Iowa State University
Digital Repository

Retrospective Theses and Dissertations

Iowa State University Capstones, Theses and Dissertations

1979

Longitudinal effects of a self-paced instructional method in an introductory university biology course

Mohammad Hossain Najmaie
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

Part of the Curriculum and Instruction Commons, Educational Methods Commons, Instructional Media Design Commons, and the Science and Mathematics Education Commons

Recommended Citation
Najmaie, Mohammad Hossain, "Longitudinal effects of a self-paced instructional method in an introductory university biology course" (1979). *Retrospective Theses and Dissertations*. 7298.
https://lib.dr.iastate.edu/rtd/7298

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

المنارة للاستشارات

www.manaraa.com

INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help you understand markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure you of complete continuity.

2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed the photographer has followed a definite method in "sectioning" the material. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.

University Microfilms International

300 N. ZEEB ROAD, ANN ARBOR, MI 48106
18 BEDFORD ROW, LONDON WC1R 4EJ, ENGLAND

8010249

NAJMAIE, MOHAMMAD HOSSAIN

LONGITUDINAL EFFECTS OF A SELF-PACED INSTRUCTIONAL METHOD IN AN INTRODUCTORY UNIVERSITY BIOLOGY COURSE

Iowa State University PH.D. 1979

University Microfilms International 300 N. Zeeb Road, Ann Arbor, MI 48106 18 Bedford Row, London WC1R 4EJ, England

Longitudinal effects of a self-paced instructional method in an introductory university biology course

by

Mohammad Hossain Najmaie

A Dissertation Submitted to the

Graduate Faculty in Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Department: Professional Studies in Education
Major: Education (Research and Evaluation)

Approved:

Signature was redacted for privacy.

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1979

TABLE OF CONTENTS

Page

CHAPTER I. INTRODUCTION

Educators have long been concerned with employing the most effective methods of teaching in schools. "Indeed in many ways a history of education is a history of methodological revolutions and innovations" (Ryan, 1974, p. 1). However, an exception to employing the most effective educational procedures may be the instructional practices followed in universities. In these institutions, lecture has been the predominant mode of instruction since the 12th century. The adoption of the lecture method in those early days can be easily understood given the difficulties associated with the production of textbooks and other educational material before the advent of the printing press (Haskin, as cited in Ryan, 1974).

The reasons why universities have been resistant to the changes and innovations in instructional methods, while other educational institutions have been willing to change their procedures, are not at all evident. One probable explanation may be that those responsible for university teaching have never really been convinced that they are engaged in instruction, but rather in some more important activity with the main purpose being promoting the general powers of the mind (Ryan, 1974).

Since the beginning of the 20th century, the function and the role of the university in the society has dramatically changed. Unlike in past centuries, the average person today has access to higher education so that universities influence the social life of almost everyone in society. This society is more aware of what is happening in higher education and is beginning to demand that the university be "accountable" in more or less the same fashion that has been required of the other levels in the educational system.

In response to society's demands, universities in many parts of the world, have begun to conduct programs of self-evaluation. In these introspective programs the instructional methods employed have received much attention. A sincere search has begun for techniques that are more effective than the traditional lecture method.

Prior to the 1950's most research on college teaching dealt with comparison of traditional methods and independent study strategies (Robin, 1976). Dubin and Taveggia (1968) reviewed 91 of these studies and concluded that there were no significant differences in the performance of students taught by any of these methods as compared to the traditional methods.

In the 1950's the experimental psychologists started to apply learning principles to individualized instruction (Skinner, 1954, 1958). Skinner's methods depended on the use of teaching machines and computers but high initiation and maintenance costs prevented general adoption. Elimination of human interaction from the teaching process was also viewed as an additional disadvantage. In recent years, a strategy of college teaching (Keller, 1966, 1968; Sherman, 1974a) has been developed which does not eliminate human interaction and does not place reliance upon high cost machines.

After its introduction it has been labeled variously: the Keller plan (Green, 1971), the operant approach (Myers, 1970), the behavioral approach (Johnston & Pennypacker, 1971), Personalized System of Instruction (PSI) (Born, Gledhill & Davis, 1972; Born & Herbert, 1971; Keller, 1966, 1968), individualized instruction (Ferster, 1968; Witters & Kent, 1972), self-paced programmed instruction (Koen, 1970, Lloyd & Knutzen, 1969; Sheppard and MacDermot, 1970), continuous progress concept (Moore, Mahan & Ritts, 1969) and contingency management (Cooper & Greiner, 1971; Malott & Svinicki, 1969; McMichael & Corey, 1969). Hereafter, it will be called PSI for brevity. PSI has attracted the attention of educators around the world in many disciplines since its introduction at the new university of Brazilia (Vander Klaw & Plomp, 1974). It has also attained the stature of an educational movement, with a national center for PSI, professional publications (Personalized System of Instruction Newsletter, 1971-1975; Journal of Personalized Instruction), textbooks (Keller & Sherman, 1974; Sherman, 1974a; Johnston, 1975), national conferences (Johnston, 1975; Johnston & O'Neill, 1975, Ruskin & Bono, 1974; Keller, 1971; Dessler, 1972; Eliason & Munsee, 1972) and literature reviews (Kulik, Kulik & Carmichael, 1974;

Kulik & Kulik, 1975; Kulik, 1976; Ryan, 1974; Robin, 1976; Hursh, 1976).

A Brief History

The idea of self-pacing is not new. Sherman (1974b) indicated that the concept was in practice when Frederick Buck developed one of the first systems of individualized instruction at the San Francisco State Normal School in 1912. Buck and his colleagues prepared courses of study to permit learners to advance at their own rate. In 1917, his activities were stopped by a ruling from the California attorney general, as the power to publish textbooks or printed instructional materials was judged to rest entirely with the State Board of Education in that state.

The recent development of the PSI had three phases (Sherman, 1974b). It started in the 1950's when B. F. Skinner first introduced his behavioral theory of learning. In those days Keller and Schoenfeld earned a reputation for innovation from their textbook and their introductory laboratory course in psychology. Sherman was a graduate student at Columbia University and was exposed to the idea of programmed instruction. The second phase started in 1964 when both Keller and Sherman went to Brazil and established a psychology department in the new university of Brazilia with the aid of two Brazilians, Carolina Murtuscelli Bori and Rodolpha Azzi. With these two Brazilians they developed a new concept of individualization, namely the concept of personal individual interaction. Though founded on firm psychological principles, one of the principal reasons this system was started at Brazilia may have been language differences between the instructors and students. The final phase occurred in 1968 when they had to come back to the U.S.A., they hired proctors to serve in the system.

The Theory and Philosophy Behind PSI

The theory underlying PSI has a close connection with a branch of psychology called "behaviorism." Skinner's ideas are based upon the previous works by Fechnner, Wundt, Guthrie, Thorndike, and Watson.

In Skinner's own words: "Behavior is said to be strengthened by its consequences and for that reason, the consequences are called 'reinforcers', thus when a hungry organism exhibits behavior that produces food, the behavior is reinforced by that consequence and is therefore more likely to recur" (Skinner, 1974, p. 39).

According to Skinner, there are two types of reinforcers, positive stimuli and negative (punishing) stimuli. A positive reinforcer strengthens any behavior that produces it and thus increases the likelihood of its being repeated. Negative reinforcers, on the other hand, are those things that an individual is willing to work hard to avoid. Behaviors that precede these consequences are likely to be repeated. An important concept is that the definition of a stimulus depends on how the individual responds to it, not on the basis of inherent properties of the reinforcer itself (some students work hard for good grades, some don't).

Through different procedures, it is possible to change the frequency of behaviors. In order to increase the frequency of a "behavior" some "shaping" procedures must be utilized. Davis (1976) describes this sequence of procedures as follows:

1. A clear behavioral goal must be identified.
2. A base rate measure of existing skill or present performance level must be taken.
3. The task must be broken into steps.
4. Successive approximations of the goal are reinforced until the goal is reached (p. 20).

These "shaping" procedures of "behaviorism theory" are obviously employed in Keller's PSI. More specifically, having a clear description of what is to be learned is a direct translation of the first "shaping" procedure. The second procedure refers to importance of pretests in classes. Having small units is exactly the same as the third procedure and finally having numerous examinations and immediate feedback along with providing chances to retake the examinations is nothing but the last procedure, which is the concept of mastery found in PSI.

Another educational philosophy which played a role in developing the self-paced system was the concept of individual differences in teaching-learning activities. Why shouldn't students progress at their own speed? Advocates of this philosophy strongly disagree with blending the differences between individuals into the concept of a hypothetical average student (HAS) (Latta, Dolphin & Grabe, 1978).

Another learning theory which is also closely connected with self-paced instruction is that of mastery learning, which, according to Davis (1976) is an outgrowth of behavioral learning theory. The advocates of this theory have basic questions about the "norm-referenced" grading system. In an essay on mastery theory, Bloom (1971) noted some deficiencies in grading on the curve:

> Having become 'conditioned' to the normal distribution, we set grading policies in these terms and are horrified if a teacher recommends a new grading distribution . . . . The normal curve is not sacred. It described the outcome of a random process. Since education is a purposeful activity in which we seek to have students learn what we teach, the achievement distribution should be very different from the normal curve if our instruction is effective. In fact, our educational effort may be said to be unsuccessful to the extent that achievement is normally distributed (p. 49).

A mastery requirement is an essential part of PSI, but there are some substantial differences in Bloom's definition of mastery and that found in PSI.

Bloom's strategy is primarily designed for use with the group based instructional situation, whereas PSI is primarily concerned with individual instruction. For Bloom and Block, mastery of the parts is not synonymous with mastery of the whole, but in the PSI method mastery is required for each unit as well as the comprehensive final exam. Block (1974) claims that unit feedback instruments in Bloom's method provide more detailed information about the student's performance than Keller's method does. In other words, he claims they are more diagnostic in nature. Standard Keller's plan tests require almost perfect mastery, while under Bloom's strategy the mastery rate is not as high.

Basic Features of the Method

Although there are a wide variety of PSI methods, there are common features which include the following: detailed instructional objectives; small units or blocks; frequent tests; student proctors; subject matter mastery; and student determined progress.

There are no empirical data concerning variations in phrasing of objectives except those of Green (1971) and Johnston & Pennypacker (1971) who have short discussions about what might be expected in a good set of objectives.

Unit construction is an important variable in the method. In the studies published thus far the number of units has ranged from 8 (Alba & Pennypacker, 1972) to 30 (Witters & Kent, 1972) with a median of 13 units per academic period, which may be either a quarter or a semester depending upon the university.

Several users structured the material so that students study the subject at a rate of one unit per week. As a whole, the structure of the textbook used determines the number of units. Having more units allows frequent examinations which to some proponents of PSI is more effective than having less frequent tests.

Frequent testing is an important component of the method. Some investigators in the field feel this is the single most important factor responsible for the effectiveness of PSI (Abbott & Falstorm, 1977). There are no preset criteria for the nature of these tests and usually the instructor is free to design his measures to conform to his circumstances and objectives. Tests may be either written or verbal or a combination of both, but in practice they are of the objective and easily scored variety. However, in theory, the only requirement is that the tests measure the instructional objectives of the course.

Employing student proctors in a PSI course has been the subject of some extensive investigations. The majority of proctors used in PSI courses are undergraduates; however, in some studies graduate students have been employed as well. One study (Born & Herbert, 1971) reported a comparison of graduate and undergraduate proctors effectiveness and concluded that students' satisfaction with their proctors was quite independent of the proctors' level of training.

The PSI method depends on a set of mastery criteria. The degree of mastery is different in various studies and ranges from almost perfect mastery to some lower level of achievement, e.g., 60% mastery in some versions of the method.

PSI courses have been used with any number of students ranging from 20 (Green, 1971) to 1000 (Malott & Svinicki, 1969). However, Sherman (1972) has expressed skepticism about the effectiveness of a PSI course with more than 150 to 200 students. One way to handle the large number of students in a PSI course is the use of undergraduate or graduate student proctors to assist in the test administration, grading, interviewing, and record keeping. In some recent versions of the method, computer techniques have been employed to do some of the above mentioned tasks.

Some Procedural Variations

Since its introduction, PSI has been modified by many authors to cope with their particular philosophy and situations. Some of these modifications have been quite extensive, while many of them have imposed minor changes, mainly by modifying one or two components or eliminating some features of the method. It is useful to review the "standard" PSI procedure before introducing some important versions of the method.

In the "standard" Keller's procedure, students are given a study guide for each unit; after a study period, they return to

Several users structured the material so that students study the subject at a rate of one unit per week. As a whole, the structure of the textbook used determines the number of units. Having more units allows frequent examinations which to some proponents of PSI is more effective than having less frequent tests.

Frequent testing is an important component of the method. Some investigators in the field feel this is the single most important factor responsible for the effectiveness of PSI (Abbott & Falstorm, 1977). There are no preset criteria for the nature of these tests and usually the instructor is free to design his measures to conform to his circumstances and objectives. Tests may be either written or verbal or a combination of both, but in practice they are of the objective and easily scored variety. However, in theory, the only requirement is that the tests measure the instructional objectives of the course.

Employing student proctors in a PSI course has been the subject of some extensive investigations. The majority of proctors used in PSI courses are undergraduates; however, in some studies graduate students have been employed as well. One study (Born & Herbert, 1971) reported a comparison of graduate and undergraduate proctors effectiveness and concluded that students' satisfaction with their proctors was quite independent of the proctors' level of training.

The PSI method depends on a set of mastery criteria. The degree of mastery is different in various studies and ranges from almost perfect mastery to some lower level of achievement, e.g., 60% mastery in some versions of the method.

PSI co [redacted] with any number of students rangi [redacted] (Malott & Svinicki, 19 [redacted] ssed skepticism about [redacted] than 150 to 200 [redacted] of students in a [redacted] uate student [redacted] grading, inter- [redacted] rsions of the [redacted] to do some of [redacted] modified by many author [redacted] r philosophy and situations. Some of these [redacted] have been quite extensive, while many of them have imposed minor changes, mainly by modifying one or two components or eliminating some features of the method. It is useful to review the "standard" PSI procedure before introducing some important versions of the method.

In the "standard" Keller's procedure, students are given a study guide for each unit; after a study period, they return to class on any of the scheduled class periods (five to six sessions in a week). Students take a unit exam which is immediately corrected by a proctor and returned to them. If the student fails to meet a specified standard of mastery on that unit, his problem will be diagnosed and he is asked to return after appropriate study to take a parallel form of the same test. There is no penalty for failing a unit except that it takes the student longer to pass the unit. If the student passes the test at the mastery level, he is immediately given the study guide for the next unit and the preceding cycle of activity is repeated. In this way, the student is able to proceed throughout the course at his own speed, but must follow an instructor's sequence. Under Keller's (1966) original procedure, the students' final grade in the course was based entirely on the number of units passed. But as practiced at many schools, the final grade is based upon the unit success rate in combination with a comprehensive final examination. The lecture in Keller's original plan was not the primary source of information; rather it was motivational and supplementary to the text material and students were left free to attend or not.

One of the most widely employed variations of the Keller's original plan has been the use of interviews for evaluation (Ferster, 1968; Ferster & Perrott, 1968; Sheppard & MacDermot, 1970; Wood, 1972). A proctor conducts an interview and each student has to successfully demonstrate his familiarity with the unit concepts before taking the written test. The interviewer is not supposed to instruct or teach, but merely to serve as a focal point for the verbalizations of the student being tested.

Born et al. (1972) developed a new version of PSI in which the student determined the size of the units in his program. The text used (Ferster & Perrott, 1968) had a large number of small five or six-page sections and they made short tests based on each of these small sections. The student reported the number of sections which she or he wished to take and the proctor would select the appropriate test pages from the master test file as a personalized test for the student.

Another variation is concerned with the testing schedule (Cooper & Greiner, 1971). They formalized the testing sequence within the Monday-Wednesday-Friday, 50-minute class periods. There were several varieties of this version in the literature (Whitehurst, 1972; Malott & Svinicki,1969).

Johnston & Pennypacker (1971) introduced the idea of "performance sessions," in which there was discourse between proctor and student with an emphasis on the rate of the student's correct response to test questions. In these sessions the proctor randomly selected questions from a master file and the student read the question aloud, completing the blank or indicating that he did not know the answer; the correct answer was immediately given to him. The proctor measured the response time while the student was reading and answering the question and informed the student whether the performance rate was acceptable or not.

Another version of behavioral instruction used "Group Remediation" (Bostow & Blumenfeld, 1972; Cooper & Greiner, 1971; Malott & Svinicki, 1969). It involved teacher-paced, group-administered quizzes, and immediate feedback to the group by the instructor. There were two tests each week, the first one covering the new material and the second being a make-up test. If the student failed the first examination, he had to pass the second in order to receive credit for the unit.

Lloyd (1971) provided many activities for students in a PSI course in order to involve students in the subject matter. These activities included: class attendance, class participation, text review quizzes, book reviews, movie reviews, tape reviews, discussion with other faculty about physiological research, animal labs, field trips and so on. These activities were optional and students had to choose a minimum number of them. A student grade was based on accumulation of points for various combinations of activities.

The Method Under Investigation

In 1969 Iowa State University offered an introductory course in biology called "Principles of Biology." Nearly 600 students were enrolled in lecture sessions. Pedagogical problems were obvious as a result of this large number of students. Dolphin, Franke, Covert & Jorgensen (1973) best described the situation this way:

> The lectures were necessarily impersonal; Socratic methods were impossible and even disrupting, and personal acquaintance with the instructor was improbable. Lectures and exam were directed at a mythical 'average' student, and as a result, the needs of either rapid or slow learners were seldom met (p. 24).

In order to solve these and other problems, Dolphin and his colleagues designed an instructional strategy called Phase Achievement System (PAS). In their design efforts they tried to maintain the main economical feature of the course, large enrollment, while creating a system which would allow for individual differences in rate of learning.

This system has some similarities to personalized system of instruction or Keller plan and Bloom's mastery learning strategy. However, it also differs from these two instructional systems in some of its basic elements.

As in other self-paced systems, students are provided with a clear description of objectives which are organized by units and serve as a study guide. The role of lecture in PAS is quite different from PSI and Bloom's mastery method. In PSI the primary role of lecture is a motivational role and the emphasis is more on the written materials, while in PAS the instructor discusses all material in his lectures. The role of lectures in PAS differs from Bloom's since it is optional for students to attend. Bloom's lectures are closer to traditional method than PAS or PSI method.

The course content is divided into 8 units. The unit size of PAS is somewhat between PSI and mastery learning (PSI courses tend to have small units but those utilizing Bloom's methods have bigger units). Learning in PAS is self-paced, *i.e.* the students can progress at their own rate using the study guide. This feature of PAS is the same as PSI and the opposite of Bloom's method in which learning is teacher-paced, but PAS can be teacher paced through the lectures if a student opts not to use the study guide to pace himself. Thus, within one lecture section alternative pathways for learning are provided.

Mastery requirements in PAS have some unique features which differ from both PSI and Bloom's strategy. It differs from Bloom's since it requires mastery in each unit; meanwhile it is also different from PSI in two aspects (1) the mastery rate is less in PAS than it is in the standard PSI (similar to Bloom's strategy) and, (2) PAS introduces two thresholds in mastery requirement namely breadth of comprehension and the depth of comprehension.

Students' final grades in PAS are based on passing all units and their average is judged against preset standards for grading. Past experiences indicated what average percent levels corresponded to grades. Another difference between PAS and PSI is found in the sequence of taking tests by students. In PSI, students can't take a test on unit 2 before passing unit 1, but PAS students can pass units in any sequence they choose. There is no obligatory hierarchy for taking the tests.

In the standard Keller's procedure, the problem of administering, scoring and record keeping of examinations is solved by student proctors; while in PAS this problem is solved by utilizing a computer system which randomly generates the tests from a pool of 3000 multiple-choice items, scores tests, keeps records on individual students and provides statistical information within the modular framework of PAS. The examinations were given at two-week intervals during the academic period and students might retake any examination over any phase not previously passed or passed with a personally unsatisfying score.

An audio tape library was also available for students to use. The system is presently being modified to include a video tape library of lectures over the content of each unit to be available to each student upon demand.

In 1974, Dr. Dolphin and his colleagues evaluated the effectiveness of the PAS at Iowa State University. The results of that study revealed positive changes in both the cognitive and affective domains (Latta et al., 1978).

The Statement of the Problem

In Fall 1974, Dolphin taught two sections of Biology 101, a PAS section and a traditional section. The cognitive and affective outcomes from these experimental and control teaching methods were compared. The results indicated that PAS was a highly beneficial learning experience for highly test anxious female students.

After four years the impact of this experiment should be explored to determine if there have been any long-term effects. Utilizing a data base gathered by Dolphin and his colleagues, the purpose of the present investigation is to explore the impact of PAS upon the later performance and academic careers of the students in PAS compared to students in traditional method. Therefore this study focuses on examining the impact of the PAS method in the following areas:

1) Differences in grades achieved in the related courses taken after Fall, 1974. Any course taken in Biology, Zoology, Botany, Bacteriology, Biochemistry and Biophysics, Animal Ecology, Food and Nutrition, Entomology, Genetics, Environmental Studies, Agronomy, Horticulture, Animal Science and Forestry will be considered as a related course.

2) Differences in number of credits the students completed in the above mentioned areas.

3) Facilitative or concurrent effects of PAS compared to traditional method.

Since the students were exposed to PAS only in one quarter and in one course, the present study is going to examine its impact for only three subsequent quarters. It is improbable but unknown that one course in PAS at the freshman level would have a substantial impact upon the total academic career of the students. Consequently, the present study will consider the available data up to the end of Fall, 1975 (i.e. quarters W'75, Sp'75, F'75).

Variables Under Consideration

Drs. Latta, Dolphin and Grabe (1978) in their "individual differences study" collected a number of measures on independent variables which were readily available on file. Since that study revealed the significance of those independent variables in predicting achievement, this study will use the same independent variables. These are as follows:

1) Teaching method (PAS vs. traditional).
2) Gender (Male vs. Female).
3) Scores on Minnesota Scholastic Aptitude Test (MSAT).
4) High school graduation ranking (HSR).
5) A composite measure called High school Background (HSBKGD). Number of high school credits in physics, chemistry, biology and mathematics comprised this measure.
6) Scores on a test anxiety questionnaire (TAQ) which was designed by Mandler and Sarasan (1952).

In addition to these independent variables, there are other independent variables which will be used to statistically control the results of this study. These include:

7) Major, which is categorized as life-science or non-life-science.

8) Year in university, which is classified as freshman and upperclass.

9) Cumulative university GPA by the end of September 1974 (CUMGPA), excluding the grade of Biology 101 and concurrent credits in life-science.

10) Number of university credits up to September 1974 (NCUS).

11) GPA of university credits up to September, 1974 in mathematics (GPAMS), physics (GPAPS), chemistry (GPACS), and life-science (GPALS) courses.

12) Number of the university credits up to September 1974 in mathematics (NMCS), physics (NPCS), chemistry (NCCS) and life-science (NCLS).

The dependent variables under study include:

1) GPA in life-science courses taken during the one-year period (GPALQ).

2) Number of university credits in the field taken during the one-year (NCLQ).

3) GPA of university credits in life-science taken concurrently with Biology 101 in fall 1974 (GPACCL).

All variables not obtained from the original file were obtained from student transcripts. Five subjects were dropped from the study because of inadequate information in their transcripts.

Subjects and the Original Study

Since part of the data for this study came from a previous study (Latta _et al_., 1978), the design of that study and the conditions under which those data were gathered should be described. The subjects were 99 males and 92 females enrolled in a PAS section and 102 males and 92 females in a traditionally instructed section. Efforts were made to control the study, experimentally. Although the students could choose any section they wanted, they had no way of knowing which section was experimental and which was control prior to the beginning of the experiment. After registration the groups were randomly assigned to PAS (experimental) and traditional (control) methods. In order to avoid the differential effects due to instructors, both sections were taught by the same instructor (Dr. Dolphin). Both sections used the same course objectives and were lectured to as equally as possible. The primary differences were in the test taking procedures and feedback procedures for PAS group. The PAS group had at least five chances to take or retake tests under a criterion referenced grading plan, while the traditional group had three examinations and a final under a norm referenced testing procedure.

The drop out rate during the time the class was given, was about 10% for both PAS and traditional groups.

The Nature of the Study

This study is investigatory rather than hypothesis testing. The primary interest is in exploring the potential long-term impact of PAS rather than testing a prior hypothesis. Models are sought which can show the effects of different factors of the study upon each dependent variable, e.g. grades in subsequent life-science courses. Sensitivity and power of these models will be emphasized. Special attention is placed on individual student differences and attribute by treatment interaction. This feature allows one to answer the question of "what type of student benefits most from what type of instructional method?". Furthermore, the development of such models will allow an assessment of any long-term effects of self-pacing which influences events after students have taken the course. The general procedures developed here should be applicable in other educational settings.

Limitations

Since some of the independent variables used in the present study are adopted from an earlier study on the method (Latta et al., 1978), the results will be dependent on the accuracy and reliability of those measures. Due to missing information on some variables, it will not be possible to do the data analysis with all students involved in the original study. The completion rate was 70, 85 and 50 percent for the three dependent variables, respectively. Loss of subjects always causes problems and leaves more skepticism about the results. The loss of subjects for the third dependent variable, *i.e.*, grades on concurrent credits (GPACCL) is rather substantial; therefore the results of this section should be considered cautiously. Another possible limitation of the study is the usage of rather large and complex statistical models which requires extensive assumptions to be reliable. And finally, a "caveat" to the reader is appropriate: this study is hypothesis forming rather than hypothesis testing. Any attempt to draw inference from the results of this study should be done in a separate experiment with preset hypotheses.

CHAPTER II. LITERATURE REVIEW

Although the purpose of this investigation is to explore the long-term influences of an instructional strategy similar to PSI, studies dealing with immediate effects are useful because later effects are obviously related to those which occurred earlier. There are hundreds of such studies and it would require Herculean effort to discuss all of them. Therefore three major literature reviews already published will be covered to indicate general trends. Since the latest reviews only cover studies published up to 1975, the literature will be updated to cover studies on immediate effects of PSI which have been published since then. The immediate effects of instruction will be divided into two broad categories: cognitive and affective effects.

Immediate Effects

Immediate cognitive effects

Academic performance The most extensive review has been done by Robin (1976) who reviewed 39 studies on the effectiveness of behavioral instruction.[1] This review only included studies which utilized achievement measures (other than final grades) common to both experimental and control

[1]This is a general name which Robin (1976) applied to all variations of personalized instructional systems.

groups. Thirty of these studies showed significant differences favoring behavioral instruction and six reported equal performance. Two studies employed multiple comparison within the same study, finding significant differences in favor of the experimental group in one part of each study. There was only one single study which showed better performance for the control group. Behavioral instruction apparently results in cognitive gains.

Kulik et al. (1974) reviewed 15 studies on cognitive effectiveness of the PSI. Eight of these fifteen studies were common with Robin's selected studies. All of these eight common studies were the ones with results in favor of PSI. From the seven remaining studies, four revealed results favoring PSI and three found no significant differences between experimental and control groups in terms of academic performance.

Hursh (1976) reviewed 23 studies concerning the cognitive effectiveness of PSI. Ten of these selected studies were common with the two previously mentioned reviews and were the ones which showed significantly higher performance for PSI students. Seven of his reviewed studies reported the results of using PSI without attempting to experimentally compare the results with those from other teaching methods. The remaining six studies uniformly suggested that PSI produced higher examination scores.

If all three of these reviews are combined, 59 studies were described (counting the common studies in just one review). Forty-two studies (72%) reported results suggesting significant differences favoring PSI, while nine (15%) showed the equal performance and one (2%) resulted in higher achievement for the traditional method and seven studies (12%) contained no comparisons.

All three reviews indicated some shortcomings of research in this field and made suggestions for improving the designs. Robin (1976) describes three critical issues relative to 39 selected studies: methodology; the nature of the attitudinal responses; and high student withdrawal rates. Subject assignment, initial equivalence of groups and objective evaluation of dependent measures are the more important issues. Twelve of the studies he considered employed random assignment, while 14 relied upon administrative assignment. Ten of the latter fourteen reported between group equivalence on at least one of the following variables: grade point average, standardized ability tests, age, major, course pre-test, attitudes toward the academic discipline and the number of previous courses in the area. Seven studies were done across academic periods, six of them did not demonstrate group equivalence. Two studies let the students select the preferred instructional method. Obviously nonrandom assignment or (at least) not showing the between group equivalence can jeopardize the results of the studies and challenges the validity of the data. Lack of unbiased data on dependent measures in studies which used essay type questions or fill in items on the test is another problem identified by Robin.

Achievement results can be further confounded by differential withdrawal rates. Usually, the withdrawal rate for PSI courses was higher than that found in traditional courses (Robin, 1976; Ryan, 1974). This may be a significant factor in studies which favor PSI. Robin suggested two ways to handle this problem: (a) test for the equivalence of those students who drop out versus those who complete; and (b) statistically control for any obtained discrepancies. Robin found 14 of 39 studies reported the withdrawal rates and only two of these 14 performed the appropriate tests and statistical controls (Born *et al.*, 1972; Sheppard & MacDermot, 1970). The results of these controlled studies still favored PSI.

Hursh (1976) round that the major shortcoming in the selected studies was the lack of enough experimental control and reliable dependent measures. He named six factors to be controlled in any PSI versus control comparison studies: differences in instructors between personalized and control classes; curriculum materials; grading criteria; testing formats; student selection; and student expectations. He reported that in his 23 selected studies, the median study controlled for only four of the six above mentioned confounding factors (range: 0-6).

The views of Kulik _et al_. (1974) on control were similar to those of the other two reviewers. However, in addition he added a question: "Are students taught the test in the Keller section?". Since clearly specified objectives are one of the major features of Keller's courses, students know exactly what to study for the examinations, whereas the students in conventional methods do not have this privilege.

Study time Another immediate effect is the amount of study time students have spent preparing in the Keller system versus that spent in traditional methods. Two reviews (Robin, 1976; Kulik _et al_., 1974) considered study time. In their reviews, there were 15 studies (counting the common studies in just one review) which quantified study time through various techniques, _e.g._, end-of-the-semester self-reports, weekly-self-reports and objective recoding. Fourteen out of these 15 studies (93%) have shown that the students under PSI have worked significantly longer than students in traditional courses. The remaining study showed no significant differences in time spent studying.

Withdrawal rates Probably the most negative aspect of PSI was the relative higher withdrawal rates found for students in PSI courses. This fact is reported in all of three major literature reviews. Robin (1976) reported that 14 of his 39 selected studies reported withdrawal rates for both PSI and traditional groups. His conclusion was that PSI resulted in a 14% withdrawal rate which is 40% higher than 10% withdrawal rate for traditional, lecture-discussion conditions. Kulik et al. (1974) reported that withdrawal rates in PSI courses were three to four times higher than those for the conventional courses, though they found one case (McMichael & Corey, 1969) which had a lower withdrawal rate for the PSI course than the conventional course. Hursh's review (1976) also mentioned the high withdrawal rates from PSI and considered it as a disadvantage of the system. High withdrawal rates have made the evaluation of the PSI difficult and, consequently both affective and cognitive results favoring behavioral instruction should be considered with caution, since they may be influenced by differential withdrawal rates.

An explanation for this high rate could be that the immediate feedback and numerous quizzes allow students in PSI to see their progress in a course clearly. Students with less academic skill see their low standing earlier and better than the same students in a conventional course where there are one or two exams during the whole period of the enrollment. This hypothesis is supported by the higher percentage of the students who earn "F" grades in conventional courses compared to PSI courses (Keller, 1968). Robin (1976), after analyzing numerous studies which systematically and experimentally studied the withdrawal rates, concluded that procrastination and poor academic ability were the contributing factors for high withdrawal rates in PSI courses.

Immediate affective effects

Although there is solid evidence in favor of PSI in cognitive aspects, this advantage is small when compared to the very high rating given such courses by students.

Sixteen of the studies reviewed by Robin (1976) included attitude measures. These were typically gathered using a teacher-made, unstandardized self-report questionnaire, on a single occasion at the end of the semester. Fourteen of these studies reported results significantly favoring PSI courses. The remaining two studies have shown equal and positive attitudes toward both PSI and traditional courses.

Kulik et al. (1974) reviewed 10 studies concerned with attitude surveys, which lacked control groups. In all of 10 studies the PSI courses were rated very high by the majority of the students enrolled in such courses. Hursh (1976) reported similar results in his review.

There are some deficiencies in studies dealing with attitudes toward PSI and all three reviewers discussed them. Robin's review is the most extensive and critical in this regard. Considering the problem of response set and expectation biases as the most important, he found that few studies attempted to minimize these distortions (Wodarski & Buckholdt, 1975) and gave examples of how the students may be biased in their evaluation of PSI courses under the criterion referenced grading system. Students in PSI courses are more likely to know that they have attained a high grade, and knowing the final grade, may differentially inflate their attitudes in favor of PSI. Novelty of the method and instructor's enthusiasm could also inflate the responses.

Robin (1976) cites a study (Sheldon, Sherman, Wolf, Minkin & Minkin, 1975) that found that attitudes toward Keller plan were less positive on a university wide, standardized questionnaire than on a teacher-made questionnaire. This clearly shows the need for more reliable, standardized instruments. When the standardized questionnaire was used students objected to four aspects of their PSI courses: (a) PSI encourages memorization rather than understanding and application of course materials; (b) there was too little interaction between instructor and student; (c) there was a lack of class discussion; and (d) lack of individual attention to students was detrimental. The results of a number of studies (e.g., McLaughlin, Bushell & Semb as cited by Robin, 1976), have shown that instructors can significantly influence students attitudes. Therefore, another factor to be controlled in attitudinal research would be the instructors' behaviors in courses.

Updating the Literature

Since these review articles were at least three years old, an attempt was made to review the current literature. A computer search utilizing the ERIC data base, Psychology Abstracts and International Dissertation Abstracts was made through the Iowa State University Library and all the major Educational Journals of Social and Physical Science in the years 1976, 1977 and 1978 were reviewed. Two major criteria were used in selecting papers for inclusion in this review: having a common final exam and employing Keller's plan of strategy or some version of this method. As a result, 16 studies were selected. The major features of these studies are summarized in Table 1. The table is as complete and self-explanatory as possible, but it would be instructive to discuss the format and summarize some of the findings in text form.

Strategy

In this section, various features of the method employed by the authors are indicated in table form.

Detailed objectives   In all of 16 reviewed studies there were some sorts of detailed objectives. Here as in the previous literature there was no information about the nature of these objectives.

Units   All of the selected studies have used blocking of the course content, but only eight of these studies (2,5,6,

Table 1. Selected studies on immediate cognitive and affective effects of PSI[a]

Author	Strategy[b]									No. of subjects		Experiment control[c]				
	O	U	M	P	L	SP	TP	IF	AV	Control	Exp	RS	RT	TE	T	TX
1. Boren & Foree (1977)	+	+	80%	-	-	+		-	+	64	109			-		
2. Cote (1976)	+	169%		+		+		-		36	58	-	-	+		+
3. Harrison (1977)	+	+	+	-	-	+		-		134	162	-	-	+		+
	Self-paced competency-based															
4. Hohn et al. (1977)	+	+			+	+				18	18	-	+	+		
	Self-paced instruction															
5. Kulik & Kulik (1976)	+	18	+	+	-	+		+		35	35	-		-		+
	S. Keller (No Motivational L.)															
6. Lu (1976)	+	8	+	+	-	+		+		43	44	+	-	+		+
	S. Keller (No Motivational L.)															
7. McFarland (1976)	+	+	90%	+		+				55	55	-	+			
8. Moretz (1976)	+	15	95%	-	+	+		-		83	28	-	-	+		+
9. Randels et al. (1976)	+	22			-	+			+	82	80	-	-	-	-	+
	Self-paced, programmed with AV															
10. Riedel et al. (1976)	+	10	80%	+		+		+		65	65	-	-	+		+
	S. Keller															
11. Rogers et al. (1977)	+	13	-			+		+		130	57	-	-	-		+
	Self-scheduled Instr. (No Mastery)															
12. Shelden (1978)	+	+	+	-	+	+		-	+			-	-	-		+
	Individual Instruction															
13. Siegfried & Strand (1976)	+	+	+	+	-	+		+		127	26			-		+
	S. Keller (No motivational lecture)															
14. Smiley (1976)	+	+	+			+				?	55	-	-			+
	Keller)															
15. Soper & Thornton (1976)	+	+	-	-	-	+		-		343	115	-	-	-	-	-
	A self-paced system using a standard package of program material															
16. Spector (1976)	+	8	+	+		+		+		60 Sp'74	59	-	+	+		+
	S. Keller (with oral interview)									61	54					

[a]+ means "yes, criterion has met or Experimental group > Control group on that crit
group = Control group"; - means "no, criterion has not been met or Experimental group <
criterion"; No entry means insufficient information.

ects of PSI[a]

	Experimental control[c]								Statistical control[d]					Type of Common Final Exam[e]			Results[f]		Noncomparative Attitudes Results	Resu Mul Compa
xp	RS	RT	TE	T	TX	P	E	M	P	A	GPA	EB	PV	Obj	Sub	St	P	A		+ = -
09			-			+	+	+	+			+							Favorite Results	1 1
58	-	-	+		+	+	+							+			+		81% Favorite 4% Unfavcrite	
52	-	-	+		+	-	+	-	-	-	-	-	-	+			+(.01)	+		
.8	-	+	+			+	+	+					+				=	+		
5	-		-		+	+	+			+				+			+(.01)			
4	+	-	+		+	+	+							+			+(.05)			
5	-	+				+	+	+	+	+	+								Favorite Results	2 3 1
3	-	-	+		+	-	+					+					+			4 5 0
)	-	-	-	-	+	-	+	+	+	+		+				+		+		10 0 0
	-	-	+		+	-	+							+		+	+(.002)			
	-	-	-		+	+	+	+	-	+		+		+			+			
	-	-	-		+	-	+							+			+		Favorite Results	
			-		+	+	+	+	+	+	+	+			+	+	=	+		
	-	-			+		+										+	+		
	-	-	-	-	-	+	+	+	+	+		+				+	-(.05)			0 2 1
	-	+	+		+	+	+	+								+			Favorite Results	

group on that criterion"; = means "[illegible]

Common Exam[e]		Results[f]		Noncomparative Attitudes Results	Results of Multiple Comparisons		Performance Results Studies including Multiple Experiments		Withdrawal	
ub	St	P	A		+ = -	Total	+ = -	Total	Control	Experimental
				Favorite Results	1 1	2				
		+		81% Favorite 4% Unfavorite					18%	19%
		+(.01)	+						20% 27%	13% 21%
		=	+							
		+(.01)								
		+(.05)								
				Favorite Results	2 3 1	6				
		+			4 5 0	9			14.5%	14.2%
	+		+		10 0 0	10				
	+	+(.002)								
		+							2%	19%
		+		Favorite Results						
+	+	=	+							
		+	+							
	+	-(.05)			0 2 1	4				
	+			Favorite Results			2 0 0	2		

7. McFarland (1976)	+ + 90% + +	55	55	-	+			
8. Moretz (1976)	+ 15 95% - + + -	83	28	-	-	+		+
9. Randels et al. (1976)	+ 22 - + + Self-paced, programmed with AV	82	80	-	-	-	-	+
10. Riedel et al. (1976)	+ 10 80% + + + S. Keller	65	65	-	-	+		+
11. Rogers et al. (1977)	+ 13 - + + Self-scheduled Instr. (No Mastery)	130	57	-	-	-		+
12. Shelden (1978)	+ + + - + + - + Individual Instruction			-	-	-		+
13. Siegfried & Strand (1976)	+ + + + - + + S. Keller (No motivational lecture)	127	26			-		+
14. Smiley (1976)	+ + + + Keller)	?	55	-	-			+
15. Soper & Thornton (1976)	+ + - - - + - A self-paced system using a standard package of program material	343	115	-	-	-	-	-
16. Spector (1976)	+ 8 + + + + S. Keller (with oral interview)	60 Sp'74 61	59 54	-	+	+		+

[a]+ means "yes, criterion has met or Experimental group > Control group on that criter group = Control group"; - means "no, criterion has not been met or Experimental group < Co criterion"; No entry means insufficient information.

[b]Strategy: O: Detailed Instructional objectives; U: Units of the subject matter; M: Student proctor; L: Lecture; SP: Self-pacing; TP: Teacher pacing; IF: Immediate Feedback;

[c]Experimental Control: RS: Random Assignment of Subjects; RT: Random Assignment of T Teacher; T: Same Class Time.

[d]Statistical Control: M: Initial Matching by a pretest; P: Personal Data i.e., Sex, Tests; GPA: Grade Point Average; EB: Educational Background; PV: Personality Variables.

[e]Type of Common Final Exam: obj: objective final exam; Sub: subjective common final common.

[f]Results: P: Academic Performance; A: Attitudes Survey.

	35	-		-		+	+	+			+			+			+(.01)			
	44	+	-	+		+	+	+						+			+(.05)			
	55	-	+				+	+	+	+	+	+							Favorite Results	2
	28	-	-	+		+	-	+					+				+			4
	80	-	-	-	-	+	-	+	+	+	+		+			+		+		10
	65	-	-	+		+	-	+						+		+	+(.002)			
	57	-	-	-		+	+	+	+	-	+		+	+			+			
		-	-	-		+	-	+						+			+		Favorite Results	
	26			-		+	+	+	+	+	+	+	+		+	+	=	+		
	55	-	-			+		+									+	+		
	115	-	-	-	-	-	+	+	+	+	+		+			+	-(.05)			0
74	59 54	-	+	+		+	+	+	+							+			Favorite Results	

Control group on that criterion"; = means "Experimental
t or Experimental group < Control group on that

s of the subject matter; M: Mastery criterion; P:
ng; IF: Immediate Feedback; AV: Audio Visual Aids.

; RT: Random Assignment of Treatments; Te: Same

P: Personal Data i.e., Sex, age, etc.; A: Achievement
V: Personality Variables.

ub: subjective common final exams; ST: standardized

Common Exam[e]		Results[f]		Noncomparative Attitudes Results	Results of Multiple Comparisons		Performance Results Studies including Multiple Experiments		Withdrawal	
Sub	St	P	A		+ = -	Total	+ = -	Total	Control	Experimental
				Favorite Results	1 1	2				
		+		81% Favorite 4% Unfavorite					18%	19%
		+(.01)	+						20%/27%	13%/21%
		=	+							
		+(.01)								
		+(.05)								
				Favorite Results	2 3 1	6				
		+			4 5 0	9			14.5%	14.2%
	+		+		10 0 0	10				
	+	+(.002)								
		+							2%	19%
		+		Favorite Results						
+	+	=	+							
		+	+							
	+	-(.05)			0 2 1	4				
	+			Favorite Results			2 0 0	2		

8,9,10,11,16) reported the number of units, ranging from 8 to 22 with a median of 14. From these eight studies one can conclude that compared to the previous studies, the new courses tend to have smaller units (according to Robin, 1976) the median number in his 39 reviewed studies was eight units).

Mastery   Twelve (1,2,3,5,6,7,8,10,12,13,14,16) of the studies (75%) reported the usage of mastery criteria. Five of these 12 studies (1,2,7,8,10) indicated a level of the mastery ranging from 80 to 95 percent. From the remaining four studies, two (11,15) lacked this criteria (13%) and two others (9,4) did not provide sufficient information about the mastery criteria.

Student proctors   Seven (2,5,6,7,10,13,16) studies (44%) have used student proctors in their experiments. Five papers (1,3,8,12,15) mentioned the lack of this feature in their studies and four (4,9,11,14) studies did not provide enough information concerning this feature of the Keller plan.

Lectures   In the original Keller plan, the use of optional motivational lectures is recommended. But seven (1,3,5,6,9,13,15) of the studies (44%) avoided even the motivational lectures. Three studies (4,8,12) employed content lectures in their courses, while six (2,7,10,11,14,16) studies did not clearly describe how lectures were used.

Self pacing Despite some arguments about high withdrawal rates and procrastination due to self-pacing feature of the Keller plan, all of the selected studies used self-pacing in some way.

Immediate feedback According to the "Behaviorism" theory and the original Keller plan, this feature is essential to the method, but only six (5,6,10,11,13,16) of selected studies (38%) employed this feature. Six studies (1,2,3,8,12, 15) reported the lack of this feature and the remaining four (4,7,9,14) did not write about this matter.

Audio-visual aids The standard Keller plan does not depend on the use of audio-visual aids in the system. In the recent studies there is a trend toward employing these aids. Three studies (1,9,12) indicated the use of audio-visual aids.

Experimental control

Lack of sufficient experimental control is a major shortcoming in social science research. All of the major literature reviews on the Keller plan have mentioned this problem. In the studies found for this review the problem of insufficient experimental control still remains.

Random assignment of subjects This criterion is hard to achieve in educational research. In the 16 studies only one (6) reported doing this and it should be noted that the number of subjects in this study is relatively small.

Random assignment of subjects to the treatments Unlike the previous criterion, this feature is not that hard to achieve, but only three (4,7,16) studies have paid attention to this experimental control factor. This problem could be improved by a little effort from the researchers.

Teacher Numerous studies have shown that the use of different teachers may influence the outcome of an evaluation. Seven of the 16 studies (2,3,4,6,18,10,16) employed the same teacher for both the control and experimental groups to avoid the differential effects of the teacher upon the results of the study.

Class time While several studies have indicated that different numbers of class periods for experimental and control groups may lead to different results, none of those reviewed studies have attempted the control for this factor.

Text book The choice of a textbook is a factor that can be easily controlled experimentally but only twelve (2,3,5, 6,8,9,10,11,12,13,14,16) of the selected studies (75%) used the same text book for both groups.

Academic period Academic period is another factor that may have some effects on the outcome of the experiment. There were ten (1,2,4,5,6,7,11,13,15,16) studies (63%) that conducted both the experimental and control methods in the same academic period (either same quarter or semester).

Common final exam In order to compare the academic performance of the two groups, the very least and essential requirement is to have a common final exam. Since this experimental control factor was set as a criterion to select the studies, all of the studies used, at least, this controlling factor.

Statistical controls

Experimental controls are hard to obtain in social science research. This is due to the natural settings of social research situations more than anything else. To overcome this difficulty, social science researchers have used various statistical techniques to control for confounding factors. In the recent years, univariate analyses have been gradually replaced by multivariate analyses in which several confounding factors can be statistically controlled. Of course there are still studies using simple tests to compare the results of control and experimental groups, but more studies are employing multivariate analysis instead of univariate analysis.

Matching the subjects by a teacher-made pretest Random assignment of subjects has always been a difficult task for social researchers. Matching subjects by a pretest is a statistical alternative for random assignment of subjects. In this way the researcher will have some assurance of equality of the subject's academic ability. Eight (1,4,7,9,11,13,15,16) of the selected studies (50%) used this statistical control procedure.

Personal data   Four (1,9,13,15) of the reviewed studies (25%) have considered personal data as a statistical control factor. These personal data have included one or some combination of factors like: age, sex, marital status, outside the field employment, social and economical status, parents' occupation and so on.

Standardized achievement tests   Some researchers believe that standardized achievement tests would do a better job in showing the equal ability of the students in control and experimental groups than teacher-made pretests. Six (5,7,9,11, 13,15) studies (38%) employed standardized achievement tests as a means of matching students' academic ability.

GPA   College or high school grade point average is an indication of the student's academic achievement, and can be used as readily available option to control for the academic achievement of subjects in experimental and control groups. Two (7,13) of the 16 selected studies (13%) used this factor alone or with the other statistical control factors as a means of matching the subjects in the experimental and control group.

Educational background   This term can be applied to many factors including: number of credits or the GPA of the credits in the related field, high school or college cumulative GPA, year in the college, major, high school rank, the results of some achievement or aptitude tests and so on. Six (1,8,9, 11,13,15) of the selected studies (38%) employed one or a combination of the above mentioned factors as an educational background control factor for the subjects.

Personality variables Personal characteristics of the participants in the study can play a substantial role in the results of the study. Due to the difficulties in gathering such information from the subjects, there are not too many studies using personality variables as a statistical control factor. Only one of the 16 studies (8) has used this information.

Type of common final examination Eight (1,2,3,5,6,9, 10,11) of the selected studies (50%) used teacher-made objective common final examination, as an achievement variable and two (7,13) of the selected studies (13%) employed the teacher-made essay type tests. Four (9,13,15,16) of the studies (31%) used some kind of standardized achievement tests to compare the academic performance of the experimental and control groups.

Results

Academic performance Twelve (1,2,3,5,6,8,9,10,11,12, 14,16) studies (75%) reported significant positive results favoring the experimental group, while two (4,13) studies (13%) indicated equal results and only one (15) study (6%) reported significantly positive results favoring the control group.

Attitudes results Four studies (3,4,9,13) used a comparative attitudinal survey and all of them reported results favoring the experimental group. There were also five (1,2,7, 12,16) studies which carried out noncomparative attitude surveys about the experimental method and all of them showed results favoring the experimental group.

Studies with multiple comparisons

Five studies (1,7,8,9,15) used multiple comparisons between experimental and control groups, concerning the academic performance of the subjects in the two groups. The results of these studies showed findings favoring the experimental group. There was one study (16) that conducted two experiments with the results favoring the experimental group.

Withdrawal rates

Four (1,2,3,8) studies included data regarding the withdrawal rates. The results showed approximately the same withdrawal rates for experimental and control groups.

A General Discussion

The main purpose of this updating was to see if there was an overall difference between the more recent studies on the Keller plan and previous studies as summarized in the three major literature reviews. In this general discussion, efforts will be made to compare the various features of these more recent studies with the previous studies.

Strategy

Only one of the three previous major reviews (Robin, 1976) clearly showed the strategy used by each study. In 39 reviewed studies, Robin reported 27 studies (69%), using the standard Keller procedure, while in this updating review only five (5,6,10,14,16) out of 16 selected studies (39%) used the standard Keller procedure. This shows that in the recent years, there has been a tendency to deviate more from the standard Keller plan. One probable reason for this would be the results of numerous studies on component analysis of the Keller plan, which have shown the advantages and disadvantages of the various components of the method. Another trend has been more emphasis on educational technology in the system. The introduction of various audio-visual aids and the usage of computers in keeping records of the student's progress and generating numerous forms of tests are some examples of these efforts. Integration of new educational ideas with new educational technology has always been fruitful and may yield results which accentuate the positive effects of PSI. The common features of the Keller plan used in the present 16 studies are respectively: detailed objectives (100%), small units (100%), mastery (75%), and student proctors (44%).

The number of subjects has always been a relatively controversial topic among the researchers in social science. The problem centers on statistical analysis and the reliability of the findings. Many of the researchers believe that one should have a large number of subjects before one can seriously consider the findings, while others argue that if one has enough subjects he can prove almost any hypothesis (for more technical detail about the number of subjects and testing hypothesis the interested reader is referred to Cronbach & Snow, 1977). Only Robin's (1976) literature reviews indicated the number of subjects used in each study. In his report the number of subjects utilized by the reviewed studies ranged from 17 to 266 with a mean of 78 in the experimental groups. In the control groups, this number ranged from 177 to 659 with a mean of 102. In the papers published and reviewed since then, the number of subjects in the experimental groups ranges from 18 to 162 with a mean of 65, while the number in the control groups ranges from 18 to 343 with a mean of 89. Apparently the number of subjects involved has decreased in more recent studies. One reason for this could be the increasing cost of conducting experiments with large numbers of students. Another factor may be some doubts about the effectiveness of the Keller plan with a larger number of students as expressed by some experts in the field, e.g., Sherman (1974a). Despite this decrease in number of subjects, it should be noted that the number of subjects in experimental and control group is still in a reasonable range.

Experimental control

Seven criteria were chosen as experimental control factors: random assignment of subjects into experiment, random assignment of subjects to treatments, same teacher, same class period, same textbook, same academic period and same final exam. Out of 16 reviewed studies, three controlled for five factors (2,6,16), one (4) controlled for four factors, seven (3,5,7,8,10,11,13) controlled for three factors and five (1,9, 12,14,15) controlled for two factors. The number of controlled factors ranges from two to five with a median of three controlled factors. Comparing these figures with similar data from the major literature reviews, one can conclude that these 16 more recent studies, on the average have controlled for more experimental factors (for such a comparison see Robin, 1976). Even so, experimental control still remains one of the primary shortcomings.

Statistical control

Statistical control is a definite improvement of recent studies over those published previously. Since it is difficult to control the confounding factors experimentally, researchers now are using various multivariate statistical techniques as a substitute for experimental control. There were six distinguishable statistical control factors among the 16 reviewed studies including: matching the subjects by teacher-made pretests, personal data, matching the subjects by standardized achievement tests, GPA as a matching factor, educational background and personality variables. Out of 16 reviewed studies, one (13) controlled for five factors, two (9,15) controlled for four factors, three (1,7,11) controlled for three factors, four (4,5,8,16) controlled only one factor and five (2,6,10, 12,14) studies didn't employ any of the above mentioned statistical controls. A good example among these recent 16 studies using numerous statistical controls is a study conducted by Soper & Thornton (1976). An instructive point in this study was that the authors showed how the usage of multivariate analysis (multiple regression, in this case) could help to clear up the ambiguous results obtained from a simple test comparison. In fact, if they had not controlled for a number of confounding factors, they would have had different results.

Type of common final examinations

All three cited literature reviews indicated the lack of reliable measures on dependent variables. Although the problem remains the same in more recent studies, greater usage of standardized achievement tests as a common final exam was a step towards having more reliable measures. Twenty-five percent of the present reviewed studies used some form of standardized achievement test as a common final exam in their studies.

Results

Academic performance   Comparing 59 studies in major literature reviews with the present 16 cases, one can conclude that the results are almost identical. Specifically 42 (72%) of the previous studies indicated significant positive results favoring the experimental group, nine (15%) have reported equal performance and one (2%) has shown significant positive results favoring the control group. These figures for the present review are respectively 12 (75%), two (13%) and one (6%).

Attitudes results   As in the previous reviews, the results of the present review have shown significant results favoring the experimental group in all of the 16 studies. Therefore, as before, the affective results are the most encouraging findings of the related studies.

Withdrawal rates

According to previous reviews, high withdrawal rates were the main disadvantage of the Keller plan. Only four out of the 16 present reviewed studies included withdrawal rates data, showing almost equal withdrawal rates for experimental and control groups. From the information given in those four studies, it is not possible to say why they had a lower withdrawal rate for the experimental groups than the previous reviewed studies.

Long-Term Effects

Past reviews

Long term effects of instructional innovations have always attracted the attention of researchers in education. After all, educators are interested in the impacts of the educational activities rather than simply in the results of a final exam. If an educational innovation is really effective, its impact should be apparent beyond the end of the academic period and students should demonstrate some long term superiority as well as doing better in the final exam.

Long-term effects can include several factors. Robin (1976) in his review of PSI effects mentioned "retention" and "transfer of learning" to other courses as possible factors. Retention is probably the most explored long-term effect of PSI. Robin reported seven studies which supplemented final examinations with a general follow-up study conducted with samples of the original group at 2 to 24 month intervals with a mean interval of 8.5 months (Austin & Gilbert, 1973; Breland & Smith, 1974; Cole, Martin & Vincent, 1975; Cooper & Greiner, 1971; Corey & McMichael, 1974; Moore, Hauck & Gagné, 1973; DuNann & Weber, 1976). According to Robin, the experimental group had significantly greater retention than the control group in every case.

The question of "transfer of learning" to other courses is more complex, because many other experiences occur concurrently with a PSI course. Robin reviewed two papers in this respect (Moore et al., 1973; Moore and Gagné, note 2). These authors in their first study found that students under PSI group outperformed students in the control group in a second-semester, traditionally taught physics course. However, the same authors were not able to replicate the finding for courses in religion, psychology or biology (Moore and Gagne', note 2). Consequently, these authors suggested that transfer might occur most readily in a discipline where basic material is hierarchically related to advanced material.

Kulik (1976) suggested three kinds of long term effects of the PSI method: retention, transfer and facilitation. According to him, Corey and McMichael (1974), at C.W. Post College of Long Island University, conducted a careful study on retention under PSI: These authors gave a retention test to two random samples of students who had completed an introductory psychology course some 10 months earlier. The results indicated that the experimental group performed significantly better than the control group. Kulik (1976), cited five other studies (Anderson & Artman, 1973; Austin & Gilbert, 1973; Breland & Smith, 1974; Calhoun, note 1; Nazzaro, Todorov & Nazzaro, 1972), with the same results and concluded that students in PSI do something more than "rote learning."

Transfer gain was strong in a study which showed students who finished physics under PSI group outperformed the students in the control group by about an average of one letter grade, (Anderson & Artman, 1973). Three other studies (Lubkin, 1974; McMichael, 1975; Weisberg, 1973) cited by Kulik reported similar results.

Kulik also suggested that studies of the performance of students in conventional courses taken concurrently with a PSI course would indicate whether any facilitation occurred. Schimpfhauser et al. (1974) worked with medical students enrolled in three courses, namely biochemistry, anatomy and physiology. There were two sections in biochemistry, a PSI section and a traditional one. Students in PSI not only performed significantly better than the control group in biochemistry, but they did somewhat better in the other two courses. This study was highly controlled, because the students in both groups were initially identical on such relevant variables as GPA and MCAT scores. From this study, Kulik (1976) concluded that: "while PSI makes average demands on students' time, it gives students more flexibility in organizing their schedules" (p. 6), which supposedly allows achievement improvement in other courses.

In the following sections an effort will be made to cover the long-term studies which were not included in Robin's and Kulik's reviews.

Retention

One of the problems of comparisons which involve final examinations is the differing motives for achievement in PSI and lecture finals. Such examinations usually influence grade to a greater extent in a lecture course, compared to a PSI course. Those problems may be avoided by using retention tests as the achievement measure. But these studies have their own problems as Kulik & Kulik (1975) indicated.

> . . . the investigator has to take into account the additional exposure to course material of students who become proctors; he has to measure retention intervals individually when students complete PSI courses at their own rates; and he has to make herculean efforts to locate and reassemble representative groups of subjects after courses end (p. 229).

Table 2 summarizes studies which will be reviewed in this section.

Table 2. Selected retention studies

Author(s)	Study	Results
Johnson & Walsh (1978)	Self-Scheduled Instruction in Introductory Sociology	+
Mary C. Lu (1976)	The retention of material learned by "PSI" in a Math course	+
Paul H. Lu (1976)	Human growth and development by "PSI"	+
Morris & Surber (1978)	Retention study of a self-paced and an instructor-paced courses	+

Note. (+) means that experimental group did significantly better than control group.

Johnson & Walsh (1978) described their Self-Scheduled Instruction as:

> . . . (SSI), basically entails division of course materials into small sequential units which must be mastered in order, at the student's own pace, with the instructor and student proctors serving as administrators, tutors, and sources of immediate feedback after each unit exam (p. 363).

From their own description, it is obvious that their method is nothing more than standard Keller procedure, with a different name. There were three sections of Sociology 1 involved in their original experiment, two of these sections (labeled as "Computer Lab" and "Traditional") served as control groups. The results of a retention test showed that "SSI" group significantly outperformed the other two groups while there were no significant differences in performance between the two control groups. There were some obvious shortcomings in the original study which the authors indicated:

> Inconveniently three instructors used three different text, emphasized different material in the course and did not all teach sections during the same term (p. 365).

The last problem is an especially negative point in retention studies, because if experimental and control groups do not receive their training in the same academic period, then at the time of "retention testing" the time passed after the learning experience will not be the same for the two groups. This may well invalidate the results of the "retention test."

Another special problem of this study is different emphasis on the course material by the different groups which led the retention test to be only "the lowest common denominator" of the three sections. The question becomes how well this "lowest common denominator" can be regarded as a representative of what students have learned in the different groups. As a whole, the lack of control in the experiment is of such a magnitude as to confound interpretation and prevent generalization.

According to Charles, (1972) retention can be increased by a number of factors such as psychological intention, muscular activity, sensory involvement, periodic practice, emotional involvement and over learning. Mary Lu at Lincoln Memorial University believes that all these factors are major features of PSI and conducted a study to prove that PSI can increase retention.

The retention tests in this study were given in a certain pattern. Lu (1976) indicated:

> The experimental design was adapted from the arrangement for testing retroactive inhibition . . . and consolidation in memory . . . as follows: first period: Phase 1 = Learn A1, Phase 2 = rest, Phase 3 = Recall A1, second period: Phase 1 = Recall A1, Phase 2 = Learn A2; Phase 3 = Recall A1 (p. 151).

The subjects in the two groups were given two retention tests covering material in a beginning mathematics course. Both tests were given while students were enrolled in a second mathematics course. The first test occurred after a five week vacation period while the second occurred after a nine weeks of instruction in the second course in mathematics. The PSI group did significantly better on both retention tests compared to a control group. She also found out that retention of learning is better after a period of rest than after another period of learning activities.

There are some good examples of experimental control in the original experiment of this retention study. Both sections were taught by the same teacher and the textbook was the same for both groups. The main deficiency was the self-selection of the sections by the students. Since the author didn't specify whether the students knew about the experiment in advance or not, serious problems could be introduced due to different characteristics of students who may have chosen different sections. This could cause serious problems in the validity of findings, especially if there is no evidence of initial equivalence of academic ability of students under the two methods. Though the author included personal and educational background data in a table, there was nothing to show that these were actually used to statistically control the results.

A study by Paul H. Lu (1976) is a good example of a well-controlled retention study. The original study was well-controlled experimentally by including controls for teacher, textbook, tests, classroom and random selection of subjects. The students were free to choose any section, but they were not told about the two different methods in advance. Since the selection of subjects was random, there was no attempt to prove the subjects' initial equivalence. However, it should be indicated that showing the initial equivalence is a good way to check the randomness of selection. The results of this study showed that the students under PSI method did significantly better than the students of the control group in a retention test administered five weeks after the completion of the course.

A recent study by Morris and Surber (1978) will be the last one to be reviewed in this section. The authors did not compare the PSI with a traditional group, rather they compared two different versions of PSI. The primary difference between the teaching mode of the two groups was in the test-taking procedures. One group was assigned to a self-paced section in which students were allowed to take the tests at their own rate within the semester's time. The other group was assigned to an instructor paced group in which students were asked to take the tests at a preset period of time. The subjects were freshmen, sophomores, juniors and seniors enrolled in a PSI section of an introductory child development course. Approximately half of these students were randomly assigned to self-paced and half to the instructor-paced section. Factors such as text, course objectives, tests were held constant. As a result the authors didn't find any significant differences between the two groups on withdrawal rate or the end of the semester performance. After nine months a content retention test was administered to a volunteer sample of the original students. This time the self-paced did some what better than instructor-paced group ($p = .094$), although the difference did not reach the preset significant level of the study. The main pitfall of this experiment was its failure to show the representativeness of the retention sample relative to the original students involved in the study.

Transfer

The central question in transfer studies is: "what is the performance of students under PSI in a subsequent traditional course as compared to a control group which takes the same course after having had the prerequisite by the traditional method?" A factor which complicates these studies is that the researcher has no control over the students who took the second course. Consequently, there are no assurances that PSI students who take the second course would not have done better than their associate control group, independent of the fact that they were under PSI. In the view of this reviewer there must be some evidence to show that the students who take the second course are good representatives of the students in the original study; before one can consider the findings seriously. Of course, the ideal situation occurs when all students in both groups take the second course. A second complicating factor is that the researcher has no control over the learning experiences which students in the two groups may have in the interval after completion of the experiment but before starting the subsequent course. Here, the ideal situation occurs when the second course starts immediately after the completion of the original study.

There are several hypotheses to explain the transfer effects of PSI. One hypothesis emphasizes the transfer of course content. Students in PSI will enter the second course with better knowledge of the previous material and therefore will outperform the students in control group in the second course of sequence. Another hypothesis stresses the attitudinal effects of PSI courses. Since students under PSI presumably had a pleasant first encounter with the subject, they would build a sense of competence in the discipline. And finally a third possibility is that good study habits developed by students under PSI would help them to outperform students in control group.

Table 3 is a list of transfer studies which will be reviewed in this section.

The first paper is a progress report on a PSI project at the University of Texas (Stice, 1976). This large project involved 17 PSI courses, 12 of them in the College of Engineering and five in other colleges. The project was well organized with many well known PSI experts serving as advisors to the

Table 3. Selected transfer studies

Author(s)		Results
Stice (1976)	The study is a part of a big "PSI" project conducting several experiments mainly in Engineering College, including several transfer studies.	=,=,=,+
McFarland (1976)	The author evaluated an individualized course in elementary English composition and its transfer effects.	=
Martin & Carlton (1978)	Its a follow-up study on the second year chem course of the students who had their first year course under either PSI or traditional method.	=
Mao-Cohen & Lanson (1976)	A study of the transfer and concurrent effects of PSI in psychology.	+

Note. (=,=,=,+) means that out of four comparisons in this study, there were no significant differences in three cases and one significant difference favoring PSI group.

project. There was no information about the selection process of the subjects or other experimental control factors except the common final exams. There were some statistical controls in the analysis, mainly the educational background factors to control for initial intergroup differences in aptitude or achievement.

The study was designed to answer nine questions about the different aspects of the PSI, such as the questions of "the-end" of the course performance, transfer effects, study habits, the cost of a PSI course and so on. Here, only the transfer effects of this report will be discussed. Students from both PSI and conventional groups in Engineering Statics later took Engineering Dynamics 1 or a Strength of Materials course. There was no significant difference in performance of the two groups in either subsequent course. In another experiment in the same report the PSI graduate students in a library science course were compared to their control group in a subsequent course. Again, no significant difference was found. In mathematics, there were significant long-term differences favoring the PSI group.

The findings of this report confirmed the general conclusion from a previous study on transfer effects of PSI (Moore and Gagné, note 2), which suggested that transfer might occur in a discipline where the material is hierarchically related to advanced material.

The second reviewed study (McFarland, 1976) was designed to evaluate the effects of PSI method for the lowest 10 percent of English composition students. Half of the subjects (55) were assigned to three PSI sections and other half (control group - 55 students) were randomly distributed among 22 traditional sections of the same elementary English course. Beside end of the course evaluation, the performance of the students in two groups was compared in a subsequent English course. Although the PSI group performed slightly better in the second English course, this difference was not statistically significant.

The major deficiency in this study was the formation of the control group, 55 marginal students were distributed among 22 traditional English sections. This can challenge the validity of the findings due to interactions of academic behavior of the control students with the rest of the class. The control students were marginal students distributed among ordinary students, while the experimental groups who were also marginal students, were gathered together in three sections. Lack of proper statistical analysis to show the initial equivalence of the two groups was another problem in this study. The author indicated:

> . . . the control group had a higher mean verbal SAT scores (321) than the experimental group (311, significant at .02 level). . . however, the experimental group had a higher mean GPA for the fall quarter (2.30) than the control group (2.06) significant at the .05 level. On the balance we consider the groups to have roughly equal ability (p. 128).

If the author had used multivariate analysis, he could have controlled for both factors and determined the initial equivalence of the two groups. The author did not specify what percent of the students in the original study, took the second course.

Another study dealt with a sample of students previously enrolled in either a PSI chemistry course or its associated control group at the University of Western Ontario, Ontario, Canada (Martin & Carlton, 1978). The result of a final exam on the second year chemistry course showed that there was no difference between students who had the first year course under PSI as compared to the control group. The original study in the first year chemistry course yielded similar results. The results of a chemistry assessment test indicated that the students in two groups were initially identical in their chemistry backgrounds.

The major pitfall of this study was the self-selection of the sections by the students, since they knew about the two methods in advance. There was another problem which the authors themselves were well aware of and stated:

> The chemistry assessment test shows that this group is drawn primarily from the top 25% of students entering the first year chemistry course. It is unlikely that the mode of presentation of a single course will substantially alter the performance of such a well-prepared group (p. 365).

Finally, 59 experimental and 58 control students in Queens College enrolled in an introductory psychology course were the subjects of a study in transfer and concurrent effects of PSI method (Mao-Cohen and Lanson, 1976). Both sections were taught by the same teacher but in different academic periods. The students were sophomores and their first year college GPA showed that they were identical in this respect. But the experimental group had a significant higher GPA than the control group in the next semester. The authors concluded that the introduction of a PSI course in introductory psychology was responsible for this difference. They did not provide any information about the courses which students took the next semester, but if these courses were from different disciplines, as it is usually the case, it is hard to believe that one course in psychology would have such effects on performance in the different disciplines.

Concurrent or facilitative effects

Opponents of PSI have always claimed that a PSI course is usually too demanding and takes too much of the students' time. The proponents of the method, have tried to show otherwise, by comparing performance in the other courses which PSI and "Non-PSI" students take concurrently with the PSI course. Thus, the central question of these studies is that: "What is the performance of PSI students in conventional courses taken concurrently with a PSI course, as compared to their control group?" Two effects might occur. Students might develop good study habits in the PSI course and these might generalize to other courses taken during the same academic period. Content transfer would be ruled out because students usually take courses from different disciplines during a given academic period; secondly, if students spend too much time on the PSI course, their grades in other courses may suffer.

Table 4 is a list of studies on concurrent effects of PSI method which will be reviewed in this section.

Mao-Cohen and Lanson (1976) compared the GPA of PSI and "Non-PSI" groups in concurrent courses. The PSI group had a

Table 4. Selected concurrent studies

Author(s)	Selected concurrent studies	Results
Mao-Cohen & Lanson (1976)	A study of the concurrent and transfer effects of a PSI course in introductory Psychology.	+
Taylor (1976)	A study of a personalized Instruction course in political science. Concurrent effects were explored among other effects in this course.	+
Eads (1976)	The concurrent effects of a version of PSI course in human anatomy and physiology is among other aspects of this course.	+

significant higher GPA than the "Non-PSI" section. Experimental and control groups were taught in different academic periods which could have serious effects on validity of the positive results for concurrent effects in this study (the other problems of this study was discussed in the last section).

A study by Taylor (1976) tried to evaluate, the Keller plan in an introductory political science course. There was no control group in the study and the author reported the results of three semesters which the course was offered in PSI Format. The main body of the study is a comparison of the PSI course grade with the GPA of all other courses taken concurrently with that PSI course. The results of the three semesters combined showed that the PSI grade of 74% of students was higher than their GPA for all other courses, while the grade for 10% of the students was the same and 16% of the grades was lower.

Although from these results one can not compare the students of a PSI course with the students of the same traditional course, considering each student as his own control, there is weak evidence showing the superiority of the PSI course compared to all other concurrently taken courses. Of course, having positive results with more homogenous PSI and concurrent courses will give stronger evidence about the superiority of the PSI than having the same results with courses from entirely different disciplines. For example, it is much better to compare the grade of a biology PSI course with the GPA of other concurrent life-science courses than to compare the same grade with all other courses from different disciplines.

In a report of introduction of a version of PSI, Eads (1976) described his course as:

> The separate traditional introductory Human Anatomy and Physiology course at the University of Arizona were - integrated and converted to a modularized format in the summer of 1970. The strength of Postlethwait's audio-tutorial system were incorporated into this multisensory individualized approach, and more than 3,000 students have been involved (p. 32).

The paper is a report of some positive results of this course, rather than a report of an experimental study. Four minority groups were identified and the grade of their modularized course were compared to the GPA of other nonmodularized science courses and also to their overall GPA. The data also include a minority and nonminority groups category. Almost all groups (minorities and overall sample) had a higher grade in the modularized course compared to other nonmodularized science courses and this grade nearly always approximated the overall GPA.

Like the previous study (Taylor, 1976), there was no control group in this study and no statistical test was made to see the significance of the results. But a big advantage of the present study over the former is that the author compared the results of a biology course with the other science courses. In this way the reader has some assurance of comparability of the modularized and nonmodularized courses contents. Here, again considering each students as his own control, there are some evidences of superiority of the modularized course.

In summary, long-term studies reviewed in this section revealed some positive long-term effects of the PSI method although, some of these long-term studies were not well designed.

CHAPTER III. DATA ANALYSIS

Step I: File Building

The data file for this study had basically two components: (a) a data set which was gathered and used in the original study (Latta et al., 1978) and (b) a data set which was built by the present investigator. The names of the variables in the two data sets and their descriptions appeared in Chapter I under the "variables under consideration" section.

Step II: Descriptive Statistics and Distribution of the Variables

The purpose of this step was to detect data encoding errors, to examine the descriptive statistics for each variable, and to inspect the distributions of different dependent and independent variables.

By a frequency subroutine in Statistical Package for Social Sciences (SPSS), data which were mispunched were detected and appropriate actions were taken to correct them or to deal with missing values for different variables. The mean, mode, standard deviation, variance, skewness, and kurtosis of all continuous dependent and independent variables were computed. Also the frequency histograms of all continuous variables were graphed in the following groupings:

A. Total group

B. Subtotal groups (PAS and TRAD; males and females)

C. PAS only (PAS males and PAS females)

D. TRAD only (TRAD males and TRAD females)

Inspection of the descriptive statistics and frequency histograms indicated that all the variables were approximately normally distributed except the following (see Appendix B for the frequency bar charts of the independent variables): High School Graduation Rank (HSR), Number of University Credits up to September, 1974 (NCUS), Number of University Credits in Mathematics up to September, 1974 (NMCS), Number of University Credits in Physics up to September, 1974 (NPCS), Number of University Credits in Chemistry up to September, 1974 (NCCS), Number of University Credits in Life-Science up to September, 1974 (NCLS) and Number of University Credits in Life-Science in quarters - Winter, 1975, Spring, 1975, and Fall, 1975 (NCLQ).

Following transformation guidelines recommended by Mosteller and Tukey (1977), the above named skewed variables were re-expressed and the following decisions were made: (a) independent variable (HSR) was transformed by a special square root (SQRT) function ($i_{transformed}$ = SQRT (i - 3) - SQRT (n + 1 - i - 3), where i = the rank and n = number of subjects) and (b) dependent variable (NCLQ) was re-expressed by square root transformation ($X_{transformed}$ = SQRT(X) + 1). The remaining variables could not be made approximately normal by any transformation recommended by Mosteller and Tukey. The main reason for this is that about 80% of the subjects were freshmen in Fall 1974 and did not have any University credits. Therefore, the majority of scores are zero and this would cause a highly skewed distribution which could not be smoothed by any transformation. Since the skewed variables are university educational background control variables, it was hypothesized that these variables would be correlated with the high school educational background and some other variables in the file. A correlational matrix revealed that this was actually the case with the above named variables being highly correlated with one or more control variables in the file. (See Appendix A, Table A for the correlational matrix). Since their inclusion would not add significantly to the different regression models, a decision was made not to include them in the multivariate data analysis.

Step III: Univariate Data Analysis

In this step the Pearson Moment Correlation Coefficients of different dependent variables with independent variables were computed in six categories. (All male and all female, PAS females, PAS males, traditional females and traditional males.) The main purpose of this step was to see the univariate relationships between dependent variables and different independent variables. Looking at the zero order correlation would allow the researcher to see the potential importance of each independent variable on various dependent variables and also would make it possible to detect potential interactions between independent variables (see Appendix A, Tables B, C and D for the correlational matrices).

Step IV: Multivariate Data Analysis

Introduction

In educational research, investigators must often work with students in natural settings, which means that experimental control is often marginal or lacking. Experimental control can take various forms such as: random selection of subjects, random assignment of subjects into treatment levels, stratification of subjects into homogenous blocks and the refinement of techniques for measuring the dependent variable. An alternative approach to reduce the experimental error in order to obtain a less biased estimate of treatment effects involves the use of statistical control.

Statistical control enables an experimenter to remove potential sources of bias from an experiment, biases that are often difficult or impossible to eliminate by experimental control. Statistical control, as described above, can take the form of analysis of covariance which combines the advantages of regression analysis with analysis of variance. The procedure involves measuring one or more concomitant variates (also called covariates) in addition to the dependent variate. By using this analysis the dependent variable can be adjusted so as to remove the effects of uncontrolled sources of variation represented by covariates.

Selection of covariates   Covariates in analysis of covariance should be selected with care. Kirk (1968) presented the conditions of appropriate selection of covariates:

> Covariance adjustment is appropriate for experiments that meet, in addition to the assumptions (of analysis of covariance), the following conditions:
> 1. The experiment contains one or more extraneous sources of variation believed to affect the dependent variable and considered irrelevant to the objectives of the experiment.
> 2. Experimental control of the extraneous sources of variation is either not possible or not feasible.
> 3. It is possible to obtain a measure of the extraneous variation that does not include effects attributed to the treatment. Any one of the following situations will generally meet this third condition:
> (a) The concomitant observations are obtained prior to presentation of the treatment levels,
> (b) the concomitant observations are obtained after the presentation of the treatment levels but before the treatment levels have had an opportunity to affect the concomitant variate, or
> (c) it can be assumed that the concomitant variable is unaffected by the treatment (p. 458).

Main factors and control variables in the models

To cope with the complexity of the situation and using the available data, a general linear mode of multiple regression along with the multiple covariance analysis was adopted for the data analysis. An ad hoc name (analysis of multiple covariance for factorial design) was chosen for the procedures.

There are four classificatory factors in the models, each with two levels: Teaching Method (PAS vs. traditional), Gender (female vs. male), Major (life-science vs. nonlife-science and year in University (Freshmen vs. upperclass). Latta and his colleagues did some preliminary data analysis and found some important variables influencing students' performance in Biology 101 (experimental course). Some of these variables were used in their study (Latta et al., 1978) while others were of a preliminary investigation nature. Two measures of academic ability, a measure of high school educational background and a personality variable were chosen from the above mentioned study. These measures are: high-school graduation rank (HSR) which was obtained from the student's record; scores on a scholastic aptitude test, the Minnesota Scholastic Aptitude Test (MSAT), administered to the students before the experiment started; a composite measure of student's high school credits in mathematics, chemistry, physics and biology (HSBKGD) which was obtained from the student's record; the scores from a modified Mandler and Sarason (1952) Test Anxiety Questionnaire (TAQ) survey administered in Fall 1974; and finally, cumulative GPA of university credits earned by the end of Fall 1974 quarter (excluding the grade in Biology 101 and grades in other life-science courses taken concurrently with Biology 101).

Assumptions of the models

In order for the different F ratios computed in the models to be distributed as the F distribution the assumptions of factorial design should be met. In addition the following assumptions must be tenable: (1) population within-group regression coefficients (β') are homogenous. For example in the context of this model we should have:

$\beta_1' = \beta_2'$ for factor (Teaching Method)

$\beta_1' = \beta_2'$ for factor (Gender)

$\beta_1' = \beta_2'$ for factor (Major)

$\beta_1' = \beta_2'$ for factor (Year in University)

where, for example, β_1' is the population regression coefficient of the first covariate with the first level of the factor Teaching Method (i.e. PAS group) and β_2' is the population regression coefficient of the same covariate with the second level of the same factor (i.e. traditional group). In other words, the regression lines of each covariate for the different level of the four factors should be parallel and (2) the residuals are normally independently distributed (NID) with the mean = 0 and a common variance.

In general, tests of significance in analysis of covariance are robust against the violation of the normality and homogeneity of the residual variance (Kirk, 1968). These tests are also robust against violation of the assumption of homogeneity of within group regression coefficients. For a more detailed discussion about the assumptions of factorial designs and their approximations in the present study, see the Discussion chapter.

The General Approach of the Multivariate Analysis

Although it has been said that F tests in analysis of covariance are robust against the violation of the assumption of the homogeneity of the within group regression coefficients (Kirk, 1968), a conservative approach is that of assuming them to be nonhomogenous and entering the different interactions between covariates and factors into the model. By doing this, one can have a more sensitive and powerful model. Table 5 shows the different terms which were entered into the model.

Entering the interactions between covariates and factors into the model means that regression slopes of different covariates for different level of factors are not equal. Barr and Goodnight (1976) suggested that one might test the significance of these interactions by entering them into the model and running a General Linear Model procedure on Statistical Analysis System (SAS) package. Although a majority of these interactions may not be significant (as it is the case for these models) when they are considered altogether, they can be highly significant and can add significantly to the power of the model. In order to have a more powerful and sensitive model it was decided to leave

Table 5. Terms in the models

Group	Terms
1. (Covariates)	X_1 = TAQ; X_2 = HSR; X_3 = MSAT; X_4 = CUMGPA; X_5 = HSBKGD
2. (Factors)	A = Teaching Method; B = Gender; C = Major; D = Year in University
3. (Interactions of covariates and factors)	X_1*A; X_1*B; X_1*C; X_1*D (TAQ with factors)
	X_2*A; X_2*B; X_2*C; X_2*D (HSR with factors)
	X_3*A; X_3*B; X_3*C; X_3*D (MSAT with factors)
	X_4*A; X_4*B; X_4*C; X_4*D (CUMGPA with factors)
	X_5*A; X_5*B; X_5*C; X_5*D (HSBKGD with factors)
4. (Interactions of factors)	A*B; A*C; A*D; B*C; B*D; C*D (two way interactions)
	A*B*C; A*B*D; A*C*D; B*C*D (three way interactions
	A*B*C*D (four way interaction)

Note. TAQ = Test Anxiety Questionnaire; HSR = High School Graduation Rank; MSAT = Minnesota Scholastic Aptitude Test; CUMGPA = Cumulative University GPA; HSBKGD = High school Science and Math Background.

these interaction terms in the model. For simplicity, a decision was made to put all these interaction terms together and present them in one term called "Interactions." Table 6 shows the format of the models which were used for multivariate analysis.

Table 6. Format of the models

Source	df
Group I (covariates, individually)	5
Group III (interactions of covariates and factors, altogether)	20
Group II (factors, individually)	4
Group IV (all possible interactions of factors, individually)	11

Note. All the terms in Table 6 are the same as those of Table 5. X_1*A, X_1*B . . . X_5*D are pulled together in one term labeled as "Interactions" with 20 degrees of freedom (there are 20 of such interaction terms, i.e., 4 for each covariate).

Analysis of main effects and interactions

Table 6 shows the format of the models which were used to perform multivariate analysis of data in the context of "multiple covariance analysis for the factorial design." In this way one could test the effect of the various factors after adjusting for all covariates. For example, one could test whether, after adjusting for the covariates, teaching method (factor A in the table) had any significant effect on the criterion measure. In other words, analysis of main effects would indicate whether there were any significant differences in mean performance of the students who had the beginning biology under different teaching methods, after adjusting for the covariates.

Analysis of the main effects in factorial designs is not always enough. One can have a better understanding of the data by analysis of interactions among the factors. By definition, two factors interact when the effect of one factor changes at different levels of the second factor. In the context of the present models, by analyzing the interaction between two factors, say, Teaching Method and Gender (factors A and B in Table 1) it is possible to see if these two factors interact after adjusting for all covariates.

Analysis of simple main effects

After testing an (adjusted) interaction and finding it significant, the next step is to study the pattern of that interaction. This can be done by comparison of the simple main effects and performing tests of significance for them. By definition, a simple main effect is the effect of one factor at a specific level of the other factor. For example, in the present models let us consider the two factors, Teaching Method and Gender. Teaching Method has two levels (PAS vs. TRAD) and Gender also has two levels (FEMALE vs. MALE). Therefore there are four criterion measure means and each mean is referred to as a simple main effect. Table 7 shows these groups.

Since the nature of this study was an investigatory rather than hypothesis testing, there were no preset hypothesis tested. An attempt was made to explore the data as much as

Table 7. Simple main effects of factors A and B

	FACTOR A (Teaching Method)	
Factor B (GENDER)	PAS	TRAD
MALE	Group 1	Group 3
FEMALE	Group 2	Group 4

possible. Also, for the same reason, there was no preset α level, instead all the probabilities (p values) were reported.

CHAPTER IV. RESULTS

Long Term Effects of PAS on Students' Achievement

The first question to be answered was: "How did the students from the PAS section do in subsequent life-science courses as compared to the students from traditional section?" These courses were taken during Winter, Spring and Fall Quarters of 1975, *i.e.*, an interval of one academic year after taking the experimental course. The Grade Point Average for life-science courses taken within the three quarters (GPALQ) was chosen as the criterion measure. Differences among individual students in the two sections were controlled by entering selected educational background and academic ability and personality variables discussed in the introduction chapter. A general linear model of multiple regression was developed. The General Linear Model procedure of the Statistical Analysis System (SAS), was used to obtain Tables 8 and 9.

Total variance accounted

The first thing to be noticed in Table 9 is that 54 percent of the total variance of the criterion measure is accounted for by the model ($R^2 = .54$) which is relatively high for educational research. This means that using all terms in the model, it was possible to predict more than half of the variability in the criterion measure while less than half of

Table 8. Analysis of multiple covariance for dependent variable: Grade Point Average of Subsequent Life-Science Courses (GPALQ)

Source	df	Partial sum of squares	F	P
TAQ	1	12870.09	3.56	.06
HSR	1	2748.49	.76	.38
MSAT	1	198.83	.06	.81
CUMGPA	1	741.87	.21	.65
HSBKGD	1	740.06	.20	.65
Interactions	20	121389.37	1.68	.20
Teaching Method (A)	1	10764.86	2.98	.08
Gender (B)	1	131.48	.04	.89
Major (C)	1	10432.89	2.89	.09
Year in School (D)	1	39.48	.01	.91
A*B	1	18868.69	5.22	.02
A*C	1	3519.01	.97	.32
A*D	1	17641.12	2.11	.15
B*C	1	5554.51	1.54	.22
B*D	1	852.21	.24	.63
C*D	1	6649.51	1.84	.18
A*B*C	1	425.45	.12	.73
A*C*D	1	1524.32	.42	.52
A*B*D	1	8943.65	2.47	.12
B*C*D	1	995.42	.28	.60
A*B*C*D	1	392.70	.11	.74

Note. TAQ = Text Anxiety Questionnaire; HSR = High School Graduation Rank; MSAT = Minnesota Scholastic Aptitude Test; CUMGPA = cumulative GPA; HSBKGD = High School Background (high school credits in science and math); Interaction = the interactions of the five covariates with the four factors, pulled together.

Table 9. Analysis of variance for the model

Source	df	Sum of squares	Mean squares	F	P	R^2
Model	40	1115193.79	27879.85	7.71	.0001	.54
Error	261	943289.95	3614.14			
Corrected total	301	2058483.74				

Note. P = Probability of the F test, i.e., the significance level of the test of the model: R^2 = total variance accounted.

the variability in the criterion remained unpredictable due to some uncontrolled variables.

Analysis of main effects

Analysis of main effects of a factor in the context of an analysis of multiple covariance reveals the effect of that factor after adjusting for all covariates. By a "Least-Square Means" option within General Linear Model Procedure of the "SAS" package the adjusted main effects of the four factors involved in this model were obtained. These results along with the P-values (significance level) for comparisons between the different levels of the factors are given in Table 10.

From Table 10 the following conclusions may be drawn: (1) students from the PAS section had a higher mean performance in subsequent life-science courses after adjusting for the

Table 10. Adjusted main effects of the factors

Level	Adjusted Criterion means	Prob·level (P)
(1) TEACHING METHOD		
PAS	2.65 (N = 156)	.08
TRAD	2.40 (N = 146)	
(2) GENDER		
MALE	2.54 (N = 144)	.85
FEMALE	2.51 (N = 158)	
(3) MAJOR		
Nonmajor	2.40 (N = 156)	.08
Majors	2.65 (N = 146)	
(4) YEAR IN UNIVERSITY		
Freshmen	2.52 (N = 268)	.91
Upperclass	2.54 (N = 34)	

covariates, the P-value for this comparison being .08; (2) the difference between males and females performance in subsequent courses seems to be trivial (P = .85); (3) students who were majoring in life-science did better than nonmajors in subsequent courses (P = .08) and (4) the difference between the freshman and upperclass students is due to a chance event (P = .91).

Analysis of interactions

Often, in factorial designs a better understanding of the data can be obtained by analyzing the interactions between factors. Therefore, all possible interactions (two-way, three-way and four-ways) between the four factors were entered into the model. The results are reported in Table 8. The only interaction which is significant in this table is the interaction between Teaching Method and Gender (P = .02). The significance levels (P-values) of other interactions range between .12 and .74. A significant interaction between Teaching Method and Gender in the context of the analysis of multiple covariance means that after adjusting for the control variables, the effect of the Teaching Method on subsequent performance was not the same for males and females. To clarify the pattern of this interaction statistically, comparisons between the adjusted simple main effects of these two factors were performed.

Table 11 shows the simple main effects of the two factors, Teaching Method and Gender.

A graph can better conceptualize the pattern of these adjusted means (Fig. 1).

Table 11. Adjusted simple main effects of Teaching Method and Gender

Groups	N	Adjusted Criterion Means
(1) PAS-MALES	76	2.46
(2) PAS-FEMALES	80	2.84
(3) TRAD-MALES	68	2.62
(4) TRAD-FEMALES	78	2.17

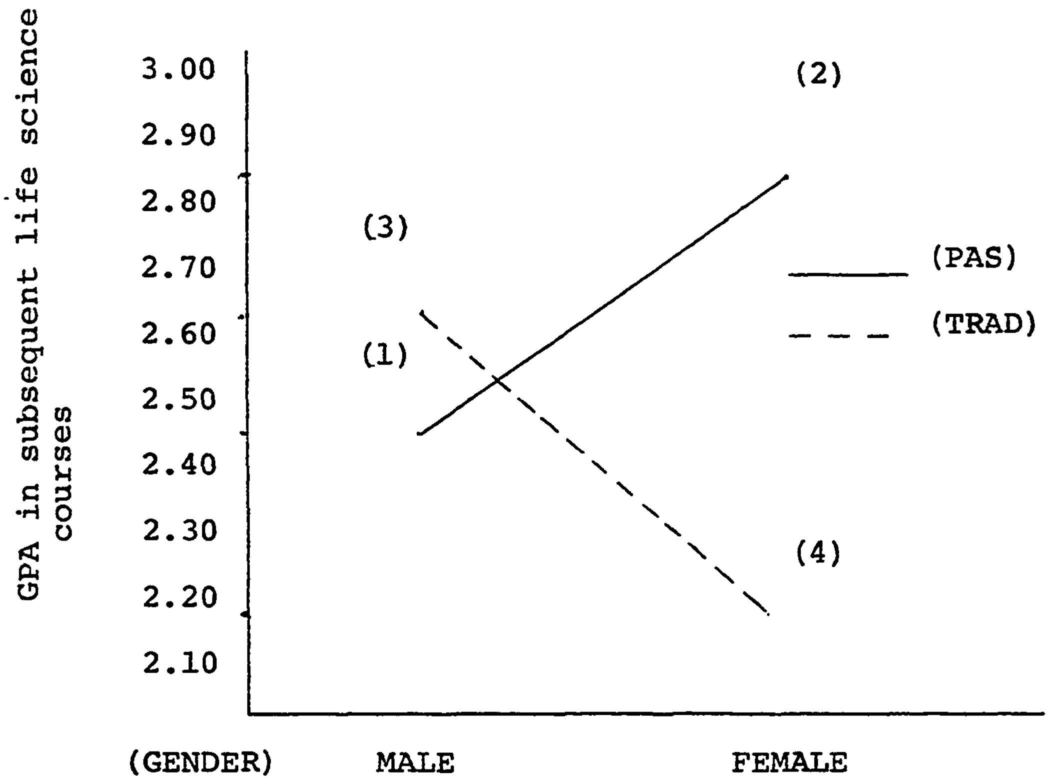


Figure 1. Adjusted simple main effects of factors Teaching Method and Gender

It is apparent from Figure 1 that female students from the PAS section (group 2) had the highest adjusted GPA in subsequent life-science courses. The difference between the adjusted mean performance of female students from PAS method and female students from Traditional section (group 2 vs. group 4) is significant at .001 level (*P* = .001). The difference between the adjusted mean performance of the male students from the two sections (group 1 vs. group 3) is trivial (*P* = .53). Other interesting aspects of these results are the gender differences in performance as a result of the treatments. Male students from traditional section did better than the females from the same section (group 3 vs. group 4) but the *P*-value for this comparison is only .18 which can be regarded as a non-significant probability. The same comparison on students who had PAS showed that females outperformed the male students (group 2 vs. group 1) in subsequent courses with a probability of .01 (*P* = .01).

Simple main effects of factors Teaching Method and Major

Although the interaction between the factors, Teaching Method by major, is significant at only the .32 level, comparison of pairs of adjusted mean may give some insights about the data. Table 12 shows the adjusted simple main effects of these factors.

Table 12. Adjusted simple main effects of factors Teaching Method and Major

Groups	N	Adjusted Criterion Means[a]
(1) PAS-NONMAJORS	79	2.60
(2) PAS-MAJORS	77	2.70
(3) TRAD-NONMAJORS	77	2.20
(4) TRAD-MAJORS	69	2.59

[a]Figure 2 shows the above adjusted means graphically.

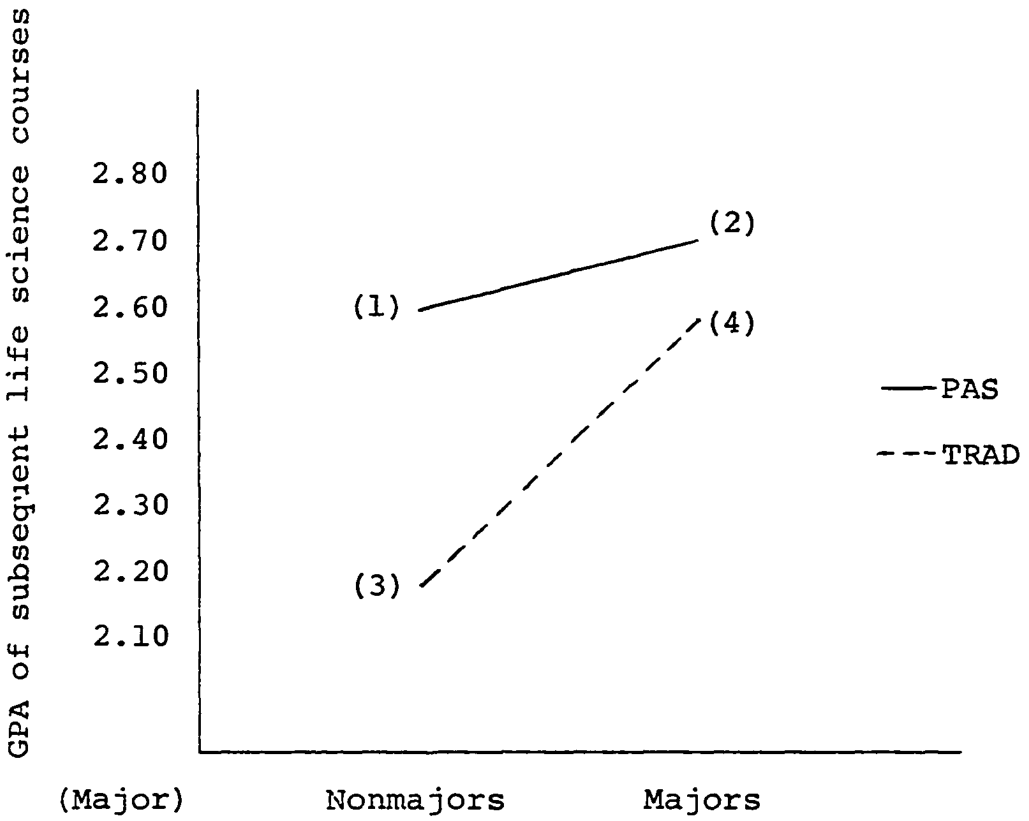


Figure 2. Adjusted simple main effects of factors Teaching Method and Major

From Figure 2 it is apparent that nonmajor students in PAS (group 1) did much better than TRAD nonmajors (group 3). The significance level associated with this comparison is .01. The difference between major students in the two methods (group 2 vs. group 4) is trivial ($\underline{P}$ = .67). Within methods this comparison shows another interesting point: while the difference between nonmajors and major students from PAS (group 1 vs. group 2) is trivial ($\underline{P}$ = .58) the same comparison for students from the traditional method (group 3 vs. group 4) shows a significance level of .09. This fact clearly shows that nonmajor students did better as a result of having had Biology 101 by the PAS method. Both major and nonmajor students from PAS did better in the subsequent courses than students from traditional courses.

Simple main effects of factors Teaching Method, Gender and Major

Although the three-way interaction between Teaching Method, Gender and Major is not significant, the two-way interaction between Teaching Method and Gender is highly significant. Consequently, insights may be gained by examining pairs of adjusted mean. There are eight groups involved in this analysis and Table 13 shows the adjusted simple main effects of the involved eight groups.

Table 13. Adjusted simple main effects of the factors Teaching Method, Gender and Major

Groups	N	Adjusted Criterion Means
(1) PAS-MALE-NONMAJORS	42	2.53
(2) PAS-MALE-MAJORS	28	2.39
(3) PAS-FEMALE-NONMAJORS	31	2.67
(4) PAS-FEMALE-MAJORS	49	3.00
(5) TRAD-MALE-NONMAJORS	44	2.49
(6) TRAD-MALE-MAJORS	24	2.75
(7) TRAD-FEMALE-NONMAJORS	33	1.91
(8) TRAD-FEMALE-MAJORS	45	2.44

Notice that the highest adjusted mean performance (3.00) belongs to female students who were majoring in life-science and experienced the PAS method (group 4). Nonmajor female students from traditional section (group 7) had the lowest adjusted mean performance in subsequent courses. The trends of the above data are easier to understand graphically. These data are depicted in Figure 3.

The first trend to notice is the difference between the adjusted mean performance of male and female students who were majoring in life-science and compare this difference in the two methods (group 2 vs. group 4 in PAS and group 6 vs. group 8 in TRAD). Among the students who had PAS, major females

Figure 3. Adjusted simple main effects of factors Teaching Method, Gender and Major

did much better than male majors (P = .01). Among those who had the Traditional method, male majors had a higher mean performance than female majors but the difference can be due to chance happening (P = .55). Comparison of major female students from the two methods (group 4 vs. group 8) showed that students from PAS had a better adjusted mean performance than the same students from Traditional method (P = .06). Performing the same comparison for the male majors (group 2 vs. group 6) revealed that although this group had a higher mean performance as a result of the Traditional method, the difference is somewhat trivial (P = .42).

By performing the same comparisons for the nonmajor students from the two methods one can conclude that: (1) in the Traditional group, male students outperformed the females P = .03; (2) in PAS, females had a higher mean than males but the difference is trivial (P = .49); (3) females in PAS outperformed females from Traditional (group 3 vs. group 7) with a P-value of .001 and (4) the mean performance in subsequent courses of the male students from PAS was almost identical with the mean performance of those from the Traditional method (P = .85).

From these trends one may conclude that in both major and nonmajor categories, females from PAS outperformed females from Traditional section. While within the PAS method, females did much better than male students, within the traditional group the opposite was found. Female students majoring in life-science from PAS had the highest adjusted mean performance and it seems that this group gained the most long-term benefits from taking the introductory core course, Biology 101, by the PAS method.

PAS Effects on Students' Persistence in the Field

The second question of interest was to see if exposure to the PAS method had any effect on students' persistence in the field of life-science. The number of credits of life-science taken in the one academic year period (quarters W '75, Sp '75 and F '75) was chosen as the criterion measure for this analysis. Here again, a decision was made to enter educational background, academic ability, and personality variables in order to control for the individual differences control variables as discussed in introductory chapter. A General Linear Model Procedure of "SAS" package was used. The results are reported in Table 14. An analysis of variance for the model is also included in Table 15.

Total variance accounted

The total variance ($\underline{R}^2$) accounted for by the model in Table 15 is only .28. In other words, the data only accounted for 28 percent of the variation in the criterion measure. This means that information is missing and that 72% of the variability of the criterion measure cannot be accounted for by the model. This low $\underline{R}^2$, probably accounts for the nonsignificant results obtained with this model.

Table 14. Analysis of multiple covariance for dependent variable Number of Life-science Credits Taken in the One Year Period (NCLQ)

Source	*df*	Partial sum of squares	*F*	*P*
TAQ	1	.73	.61	.44
HSR	1	4.00	3.37	.07
MSAT	1	.00	.00	.99
CUMGPA	1	11.02	9.26	.002
HSBKGD	1	3.11	2.62	.11
Interactions	20	27.96	1.17	.30
Teaching method (A)	1	.01	.00	.98
Gender (B)	1	2.96	2.49	.11
Major (C)	1	36.41	30.58	.0001
Year in University (D)	1	1.36	1.14	.29
A*B	1	.39	.33	.57
A*C	1	.11	.10	.76
A*D	1	.01	.01	.92
B*C	1	1.14	.96	.33
B*D	1	.06	.05	.82
C*D	1	2.85	2.39	.12
A*B*C	1	.83	.70	.40
A*C*D	1	.06	.05	.82
A*B*D	1	2.88	2.42	.12
B*C*D	1	1.84	1.64	.22
A*B*C*D	1	3.08	2.59	.11

Note. TAQ = Test Anxiety Questionnaire; HSR = High school Graduation Rank; MSAT = Minnesota Scholastic Aptitude Test; CUMGPA = Cumulative GPA; HSBKGD = high school background; Interactions = 20 interactions of covariates with the factors (see statistical procedure chapter).

Table 15. Analysis of variance for the model

Source	df	Sum of squares	Mean square	F	P	R^2
MODEL	40	152.70	3.82	3.21	.0001	.28
ERROR	329	391.70	1.19			
CORRECTED TOTAL	369	544.40				

Analysis of main effects

By a "Least-Square Means" option within General Linear Model procedure of "SAS" package the adjusted main effects of the four involved factors were calculated and are reported in Table 16.

From Table 16 one may conclude that: (1) after adjusting for the covariates the average number of life-science courses taken by students from PAS section within the one year period is almost identical with the average of those students in traditional section ($P = .97$); (2) after controlling for the covariates, females took more life-science courses than males ($P = .11$); (3) as expected students majoring in life-science had more credits in the field than nonmajor students ($P = .0001$) and (4) freshmen and upperclass students did not differ as far as the criterion measure was concerned ($P = .29$).

Table 16. Adjusted main effects of the four factors

Level	Adjusted Criterion Means	Sig·level (P)
(1) TEACHING METHOD		
PAS	8.23 (N = 184)	.97
TRAD	8.27 (N = 186)	
(2) GENDER		
MALE	7.14 (N = 172)	.11
FEMALE	9.45 (N = 198)	
(3) MAJOR		
NONMAJOR	4.97 (N = 215)	.0001
MAJORS	12.28 (N = 155)	
(4) YEAR IN UNIVERSITY		
FRESHMAN	7.58 (N = 316)	.29
UPPERCLASS	8.92 (N = 54)	

Analysis of interactions

All possible interactions between the four factors were entered into the model and reported in Table 8. Their significance levels range from .11 to .92. None of these P-values were considered important enough to perform analysis of interactions.

Facilitative Effects of PAS on Students' Performance in Concurrent Life-Science Credits

The next question of interest was to see whether the introduction of PAS method had any effects on other life-science courses taken concurrently with Biology 101 (experimental course) in Fall 1974. In other words the question becomes: "Did the PAS method help students to learn more in concurrent life-science courses?" Grade Point Average of the concurrent life-science credits (GPACCL) was chosen as the criterion measure for this analysis. As in the other two analyses, the same educational background academic ability and personality variables were controlled. A General Linear Model Procedure of "SAS" package, was used and the following results were obtained and are reported in Tables 17 and 18.

Total variance accounted

In this analysis, the total variance accounted for by the model ($\underline{R}^2$) was .49 or almost half of the variability in the students' performance on concurrent life-science credits. The other nonaccounted half is apparently due to some other uncontrolled variables which were not included in the present data base.

Analysis of main effects

By a "Least-Square Means" option within the General Linear Model procedure of the "SAS" package, it was possible to calculate

Table 17. Analysis of multiple covariance for dependent variable: Grade Point Average of Concurrent Life-Science Credits (GPACCL)

Source	df	Partial sum of squares	F	P
TAQ	1	3397.35	.68	.41
HSR	1	1071.85	.22	.64
MSAT	1	2959.91	.59	.44
CUMGPA	1	18341.73	3.68	.06
HSBKGD	1	864.97	.17	.68
Interactions	19	190877.66	.52	.46
Teaching method (A)	1	1437.78	.29	.59
Gender (B)	1	25660.41	5.15	.02
Major (C)	1	6389.90	1.28	.25
Year in university (D)	1	17987.26	3.61	.06
A*B	1	159.17	.03	.85
A*C	1	244.51	.05	.82
A*D	1	649.53	.13	.72
B*C	1	353.12	.07	.79
B*D	1	21216.83	4.26	.04
C*D	1	4398.00	.88	.34
A*B*C	1	4678.76	.94	.33
A*C*D	1	840.76	.17	.68
A*B*D	1	31.83	.01	.94
B*C*D	1	6505.66	1.31	.25
A*B*C*D	1	712.25	.14	.71

Note. TAQ = Test Anxiety Questionnaire; HSR = High school Graduation Ranking; CUMGPA = cumulative GPA; MSAT = Minnesota Scholastic Aptitude Test; HSBKGD = High school Background (credits in high school science and math); Interactions = interactions of 5 covariates and 4 factors (see the statistical procedures chapter).

Table 18. Analysis of variance for the model

Source	df	Sum of squares	Mean squares	F	P	R^2
MODEL	39	748965.18	19204.24	3.86	.0001	.49
ERROR	156	776547.98	4977.87			
CORRECTED TOTAL	195	1525513.16				

with adjusted main effects of factors involved in this multiple covariance model, and see that, for example, after adjusting for all covariates what would be the criterion measure mean performance of students in the two sections? Table 19 shows the adjusted main effects of all four factors in the model.

From Table 19 one may conclude that: (1) although the PAS students may have done better in concurrent life-science courses, after adjusting for covariates the P-value for this comparison was only .59; (2) female students outperformed males at P = .02 level; (3) major students did better than nonmajors (P = .25) and (4) freshman did better than upperclass students (P = .06).

Analysis of interactions

From Table 17 the only promising interaction is the interaction between Gender and Year in University (P = .04).

Table 19. Adjusted main effects of factors of the model

Level	Adjusted Main Effect	Sig·level (P)
(1) TEACHING METHOD		
PAS	2.30 (N = 95)	.59
TRAD	2.18 (N = 101)	
(2) GENDER		
MALES	1.91 (N = 77)	.02
FEMALES	2.57 (N = 119)	
(3) MAJOR		
NONMAJOR	2.11 (N = 86)	.25
MAJOR	2.37 (N = 110)	
(4) YEAR IN UNIVERSITY		
FRESHMAN	2.49 (N = 171	.06
UPPERCLASS	2.00 (N = 25)	

Table 20 shows the adjusted simple main effects of factors Gender and Year in University.

From Table 20 one may conclude that male students in upper classes had the lowest mean performance in concurrent life-science credits. However, the adjusted criterion mean of upper-

Table 20. Adjusted simple main effects of factors Gender and Year in University

Groups	N	Adjusted Criterion Means
(1) FRESH-MALE	70	2.46
(2) UPPER-MALE	7	1.35
(3) FRESH-FEMALE	101	2.51
(4) UPPER-FEMALE	18	2.63

Note. FRESH = students who are freshmen. UPPER = students who are in upper classes of the university.

class male students in Table 20 could be misleading because of small number of the students in this group (for a technical discussion of this point see Cronbach and Snow, 1977). For the mentioned reason and also because this analysis will not serve to clarify the central question of this section, it was decided not to compare the pairs of adjusted means of Table 20.

A Summary of Results

There were three main questions explored in this part: (1) long-term effects of PAS on students' achievement; (2) PAS effects on students' persistence in the field of life-science and (3) facilitative effects of PAS on students' performance in concurrent life-science credits. Three multiple covariance models were developed to answer these questions. Significant results was found for the first question where female students who had Biology 101 under the PAS method had significantly better long term achievement results in subsequent life-science courses. Models developed for the second and third questions revealed no significant results to answer these two questions.

CHAPTER V. DISCUSSION

The Original Study and Guidelines of the Present Study

Since part of the data for this investigation came from an earlier study (Latta et al., 1978), it is necessary to examine the original experiment in juxtaposition with the criteria developed to judge the studies included in the literature review chapter.

Latta and his colleagues used approximately 200 students in each treatment. Having a large number of subjects is especially helpful in studies with complex evaluation models and it contributes to the reliability of the employed statistical procedures. Although the selection of subjects was not random, to make the design at least quasi-experimental, the teaching method was randomly assigned to the two sections and the equivalency of sections was demonstrated. The most interesting feature of this study was its attention to individual differences by recognizing academic ability, educational background and personality variables as factors and entering them into the evaluation models. This feature helped the authors to explore an aptitude-treatment interaction and made it possible to be specific about the question of "What type of students benefit most from what type of instruction method?" The aptitude-treatment interaction is one of the most important problems which should be considered in future educational research (Cronbach & Snow, 1977). Factors like instructor, text and course objectives were the same for both groups to prevent their confounding effects upon the treatment results. The major shortcoming of this study was its criterion measure (Grade). Although there was a common comprehensive examination, this examination was differentially weighted for the two sections and could not be used. In PAS the common examination had little weight in determining the student's grade, while in the traditional system, its result comprised one-fourth of the student's grade. This differential weighting favored the control group though there were no significant differences on score between methods.

The present study also adopted an individual differences model but explored the effects of the self-paced testing method beyond the end-of-the-quarter examination. The goal was to find out that what type of students benefitted from the PAS or the traditional methods over a longer period of time. Although the general goals of the present study were the same as the earlier study, the criterion measures were different. These criterion measures included (1) students' performance in subsequent life-science courses, taken during the next academic year period; (2) number of life-science credits taken by the students in the two sections during the next academic year period and (3) experimental method effects on grades earned in life-science courses taken concurrently with Biology 101.

Some Notes About the Statistical Procedures

In an investigatory study, such as the present one, the available data often determine the data analysis procedures. A general linear mode of data analysis was adopted because no other analysis options would cope with the complexity of the situation and be statistically proper. For reference purposes the *ad hoc* name, "multiple covariance analysis for factorial design," was given to the procedures used.

Two measures of academic ability (HSR and MSAT), a measure of high school educational background (HSBKGD) along with a personality variable (TAQ) were chosen from the available data because of their demonstrated importance in prediction of academic performance as shown by the authors of the original study. A measure of university educational ability (CUMGPA) was formed by the present investigator and also added to the models. These variables were used as covariates and mainly served to control for individual differences between students and consequently to reduce the error terms of the models in the study. Other terms in the models included an experimental factor, *i.e.*, the two sections of the course. Gender of the students was added as a classification factor since gender-treatment interaction was found to be significant in the original study and also sex differences in academic performance are well proven in educational research. Major was used as another classificatory factor because it was rationalized that students majoring in life-science fields would perform better than nonmajor students. Finally, year in the university was also considered because the students with more university experience tend to perform better than students with less university experience. The main concern in choosing the co-variates and classification factors was to select variables which would increase the precision of the models through the reduction of the error term.

Complex statistical models require rather rigorous assumptions to be statistically reliable. In fact, the researcher is faced with a dilemma. If simple statistical procedures and models are chosen which need few assumptions, information is lost and even erroneous results may be obtained. Alternatively, if complex models are chosen which are more meaningful and represent the complexity of educational activities, then extensive and hard to achieve assumptions must be met. To cope with the complexity of the situation and using available data, the present investigator chose the second option and did his best to approximate some of the required assumptions. Fortunately the procedures used are robust against slight violations of the most of the required assumptions (Cochran, as cited by Kirk, 1968). In the following section these basic assumptions are stated along with the steps taken to approximately meet the requirements.

Employing a linear model for data analysis implies that the relationship between the dependent variables and different independent variables are, in fact, linear and additive. In social sciences, linear models are adopted for approximate explanations of the data, though it is known that model requirements are seldom met. Adding interaction among factors to the models, however, helps to approximate the requirement of additivity.

Independence of errors is another basic assumption underlying the general linear models. It may be assumed that these errors are independent if the subjects are randomly assigned to the treatment levels. This condition was met when the original experiment was conducted (Latta et al., 1978). Further, it is assumed that all observations are normally distributed within each treatment population. To approximate this assumption a series of transformations were carried out for highly skewed variables. The last common requirement of the general linear model is that the variance due to experimental error within each treatment population is homogenous. Again, due to nature of measurement scales in social science, this assumption is difficult to meet, but fortunately the F-tests in the models are considered robust against violation of this feature (Kirk, 1968).

In addition to the above common basic assumptions of general linear models, there is a facet of the additivity requirement which is special for the models which use covariance. This is the assumption of equal slopes. It is assumed that the regression slopes of each covariate for the different levels of each factor are equal. In other words, the regression lines of each covariate for the different levels of each factor are parallel. The interactions of each covariate with each factor were entered into the models to fulfill this requirement by recognizing unequal slopes as variables. Inclusion of these interactions (unequal slopes) improved the sensitivity and power of the models, but added 20 terms to the models. It is possible that these interactions, while individually not significant, collectively reduce the error term significantly.

Another difficulty associated with the large statistical models is a phenomenon called "collective alpha." In the simplest terms, this means that performing too many significant tests in a given model will increase the probability of making a type I error, that is, rejecting a hypothesis which is actually true. This is a serious problem in studies with a series of preset hypotheses, but not in a explanatory investigation such as the present one in which most of the terms in the model serve as controls and there is no intention of testing hypothesis about these control terms. Most of the terms in the models of the present study, including the five covariates and the 20 interactions of these covariates with the factors are comparable to the blocks of a randomized block design (repeated measures design). In this design, each subject or a number of homogeneous subjects is(are) considered as a block to minimize the individual differences and increase the power of the design. Like the present models, these designs usually have large degrees of freedom.

Another point should be mentioned about the analysis of covariance employed. The covariates were chosen to control for individual student differences in the two sections rather than equating the initial differences of the subjects since the groups were equated through random assignment. Wolins (1976) in an article showed that the usage of covariates for equating purposes in social science could result in erroneous conclusions. His main argument centered around the nonpreciseness of the measurement scales in social sciences.

Yet there is another statistical justification for large models. Due to the relatively large number of subjects (302, 370 and 196 respectively for the three models of the study), there would be enough degrees of freedom left for the error terms to be statistically reliable despite 40 degrees of freedom used for the models. Of course models with large degrees of freedom and small number of subjects are not statistically reliable because there would not be enough degrees of freedom left for the error term.

In summary, the limitations of the statistical techniques used for the data analysis were recognized and steps were taken to approximate the required assumptions of the procedures. Many of these steps contributed to the complexity of the models which to some readers may foster doubts about the reliability and meaningfulness of these models. It should be noted that in developing the statistical models, individual differences, sensitivity of the models and approximation of the underlying assumption have been given equal consideration in the investigator's mind. Consequently the resultant models were felt to best fill the needs of all of these considerations.

Summary and Comparisons of the Results

Before examining the results there is one point which should be clarified. These results were heuristic in nature which means they were hypothesis forming rather than hypothesis testing. The results of the three main questions will be discussed respectively in the following section.

Transfer effects

The results show that students who had biology under PAS as a whole did better than traditionally taught students in subsequent life-science courses ($\underline{P} = .08$). This finding is comparable to the findings of other studies in the literature. Robin (1976) cited two studies in transfer effects of PSI (Moore, Hauck & Gagné, 1973; Moore & Gagné, note 2). These authors, in their first study found that students from PSI outperformed students from their control group in a second semester traditionally taught physics course. However, two of these authors (Moore & Gagné, note 2) were not able to replicate the findings for courses in religion, psychology or biology. Kulik (1976) discussed several studies in this respect (Anderson & Artman, 1973; Lubkin, 1974; McMichael, 1975; Weisberg, 1973), all which showed favorable results for experimental groups. Some additional transfer studies were found by the present investigator. One of these studies (Mao-Cohen & Lanson, 1976) in psychology showed significant positive transfer effects favoring the experimental group, while two studies (McFarland, 1976; Martin & Carlton, 1978) reported equal results. Finally, a progress report of a large PSI project with several experiments mainly in Engineering College of the University of Texas (Stice, 1976) was reviewed. One of the reported experiments resulted in positive transfer effects for the experimental group, while three other experiments showed no differences in students' later performances.

None of these studies went beyond the comparison of the two sections as a whole group nor did they control for individual student differences. In the present study, by taking advantage of multivariate statistical procedures, it was possible to analyze the transfer effects of the method upon particular subgroups of the students, *e.g.*, female students who were majoring in life-science and who had the PAS method, while controlling for educational background, ability and personality variables. The result of employing these statistical techniques was to reveal that PAS had long-term transfer effects for the female students involved in the original experiment. The mean performance of the female students from PAS method in subsequent life-science courses was significantly higher than performance of similar students from traditional sections. This result is interesting because in the original study of the immediate effect of the PAS method the group that benefitted most was the female student, especially with high test anxiety (Latta et al., 1978). This consistency between immediate and long-term effects of PAS, demonstrated the long-lasting effects of PAS method because one can reason that female students in PAS who took advantage of the experimental method, were able to transfer the gains into subsequent life-science courses, taken one year after the experiment. Exactly why this happened can not be determined from the present data analysis, but some hypotheses can be made to explain this improvement. Latta and others (1978), by establishing casual relationships among variables through a path-analysis model, showed that the self-paced testing method reduced the debilitating effects of test-anxiety for female students through increasing study effort and consequently they did better in the course. This study showed similar benefits lasting for at least one academic year.

Possibly female students who originally compensated for their test-anxiety under PAS gained confidence in their academic performance in the field and did better in subsequent courses. An alternative explanation is that they may have been able to learn more than their traditional counterpart, and consequently did better in subsequent courses for which beginning biology is a prerequisite.

Reduction of test anxiety due to self-paced testing procedures has been reported in other papers as well. Burnard (1978) reported on a special version of PSI used in general biology. The author believes, however, it is "content free" and the method can be applied to a wide variety of disciplines. The method has common elements with the PAS method including a self-paced testing procedure. The study is descriptive and reported reduction of the test-anxiety over a period of four quarters from 327 students who enrolled in those four quarters. The author described the findings in the following way:

> Students were asked the following question at the beginning of the quarter, 'How would you rate your anxiety level just before and during exam?'
>
> Response rates: 29% (Extremely high, I hate taking tests and get all shook up before and during tests); 39% (Moderately high); 25% (Mild apprehension only); 3% (No anxiety at all; and 3% (Enjoy taking tests).
>
> However, by the end of the quarter the students, as a whole, felt considerably different when asked the same question, as indicated by the following percentages: 2% (Extremely high, I still hate taking tests and get all shook up before and during tests; 9% (Moderately high); 35% (Mild apprehension

> only); 23% (No anxiety at all; 30% Enjoy taking tests).
> Data collected indicated that 76% of the students felt they became apprehensive about test taking, 18% noticed no change and 6% felt more apprehensive (p. 166).

Attenuation of test anxiety due to the PSI method was reported by Dziadost et al., (1977). Their design included both a within-subjects, in which each subject served as his own control, and a between-subjects comparison, when course and examination content were the same for both experimental and control groups. Both analyses showed that PSI teaching method could attenuate significantly the reported test anxiety level of PSI students as compared to the traditionally instructed students. Although both of these studies on test anxiety and self-paced testing procedures had their own shortcomings, their findings along with the results of the Latta study indicate that this is not an isolated phenomenon. Findings of this study, while not directly investigating test anxiety, are not inconsistent with the hypothesis that students who have high test anxiety will do better in a self-paced course. There is another hypothesis to explain the better performance of the female students who had Biology under the PAS method. The results of some recent international studies (Kelly, 1978) showed that in the general educational environment, males tend to do better in science. If, however, the students are instructed on exactly what to study (through detailed objectives for the course and repetitive testing, both features of PAS), then females will do better. This happened in the Latta et al. (1978) study and one can speculate that after the first success in the experimental course the female students gained competency and confidence in the field and therefore outperformed the males in the subsequent courses. However, the present study has no empirical proof for these findings. One can only say from the present results that with a certain probability the female students in self-paced class, outperformed all other groups in the study in subsequent life-science courses.

Another result is that the male students from the traditional section did better in subsequent courses than the females. But females from PAS section outperformed the males from the same section. Therefore, the females from PAS not only putperformed females from traditional section but also did better than males in the same section. One might ask that what about the male students from the two sections? Did PAS section have some disadvantages for them? The results revealed that this is not the case and the difference between the mean performance of the males from the two sections can be considered as a chance event.

Persistence

Persistence in the field can be measured by looking at the credits in the field taken during an academic period. Prediction of this dependent variable is more difficult than previous dependent variable, because many factors are involved in the selection of courses by students and it is almost impossible to control for all of them. The present data analysis showed that compared to performance in the subsequent courses, a small part of the variability in this dependent variable was predicted by all the independent and control variables (28%) and teaching method had almost no predictive power in the present criterion measure indicating that students from the two sections were identical, as far as the number of credits taken in the field is concerned. As the results showed and logic confirms the curriculum or major of a student is the most important factor in selection of subsequent life-science courses. Therefore, it is hard to conceive that introduction of a new method in any course would have a significant effect on students' choice of subsequent courses in the field.

There was only one other study which addressed the question of the number of hours in a field taken by students after having had a self-paced course (DuNann & Weber, 1976). Their results are similar to those of the present study. They found that:

> Groups of students who two years previously had taken introductory psychology under the different teaching method reported no differences in how many psychology books they had since read, in how many psychology courses they had taken, in whether they become majors or minors in psychology, in their evaluation of the interest and importance of the psychology or in how much they perceived that introductory course helped them in later psychology or nonpsychology courses (p. 376).

Concurrent effects

Although concurrent effects are not long-term effects, they can be considered as a special case involving immediate transfer. Kulik (1976) cited a study by Schimpfhauser and others (1974) which considered the concurrent effects of a PSI course in biochemistry. The medical students of this study not only performed significantly better than the control group in biochemistry, but they did somewhat better in two other concurrent courses, namely, anatomy and physiology. The literature review chapter of the present study contains some additional studies in this regard (Mao-Cohen & Lanson, 1976; Taylor, 1976 and Eads, 1976) in psychology, political science and anatomy, respectively. The results of those studies uniformly suggested favorable findings for the self-paced method. However, the present results did not show any significant differences in performance of life-science courses, taken concurrently with Biology 101, by the students in the two methods. The design of those earlier studies may have been confounded however, as discussed in the literature review.

Counseling Students

One of the big problems facing educators in today's college educational environment is that of counseling. With increasing innovations in teaching methods the question of "What group of students will benefit most from what instructional system?" has become important. Latta and colleagues (1978) pointed to this question and raised another interesting question:

> Unfortunately, most students in large universities are never adequately counseled about a course of study. Many advisors never see their advisees. Given this state of affairs, it would be interesting to determine what would happen if students were allowed to choose the system of instruction for themselves. Janisse (1973) has provided some data concerning this question. Janissee found that when given a choice, low-test-anxious students will choose the traditional system more frequently than high-test-anxious students. A psychologist familiar with learning theory would explain this by saying that high-test-anxious students are avoiding a situation that may increase their test anxiety (p. 968).

These authors suggested that self-paced method should be recommended for low-ability, high test anxious students. This recommendation, especially held for low ability, highly test-anxious female students. The present study looked at the long-term effects of PAS and concluded that this recommendation is also valid when considering not only the course in which the self-paced system operates, but also for subsequent achievement in courses in the field.

Suggestion for Further Studies

Since 1974 PAS at Iowa State University has been modified to include a video tape library of lectures covering the content of each unit available to each student upon demand. Preliminary data analysis has shown potential benefits from this modification. It would be very useful to conduct a follow-up study on this modified version of PAS which would address the problem of retention as well as the questions of this study. Retention is one of the most important long-term effects of an instructional system. Unfortunately the present study didn't gather, strictly, retention data on involved students, but performance in higher level course for which Biology 101 is a prerequisite is at least an indirect measure.

Another shortcoming of the present study is its failure to provide empirical data to explain the finding of the study. It would be very helpful to show empirically why females from PAS section received the most benefit from the system. This is not an easy task, but it can be done by collection of appropriate data and establishing casual relationships between variables through the proper data analysis approaches like path analysis. An example would be studies which establish casual relationships among variables to explain the findings. Latta and his colleagues through a path analysis model showed that attenuation of the test-anxiety for the high test-anxious females in PAS method was responsible for the better performance of these students.

Another possibility for further investigation is to test hypotheses derived from the results of the present study by conducting a separate experiment. An example of these hypotheses formed by the present study would be "What is the long-term effects of PAS for highly-test-anxious student relative to its effects on less-anxious students?" or similarly the concept of gender-treatment interaction could be emphasized in the hypothesis and tested.

Some General Final Words

It is more than a decade since PSI introduction by Keller and his colleagues. Since then, numerous experiments have been conducted to explore the different aspect of the method. Most of these studies suggested that PSI resulted in better immediate and long-term performance, positive students' attitudes, high withdrawal rate and procrastination. As research in any other field of social sciences, studies in this field had their own problems and shortcomings. Lack of experimental and statistical control, ignoring attitude-treatment interaction (Tobias, 1976) and nonreliable criterion measures were among the most serious problems of these studies.

The present study looked at the long-term effects of a version of PSI and found out a treatment-gender interaction which may be due to some attribute-treatment interaction. The study tried to avoid some of the major shortcomings of the previous research in the field, but had its own pitfalls and limitations which were discussed earlier.

Naturally, the system has also been criticized by many educators in the field. Some of these criticisms follow. The implementation of the method needs a lot of initial effort and time. The method tends to train test takers which is not the ultimate goal of the higher education. Many of the users try to arrange chapters of the text into units of study which leads to text teaching which again is not the philosophy of higher education. Further, some authorities believe that PSI causes fact learning rather than deep understanding. High withdrawal rates and procrastination are other disadvantages of the method which have been shown empirically. It has also been shown that withdrawal rates and procrastination depend primarily on the students' motivation rather than different features of the method. Professional role in the method is also subject to criticism. In practice, in many PSI courses, the instructor becomes a manager of proctors rather than managing students' educational activities and progress. In addition since there is often little direct interaction between students and instructor, the instructor often loses interest in the course after one or two academic periods.

On the other hand, the proponents of the method have tried to answer some of the criticisms by conducting different experiments, comparing PSI students with those of traditional courses in different respects. For example, they have tried to show, experimentally, that learning in PSI is not just "rote learning" by showing the superiority of the PSI students in retention tests. Another concern of the proponents has been to show that PSI students are better even in application of the facts and they compete with lecture students in higher level of understanding of the material. A general note about research on any educational innovations should be considered. Usually the negative or unfavorable results bearing on an attractive hypothesis do not appear in print because of either journals' editors refusal to publish nonsignificant results or the authors unwillingness to report their failure. Of course, there are some exceptions in this regard.

PSI and PAS

Iowa State University's version of PSI (PAS) has some common elements with both PSI and Bloom's mastery learning theory, including detailed objectives, mastery criterion and a self-paced frequent-testing procedure. It lacked some features like student proctors and immediate feedback by the proctors. Probably the most noticeable difference between the two methods is that implementation of PAS needs fewer changes in the course format, requires less initial effort and presumably would be faced with less resistance from the administrative authorities.

Component analysis of methods has received increasing attention by the PSI researchers in the field. Although there is some disagreement among the results of these studies, Williams (1976) summarized the most promising features of the method to be: (1) frequent quizzing over small units; (2) written study objectives and (3) mastery criterion. Also, some evidence exists to suggest that the immediate feedback and student proctors may improve students' performances. Little or no evidence indicates that self-pacing or optional lectures have any effect on students' academic achievement (Williams, 1976). The results of a more recent study (Abbott & Falstorm, 1977) revealed some interesting facts in this regard. The study is highly controlled both experimentally and statistically. The authors concluded that the addition of content unitization and detailed objectives to a lecture/mid term course resulted in students' performance below that found with the PSI course. The findings were replicated in two successive semesters in two separate experiments. However, the addition of frequent testing, to the previous modified lecture/mid term course resulted in achievement at the same level as the result of the standard PSI course. These findings were also replicated in two independent semesters. The present study showed similar results, since the PAS method lacked some components of PSI, e.g., students' proctors and immediate feedback; yet the long-term effects of the PAS are comparable to the similar results of a standard PSI course. Although there are not yet enough evidence to say definitely that it is the frequent-testing procedure which is responsible for the favorable results of the method, these findings are strong enough to raise the question that "Is it really the personalized system of instruction which results in superior achievement or is it personalized system of testing which caused greater achievement?".

The fate of PSI

The fate of educational innovations usually is not a happy one. Often these innovations seem to be very promising in the hand of the inventors and a few of their followers. But as the time passes and the methods spread far from the initial users, their novelty goes away and they become mechanical and unappealing.

PSI has survived more than a decade which is relatively long compared to other innovations. Its fate is in the hands of the users. Blind usage of the method which is firmly rooted in behavior theory principles, can bring the same fate for the method as the fate of other innovations. The fact that the method "works" should not be the sole reason for its implementation. Evaluation of the specific educational environment and its possibilities for employing the method along with a deep understanding of the method and the theories behind it should be considered seriously before any attempt to adopt the method.

BIBLIOGRAPHY

Abbott, R. D. & Falstorm, P. Frequent testing and personalized system of instruction. Contemporary Educational Psychology, 1977, 2, 251-257.

Alba, E., & Pennypacker, H. S. A multiple change score comparison of traditional and behavioral teaching procedures. Journal of Applied Behavior Analysis, 1972, 5, 121-124.

Anderson, O. T. & Artman, R. A. A self-paced, independent study, introductory physics sequence-description and evaluation. American Journal of Physics, 1973, 41, 12-18.

Austin, S. N., & Gilbert, K. E. Student performance in a Keller Plan course in introductory electricity and magnetism. American Journal of Physics, 1973, 41, 12-18.

Barr, A. J., Goodnight, J. H., Sall, J. P. & Helwig, J. T. A user's guide to SAS 76. Raleigh, North Carolina: SAS Institute, Inc., 1976.

Block, J. H. (Ed.) Schools, society and mastery learning. New York: Holt, Rinehart & Winston, Inc., 1974.

Bloom, B. S. Mastery Learning. In J. H. Block (Ed.), Mastery learning: Theory and practice. New York: Holt, Rinehart & Winston, 1971.

Born, D. G. & Herbert, W. A. A further study of personalized instruction in large university classes. Journal of Experimental Education, 1971, 40, 6-11.

Born, D. G., Gledhill, S. N., & Davis, M. L. Examination performance in lecture-discussion and personalized instruction courses. Journal of Applied Behavior Analysis, 1972, 5, 33-44.

Boren, A. R., & Foree, S. B. Personalized instruction applied to food and nutrition in higher education. Journal of Personalized Instruction, 1977, 2, 39-42.

Bostow, D. E., & Blumenfeld, G. J. The effects of two test-retest procedures on the classroom performance of undergraduate college students. In G. Semb (Ed.), Behavior Analysis and Education - 1972. Lawrence, Kansas: The University of Kansas Support and Development Center for Follow Through and the Department of Human Development, 1972.

Breland, N. S., & Smith, M. D. A comparison of PSI and traditional methods of instruction for teaching introduction to psychology. In R. S. Ruskin & S. F. Bono (Ed.), Proceedings of the First National Conference. Washington, D.C.: Center for Personalized Instruction, 1974.

Burnard, R. K. Computer-managed instruction with individualized instruction in General Biology. Journal of Personalized Instruction, 1978, 3, 165-167.

Charles, C. M. Educational Psychology: The instructional endeavor. St. Louis, Missouri: C. V. Mosby Co., 1972.

Cole, C., Martin, S. & Vincent, J. A. Comparison of two teaching formats at the college level. In J. Johnston (Ed.), Behavior research and technology in higher education. Springfield, Ill.: Charles C. Thomas, 1975.

Cooper, J. L., & Greiner, J. M. Contingency management in an introductory psychology course produces better retention. The Psychological Record, 1971, 21, 391-400.

Corey, J. R., & McMichael, J. S. Retention in a PSI introductory psychology course. In J. G. Sherman (Ed.), Personalized system of instruction 41 germinal papers. Menlo Park, Calif.: W. A. Benjamin, 1974.

Cote, J. D. Biology by PSI in a community college. In B. A. Green (Ed.), Proceedings of the Second National Behavioral Instruction Conference. Washington, D.C.: Center for personalized instruction, 1976.

Cronbach, L. J., & Snow, R. Aptitudes and instructional methods. New York: Irvington, 1977.

Davis, J. R. Teaching strategies for the college classroom. Boulder, Colorado: Westview Press, Inc., 1976.

Dolphin, W. D., Franke, R. G., Covert, F. G., & Jorgensen, C. D. The Phase Achievement System: An instructional management system for large enrollment lecture sections. AIBS Education Review, April 1973, 2, 24-27.

Dubin, R., & Taveggia, T. C. The Teaching-Learning Paradox. Eugene, Oregon: University of Oregon Press, 1968.

DuNann, D. H., & Weber, S. J. Short- and long-term effects of contingency managed instruction on low, medium and high G.P.A. students. Journal of Applied Behavior Analysis, 1976, 9, 375-376.

Dziadost, T. H., Curran, J. P., & Santogrossi, D. A. Personalized instruction and the attenuation of test anxiety. Journal of Personalized Instruction, 1977, 2, 194-198.

Eads, J. W. Student performance in a modularized Human Anatomy and Physiology course. Journal of Personalized Instruction, March 1976, 1, 32-35.

Ferster, C. B. Individualized instruction in an introductory psychology course. The Psychological Record, 1968, 18, 521-532.

Ferster, C. B., & Perrott, M. C. Behavior principles. New York: Appleton-Century-Crofts, 1968.

Green, B. A. Physics teaching by the Keller Plan at MIT. American Journal of Physics, 1971, 39, 764-775.

Harrison, W. L. Teaching introductory data processing through competency based testing. Journal of Educational Research, 1977, 71, 45-47.

Hohn, R. L., Des Lauriers, M. & Deaton, W. Learners' - characteristics and performance effects in self-paced instruction. Psychological Reports, 1977, 40, 1011-1012.

Hursh, D. E. Personalized system of instruction: What do the data indicate? Journal of Personalized Instruction, September 1976, 1, 91-105.

Johnson, M. P., & Walsh, E. J. Grade inflation or better comprehension. Teaching Sociology, 1978, 5, 363-373.

Johnston, J. M., & Pennypacker, H. S. A behavioral approach to college teaching. American Psychologist, 1971, 26, 219-244.

Johnston, J. M. (Ed.). Behavior research and technology in higher education. Springfield, Ill.: Charles C. Thomas, 1975.

Keller, F. S. A personal course in psychology. In. R. Ulrich, T. Stochnikg & J. Marbry (Eds.), Control of human behavior (Vol. 1). New York: Scott, Foresman & Company, 1966.

Keller, F. S. "Goodby, teacher . . . ." Journal of Applied Behavior Analysis, 1968, 1, 79-89.

Keller, F. S., & Sherman, J. G. The Keller Plan handbook. Menlo Park, Calif.: W. A. Benjamin, 1974.

Kelly, A. Girls and Science: An international study of sex differences in school science achievement. Stockholm, Sweden: Almquist & Wiksell International, 1978.

Kirk, R. E. Experimental Design: Procedures for the Behavioral Sciences. Belmont, California: Brooks/Cole Publishing Co., 1968.

Koen, B. Self-paced instruction for engineering students. Engineering Education, 1970, 60, 735-736.

Kulik, J. A., Kulik, C. L., & Carmichael, K. The Keller plan in science teaching. Science, 1974, 183, 379-383.

Kulik, J. A., & Kulik, C. L. Effectiveness of the Personalized System of Instruction. Engineering Education, Dec. 1975, 66, 228-231.

Kulik, C. L., & Kulik, J. A. PSI and the mastery model. In B. A. Green (Ed.), Proceedings of the Second National Behavioral Instruction Conference. Washington, D.C.: Center for Personalized Instruction, 1976.

Kulik, J. A. PSI: A formative evaluation. In B. A. Green (Ed.), Proceedings of the Second National Behavioral Instruction Conference. Washington, D.C.: Center for Personalized Instruction, 1976.

Latta, R. M., Dolphin, W. D., & Grabe, M. Individual differences model applied to instruction and evaluation of large college classes. Journal of Educational Psychology, 1978, 70, 960-970.

Lloyd, K. E., & Knutzen, N. J. A self-paced programmed undergraduate course in the experimental analysis of behavior. Journal of Applied Behavior Analysis, 1969, 2, 125-133.

Lloyd, K. E. Contingency management in university courses. Educational Technology, 1971, 11, 18-23.

Lu, M. C. The retention of material learned by PSI in a mathematics course. In B. A. Green (Ed.), Proceedings of the Second National Behavioral Instruction Conference. Washington, D.C.: Center for Personalized Instruction, 1976.

Lu, P. H. Teaching human growth and development by the personalized system of instruction. Teaching of Psychology, 1976, 3, 127-128.

Lubkin, J. L. Engineering studies: A Keller-Plan course with novel problems and novel features. In R. S. Ruskin & S. F. Bono (Eds.), Personalized instruction in higher education: Proceedings of the first national conference. Washington, D.C.: Center for Personalized Instruction, 1974.

Malott, R. W., & Svinicki, J. G. Contingency management in an introductory psychology course for one thousand students. The Psychological Record 1969, 19, 545-556.

Mandler, G., & Sarasan, S. B. A study of anxiety and learning. Journal of Abnormal and Social Psychology, 1952, 47, 166-173.

Mao-Cohen, D., & Lanson, R. N. Effects of the PSI experience beyond the PSI classroom. Journal of Personalized Instruction, September 1976, 1, 76-78.

Martin, R. R. and Carlton, T. R. P.S.I. students in the second year: A follow-up study. Journal of Chemical Education, June 1978, 55, 365.

McFarland, B. An individualized course in elementary composition for the marginal students. In B. A. Green (Ed.), Proceedings of the second national behavioral instruction conference, Washington, D.C.: Center for Personalized Instruction, 1976.

McMichael, J. S. Following-up PSI students. PSI Newsletter, 1975, 3, 7.

McMichael, J. S., & Corey, J. R. Contingency management in an introductory psychology course produces better learning. Journal of Applied Behavior Analysis, 1969, 2, 79-83.

Moore, J. W., Hauck, W. E., & Gagné, E. D. Aquisition, retention, and transfer in an individualized college physics course. Journal of Educational Psychology, 1973, 64, 335-340.

Moore, J. W., Mahan, J. H., & Ritts, C. A. Continuous progress concept of instruction with university students. The Psychological Reports, 1969, 25, 887-892.

Moretz, H. L. PSI improves individualized interaction in a two-year college. In B. A. Green (Ed.), Proceeding of the Second National Behavioral Instruction Conference. Washington, D.C.: Center for Personalized Instruction, 1976.

Morris, E. K., Surber, C. F. Self-pacing versus instructor-pacing: Achievement, evaluation, and retention. Journal of Educational Psychology, 1978, 70, 224-230.

Mosteller, F., & Tukey, J. W. Data Analysis and Regression. Reading, Massachusetts: Addison-Wesley Publishing Co., 1977.

Myers, W. A. Operant learning principles applied to teaching introductory statistics. Journal of Applied Behavior Analysis, 1970, 3, 213-220.

Nazzaro, J. R., Todorov, J. C., & Nazzaro, J. N. Student ability and individualized instruction. Journal of College Science Teaching, 1972, 2, 523-530.

Randels, P. M., Kilpatrick, D. G., McCurdy, L. and Saunders, P. J. Comparison of the psychiatry learning system and traditional teaching of psychiatry. Journal of Medical Education, September 1976, 51, 751-757.

Riedel, R. C., Harney, B., & LaFief, W. Unit test scores in PSI versus traditional classes in beginning psychology. Teaching of Psychology, April 1976, 3, 76-78.

Robin, A. L. Behavioral instruction in the college classroom. Review of Educational Research, Summer 1976, 46, 313-354.

Rogers, S. J., Satariano, W. H., & Rogers, E. D. Self-scheduled introductory sociology: An evaluation. Teaching Sociology, 1977, 4, 129-140.

Ryan, B. A. PSI, Keller's Personalized System of Instruction: An Appraisal. Washington, D.C.: The American Psychological Association, 1974.

Schimpfhauser, F., Jorrocks, L., Richardson, K., Alben, J., & Sprecher, H. The personalized program of instruction as an acceptable alternative within the traditional structure of medical basic science. In R. S. Ruskin & S. F. Bong (Eds.), Personalized instruction in higher education: Proceedings of the first national conference. Washington, D.C.: Center for Personalized Instruction, 1974.

Siegfried, J. J., & Strand, S. H. An evaluation of the Vanderbilt-JCGE experimental PSI course in elementary economics. The Journal of Economic Education, Fall 1976, 8, 9-25.

Shelden, H. R. Individualized instruction in introductory chemistry. Journal of Chemical Education, June, 1978, 55, 379.

Sheldon, J., Sherman, J. A., Wolf, M. M., Minkin, B. L., & Minkin, N. "Hello teacher . . ." or optional lectures, discussions with instructors, and assigned proctors: Their effects on students' ratings of a PSI course. In J. Johnston & G. W. ONeil (Eds.), Research and technology in college and university teaching. Atlanta, Ga.: Georgia State University Urban Life Center, 1975.

Sheppard, W. C., & MacDermot, H. G. Design and evaluation of a programmed course in introductory psychology. Journal of Applied Behavior Analysis, 1970, 3, 5-11.

Sherman, J. G. PSI: Some notable failures. In J. L. Sayre & J. J. Knightly (Eds.), The personalized system of instruction in higher education: Reading on PSI-The Keller Plan. Enid, Okla.: Seminary Press, 1972.

Sherman, J. G. (Ed.). Personalized system of instruction: 41 germinal papers. Menlo Park, Calif., W. A. Benjamin, 1974. (a)

Sherman, J. G. The theory behind PSI. In J. G. Sherman (Ed.), Personalized system of instruction: 41 germinal papers. Menlo Park, Calif., W. A. Benjamin, 1974. (b)

Skinner, B. F. The science of learning and the art of teaching. Harword Educational Review, 1954, 24, 86-97.

Skinner, B. F. Teaching machines. Science, 1958, 128, 969-977.

Skinner, B. F. About behaviorism. New York: Alfred A. Knopf, 1974.

Smiley, E. E. PSI and the open Admissions student. In B. A. Green (Ed.). Proceedings of the Second National Behavioral Instruction Conference, Washington, D.C.: Center for Personalized Instruction, 1976.

Soper, J. C., & Thornton, R. M. Self-paced economics instruction. A large-scale disaggregated evaluation. The Journal of Economic Education, Spring 1976, 7, 81-91.

Spector, L. C. The effectiveness of personalized system of instruction in economics. Journal of Personalized Instruction, September 1976, 1, 118-122.

Stice, J. E. Seventeen PSI Projects at the University of Texas at Austin. In B. A. Green (Ed.), Proceedings of the Second National Behavioral Instruction Conference. Washington, D.C.: Center for Personalized Instruction, 1976.

Taylor, P. Personalized instruction: The introductory political science course. Teaching Political Science, April 1976, 3, 227-248.

Tobias, S. Achievement treatment interactions. Review of Educational Research, 1976, 46, 61-74.

Vander Klaw, C. F., & Plomp, T. J. Individualized study system in theory and practice. Higher Education, 1974, 3, 213-230.

Weisberg, H. L. PSI: Innovation in the engineering classroom. Tennessee Engineer, 1973, 28-31.

Whitehurst, G. Academic responses and attitudes engendered by a programmed course in child development. Journal of Applied Behavior Analysis, 1972, 5, 283-291.

Williams, R. L. Personalized system of instruction: Future research areas. Journal of Personalized Instruction, September, 1976, 1, 106-112.

Witters, D. R., & Kent, G. W. Teaching without lectures. Evidence in the case for individualized instruction. The Psychological Record, 1972, 22, 169-175.

Wodarski, J. S., & Burkholdt, D. Behavioral instruction in the college classroom. A review of methodological procedures. In J. Johnston (Ed.), Behavior research and technology in higher education. Springfield, Ill.: Charles C. Thomas Co., 1975.

Wood, R. W. A student-maintained filing and scheduling system for the interview classroom technique. Psychological Record, 1972, 22, 491-496.

Wolins, L. Secondary analysis: In published research in the behavioral science. New Directions for Program Evaluation, 1978, 4, 45-54.

REFERENCE NOTES

Calhoun, J. F. Elemental analysis of the Keller Method of Instruction. Paper presented at the annual meeting of the American Psychological Association, Montreal, Canada, September, 1973.

Moore, J. M., & Gagné, E. D. The effects of continuous progress instruction on acquisition, transfer, and attitudes in a college religion course. Unpublished manuscript, Bucknell University, 1973.

ACKNOWLEDGMENTS

First of all, this investigator would like to express his deepest respect and acknowledgment to two individuals whom without their assistance, this study may have never been completed - Dr. Warren Dolphin, Professor of Zoology at Iowa State University, co-major professor, who provided me valuable assistance and guidance in every aspect of this study and Dr. Leroy Wolins, Senior Professor of Psychology and Statistics at Iowa State University, who clarified my thinking about the statistical procedures used in this study. Although not serving as a committee member, Dr. Wolins spent long hours of his valuable time to discuss the various features and limitations of statistical procedures. I can never put my appreciation and respect toward these individuals in words.

I want to thank Dr. Anton J. Netusil, co-major professor, for his assistance. Drs. Rex Thomas, Richard Warren, Frederick Brown and Trevor Howe who served as my committee members have provided valuable comments and guidance which improved this study.

I also want to express my appreciation for my wife Minoo, my son Pouyā and my daughter Heda who have enriched my life and allowed me to complete this work by their understanding and devotion.

Portions of this work were supported by a grant from the National Science Foundation CAUSE Program (Ser. No. 76-16100) under direction of Dr. Dolphin and by funds from Iowa State Research Institute for Studies in Education (RISE) under the direction of Dr. Richard Warren.

At last but not the least, I want to indicate my debts to the heroic people of my beloved country, Iran, who have supported me from the first day of school to the last stage of my education. I am desirous to go back and devote my life to the people whom I love so much.

APPENDIX A. TABLES

Table A. Zero order correlation matrix of the deleted independent variables with independent variables of the models[a]

	NMCS	GPAMS	NPCS	GPAPS	NCCS	GPACS	NCUS	CGPAS
TAQ	-.14(415)*	.03(59)	-.10(415)	-.03(15)	-.06(415)	-.29(50)	-.29(415)*	-.16(93)
HSR	-.05(414)	-.27(59)	-.08(414)	-.76(15)*	-.08(414)	-.58(49)*	-.02(414)	-.48(92)*
MSAT	.08(371)	.16(54)	.15(371)*	.72(14)*	.11(371)	.58(45)*	.14(371)*	.44(77)*
CUMGPA	-.07(219)	.72(27)*	.01(219)	.58(8)	-.03(219)	.76(23)*	-.05(219)	.97(44)*
HSBKGD	.17(413)*	.22(59)	.12(413)	.19(15)	.09(413)	.39(49)*	.01(413)	.19(92)
COURSE	.05(415)	-.16(59)	.02(415)	.03(15)	-.06(415)	-.14(50)	-.07(415)	-.19(93)
GENDER	.21(415)*	-.05(59)	.14(415)*	-.26(15)	-.01(415)	.05(50)	.04(415)	-.07(93)
MAJOR	-.02(415)	-.05(59)	-.05(415)	-.28(15)	-.02(415)	.11(50)	-.14(415)*	-.02(93)
YR	.60(413)*	.49(59)*	.38(413)*	.36(15)	.64(413)*	.38(50)*	.85(413)*	.32(93)*

Note. NMCS = Number of Math credits up to September '74; GPAMS = Grade Point Average of Math Credits up to September '74; NPCS = Number of Physics Credits up to September '74; GPAPS = Grade Point Average of Physics Credits up to September '74; NCCS = Number of Chemistry Credits up to September '74; GPACS = Grade Point Average of Chemistry Credits up to September '74; NCUS = Number of University Credits up to September '74; CGPAS = Cumulative Grade Point Average of University Credits up to September '74; TAQ = Test Anxiety Questionnaire; HSR = High-School Graduation Rank; MSAT = Minnesota Scholastic Aptitude Test; CUMGPA = Cumulative Grade Point Average of University credits up to December '74 (excluding Biology 101 and concurrent credits in Life-science); HSBKGD = High-School Background; YR = Year in University.

[a]Asterisks indicate the significant of a correlation beyond .01. Numbers in parentheses show the number of cases involved in calculation of a given correlation.

Table B. Zero order correlational matrix of dependent variable GPALQ with independent variables of the models[a]

	Male[b]	Female[b]	PAS Male[c]	TRA Male[c]	PAS Female[c]	TRA Female[c]
TAQ	.02(176)	-.12(159)	.04(90)	.01(86)	-.20(85)	-.06(74)
HSR	-.45(176)*	-.52(158)*	-.54(90)*	-.33(86)*	-.46(84)*	.57(74)*
MSAT	.41(159)*	.37(144)*	.53(80)*	.31(79)*	.36(76)*	.38(68)*
CUMGPA	.74(176)*	.65(159)*	.72(90)*	.77(86)*	.71(85)*	.59(74)*
HSBKGD	.27(175)*	.32(158)*	.31(90)*	.23(85)*	.20(84)*	.44(74)*
MAJOR	.12(176)	.17(159)*	.03(90)	.20(86)*	.19(85)*	.15(74)
YR	.10(174)	.02(159)	.17(90)	-.00(84)	.03(85)	-.08(74)

Note. GPALQ = Grade Point Average of Life Science Credits taken in the period of the study; TAQ = Test Anxiety Questionnaire; HSR = High-School Graduation Rank; MSAT = Minnesota Scholastic Aptitude Test; CUMGPA = Cumulative Grade Point Average; HSBKGD = High School Background; YR = Year in University.

[a]Asterisks indicate the significant of a correlation at .01. Numbers in parentheses show the number of cases involved in calculation of a given correlation.

[b]All students are divided into two categories: Male students and female students.

[c]Students are divided into four categories in respect of their Gender and received method of teaching, e.g., PAS-MALE means male students who received the PAS method of teaching.

Table C. Zero order correlation matrix of dependent variable NCLQ with independent variables of the model[a]

	Male[b]	Female[b]	PAS-MALE[c]	TRA-MALE[c]	PAS-FEM[c]	TRA-FEM[c]
TAQ	.11(224)*	-.04(191)	.17(110)	.05(114)	.02(100)	.13(91)
HSR	.03(224)	-.18(140)*	-.01(110)	.09(114)	-.22(99)*	-.14(91)
MSAT	.00(199)	.06(172)	-.03(97)	.05(102)	.02(87)	.16(85)
CUMGPA	.18(224)*	.28(191)*	.11(110)	-.23(114)*	.27(100)*	.27(91)*
HSBKGD	.05(223)	.20(190)*	.13(110)	-.02(113)	.17*	.22(91)*
MAJOR	.42(224)*	.44(191)*	.40(110)*	.43(113)*	.43(100)*	.45(91)*
YR	-.09(222)	-.03(191)	-.09(110)	-.10(114)	-.09(100)	.02(91)

Note. NCLQ = Number of Credits in Life-Science Taking During the Three Quarters; TAQ = Test Anxiety Questionnaire; HSR - High School Graduation Rank; MSAT = Minnesota Scholastic Aptitude Test; CUMGPA = Cumulative Grade Point Average; HSBKGD = High School Background; YR = Year in University.

[a]Asterisks indicate the significant of a correlation at .01. Numbers in parentheses indicate the number of cases involved in calculation of a given correlation.

[b]All students divided into two categories: Male students and female students.

[c]Students divided into four categories in respect of their gender and the received method of instruction, *e.g.*, PAS-MALE means male students who received the PAS method of teaching.

Table D. Zero order correlation matrix of dependent variable GPACCL with independent variables of the model[a]

	Male[b]	Female[b]	PAS-MALE[c]	TRA-MALE[c]	PAS-FEM[c]	TRA-FEM[c]
TAQ	-.10(223)	-.39(191)*	-.07(110)	-.13(113)	-.37(100)*	-.46(91)*
HSR	-.56(223)*	-.61(190)*	-.68(110)*	-.47(113)*	-.56(99)*	-.63(91)*
MSAT	.51(198)*	.56(172)*	.60(97)*	.44(101)*	.52(87)*	.60(85)*
CUMGPA	.94(223)*	.94(191)*	.94(110)*	.93(113)*	.93(100)*	.95(91)*
HSBKGD	.32(222)*	.29(190)*	.41(110)*	.24(112)*	.25(99)*	.33(91)*
MAJOR	.09(223)	.13(191)*	.02(110)	.14(113)	.10(110)	.14(91)
YR	.22(221)*	.08(191)	.34(110)*	.10(115)	.10(100)	.04(91)

Note. GPACCL = Grade Point Average of Concurrent Credits in Life-Science; TAQ = Test Anxiety Questionnaire; HSR = High School Graduation Rank; MSAT = Minnesota Scholastic Aptitude Test; CUMGPA = Cumulative Grade Point Average; HSBKGD = High School Background; YR = Year in University.

[a]Asterisks correlations are significant at .01. Numbers in parentheses indicate the number of the cases involved in calculation of a given correlation.

[b]All students are divided into two categories: Male students and female students.

[c]Students are divided into four categories in respect of their gender and the received method of instruction, e.g., PAS-MALE means male students who received the PAS method of instruction.

APPENDIX B. FIGURES

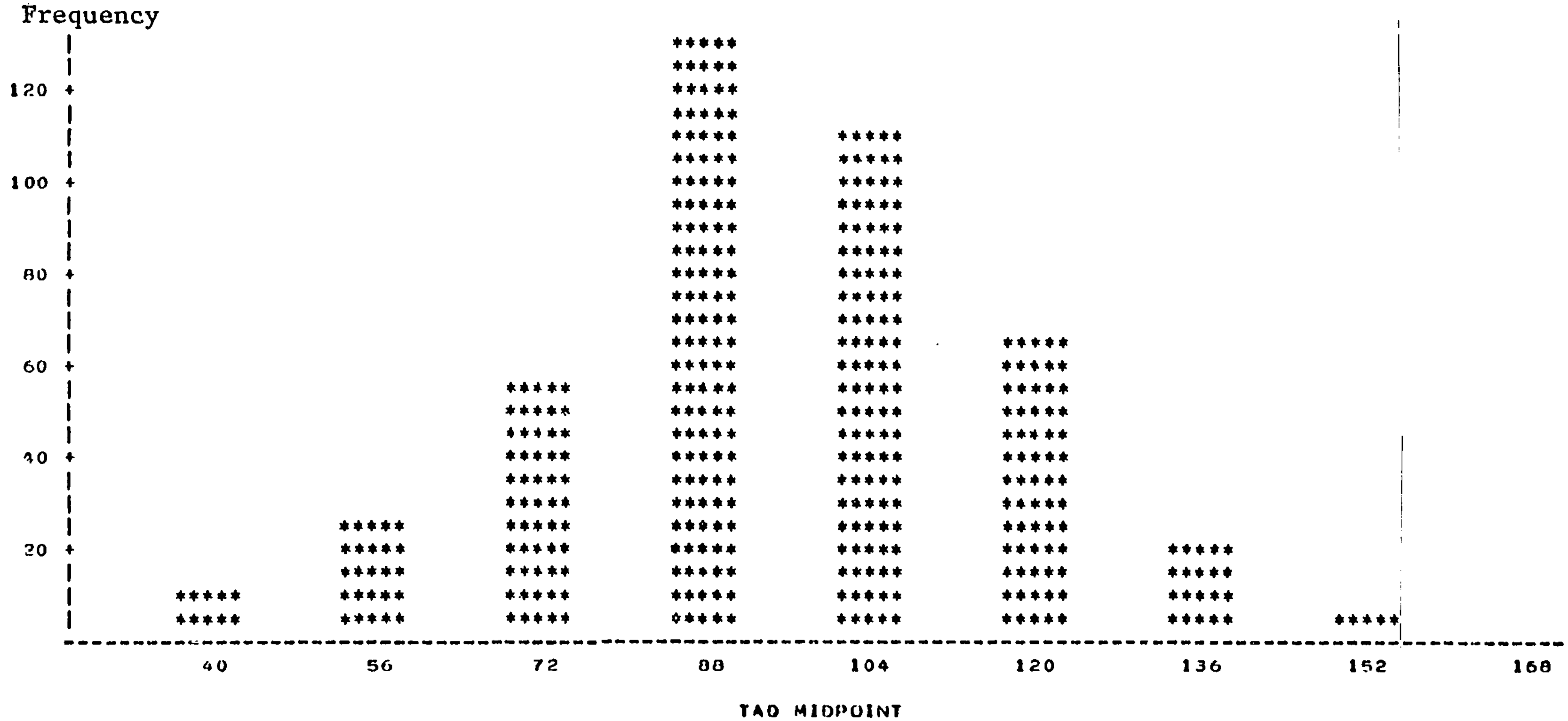


Figure B1. Frequency bar chart of independent variable: Test Anxiety Questionnaire (TAQ)

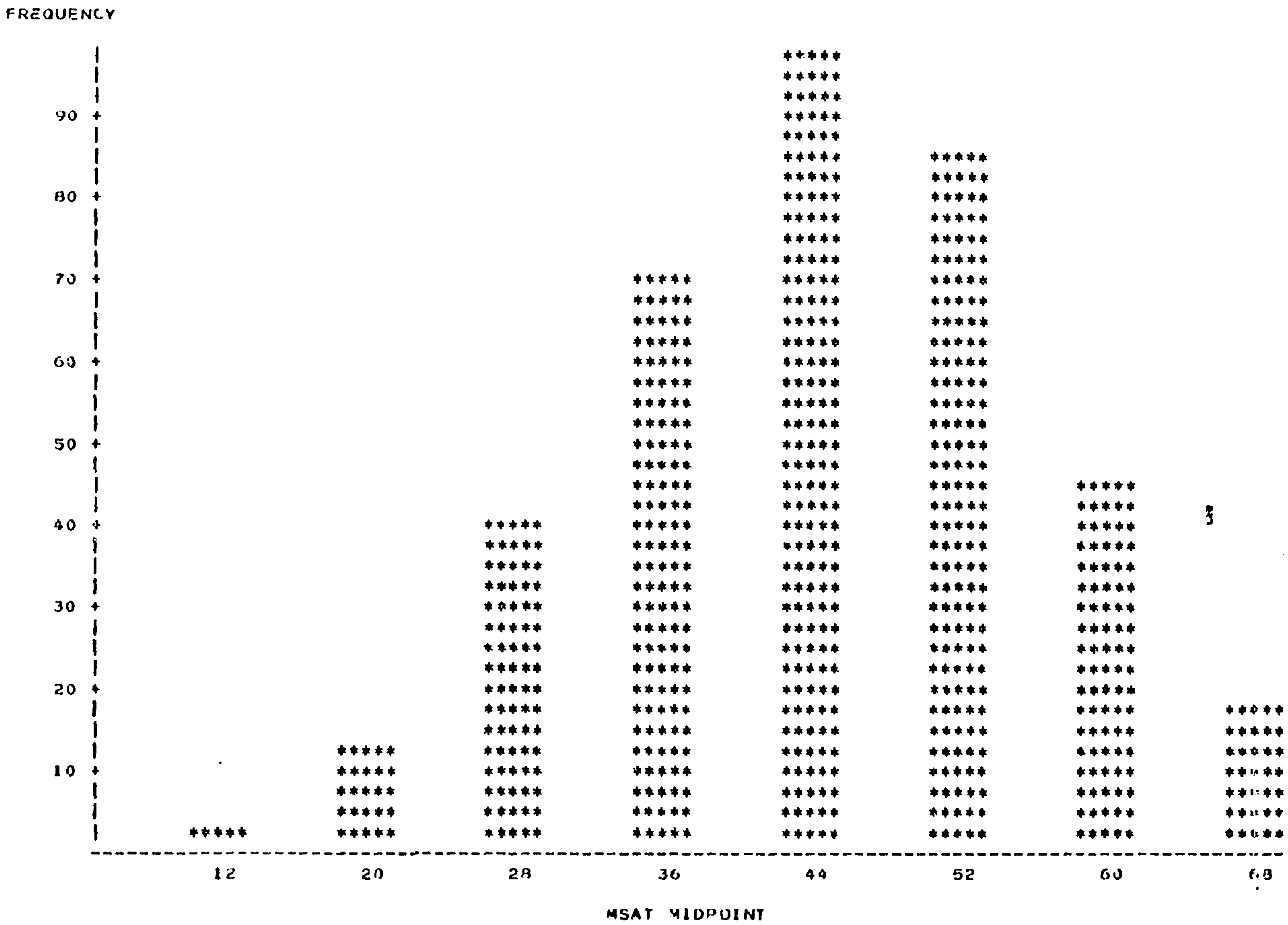


Figure B2. Frequency bar chart of independent variable: Minnesota Scholastic Aptitude Test (MSAT)

Figure B3. Frequency bar chart of untransformed independent variable: High-School Graduation Rank (HSR)

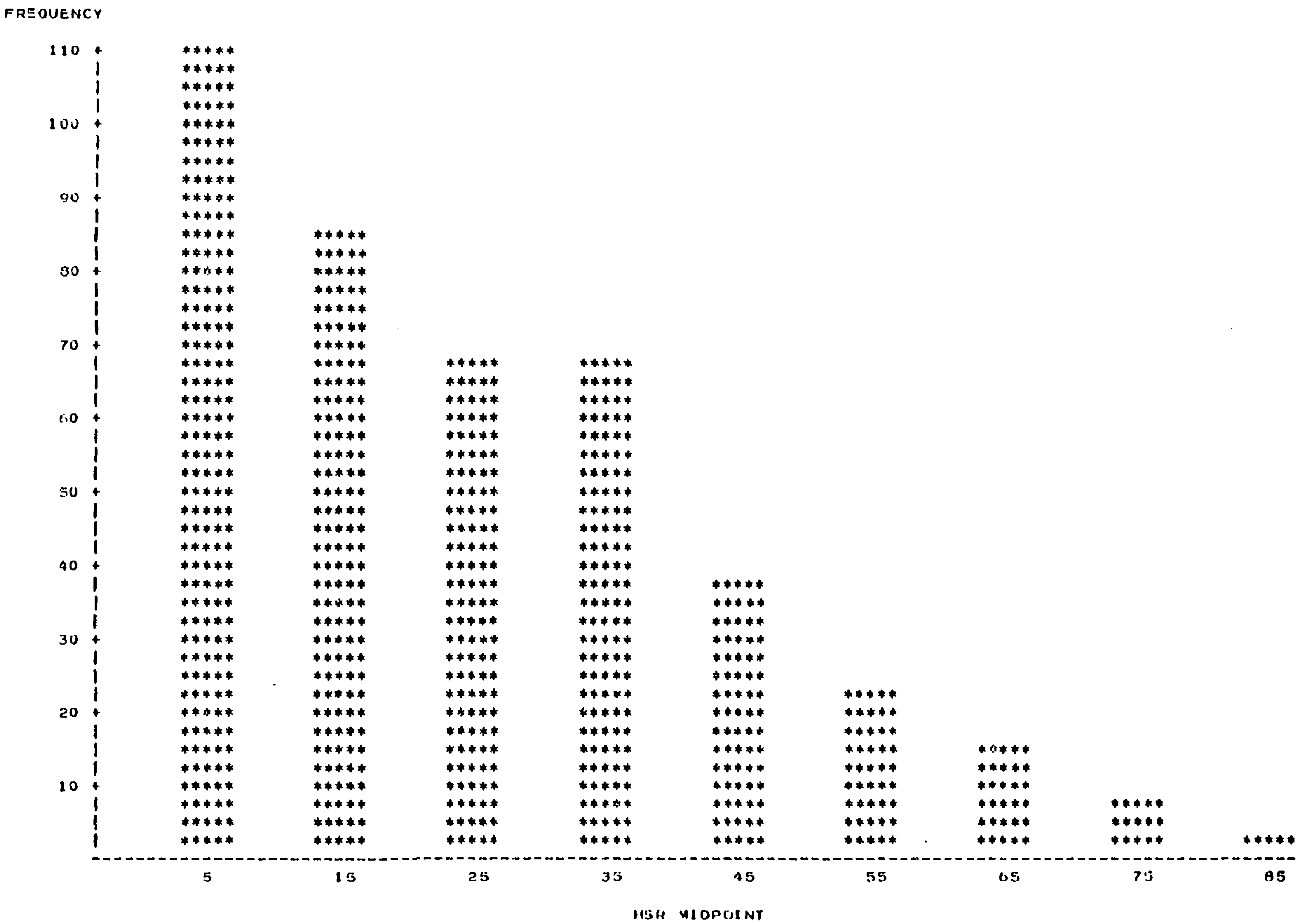
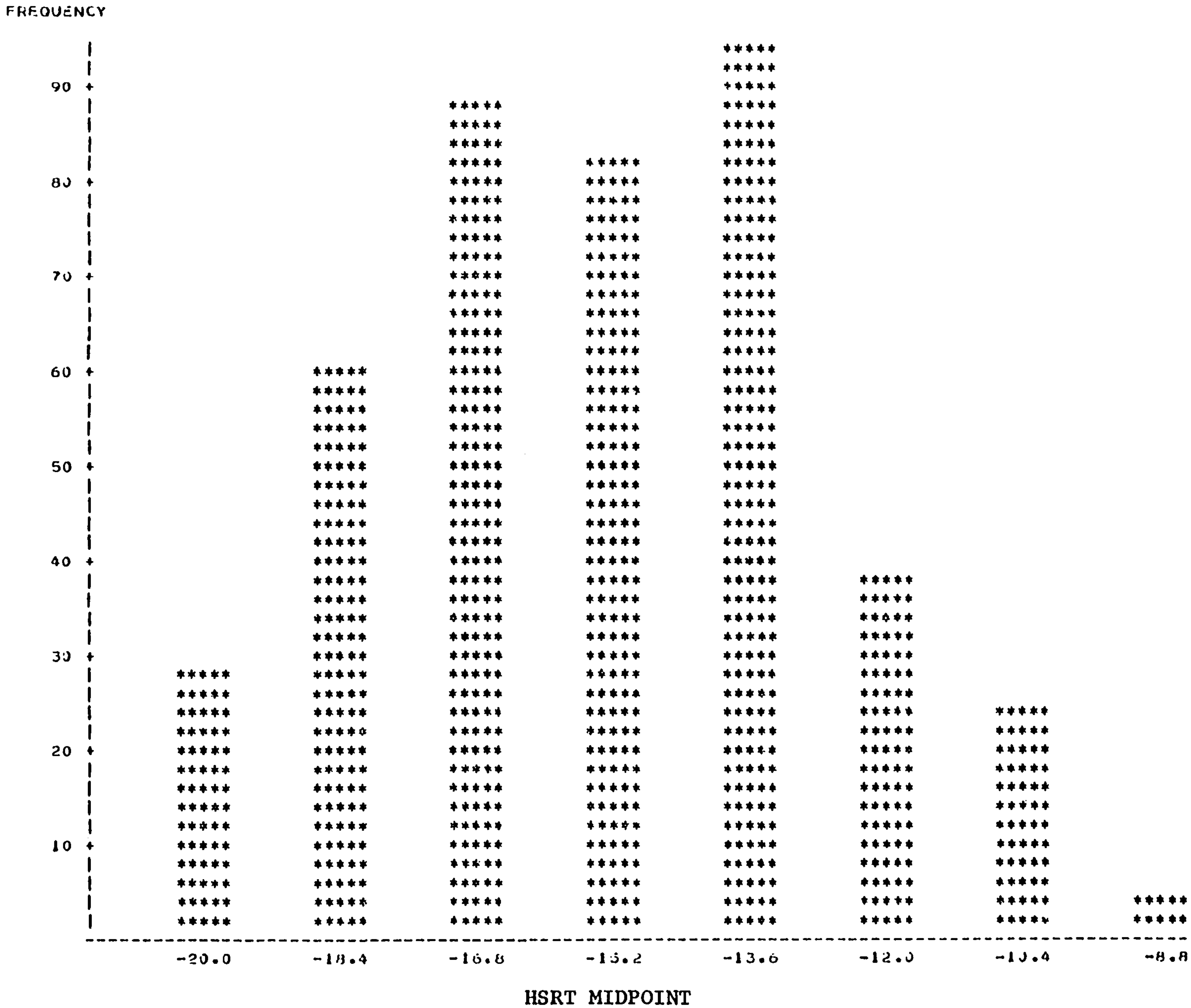

FREQUENCY
110
100
90
80
70
60
50
40
30
20
10
5
15
25
35
45
55
65
75
85
HSR MIDPOINT

Figure B4. Frequency bar chart of transformed independent variable: High-School Graduation Rank (HSRT)


FREQUENCY
90
80
70
60
50
40
30
20
10
-20.0
-18.4
-16.8
-15.2
-13.6
-12.0
-10.4
-8.8
HSRT MIDPOINT

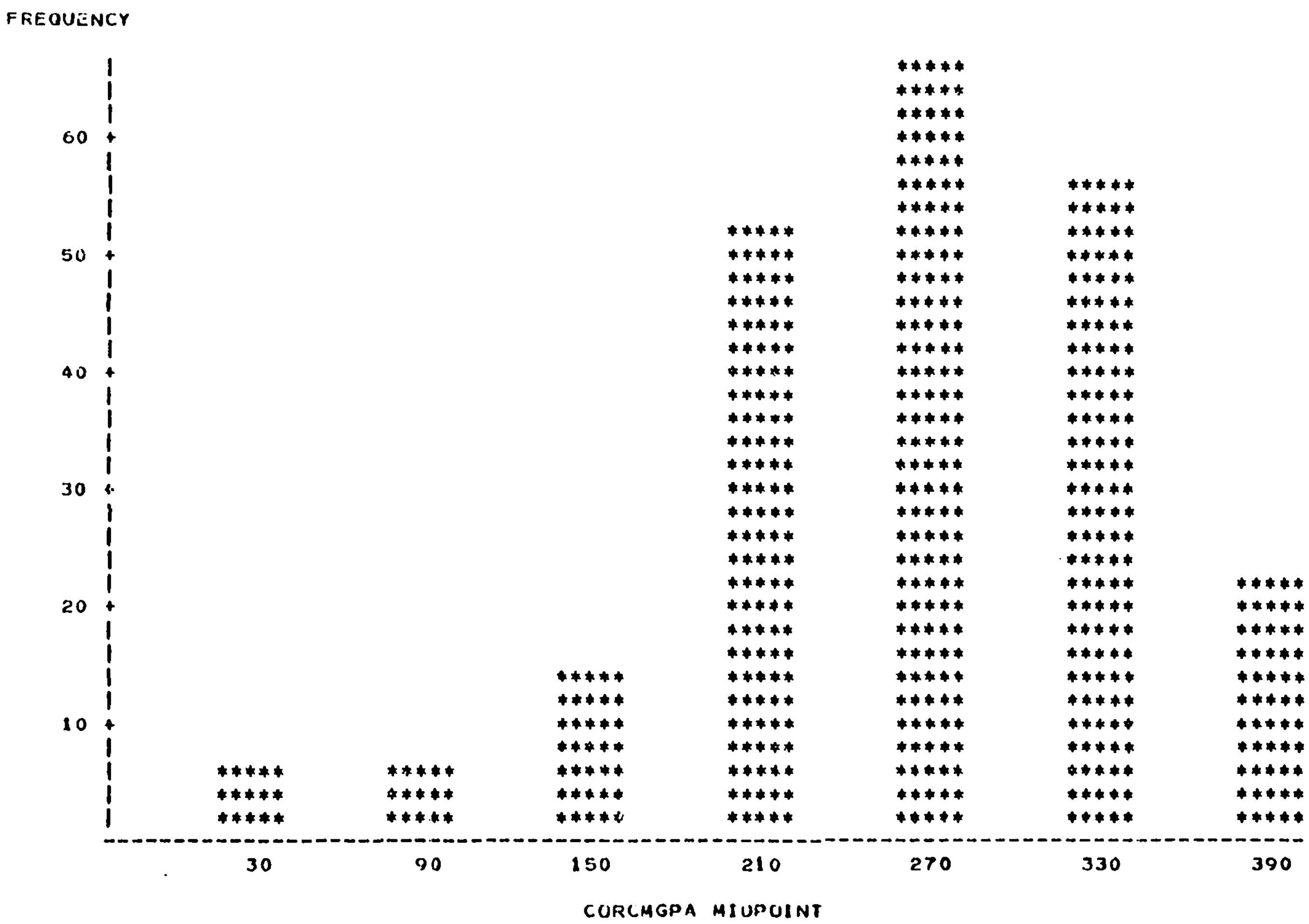


Figure B5. Frequency bar chart of independent variable: Cumulative Grade Point Average of credit earned up to the end of quarter Fall 1974 excluding Biology 101 and Concurrent life-science credits (CORCUMGPA)

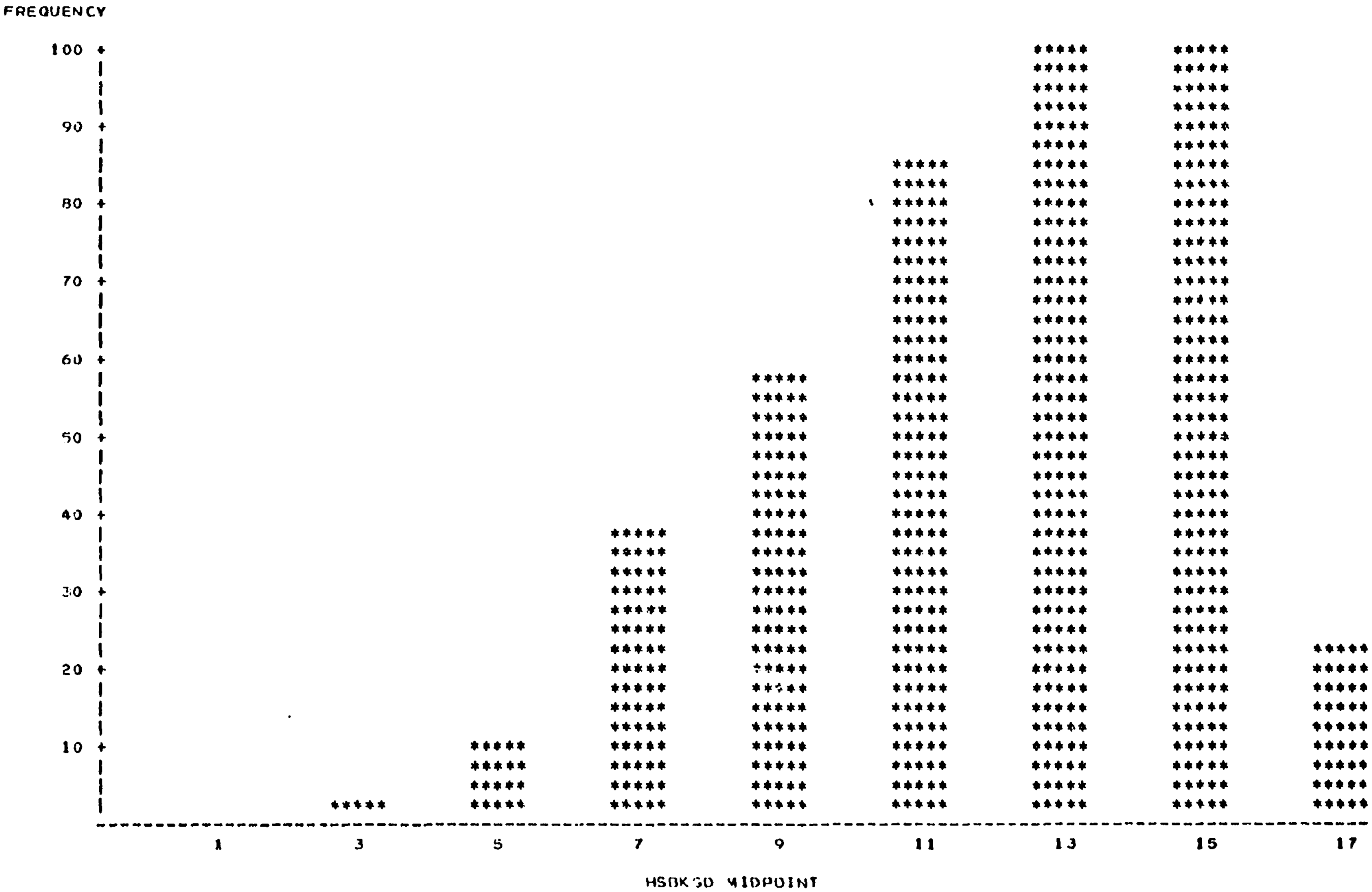


Figure B6. Frequency bar chart of independent variable: High-School Background (HSBKGD)

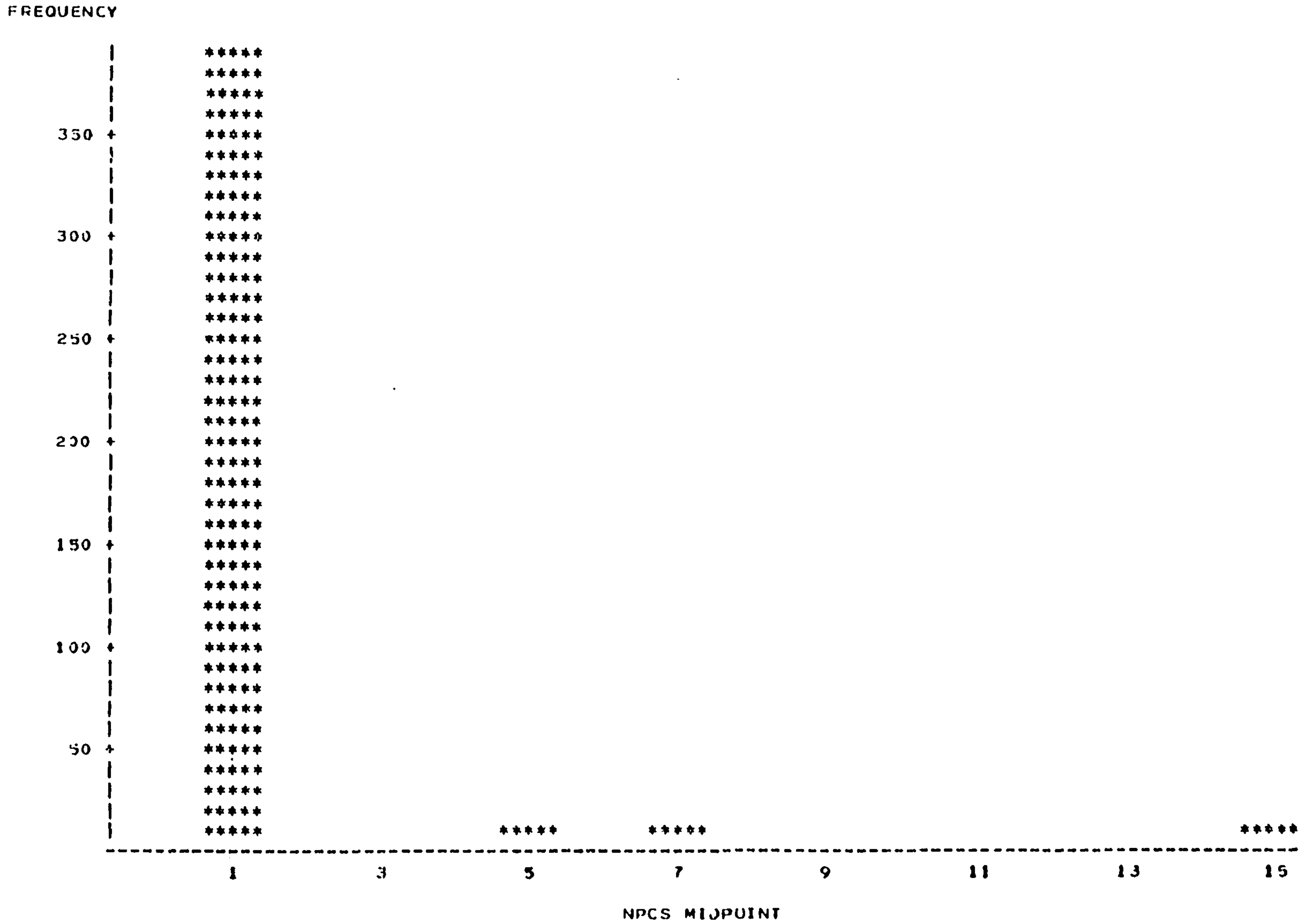


Figure B7. Frequency bar chart of deleted independent variable: Number of Credits in Physics up to September 1974 (NPCS)

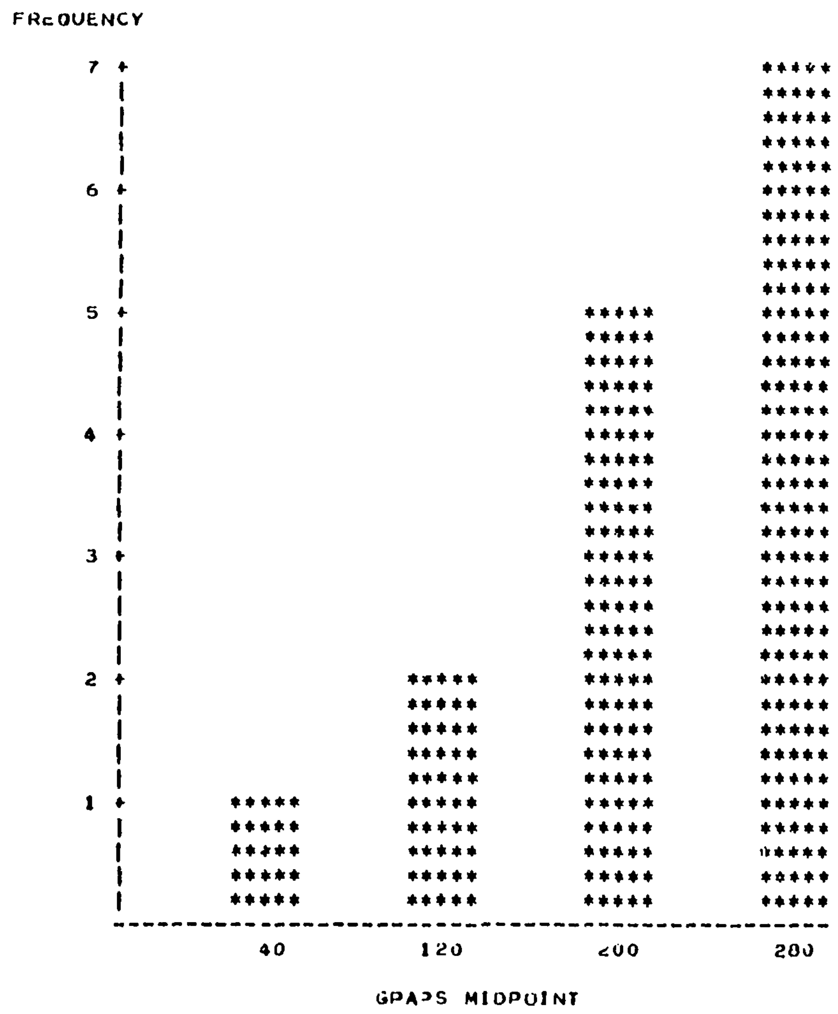
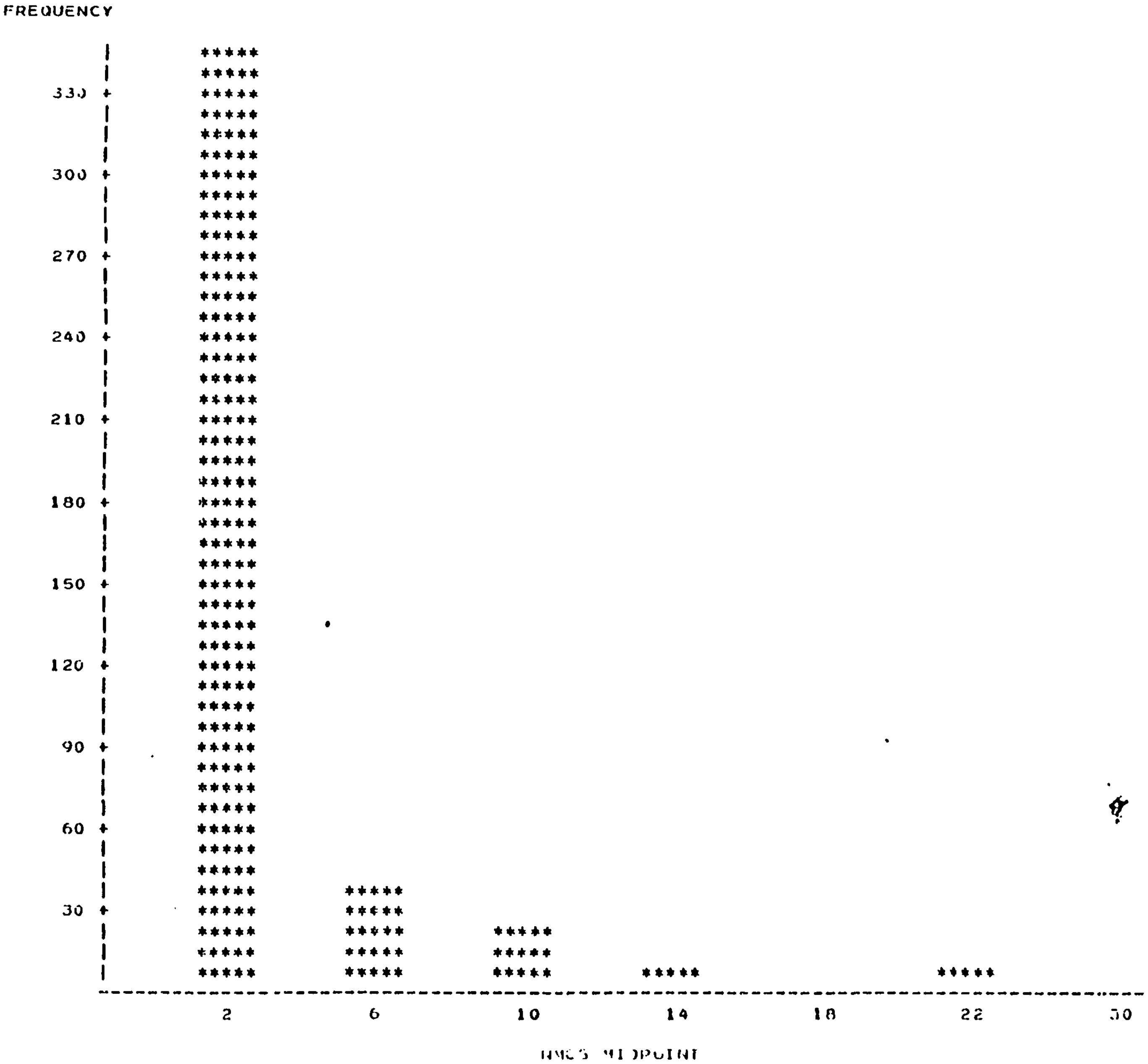


Figure B8. Frequency bar chart of deleted independent variable: Grade Point Average of Physics Credits earned up to September 1974 (GPAPS)

Figure B9. Frequency bar chart of deleted independent variable: Number of Mathematics Credits earned by September 1974 (NMCS)


FREQUENCY
330
300
270
240
210
180
150
120
90
60
30
2
6
10
14
18
22
30
MIDPOINT

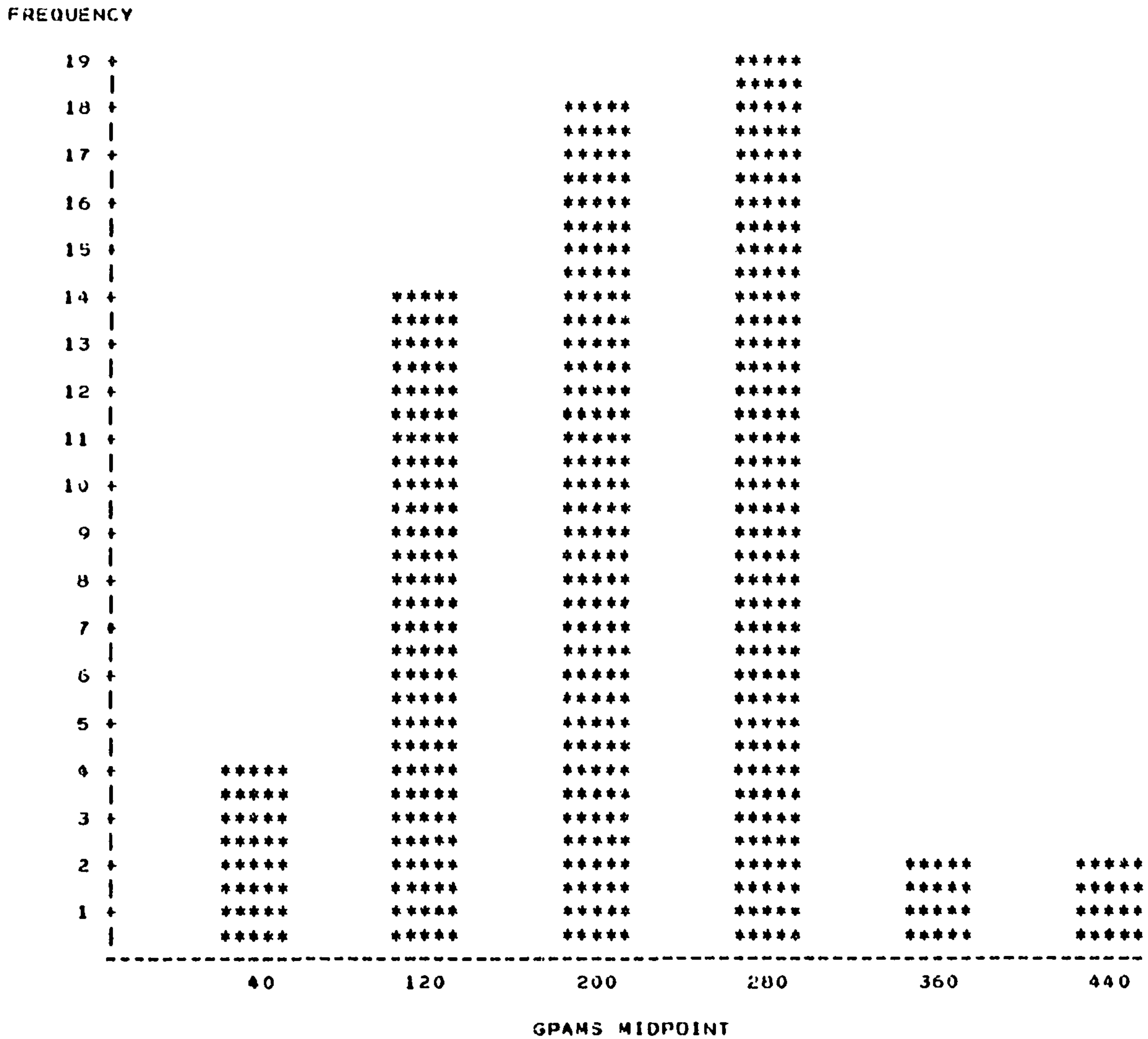
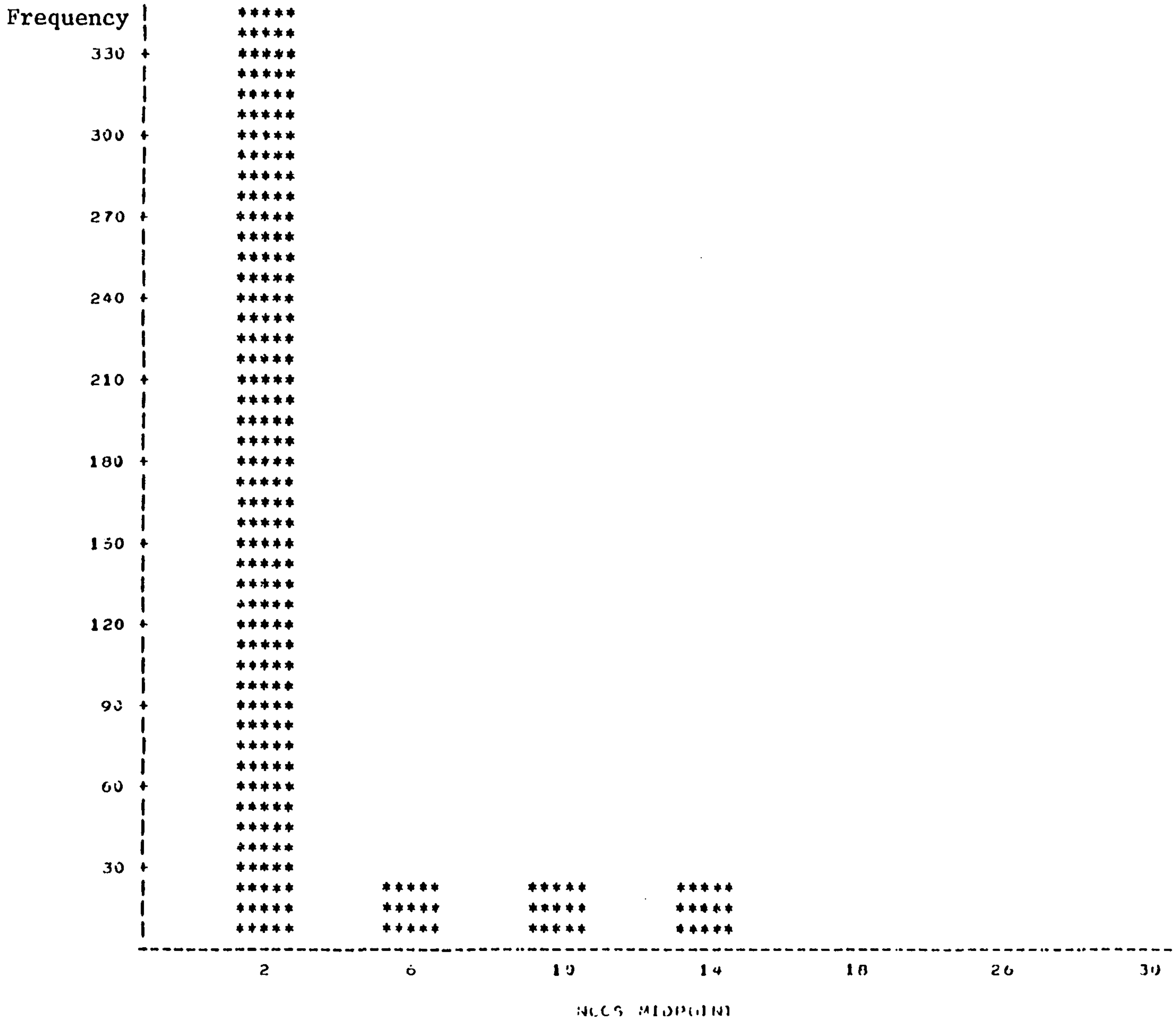


Figure B10. Frequency bar chart of deleted independent variable: Grade Point Average of Mathematics Credits earned by September 1974 (GPAMS)

Figure B11. Frequency bar chart of deleted independent variable: Number of Chemistry Credits earned by September 1974 (NCCS)


Frequency
330
300
270
240
210
180
150
120
90
60
30
2
6
10
14
18
26
30
NCCS MIDPOINT

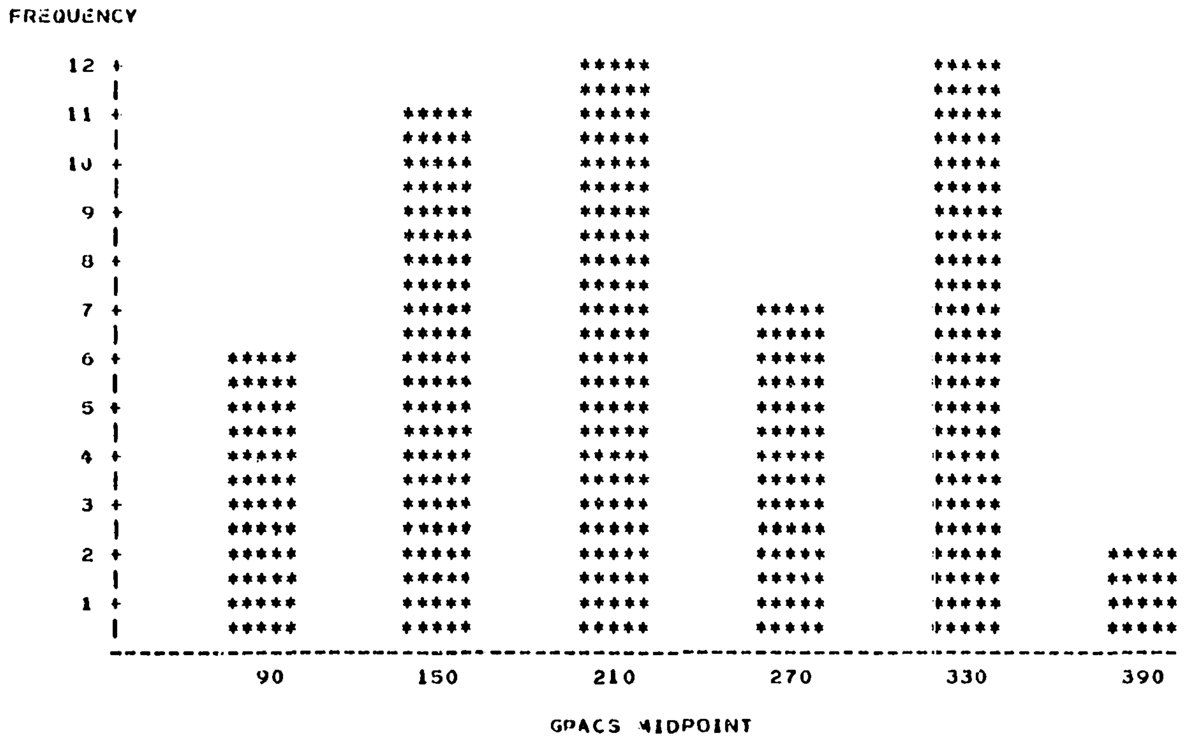


Figure B12. Frequency bar chart of deleted independent variable: Grade Point Average of Chemistry Credits earned by September 1974 (GPACS)

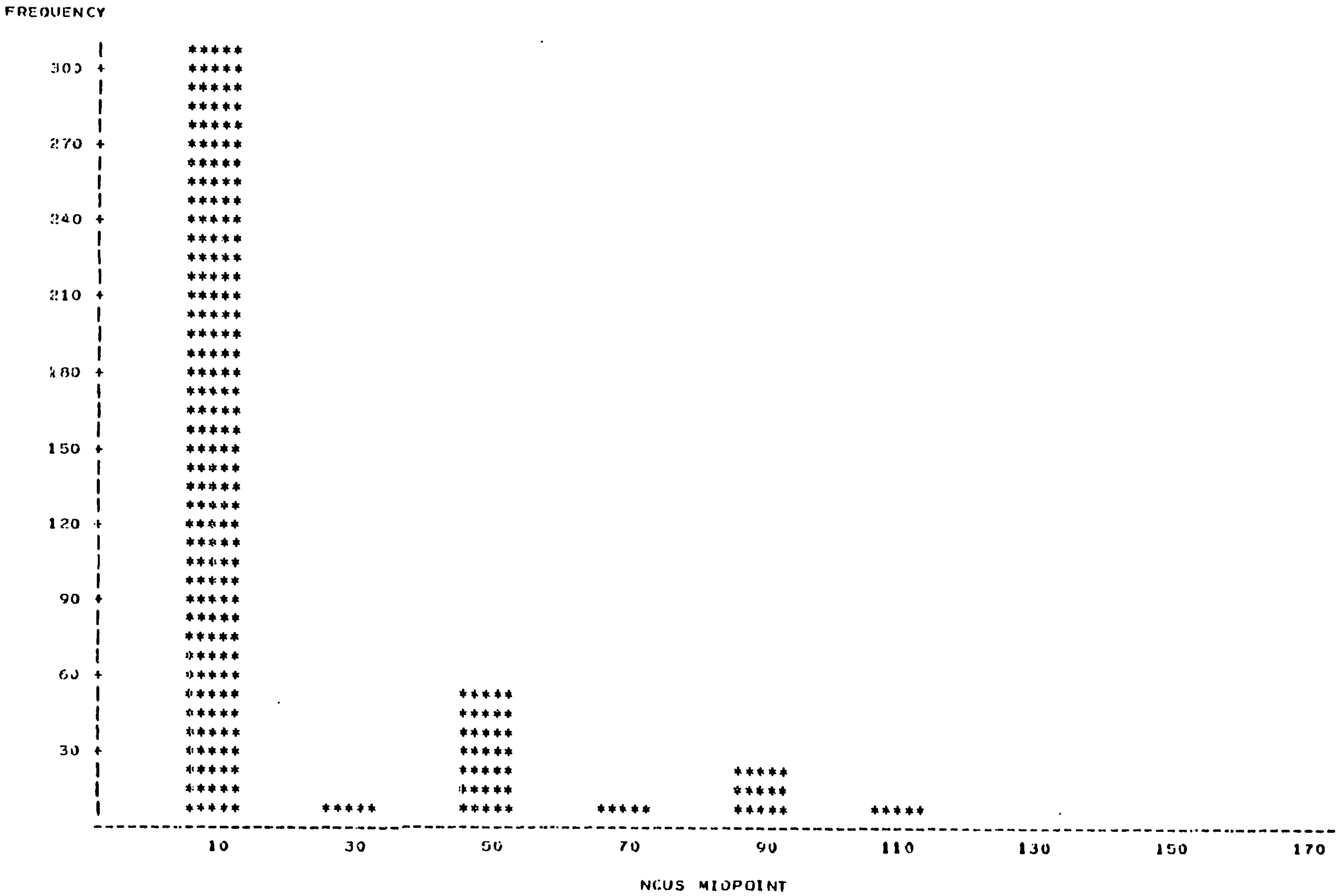


Figure B13. Frequency bar chart of deleted independent variable: Number of University Credits earned by September 1974 (NCUS)

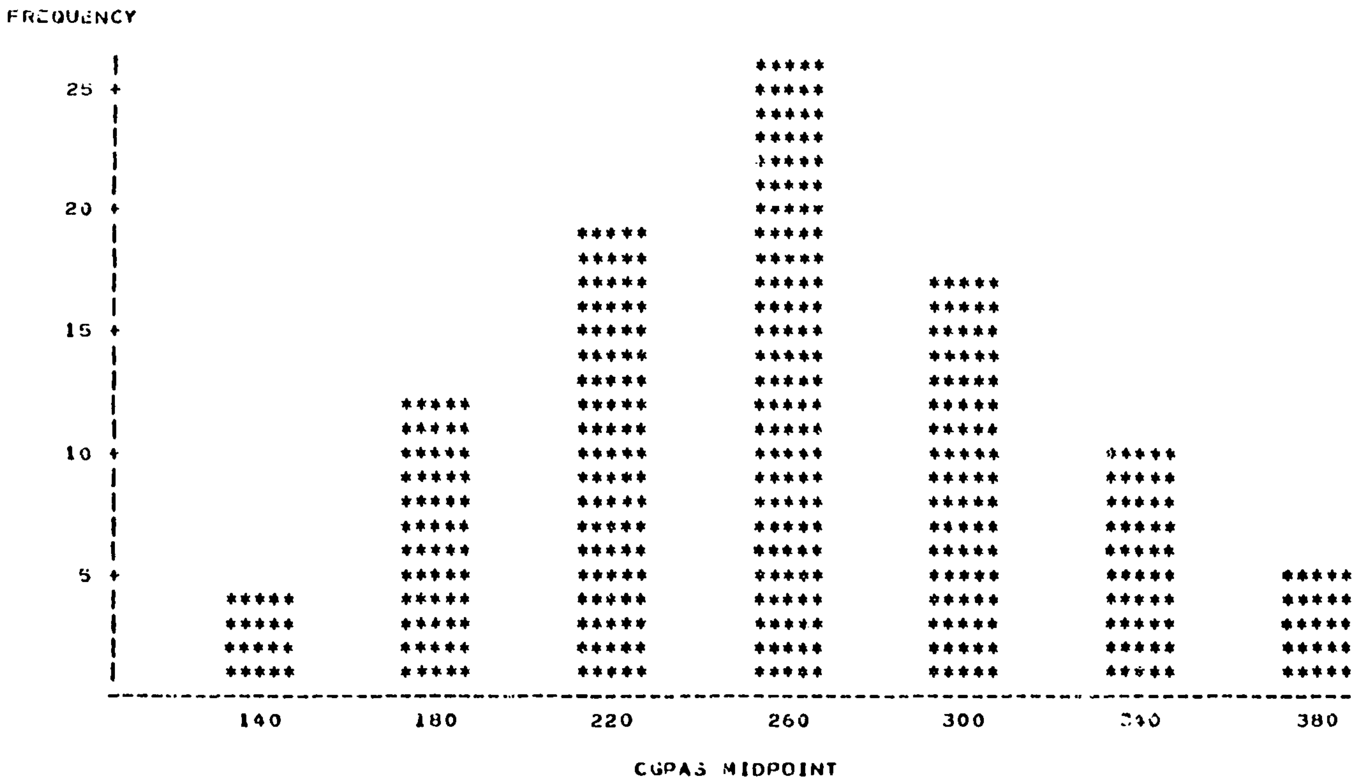


Figure B14. Frequency bar chart of deleted independent variable: Cumulative Grade Point Average of University Credits earned by September 1974 (CGPAS)

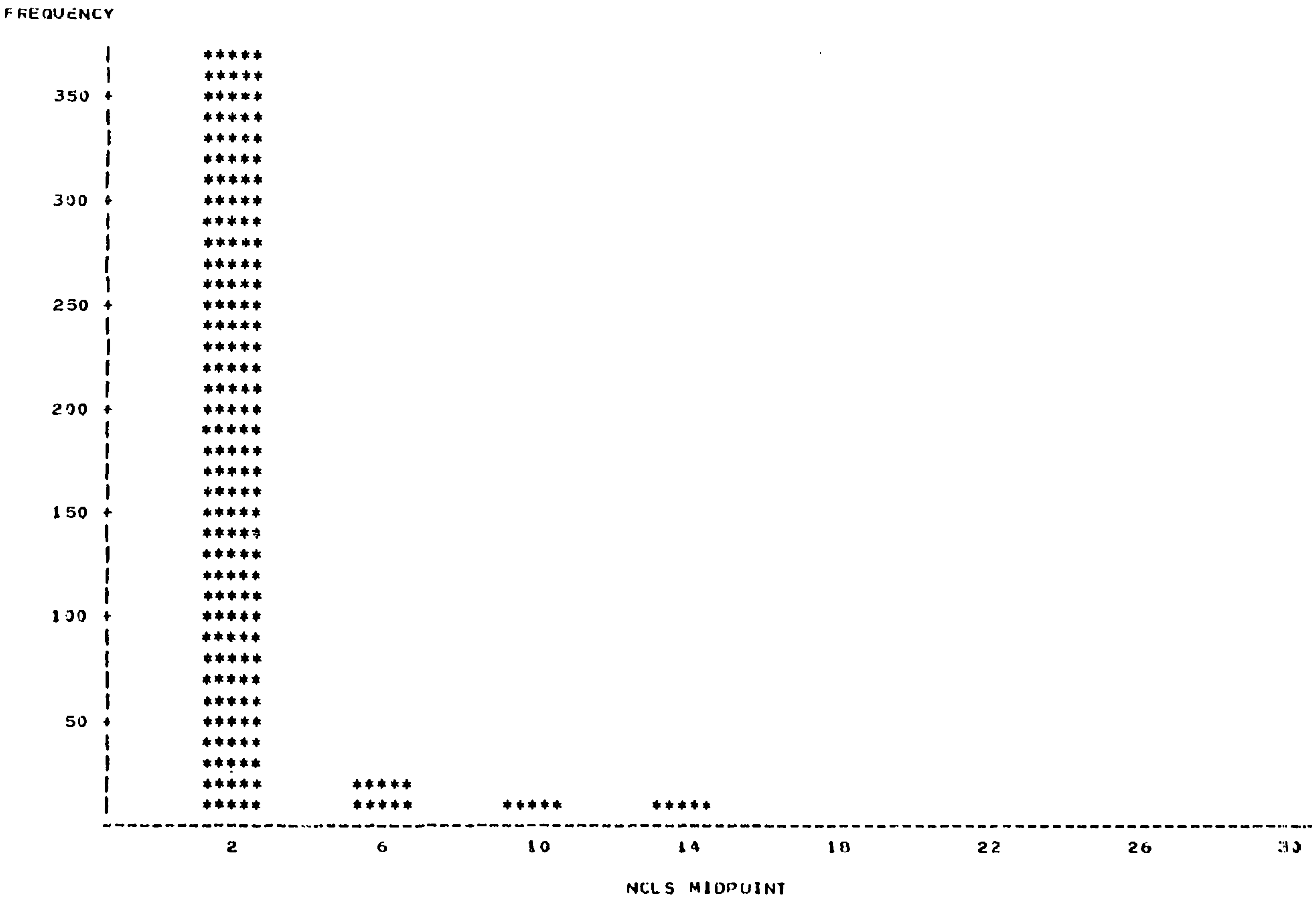


Figure B15. Frequency bar chart of deleted independent variable: Number of Credits in Life-Science earned by September 1974 (NCLS)

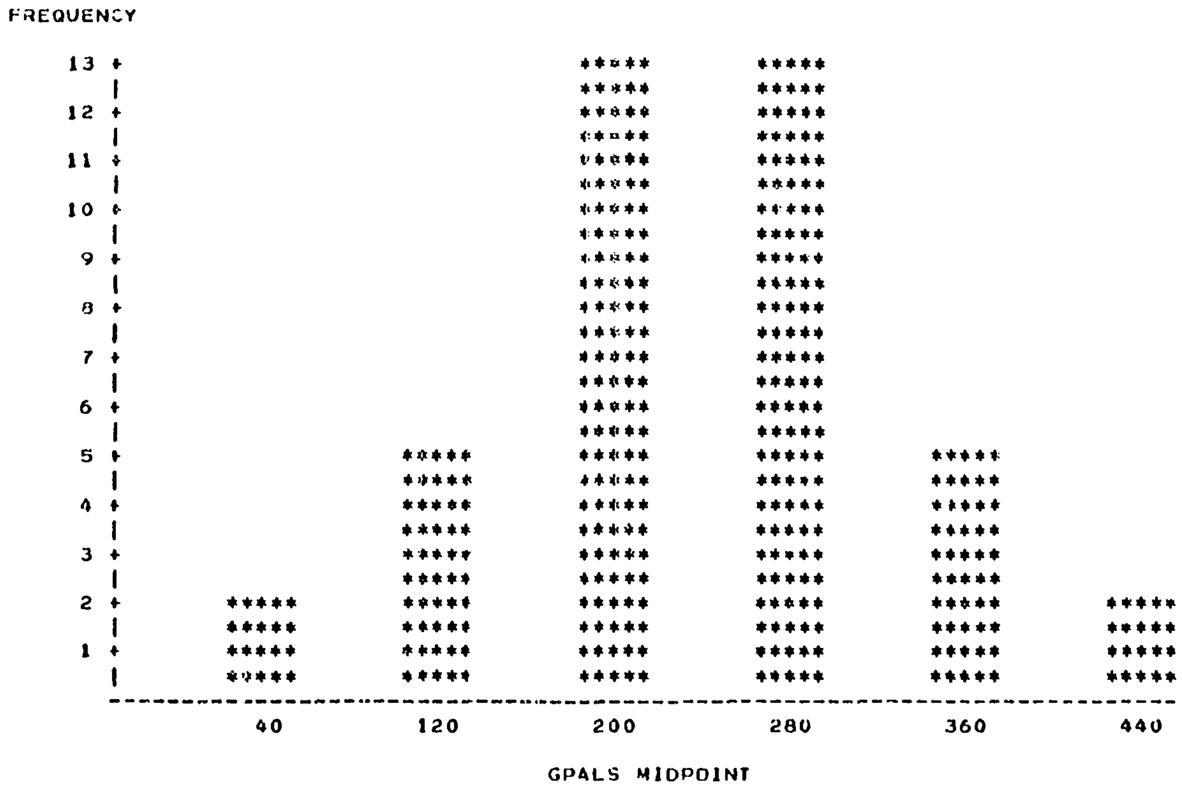


Figure B16. Frequency bar chart of deleted independent variable: Grade Point Average of Life-Science Credits earned by September 1974 (GPALS)

APPENDIX C: THE APPROVAL OF HUMAN SUBJECTS COMMITTEE

The Iowa State University Committee on the Use of Human Subjects in Research reviewed this project and concluded that the rights and welfare of the human subjects were adequately protected, that risks were outweighed by the potential benefits and expected value of the knowledge sought, that confidentiality of data was assured and that informed consent was obtained by appropriate procedures.