammatical and idiomatic structure of the foreign language. Females belonging to Group I received better grades than male students which confirmed research at the secondary school level (20).

(8) Course load was positively correlated with final achievement ($r = .17$) for a sample of 895 students which means that the better students carried higher course loads.

(9) Interest did not explain a significant portion in the variation of achievement for Group III. This is a confirmation of previous research (20).

(10) College as an independent variable was insignificant.

(11) Motivation, only in combination with course load, explained a significant portion of the variation in achievement in elementary German. The correlation of motivation with course load was negative which means that as course load increased, motivation decreased for a sample of 72 students.

(12) Achievement on the objective part of the final examination in German 131 was a reliable predictor of achievement in German 132 ($r = .63$).

Considering the evidence found, it is recommended:

a. that elementary foreign language courses, or any subject involving cumulative mastery, not be made available under the pass-fail system with a D-passing level;

b. that more emphasis in evaluating student achievement be given to the comprehensive final examination instead of weighting all the quizzes and tests given during the academic term;

c. that, in order to advance research in foreign language teaching, and to improve the learning situation for the students in elementary foreign language courses, instructors of elementary course sequences agree as to common precisely defined learning tasks, also known as behavioral goals, and use uniform methods of evaluation while being given complete freedom of how to achieve the objectives; and,

d. that the possibilities inherent in evaluation based on self-pacing be further investigated.

BIBLIOGRAPHY

1. Airasian, Peter W. An application of a modified version of John Carroll's model of school learning. Unpublished Master's thesis. Chicago, Ill., Library, University of Chicago. 1967.

2. Airasian, Peter W. The role of evaluation in mastery learning. In Block, James, H., ed. Mastery learning. Pp. 77-88. New York, Holt, Rinehart and Winston, Inc. 1971.

3. Anderson, Scarvia B. Estimating grade reliability. Journal of Applied Psychology 37:461-464. 1953.

4. Babbott, Edward F. Effect of high school pass/fail courses on college admissions. National ACAC Journal 14:25-27. 1969.

5. Banathy, Bela H. Current trends in college curriculum: a systems approach. Objectives and learning tasks. In Birkmaier, Emma M., ed. Britannica Review of Foreign Language Education. Vol. 1. Pp. 114-155. Chicago, Encyclopaedia Britannica, Inc. 1968.

6. Bending, A. W. The reliability of letter grades. Educational and Psychological Measurement 13:311-321. 1953.

7. Biehler, Robert F. A first attempt at a 'learning for mastery' approach. Educational Psychologist 7:7-9. 1970.

8. Birkmaier, Emma Marie, ed. Britannica review of foreign language education. Vol. 1. Chicago, Encyclopaedia Britannica, Inc. 1968.

9. Block, James H. The effects of various levels of performance on selected cognitive, affective, and time variables. Unpublished Ph.D. dissertation. Chicago, Ill., Library, University of Chicago. 1970.

10. Block, James H. Introduction to mastery learning: theory and practice. In Block, James H., ed. Mastery learning. Pp. 2-12. New York, Holt, Rinehart and Winston, Inc. 1971.

11. Block, James H., ed. Mastery learning. New York, Holt, Rinehart and Winston, Inc. 1971.

12. Block, James H. Operating procedures for mastery learning. In Block, James H., ed. Mastery learning. Pp. 64-76. New York, Holt, Rinehart and Winston, Inc. 1971.

13. Bloom, Benjamin S. Affective consequences of school achievement. In Block, James H., ed. Mastery learning. Pp. 13-28. New York, Holt, Rinehart and Winston, Inc. 1971.

14. Bloom, Benjamin S. Mastery learning. In Block, James H., ed. Mastery learning. Pp. 47-63. New York, Holt, Rinehart and Winston, Inc. 1971.

15. Borland, David T. Effects of midterm grades on the academic performance of college freshmen. College and University 45:181-185. 1970.

16. Bramer, George R. Grading and academic justice. Improving College and University Teaching 18:63-65. 1970.

17. Breazier, Eldon and Aven, Samuel D. Pass-fail credit for student teachers. School and Community 55:12. 1969.

18. Cady, J. W. How important are college grades? Journal of Higher Education 37:441-445. 1966.

19. Carroll, John B. The prediction of success in intensive foreign language training. In Glaser, Robert, ed. Training research and education. Pp. 87-136. Pittsburgh, University of Pittsburgh Press. 1962.

20. Carroll, John B. Problems of measurement related to the concept of learning for mastery. Educational Horizons 48:71-80. 1970.

21. Carroll, John B. Research on teaching foreign languages. In Gage, N. L., ed. Handbook of research on teaching. Pp. 1060-1100. Chicago, Rand McNally. 1963.

22. Carroll, John B. and Sapon, S.M. Manual, modern language aptitude test. (1959 ed.). New York, Psychological Corporation. 1959.

23. Chatfield, Walter L. and Phillips, Dennis R. Foreign language study without failure. Mount Vernon, Iowa, Iowa Foreign Language Bulletin Vol. 9, No. 3. April, 1971.

24. Clark, Edward C. Reliability of grade-point average. Journal of Educational Research 57:428-430. 1964.

25. College grading systems. A journal symposium. Journal of Higher Education 35:89-103. 1964.

26. Conklin, Kenneth R. Educational evaluation and intuition. Educational Forum 34:323-332. 1970.

27. Creamer, Don G. An evaluation of mid-semester deficiency grading practices. ED 034 543/JC 690 433 ERIC. 1969.

28. Dale, Wesley J. Concerning grading and other forms of student evaluation. Washington, D.C., Council of Graduate Schools in the U.S. 1969.

29. DePue, Palmer. On a uniform standard for marking. The Journal of Experimental Education 33:231-235. 1965.

30. Downing, Majorie. Variations in college grading systems. In College grading systems: A journal symposium. Pp. 98-100. Journal of Higher Education 35:89-103. 1964.

31. Dressel, Paul L. and Nelson, Clarence H. Evaluation in higher education. Boston, Houghton Mifflin Co. 1961.

32. Dykstra, John W. We reluctant appraisers. Improving College and University Teaching 18:61-62. 1970.

33. Edling, Jack V. Educational objectives and educational media. Review of Educational Research 38:177-189. 1968.

34. Elbow, Peter H. More accurate evaluation of student performance. Journal of Higher Education 40:219-230. 1969.

35. Fearing, Percy. Nongraded foreign language classes. (ERIC Focus Reports on the Teaching of Foreign Languages #4) American Council on the Teaching of Foreign Languages. New York, N.Y., Modern Language Association of America. 1969.

36. Froehlich, Clifford P. and Hoyt, Kenneth B. Guidance testing. Chicago, Ill., Science Research Associates, Inc. 1959.

37. Gardner, M. E. Study of subject loads and marks. Clearing House 35: 145-146. 1960.

38. Gardner, Robert C. and Lambert, Wallace E. Motivation variables in second-language acquisition. Canadian Journal of Psychology 13: 266-272. 1959.

39. Gronlund, Norman E. Stating behavioral objectives for classroom instruction. London, The Macmillan Company. 1970.

40. Haskell, John M. Pass-fail? A system worth trying. The Clearing House 42:172-173. 1967.

41. Horst, Paul. A generalized expression for the reliability of measures. Psychometrika 14:21-31. 1949.

42. Hoyt, C. Test reliability obtained by analysis of variance. Psychometrika 6:153-160. 1941.

43. Hyman, Lawrence W. Grades and academic standards. New York, City University of New York, Brooklyn College. 1969.

44. Ingram, John Allen. Factors affecting the success of transfer students at Drake University. Unpublished Ph.D. dissertation, Iowa State University. 1967.

45. Iowa State University Bulletin. The Chart. Student Handbook 1969-1971.

46. Iowa State University. Student Counseling Service. Prediction of academic performance at Iowa State University. Student Counseling Report 69-06.

47. Johnson, Icie F. Injustice of grades. School and Community 54: 24-25. 1967.

48. Johnson, Jack T. Grading options: do they pass or fail? College and University Business 49:77-78. 1970.

49. Johnson, Ray A. and Lauroesch, William. Pass-fail system with option for the learner. American Vocational Journal 44:16-18. 1969.

50. Kanun, Clara. A comparison of grade point averages and credit loads. ERIC. ca. 1969.

51. Kelly, Eldon G. Study of consistent discrepancies between instructor grades and term end examination grades. Journal of Educational Psychology 49:328-334. 1958.

52. Kersh, Mildred E. A strategy for mastery learning in fifth-grade arithmetic. Unpublished Ph.D. dissertation, University of Chicago. 1970.

53. Kim, Hogwon, et al. A study of Bloom strategies for mastery learning. Seoul: Korean Institute for Research in the Behavioral Sciences. Abstract in Block, James H., ed., Mastery learning. Pp. 123-124. New York, Holt, Rinehart and Winston, Inc. 1971.

54. Lambert, Wallace E. Motivation and language learning: psychological aspects. In Dimension: Language 68. Proceedings of the Fourth Southern Conference on Language Teaching, New Orleans, February, 1968. Spartanburg, S.C., Southern Conference on Language Teaching, Converse College. 1968.

55. Lange, Dale L., ed. Britannica review of foreign language education. Vol. 2. Chicago, Encyclopaedia Britannica, Inc. 1970.

56. Latimer, John F. Motivation for the Study of Latin. In Newell, Sanford, ed. Dimension: Language 68. Proceedings of the Fourth Southern Conference on Language Teaching, New Orleans, February, 1968. Spartanburg, S.C., Southern Conference on Language Teaching, Converse College. 1968.

57. Law School Admission Test Council. Mimeographed statement on pass-fail grading system as endorsed by the Council of the Section of Legal Education and Admissions to the Bar of the American Bar Association. 1970.

58. Mannello, George. College teaching without grades. Journal of Higher Education 35:328-334. 1969.

59. Marshall, Max. This thing called evaluation. Educational Forum 23: 41-53. 1958.

60. Marshall, Max. To grade or not to grade. Journal of Higher Education 23:264-273. 1952.

61. Maxson, Wilbur B. Grading, a serious matter. NEA Journal 53: 56-57. 1964.

62. McQueen, Mildred. Motivation to learn. Chicago, Ill., Science Research Associates, Inc. 1966.

63. Melville, George L. The pass-fail system and the change in the accounting of grades on comprehensive examinations at Knox College. Office of Institutional Research. 1967.

64. Modu, Christopher C. Affective consequences of cognitive changes. Unpublished Ph.D. dissertation. Chicago, Ill., Library, University of Chicago. 1969.

65. Moulton, William G. Study hints for language students. Boston, Massachusetts, Houghton Mifflin. 1970/71.

66. Mueller, Ronald H. Is self-grading the answer? The Journal of Higher Education 41:221-224. 1970.

67. Mulholland, Walter E. Are academic failures preventable? Educational Forum 31:79-81. 1966.

68. Munday, Leo. Predicting college grades using ACT data. Education and Psychology 27:401-406. 1967.

69. Newell, Sanford, ed. Dimension: Language 68. Proceedings of the Fourth Southern Conference on Language Teaching, New Orleans, February, 1968. Spartanburg, S.C., Southern Conference on Language Teaching, Converse College. 1968.

70. Nida, Eugene A. Motivation in second language learning. Language Learning 7:11-16. 1956-57.

71. Oberteuffer, Margaret. College reactions to a pass-fail threat. The Journal of the National Association of College Admissions Counselors 15:8-9. 1970.

72. Ortega y Gasset, José. Mission of the university. New York, W. W. Norton and Company, Inc. 1966.

73. Ostle, Bernard. Statistics in research. Ames, Iowa, Iowa State University Press. 1963.

74. Perry, Louis B. College grading: a case study and its aftermath. Educational Record 49:78-84. 1968.

75. Phi Beta Kappa. Pass-Fail Study Committee Report to the Senate. Mimeographed report. December 6, 1969.

76. Pimsleur, Paul. Pimsleur language aptitude battery. New York, Harcourt, Brace and World, Inc. 1966.

77. Pimsleur, Paul, Mosberg, Ludwig and Morrison, Andrew L. Student factors in foreign language learning: a review of literature. Modern Language Journal 46:160-170. 1962.

78. Pimsleur, Paul, Sundland, Donald M., and McIntyre, Ruth D. Underachievement in foreign language learning. New York, Modern Language Association of America. 1966.

79. Politzer, Robert L. Linguistics and motivation. In Newell, Sanford, ed. Dimension: Language 68. Proceedings of the Fourth Southern Conference on Language Teaching, New Orleans, February, 1968. Spartanburg, S.C., Sourthern Conference on Language Teaching, Converse College. 1968.

80. Politzer, Robert L. and Weiss, Louis. Improving achievement in foreign language. A Series in Applied Linguistics. Vol. 4. Philadelphia, Pa., The Center for Curriculum Development, Inc. ca. 1970.

81. Politzer, Robert L. Student motivation and interest in elementary language courses. Language Learning 5:15-21. 1953/54.

82. Politzer, Robert L. and Dwight, Allen W. Flexible scheduling and foreign language instruction. The Modern Language Journal 51: 275-281. 1967.

83. Politzer, Robert L. and Weiss, Louis. The successful foreign-language teacher. A Series in Applied Linguistics. Vol. 3. Philadelphia, Pa., The Center for Curriculum Development, Inc.

84. Quann, Charles J. Survey shows variations in grading trends. College and University Business 49:78-79.

85. Reichmann, Eberhard, ed. The teaching of German. Philadelphia, Penn., National Carl Schurz Association, Inc. 1970.

86. Rogers, H. H. The reliability of college grades. School and Society 45:758-760. 1937.

87. Roscoe, John T. Fundamental research statistics. New York, Holt, Rinehart and Winston, Inc. 1969.

88. Ross, Nancy L. Pass-fail systems trouble Phi Beta Kappa selection. Des Moines Tribune, December 12, 1970.

89. Scales, Eldridge E. Variability in grading practices of instructors of a multiple section natural science course. Science Education 50: 332-335. 1966.

90. Scherer, George A. C. Toward more effective individualized learning. In Reichmann, Eberhard, ed. The Teaching of German. Pp. 112-114. Philadelphia, Pa., National Carl Schurz Association, Inc. 1970.

91. Schulz, Dora, Griesbach, Heinz and von Hofe, Harold. Deutsche Sprachlehre für Amerikaner. Zweite Ausgabe. New York, Charles Scribner's Sons. 1970.

92. Seawall, Frank. A quality grading system. College and University 43: 47-51. 1967.

93. Semas, Philip W. Two-thirds of nation's campuses are using some form of 'pass-fail' grading system. The Chronicle of Higher Education February 15, 1971.

94. Sgan, Mathew R. First year of pass-fail at Branders University. Journal of Higher Education 40:135-144. 1969.

95. Simon, Sidney B. Grades must go. School Review 78:397-402. 1970.

96. Simpson, C. Computer handles a unique pass-fail system. College and University 44:700-701. 1969.

97. Smith, Lawrence G. Non-punitive grading in California junior colleges. ED 035409. 1969.

98. Snedecor, George W. Statistical methods. Ames, Iowa, Iowa State College Press. 1956.

99. Sparks, David S. Grading and student evaluation. Council of Graduate Schools in the U.S., Washington, D.C. December 5, 1969.

100. Spolsky, Bernard. Attitudinal aspects of second language learning. Language Learning 19:271-283.

101. Stallings, William M. and Leslie, Elwood K. Student attitudes towards grades and grading. Improving College and University Teaching 18:66-68. 1970.

102. Steiner, Florence. Performance objectives in the teaching of foreign languages. Foreign Language Annals 3:579-591. 1970.

103. Steiner, Florence. Behavioral objectives and evaluation. In Lange, Dale L., ed., Britannica Review of Foreign Language Education. Vol. 2. Pp. 33-79. Chicago, Encyclopaedia Britannica, Inc. 1970.

104. Stroup, Francis. Comparison of scales used by state and privately endowed universities in recording academic achievement. College and University 41:180-184. 1966.

105. Stroup, Francis. The grade-point average is obsolete. Journal of Higher Education 34:10-15. 1963.

106. Thompson, Wayne N. A study of the grading practices of thirty-one instructors in freshman English. Journal of Educational Research 49:65-68. 1955.

107. Thorshen, Kay. The relation of classroom evaluation to students' self-concepts and mental health. Unpublished Ph.D. dissertation. Chicago, Ill., Library, University of Chicago. 1969.

108. Trow, William Clark. Grades and objectives in higher education. Educational Record 49:85-91. 1968.

109. Valette, Rebecca M. Directions in foreign language testing. New York, Modern Language Association. ca. 1970.

110. Valette, Rebecca M. Testing and motivation. In Newell, Sanford, ed. Dimension: Language 68. Proceedings of the Fourth Southern Conference on Language Teaching, New Orleans, February, 1968. Spartanburg, S.C., Southern Conference on Language Teaching, Converse College. 1968.

111. Valette, Rebecca M. Testing. In Birkmaier, Emma M., ed. Britannica Review of Foreign Language Education. Vol. 1. Pp. 343-374. Chicago, Encyclopaedia Britannica, Inc. 1958.

112. von Wittich, Barbara. Prediction of success in foreign language study. Modern Language Journal 46:208-212. 1962.

113. Walker, Helen M. and Lev, Joseph. Statistical inference. New York, Henry Holt and Company. 1953.

114. Washburne, C. W. Educational measurements as a key to individualizing instruction and promotions. Journal of Educational Research 5: 195-206. 1922.

115. Weiss, Peter, Wertheimer, Michael and Groesbeck, Byron. Achievement motivation, academic aptitude and college grades. Education and Psychology 19:663-666. 1959.

116. Wert, James E., Neidt, Charles O. and Ahmann, J. Stanley. Statistical methods in educational and psychological research. New York, Appleton-Century-Crofts, Inc. 1954.

117. Wright, William. Achievement as a function of time: an analysis of selected Stanford achievement test battery results. Unpublished manuscript, University of Chicago, Department of Education. Abstract in Block, James H., ed. Mastery learning. Pp. 144-145. New York, Holt, Rinehart and Winston, Inc. 1971.

118. Yelon, Stephen L. An alternative to letter grades. The Educational Forum 35:65-70.

ACKNOWLEDGEMENTS

First, I wish to thank Dr. Ray Bryan and Dr. Trevor Howe for their guidance in organizing and editing the dissertation. I also thank all my Committee members for their time and advice, and Dr. Jean Laffoon and Dr. Bernard Skold for screening the manuscript.

To Dr. Carol E. Fuchs I am deeply indebted for her professional help with the statistical treatment of the data and for the generous amount of time she has given to this task.

Special thanks go to Dr. Leroy Wolins for his valuable advice in clarifying some of the statistical problems, and to Dr. Trevor Howe and Dr. Anton Netusil for their explanations.

I am indebted to Dr. Arthur Gowan, to Dr. Trevor Howe, to Dr. Fred Schlunz, to Walter Chatfield, and to Charlotte Whitney for source materials for the study, and to John Sjoblom for facilitating the gathering of data in the Office of Admissions and Records.

The methodological research in the area of foreign language teaching would have been impossible without Dr. Walter Morris having opened the possibilities for such research in the Foreign Language Department at Iowa State University.

# APPENDIX A: ANALYSIS OF VARIANCE AND PREDICTION EQUATION COEFFICIENTS

Table 65. Group I. Analysis of variance (N = 750) and prediction equation coefficients (N = 750)

| Model | Sum of squares d.f. | Regression | d.f. | Residual | Mean square Regression | Residual | F ratio | Multiple $R^2$ |
|---|---|---|---|---|---|---|---|---|
| Total sum of squares = 849.48 with 749 degrees of freedom | | | | | | | | |
| Analysis of variance: | | | | | | | | |
| 1-IA | 14 | 334.47 | 735 | 515.01 | 23.89 | .70 | 34.10 | .39 |
| 1a-IA | 13 | 332.30 | 736 | 517.18 | 25.56 | .70 | 36.38 | .39 |
| 2-IA | 11 | 287.65 | 738 | 561.83 | 26.15 | .76 | 34.35 | .34 |
| 2a-IA | 10 | 328.89 | 739 | 520.59 | 32.89 | .70 | 1.99 | - |
| 3-IA | 10 | 286.94 | 739 | 562.54 | 28.69 | .76 | 16.96 | - |
| 4-IA | 13 | 334.44 | 736 | 515.04 | 25.73 | .69 | 0.04 | - |
| 5-IA | 6 | 276.23 | 743 | 573.25 | 46.04 | .77 | 59.67 | .33 |
| 6-IA | 5 | 276.13 | 744 | 573.35 | 55.23 | .77 | 71.66 | .33 |
| 7-IA | 5 | 276.12 | 744 | 573.36 | 55.22 | .77 | 71.66 | .33 |
| 8-IA | 5 | 274.02 | 744 | 575.46 | 54.80 | .77 | 70.85 | .32 |
| 9-IA | 4 | 276.02 | 745 | 573.46 | 69.00 | .77 | 89.64 | .32 |
| 10-IA | 3 | 273.87 | 746 | 575.61 | 91.29 | .77 | 118.31 | .32 |
| 11-IA | 3 | 98.72 | 746 | 750.76 | 32.91 | 1.01 | 32.70 | .12 |
| 12-IA | 3 | 99.13 | 746 | 750.35 | 33.04 | 1.01 | 32.85 | .12 |
| 13-IA | 3 | 118.52 | 746 | 730.95 | 39.51 | .98 | 40.32 | .14 |
| 14-IA | 3 | 208.98 | 746 | 640.50 | 69.66 | .86 | 81.14 | .25 |
| 15-IA | 3 | 267.99 | 746 | 581.49 | 89.33 | .78 | 114.60 | .32 |
| 16-IA | 3 | 208.23 | 746 | 641.25 | 69.41 | .86 | 80.75 | .25 |
| 17-IA | 3 | 111.97 | 746 | 737.51 | 37.32 | .99 | 37.75 | .13 |
| 18-IA | 3 | 46.99 | 746 | 802.49 | 15.66 | 1.08 | 14.56 | .06 |
| 19-IA | 3 | 91.06 | 746 | 758.42 | 30.35 | 1.02 | 29.86 | .11 |
| 20-IA | 2 | 98.71 | 747 | 750.77 | 49.35 | 1.01 | 49.10 | .12 |
| 21-IA | 2 | 207.79 | 747 | 641.69 | 103.89 | .86 | 120.94 | .24 |
| 22-IA | 2 | 46.87 | 747 | 802.61 | 23.43 | 1.07 | 21.81 | .06 |
| 23-IA | 2 | 90.82 | 747 | 758.66 | 45.41 | 1.02 | 44.71 | .11 |
| 24-IA | 2 | 30.15 | 747 | 819.33 | 15.08 | 1.10 | 13.74 | .04 |

Prediction equation coefficients; dependent variable $X_1$ = course grade:

| Model | $b_o$ | $b_2$ | $b_3$ | $b_4$ | $b_5$ | $b_6$ | $b_7$ | $b_8$ | $b_9$ | $b_{10}$ | $b_{11}$ | $b_{12}$ | $b_{13}$ | $b_{14}$ | $b_{15}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1-IA | .22 | -.65 | .13 | .01 | .00 | .02 | .00 | -.41 | -.33 | -.65 | -.30 | -.51 | -.58 | -.36 | -.01 |
| 1a-IA | .42 | -.66 | | .01 | .00 | .02 | .00 | -.48 | -.44 | -.74 | -.30 | -.50 | -.58 | -.37 | -.01 |
| 2-IA | .41 | -.59 | .10 | .01 | -.01 | .02 | -.00 | -.51 | -.51 | -.62 | -.37 | | | | -.07 |
| 2a-IA | -.17 | -.65 | .18 | .01 | .00 | .02 | .00 | | | | | -.51 | -.59 | -.36 | -.03 |
| 3-IA | .48 | -.58 | .10 | .01 | -.01 | .02 | -.00 | -.48 | -.49 | -.61 | -.36 | | | | |
| 4-IA | .26 | -.65 | .13 | .01 | | .02 | .00 | -.40 | -.32 | -.64 | -.30 | -.51 | -.58 | -.36 | -.01 |
| 5-IA | .14 | -.56 | | .01 | -.02 | .02 | -.00 | | | | | | | | .03 |
| 6-IA | .18 | -.55 | | .01 | -.02 | .02 | -.00 | | | | | | | | |
| 7-IA | .15 | -.63 | | .01 | -.02 | .02 | | | | | | | | | .03 |
| 8-IA | .35 | -.57 | | .01 | -.02 | | -.00 | | | .03 | | | | | |
| 9-IA | .18 | -.63 | | .01 | -.02 | .02 | | | | | | | | | |
| 10-IA | .38 | -.62 | | .01 | -.02 | | | | | | | | | | |
| 11-IA | 1.24 | -.65 | | | .06 | | .00 | | | | | | | | |
| 12-IA | 1.17 | -.64 | | | .06 | | | | | | | | | | .05 |
| 13-IA | .58 | -.65 | | | .05 | .06 | | | | | | | | | |
| 14-IA | -.06 | | | .01 | -.01 | .01 | | | | | | | | | |
| 15-IA | .36 | | | .01 | -.02 | | -.00 | | | | | | | | |
| 16-IA | .16 | | | .01 | -.01 | | | | | | | | | | .06 |
| 17-IA | .54 | | | | .05 | .06 | -.00 | | | | | | | | |
| 18-IA | .38 | | | | .05 | .05 | | | | | | | | | -.03 |
| 19-IA | 1.16 | | | | .06 | | .00 | | | | | | | | .04 |
| 20-IA | 1.24 | -.63 | | | .06 | | | | | | | | | | |
| 21-IA | .09 | | | .01 | .01 | | | | | | | | | | |
| 22-IA | .34 | | | | .05 | .05 | | | | | | | | | |
| 23-IA | 1.21 | | | | .06 | | -.00 | | | | | | | | |
| 24-IA | .99 | | | | .06 | | | | | | | | | | -.03 |

Independent variables:

$X_2$ = grading system
$X_3$ = sex
$X_4$ = cumulative grade-point average
$X_5$ = ACT composite score
$X_6$ = course load
$X_7$ = cumulative credit hours
$X_{8-11}$ = college
$X_{12-14}$ = language
$X_{15}$ = course sequence

Table 66. Group I. Analysis of variance (N = 895) and prediction equation coefficients (N = 895)

| Model | Sum of squares d.f. | Regression | d.f. | Residual | Mean square Regression | Residual | F ratio | Multiple $R^2$ |
|---|---|---|---|---|---|---|---|---|
| Total sum of squares = 1018.64 with 894 degrees of freedom Analysis of variance: | | | | | | | | |
| 1-IB | 13 | 400.68 | 881 | 617.97 | 30.82 | .70 | 43.94 | .39 |
| 2-IB | 5 | 342.94 | 889 | 675.70 | 68.59 | .76 | 90.24 | .34 |
| 3-IB | 3 | 342.94 | 891 | 675.70 | 114.31 | .76 | 150.73 | .34 |
| 4-IB | 3 | 341.33 | 891 | 677.31 | 113.78 | .76 | 149.67 | .34 |
| 5-IB | 3 | 341.33 | 891 | 677.31 | 113.78 | .76 | 149.67 | .34 |
| 6-IB | 3 | 111.73 | 891 | 906.91 | 37.24 | 1.02 | 36.59 | .11 |
| 7-IB | 3 | 91.46 | 891 | 927.19 | 30.49 | 1.04 | 29.30 | .09 |
| 8-IB | 3 | 334.04 | 891 | 684.60 | 111.35 | .77 | 144.91 | .33 |
| 9-IB | 3 | 266.68 | 891 | 751.96 | 88.89 | .84 | 105.33 | .26 |
| 10-IB | 3 | 332.17 | 891 | 686.47 | 110.72 | .77 | 143.71 | .33 |
| 11-IB | 3 | 98.88 | 891 | 919.76 | 32.96 | 1.03 | 31.93 | .10 |
| 12-IB | 2 | 341.32 | 892 | 677.32 | 170.66 | .76 | 224.76 | .34 |
| 13-IB | 2 | 110.47 | 892 | 908.17 | 55.24 | 1.02 | 54.25 | .11 |
| 14-IB | 2 | 89.94 | 892 | 928.70 | 44.97 | 1.04 | 43.19 | .09 |
| 15-IB | 2 | 91.10 | 892 | 927.54 | 45.55 | 1.04 | 43.80 | .09 |
| 16-IB | 2 | 266.02 | 892 | 752.63 | 133.01 | .84 | 157.64 | .26 |
| 17-IB | 2 | 332.16 | 892 | 686.48 | 166.08 | .77 | 215.80 | .33 |
| 18-IB | 2 | 266.21 | 892 | 752.43 | 133.10 | .84 | 157.79 | .26 |
| 19-IB | 2 | 97.99 | 892 | 920.65 | 48.99 | 1.03 | 47.47 | .10 |
| 20-IB | 2 | 16.23 | 892 | 1002.41 | 8.12 | 1.12 | 7.22 | .02 |
| 21-IB | 2 | 77.02 | 892 | 941.62 | 38.51 | 1.06 | 36.48 | .08 |

Prediction equation coefficients; dependent variable $X_1$ = final grade:

| Model | $b_o$ | $b_2$ | $b_3$ | $b_4$ | $b_6$ | $b_7$ | $b_8$ | $b_9$ | $b_{10}$ | $b_{11}$ | $b_{12}$ | $b_{13}$ | $b_{14}$ | $b_{15}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1-IB | .35 | -.64 | .09 | .01 | .02 | .00 | -.39 | -.26 | -.33 | -.28 | -.46 | -.49 | -.33 | -.01 |
| 2-IB | -.10 | -.62 | | .01 | .01 | .00 | | | | | | | | .00 |
| 3-IB | -.10 | -.62 | | .01 | .01 | | | | | | | | | |
| 4-IB | .08 | -.63 | | .01 | | .00 | | | | | | | | |
| 5-IB | .07 | -.62 | | .01 | | | | | | | | | | .00 |
| 6-IB | 1.83 | -.69 | | | .05 | | | | | | | | | .09 |
| 7-IB | 2.60 | -.79 | | | | .00 | | | | | | | | .09 |
| 8-IB | -.16 | | | .01 | .02 | -.00 | | | | | | | | |
| 9-IB | -.20 | | | .01 | .01 | | | | | | | | | -.06 |
| 10-IB | .05 | | | .01 | | -.00 | | | | | | | | -.01 |
| 11-IB | 1.80 | | | | .05 | -.00 | | | | | | | | .07 |
| 12-IB | .08 | | | .01 | | | | | | | | | | |
| 13-IB | 1.97 | -.69 | | | .05 | | | | | | | | | |
| 14-IB | 2.76 | -.78 | | | | .00 | | | | | | | | |
| 15-IB | 2.61 | -.68 | | | | | | | | | | | | .09 |
| 16-IB | -.29 | | | .01 | .01 | | | | | | | | | |
| 17-IB | .03 | | | .01 | | -.00 | | | | | | | | |
| 18-IB | -.10 | | | .01 | | | | | | | | | | -.06 |
| 19-IB | 1.92 | | | | .05 | -.00 | | | | | | | | |
| 20-IB | 1.81 | | | | .04 | | | | | | | | | .01 |
| 21-IB | 2.60 | | | | | -.00 | | | | | | | | .08 |

Table 67a. Group III. Analysis of variance (N = 173)

| Dependent variable | Model | Sum of squares d.f. | Regression | d.f. | Residual | Mean square Regression | Residual | F ratio | Multiple $R^2$ |
|---|---|---|---|---|---|---|---|---|---|
| | | Total sum of squares = 57319.97 with 172 degrees of freedom | | | | | | | |
| $X_2$ | 1-IIIC | 16 | 25313.57 | 156 | 32006.40 | 1582.10 | 205.17 | 7.71 | .44 |
| | 2-IIIC | 13 | 23900.92 | 159 | 33419.05 | 1838.53 | 210.18 | 8.74 | .42 |
| | 3-IIIC | 12 | 22595.38 | 160 | 34724.58 | 1882.95 | 217.03 | 8.68 | .39 |
| | 4-IIIC | 11 | 24246.89 | 161 | 33073.08 | 2204.26 | 205.42 | 10.73 | .42 |
| | 5-IIIC | 1 | 680.10 | 171 | 56639.86 | 680.10 | 331.23 | 2.05 | .01 |
| | | Total sum of squares = 262.23 with 172 degrees of freedom | | | | | | | |
| $X_3$ | 6-IIIC | 16 | 117.60 | 156 | 144.63 | 7.35 | .93 | 7.93 | .45 |
| | 7-IIIC | 13 | 109.07 | 159 | 153.16 | 8.39 | .96 | 8.71 | .42 |
| | 8-IIIC | 11 | 112.33 | 161 | 149.90 | 10.21 | .93 | 10.97 | .43 |
| | 9-IIIC | 1 | 14.23 | 171 | 248.00 | 14.23 | 1.45 | 9.81 | .05 |

Table 67b. Group III. Prediction equation coefficients (N = 173)

| Model | $b_o$ | $b_1$ | $b_5$ | $b_6$ | $b_7$ | $b_9$ | $b_{10}$ | $b_{12}$ |
|---|---|---|---|---|---|---|---|---|
| | | | Dependent variable $X_2$ = raw score: | | | | | |
| 1-IIIC | 27.63 | -.08 | .22 | -.05 | -7.91 | .64 | .43 | 1.59 |
| 2-IIIC | | -1.92 | .22 | .07 | | | | .44 |
| 3-IIIC | 41.78 | .21 | .19 | .44 | -7.57 | -1.38 | -2.58 | 3.19 |
| 4-IIIC | 31.53 | -.11 | .21 | .01 | -7.17 | .78 | -1.22 | .89 |
| 5-IIIC | 112.40 | -4.24 | | | | | | |
| | | | Dependent variable $X_3$ = grade based on raw score: | | | | | |
| 6-IIIC | -2.69 | -.29 | .01 | -.02 | -.66 | -.03 | .03 | .25 |
| 7-IIIC | -2.36 | -.48 | .01 | -.01 | | | | .16 |
| 8-IIIC | -2.19 | -.30 | .01 | -.02 | -.60 | -.02 | -.08 | .18 |
| 9-IIIC | 2.76 | -.61 | | | | | | |

| $b_{13}$ | $b_{14}$ | $b_{15}$ | $b_{16}$ | $b_{17}$ | $b_{18}$ | $b_{19}$ | $b_{20}$ | $b_{21}$ |
|---|---|---|---|---|---|---|---|---|
| 20.75 | 12.44 | 11.26 | 20.76 | 1.27 | -.87 | 2.38 | .71 | -33.20 |
| 19.94 | 11.87 | 10.77 | 16.64 | -1.25 | | | | |
| | | | | -.36 | -1.19 | 4.84 | -1.99 | -25.17 |
| 19.41 | 11.01 | 9.99 | 16.97 | | | | | |
| 1.23 | .77 | .67 | 1.34 | .33 | .30 | .66 | .24 | -1.94 |
| 1.16 | .72 | .62 | 1.01 | .15 | .14 | .49 | .23 | -1.59 |
| 1.12 | .64 | .54 | 1.09 | | | | | |

Table 68. Group II. Prediction equation coefficients (N = 72) and analysis of variance (N = 72)

| Model | Dependent variable | Prediction equation coefficients: $b_o$ | $b_1$ | $b_8$ | $b_9$ | $b_{10}$ |
|---|---|---|---|---|---|---|
| 1-II | $X_2$ | -42.62 | -46.87 | 88.26 | 2.59 | 0.79 |
| 2-II | $X_3$ | 4.54 | 1.81 | 10.51 | 0.61 | 0.26 |
| 3-II | $X_4$ | -33.33 | 6.28 | 45.77 | 5.21 | 0.54 |
| 4-II | $X_5$ | -3.71 | 0.07 | 1.42 | 0.09 | 0.01 |
| 5-II | $X_6$ | -82.91 | -42.27 | 142.04 | 7.12 | 1.27 |
| 6-II | $X_7$ | -4.52 | -0.28 | 1.41 | 0.05 | 0.02 |

Dependent variables:

$X_2$ = total unit quiz score
$X_3$ = objective final test score
$X_4$ = total final test score
$X_5$ = course grade based on $X_4$
$X_6$ = total sum of scores
$X_7$ = course grade based on $X_6$

Analysis of variance:

| Sum of squares | | | Mean square | | F ratio | Multiple $R^2$ |
|---|---|---|---|---|---|---|
| Total | Regression | Residual | Regression | Residual | | |
| (Degrees of freedom: total = 71; regression = 4; residual = 67) | | | | | | |
| 4976078.75 | 286163.76 | 211444.11 | 71540.94 | 3155.88 | 22.67 | .58 |
| 90591.11 | 4483.22 | 4575.89 | 1120.80 | 68.30 | 16.41 | .50 |
| 200548.88 | 84696.17 | 115852.70 | 21174.04 | 1729.14 | 12.25 | .42 |
| 129.50 | 62.11 | 67.39 | 15.52 | 1.01 | 15.44 | .48 |
| 1150797.28 | 694092.26 | 456705.02 | 173523.07 | 6816.49 | 25.46 | .60 |
| 121.65 | 67.69 | 53.96 | 16.92 | .81 | 21.01 | .56 |

Independent variables:

$X_1$ = method of evaluation
$X_8$ = grade-point average
$X_9$ = course load
$X_{10}$ = language aptitude

Table 69. Group III. Analysis of variance (N = 56) and prediction equation coefficients (N = 56)

| Model | Sum of squares d.f. | Regression | d.f. | Residual | Mean square Regression | Residual | F ratio | Multiple $R^2$ |
|---|---|---|---|---|---|---|---|---|
| Total sum of squares = 25539.55 with 55 degrees of freedom | | | | | | | | |
| Analysis of variance: | | | | | | | | |
| 1-IIIA | 13 | 13872.81 | 42 | 11666.74 | 1067.14 | 277.78 | 3.84 | .54 |
| 2-IIIA | 10 | 11295.56 | 45 | 14243.99 | 1129.56 | 316.53 | 3.57 | .44 |
| 3-IIIA | 10 | 13276.21 | 45 | 12263.35 | 1327.62 | 272.52 | 4.87 | .52 |
| 4-IIIA | 6 | 9892.14 | 49 | 15647.41 | 1648.69 | 319.33 | 5.16 | .39 |
| 5-IIIA | 9 | 12316.03 | 46 | 13223.53 | 1368.45 | 287.47 | 4.76 | .48 |
| 6-IIIA | 7 | 11175.05 | 48 | 14364.50 | 1596.44 | 299.26 | 5.34 | .44 |
| 7-IIIA | 5 | 11375.22 | 50 | 14164.34 | 2275.04 | 283.29 | 8.03 | .45 |
| 8-IIIA | 4 | 11218.39 | 51 | 14321.17 | 2804.60 | 280.81 | 9.99 | .44 |
| 9-IIIA | 4 | 11365.99 | 51 | 14173.57 | 2841.50 | 277.91 | 10.22 | .45 |
| 10-IIIA | 4 | 10351.89 | 51 | 15187.66 | 2587.97 | 297.80 | 8.69 | .41 |
| 11-IIIA | 4 | 9802.64 | 51 | 15736.91 | 2450.66 | 308.57 | 7.94 | .38 |
| 12-IIIA | 4 | 6399.71 | 51 | 19139.85 | 1599.93 | 375.29 | 4.26 | .25 |
| 13-IIIA | 3 | 11048.14 | 52 | 14491.42 | 3682.71 | 278.68 | 13.21 | .43 |
| 14-IIIA | 3 | 10145.62 | 52 | 15393.94 | 3381.87 | 296.04 | 11.42 | .40 |
| 15-IIIA | 3 | 10254.98 | 52 | 15284.58 | 3418.33 | 293.93 | 11.63 | .40 |
| 16-IIIA | 3 | 9630.99 | 52 | 15908.56 | 3210.33 | 305.93 | 10.49 | .38 |
| 17-IIIA | 3 | 9801.18 | 52 | 15738.37 | 3267.06 | 302.66 | 10.79 | .38 |
| 18-IIIA | 3 | 9027.65 | 52 | 16511.90 | 3009.22 | 317.54 | 9.48 | .35 |
| 19-IIIA | 3 | 5916.53 | 52 | 19623.02 | 1972.18 | 377.37 | 5.23 | .23 |
| 20-IIIA | 3 | 6396.10 | 52 | 19143.46 | 2132.03 | 368.14 | 5.79 | .25 |
| 21-IIIA | 3 | 3228.61 | 52 | 22310.94 | 1076.20 | 429.06 | 2.51 | .13 |
| 22-IIIA | 3 | 4365.36 | 52 | 21174.19 | 1455.12 | 407.20 | 3.57 | .17 |
| 23-IIIA | 2 | 9572.72 | 53 | 15966.83 | 4786.36 | 301.26 | 15.89 | .38 |
| 24-IIIA | 2 | 9563.11 | 53 | 15976.45 | 4781.55 | 301.44 | 15.86 | .37 |
| 25-IIIA | 2 | 8812.45 | 53 | 16727.11 | 4406.22 | 315.61 | 13.96 | .35 |
| 26-IIIA | 2 | 9002.14 | 53 | 16537.41 | 4501.07 | 312.03 | 14.43 | .35 |
| 27-IIIA | 2 | 5726.11 | 53 | 19813.45 | 2863.05 | 373.84 | 7.65 | .22 |
| 28-IIIA | 2 | 2449.12 | 53 | 23090.43 | 1224.56 | 435.67 | 2.81 | .10 |
| 29-IIIA | 2 | 3143.40 | 53 | 2296.15 | 1571.70 | 422.57 | 3.72 | .12 |
| 30-IIIA | 2 | 3831.55 | 53 | 21708.01 | 1915.77 | 409.59 | 4.68 | .15 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 31-IIIA | 2 | 4315.91 | 53 | 21223.65 | 2157.95 | 400.45 | 5.39 | .17 |
| 32-IIIA | 2 | 1614.03 | 53 | 23925.52 | 807.02 | 451.42 | 1.79 | .06 |
| 34-IIIA | 1 | 848.02 | 54 | 17059.39 | 8480.16 | 315.91 | 26.84 | .33 |
| 35-IIIA | 1 | 1268.61 | 54 | 24270.95 | 1268.61 | 449.46 | 2.82 | .05 |
| 36-IIIA | 1 | 3762.68 | 54 | 21776.88 | 3762.68 | 403.28 | 9.33 | .15 |
| 37-IIIA | 1 | 798.94 | 54 | 24740.61 | 798.94 | 458.16 | 1.74 | .03 |
| 38-IIIA | 1 | 1598.79 | 54 | 23940.76 | 1598.79 | 443.35 | 3.61 | .06 |

Prediction equation coefficients; dependent variable = $X_2$:

| Model | $b_o$ | GPA $b_5$ | Course load $b_6$ | Sex $b_{12}$ | Year in college $b_{13}$ | $b_{14}$ | $b_{15}$ | $b_{16}$ | College $b_{17}$ | $b_{18}$ | $b_{19}$ | Language aptitude $b_{22}$ | Interest $b_{23}$ | Motivation $b_{24}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1-IIIA | -28.69 | .20 | 1.58 | -4.94 | 18.10 | 4.26 | 11.30 | 12.32 | -12.51 | -18.29 | 10.66 | .67 | 1.56 | .57 |
| 2-IIIA | 13.10 | .26 | .91 | -2.90 | 24.55 | 12.92 | 16.11 | 27.39 | -10.67 | -10.19 | -6.11 | | | |
| 3-IIIA | -35.14 | .22 | 1.34 | -4.75 | 20.60 | 7.18 | 13.55 | 23.53 | | | | .55 | 1.16 | .95 |
| 4-IIIA | 19.28 | .24 | 1.57 | -1.30 | | | | | -6.41 | -8.02 | 8.87 | | | |
| 5-IIIA | -27.36 | .20 | 2.07 | -3.36 | | | | | -7.52 | -15.41 | 15.86 | .61 | 1.15 | 1.16 |
| 6-IIIA | 2.91 | .26 | .91 | -3.35 | 24.51 | 13.29 | 15.85 | 28.62 | | | | | | |
| 7-IIIA | -2.65 | 1.98 | 1.90 | | | | | | | | | 5.15 | 2.72 | 1.92 |
| 8-IIIA | -2.75 | 2.03 | 1.91 | | | | | | | | | 5.26 | 9.40 | |
| 9-IIIA | -2.62 | 1.97 | 1.89 | | | | | | | | | 5.25 | | 2.19 |
| 10-IIIA | 4.62 | 2.25 | 1.74 | | | | | | | | | | 8.64 | 2.20 |
| 11-IIIA | 9.41 | 2.06 | | | | | | | | | | 4.45 | -1.08 | 2.01 |
| 12-IIIA | -8.20 | | 2.16 | | | | | | | | | 8.63 | -1.70 | 3.34 |
| 13-IIIA | -2.70 | 2.03 | 1.84 | | | | | | | | | 5.89 | | |
| 14-IIIA | 4.19 | 2.31 | 1.74 | | | | | | | | | | 1.65 | |
| 15-IIIA | 7.50 | 2.25 | 1.67 | | | | | | | | | | | 3.13 |
| 16-IIIA | 8.47 | 2.11 | | | | | | | | | | 4.57 | 5.89 | |
| 17-IIIA | 9.42 | 2.07 | | | | | | | | | | 4.41 | | 1.90 |
| 18-IIIA | 3.39 | 2.30 | | | | | | | | | | | 4.39 | 2.24 |
| 19-IIIA | -1.64 | | 2.18 | | | | | | | | | 8.97 | 9.94 | |
| 20-IIIA | -9.37 | | 2.17 | | | | | | | | | 8.57 | | 3.17 |

Table 69 (continued)

| Model | Sum of squares d.f. | Sum of squares Regression | Sum of squares d.f. | Sum of squares Residual | Mean square Regression | Mean square Residual | F ratio | Multiple $R^2$ |
|---|---|---|---|---|---|---|---|---|

Prediction equation coefficients; dependent variable = $X_2$ (contd):

| Model | $b_o$ | GPA $b_5$ | Course load $b_6$ | Sex $b_{12}$ | Year in college $b_{13}$ | $b_{14}$ | $b_{15}$ | $b_{16}$ | College $b_{17}$ | $b_{18}$ | $b_{19}$ | Language aptitude $b_{22}$ | Interest $b_{23}$ | Motivation $b_{24}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 21-IIIA | 6.33 | | 1.92 | | | | | | | | | | 8.10 | 4.21 |
| 22-IIIA | 4.14 | | | | | | | | | | | 8.00 | -6.25 | 3.51 |
| 23-IIIA | 1.26 | 2.39 | 1.56 | | | | | | | | | | | |
| 24-IIIA | 7.93 | 2.11 | | | | | | | | | | 4.99 | | |
| 25-IIIA | 3.35 | 2.36 | | | | | | | | | | | 1.24 | |
| 26-IIIA | 3.48 | 2.30 | | | | | | | | | | | | 2.72 |
| 27-IIIA | -9.63 | | 2.10 | | | | | | | | | | 9.64 | |
| 28-IIIA | 6.55 | | 1.94 | | | | | | | | | | 2.35 | |
| 29-IIIA | 6.59 | | 1.86 | | | | | | | | | | | 5.09 |
| 30-IIIA | 4.10 | | | | | | | | | | | 8.36 | 5.94 | |
| 31-IIIA | 4.17 | | | | | | | | | | | 7.78 | | 2.87 |
| 32-IIIA | 9.69 | | | | | | | | | | | | 3.39 | 4.31 |
| 34-IIIA | 3.76 | 2.41 | | | | | | | | | | | | |
| 35-IIIA | 8.05 | | 1.68 | | | | | | | | | | | |
| 36-IIIA | 4.05 | | | | | | | | | | | 8.78 | | |
| 37-IIIA | 9.95 | | | | | | | | | | | | 1.91 | |
| 38-IIIA | 9.76 | | | | | | | | | | | | | 4.68 |

# APPENDIX B: ANALYSIS OF VARIANCE BY SECTION, INSTRUCTOR AND LANGUAGE

Table 70. Group I. Analysis of variance by section, instructor and language - summary of data; N = 900

FRENCH

| Instr./ Section | N[a] | ΣX[b] | ΣX² | $\frac{(\Sigma X)^2}{n}$ Section | $\frac{(\Sigma X)^2}{n}$ Instructor | $\Sigma X^2 - \frac{(\Sigma X)^2}{n}$ Section | $\Sigma X^2 - \frac{(\Sigma X)^2}{n}$ Instructor | $\bar{X}$[c] |
|---|---|---|---|---|---|---|---|---|
| 1-102-1 | 24 | 49 | 117 | 100.04 | | 16.96 | | 2.04 |
| 1-102-2 | 18 | 41 | 133 | 93.39 | | 39.61 | | 2.28 |
| 1-103-3 | 16 | 42 | 136 | 110.25 | | 25.75 | | 2.63 |
| | 58 | 132 | 386 | 303.68 | 300.41 | 82.32 | 85.59 | 2.27 |
| 2-102-4 | 18 | 47 | 147 | 122.72 | | 24.28 | | 2.61 |
| 2-102-5 | 8 | 22 | 66 | 60.50 | | 5.50 | | 2.75 |
| 2-103-10 | 9 | 24 | 72 | 64.00 | | 8.00 | | 2.67 |
| | 35 | 93 | 285 | 247.22 | 247.11 | 37.78 | 37.89 | 2.66 |
| 3-102-6 | 18 | 37 | 95 | 76.06 | 76.06 | 18.94 | 18.94 | 2.06 |
| 4-103-1 | 26 | 55 | 141 | 116.35 | | 24.65 | | 2.12 |
| 4-103-2 | 21 | 37 | 85 | 65.19 | | 19.81 | | 1.76 |
| | 47 | 92 | 226 | 181.54 | 180.09 | 44.46 | 45.91 | 1.96 |
| 5-103-3 | 6 | 17 | 53 | 48.17 | 48.17 | 4.83 | 4.83 | 2.83 |
| 6-103-4 | 13 | 29 | 71 | 64.69 | | 6.31 | | 2.23 |
| 6-103-5 | 16 | 35 | 91 | 76.56 | | 14.44 | | 2.19 |
| | 29 | 64 | 162 | 141.25 | 141.24 | 20.75 | 20.76 | 2.20 |
| 7-103-7 | 15 | 33 | 81 | 72.60 | 72.60 | 8.40 | 8.40 | 2.20 |
| 8-103-8 | 18 | 41 | 105 | 93.39 | | 11.61 | | 2.28 |
| 8-103-11 | 10 | 24 | 66 | 57.60 | | 8.40 | | 2.40 |
| | 28 | 65 | 171 | 150.99 | 150.89 | 20.01 | 20.11 | 2.32 |
| 9-103-9 | 26 | 66 | 202 | 167.54 | 167.54 | 34.46 | 34.46 | 2.54 |
| Total | 262 | 599 | 1661 | 1389.05 | 1384.11 | 271.95 | 276.89 | 2.29 |

[a] N = Number of students.

[b] ΣX = sum of grade points earned by the students in one section.

[c] $\bar{X}$ = section grade-point average.

Table 70 (continued)

| Instr./ Section | N[a] | ΣX[b] | ΣX² | $\frac{(\Sigma X)^2}{n}$ Section | $\frac{(\Sigma X)^2}{n}$ Instructor | $\Sigma X^2 - \frac{(\Sigma X)^2}{n}$ Section | $\Sigma X^2 - \frac{(\Sigma X)^2}{n}$ Instructor | $\bar{X}$[c] |
|---|---|---|---|---|---|---|---|---|
| **GERMAN** | | | | | | | | |
| 10-132-1 | 23 | 46 | 120 | 92.00 | | 28.00 | | 2.00 |
| 10-133-6 | 24 | 64 | 194 | 170.67 | | 23.33 | | 2.67 |
| | 47 | 110 | 314 | 262.67 | 257.45 | 51.33 | 56.55 | 2.34 |
| 11-132-2 | 11 | 29 | 97 | 76.45 | | 20.55 | | 2.64 |
| 11-132-3 | 14 | 37 | 113 | 97.79 | | 15.21 | | 2.64 |
| 11-132-4 | 5 | 13 | 35 | 33.80 | | 1.20 | | 2.60 |
| | 30 | 79 | 245 | 208.04 | 208.03 | 36.96 | 36.97 | 2.63 |
| 12-133-1 | 14 | 37 | 119 | 97.79 | | 21.21 | | 2.64 |
| 12-133-2 | 15 | 42 | 142 | 117.60 | | 24.20 | | 2.80 |
| 12-133-4 | 7 | 20 | 70 | 57.14 | | 12.86 | | 2.86 |
| | 36 | 99 | 331 | 272.53 | 272.25 | 58.47 | 58.75 | 2.75 |
| 13-133-5 | 9 | 17 | 41 | 32.11 | | 8.89 | | 1.89 |
| 13-133-7 | 13 | 30 | 92 | 69.23 | | 22.77 | | 2.31 |
| 13-133-8 | 15 | 32 | 78 | 68.27 | | 9.73 | | 2.13 |
| | 37 | 79 | 211 | 169.61 | 168.68 | 41.39 | 42.32 | 2.14 |
| 14-133-9 | 26 | 62 | 180 | 147.85 | | 32.15 | | 2.38 |
| 14-133-10 | 14 | 31 | 85 | 68.64 | | 16.36 | | 2.21 |
| | 40 | 93 | 265 | 216.49 | 216.23 | 48.51 | 48.77 | 2.33 |
| 15-133X | 11 | 35 | 119 | 11.37 | 111.37 | 7.63 | 7.63 | 3.18 |
| Total | 201 | 495 | 1485 | 1240.71 | 1234.01 | 244.29 | 250.99 | 2.46 |
| **RUSSIAN** | | | | | | | | |
| 16-123-1 | 14 | 40 | 126 | 114.29 | 114.29 | 11.71 | 11.71 | 2.86 |
| 17-123-2 | 24 | 64 | 190 | 170.67 | | 19.33 | | 2.67 |
| 17-123-3 | 23 | 64 | 208 | 178.09 | | 29.91 | | 2.78 |
| | 47 | 128 | 398 | 348.76 | 348.60 | 49.24 | 49.40 | 2.72 |
| Total | 61 | 168 | 524 | 463.05 | 462.89 | 60.95 | 61.11 | 2.75 |

Table 70 (continued)

SPANISH

| Instr./ Section | N[a] | ΣX[b] | ΣX² | $\frac{(\Sigma X)^2}{n}$ Section | $\frac{(\Sigma X)^2}{n}$ Instructor | $\Sigma X^2 - \frac{(\Sigma X)^2}{n}$ Section | $\Sigma X^2 - \frac{(\Sigma X)^2}{n}$ Instructor | $\bar{X}$[c] |
|---|---|---|---|---|---|---|---|---|
| 18-152-1 | 25 | 78 | 268 | 243.36 | | 24.64 | | 3.12 |
| 18-152-2 | 25 | 64 | 212 | 163.84 | | 48.16 | | 2.56 |
| 18-153-1 | 24 | 77 | 273 | 247.04 | | 25.96 | | 3.21 |
| | 74 | 219 | 753 | 654.24 | 648.12 | 98.76 | 104.88 | 2.96 |
| 19-152-3 | 16 | 38 | 114 | 90.25 | | 23.75 | | 2.38 |
| 19-152-4 | 16 | 39 | 113 | 95.06 | | 17.94 | | 2.44 |
| 19-152-11 | 14 | 32 | 78 | 73.14 | | 4.86 | | 2.29 |
| | 46 | 109 | 305 | 258.45 | 258.28 | 46.55 | 46.72 | 2.37 |
| 20-152-5 | 9 | 29 | 99 | 93.44 | 93.44 | 5.56 | 5.56 | 3.22 |
| 21-153-5 | 31 | 84 | 242 | 227.61 | 227.61 | 14.39 | 14.39 | 2.71 |
| 22-153-3 | 9 | 31 | 111 | 106.78 | | 4.22 | | 3.44 |
| 22-153-4 | 17 | 43 | 125 | 108.76 | | 16.24 | | 2.53 |
| | 26 | 74 | 236 | 215.54 | 210.62 | 20.46 | 25.38 | 2.85 |
| 23-153-2 | 14 | 41 | 139 | 120.07 | | 18.93 | | 2.93 |
| 23-153-6 | 13 | 36 | 118 | 99.69 | | 18.31 | | 2.77 |
| 23-153-7 | 10 | 33 | 119 | 108.90 | | 10.10 | | 3.30 |
| | 37 | 110 | 376 | 328.66 | 327.03 | 47.34 | 48.97 | 2.97 |
| 24-153-8 | 38 | 93 | 255 | 227.61 | | 27.39 | | 2.45 |
| 24-153-9 | 37 | 87 | 225 | 204.57 | | 20.43 | | 2.35 |
| 24-153-10 | 33 | 97 | 303 | 285.12 | | 17.88 | | 2.94 |
| | 108 | 277 | 783 | 717.30 | 710.45 | 65.70 | 72.55 | 2.56 |
| 25-153-12 | 19 | 43 | 129 | 97.32 | | 31.68 | | 2.26 |
| 25-153-13 | 10 | 29 | 101 | 84.10 | | 16.90 | | 2.90 |
| 25-153-14 | 16 | 27 | 57 | 45.63 | | 11.37 | | 1.69 |
| | 45 | 99 | 287 | 227.05 | 217.80 | 59.95 | 69.20 | 2.20 |
| Total | 376 | 1001 | 3081 | 2722.29 | 2693.35 | 358.71 | 387.65 | 2.66 |

Table 70 (continued)

SUMMARY OF DATA

| Total | N[a] | k[d] | $\Sigma X$[b] | $\Sigma X^2$ | $\frac{(\Sigma X)^2}{n}$ Section | $\frac{(\Sigma X)^2}{n}$ Instructor | $\Sigma X^2 - \frac{(\Sigma X)^2}{n}$ Section | $\Sigma X^2 - \frac{(\Sigma X)^2}{n}$ Instructor | $\bar{X}$[c] |
|---|---|---|---|---|---|---|---|---|---|
| French | 262 | 16 | 599 | 1661 | 1389.05 | 1384.11 | 271.95 | 276.89 | 2.29 |
| German | 201 | 14 | 495 | 1485 | 1240.71 | 1234.01 | 244.29 | 250.99 | 2.46 |
| Spanish | 376 | 19 | 1001 | 3081 | 2722.29 | 2693.35 | 358.71 | 387.65 | 2.66 |
| Russian | 61 | 3 | 168 | 524 | 463.05 | 462.89 | 60.95 | 61.11 | 2.75 |
| | 900 | 52 | 2263 | 6751 | 5815.10 | 5774.36 | 935.90 | 976.64 | 2.51 |
| Total | | | | | 5690.19 | | 1060.81 | | |

AOV based on unweighted means

Summary of data:

SS within sections based on individuals = 1091.16

SS with sections based on means = 77.44

Harmonic mean = 14.09

M.S. for within sections based on means = .0913

[d] k = Number of sections.