This study correlates the relationship between varying methods in student evaluation and its effect on student achievement and attrition. The sample studied consisted of 230 students enrolled in three separate semesters of General Biology IIA at Pasadena City College in 1972, 1973, and 1974. The earlier students were given longer exams over three or four weeks of material. The 1974 class was given frequent quizzes over smaller amounts of material. All three courses were taught by the same instructor. Results indicate that A and B grades constituted 40 percent of all grades given in the 1972 and 1973 semesters; C and D grades constituted the remaining 60 percent. These proportions were reversed for the 1974 semester, in which A and B grades represented 60 percent and C and D grades represented 40 percent. Furthermore, attrition percentage dropped from an average of 26 percent in the 1972 and 1973 semesters to approximately 12 percent during the spring 1974 semester. From these findings it seems that one of the crucial factors influencing final grades and attrition is the method used in student evaluation. Results are graphed and are compared using chi-square as a test of significance. A brief review of the literature relating to grading and evaluation systems is presented along with a 27 item bibliography. (DC)
A COMPARISON OF TWO METHODS OF EVALUATION AND ITS EFFECT ON ATTRITION AND FINAL GRADES IN GENERAL BIOLOGY

by

Thomas J. Belzer

Pasadena City College

A PRACTICUM PRESENTED TO NOVA UNIVERSITY IN A PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF EDUCATION

NOVA UNIVERSITY

DECEMBER 30, 1974
# TABLE OF CONTENTS

- Title of Practicum ........................................... 1
- Statement of Problem ....................................... 1
- Hypothesis ..................................................... 1
- Background and Significance of Study ....................... 2
- Definition of Terms ......................................... 11
- Limitations of Study ........................................ 12
- Basic Assumptions ........................................... 13
- Procedures For Collecting Data .............................. 14
- Procedures For Treating Data ............................... 14
- Significance of the Data .................................... 15
- Conclusions and Recommendations .......................... 17
- Summary ....................................................... 19
- Data Resulting from Study .................................. 20
- Bibliography .................................................. 27
BELZER, Thomas J., "A Comparison of Two Methods of Evaluation and Its Effect on Attrition and Final Grades in General Biology."
Nova University, 1974.

ABSTRACT

This study correlates the relationship between varying methods in student evaluation and its effect on student achievement and attrition.

The sample studied consisted of three general biology classes taught during three different semesters. Two classes were given longer examinations over three or four weeks of material. One class was given frequent quizzes over smaller amounts of material.

Chi square analysis was used to compare the classes in an attempt to discover significant differences in final grades and student attrition.
LEARNING THEORY AND APPLICATIONS

TITLE OF PRACTICUM

"A Comparison of Two Methods of Evaluation and Its Effect on Attrition and Final Grades in General Biology."

STATEMENT OF THE PROBLEM

The problem of this study is to investigate and compare the effects of two methods of evaluation used in General Biology 11A at Pasadena City College, and to determine what effect these methods had on final grade achievement and attrition. Both low student achievement as measured by final grades and high attrition rates have been areas of major concern by the faculty for the past several years. This study will attempt to show what relationships may exist between different evaluation procedures and these problems.

HYPOTHESIS

The null hypotheses were based on the assumption that different evaluative techniques used in biology classes had no effect on attrition or final grade distribution.

1. There is no significant difference in final grades achieved in General Biology 11A where evaluation techniques have been changed to shorter and more frequent evaluations.
2. There is no significant difference in attrition rates in General Biology 11A where evaluation techniques have been changed to shorter and more frequent evaluations.

BACKGROUND AND SIGNIFICANCE OF STUDY

During the past ten years General Biology 11A, a course for non-science majors, has been taught in the traditional manner with three hours of lecture and three hours of laboratory weekly. The enrollment during this period has steadily increased until the last two years when it has leveled off at approximately fifteen hundred new biology students annually. When this course was rapidly expanding, the Life Science Department staff decided to use a modified team-teaching approach in order to handle the large number of students most efficiently. The lecture classes averaged about ninety students, and the laboratory classes were established at approximately thirty students in each laboratory. The staff has felt that the drawbacks to large lecture classes has been at least partly offset by the smaller laboratory classes where there are closer teacher-student relationships.

During this period of growth there have been two primary areas of concern to the Life Science teaching staff. The first problem has been the poor academic achievement by many of the students taking General Biology. There have been studies conducted by the Dean of Institutional Research at Pasadena City College in an attempt to uncover the reasons for this.
Two or three facts have surfaced as partial explanations for unsatisfactory student achievement in biology classes. One study revealed that many students had poor math and science preparation in high school, and another study revealed that many students had poor reading ability. When a student had both problems, there was little chance of success in General Biology 11A.

The second area of concern has been the high attrition rates in the General Biology 11A classes. Students with poor math and science backgrounds, with poor reading ability, and with other problems have dropped out of the course in large numbers. Some of these students attempt the course another time, but many do not. There is some evidence that student attrition college-wide is related to the changes in college withdrawal policies. In the fall of 1969, the first withdrawal policy change occurred changing the withdrawal date from the sixth week to the last three weeks of the semester. Another change made in the fall of 1970 moved the withdrawal date back further to the last Friday before final examinations. The first policy change produced a 6.73% increase in attrition college-wide during the next two semesters. The second change produced a further 2.08% attrition increase throughout the college. Since these policy changes were made, attrition in university lower division courses has increased about 10% to a mean attrition of about 25% in these courses. It is obvious from these data that high student attrition since 1969 is not a problem unique to the Life Science Department and its biology courses but a college-wide problem.
These same changes in withdrawal policies have not only influenced the W grades but also shifted the distribution for A-F grades in the college. Prior to the fall of 1969 the most frequent grade was C; after the policy changes the most frequent grade was B. Since 1969 the percentage of A and B grades has increased, and there has been a reduction in C, D, and F grades. (See Figure 1)

The General Biology 11A course has been modified over this same period in an attempt to present a more meaningful course for the greatest number of students. Lecture material has been revised completely and new texts, study guides, and laboratory manuals have been added. One lecture class has been designed for audio-tutorial instruction with attempts to individualize the instruction. These new methods and materials do not appear to have reduced attrition or raised final grades in General Biology 11A. In a study conducted in 1973 at Pasadena City College comparing audio-tutorial instruction in biology (Fall, 1972 semester) with traditional instruction it was found that attrition was actually higher in the audio-tutorial class than it was in the traditionally taught class. Grades were not significantly improved either although students in the audio-tutorial biology program had a positive attitude toward the course.

The evaluation procedure first used in the traditional team-taught course was to administer four departmental exams plus a final exam during the semester. A point system was also used with a percentage grading scale of:

\[ 8 \]
Currently the point system is still used, and the grading scale is still followed by all instructors in this course. A major change was made two years ago, however, when instructors decided to write their own examinations from a common pool of test questions. Instructors still report that student achievement on these examinations has been generally very low. It does not appear that the changes made so far in this course have reduced attrition or improved student achievement in General Biology 11A.

Before the beginning of the spring semester, 1974, one instructor decided to change the system of evaluation slightly. While using the same point system and grading scale this instructor produced eleven quizzes, each consisting of ten questions, to be administered at the end of a unit of study. A midterm examination was given also plus a final examination. These two examinations were similar to the larger examinations given by all instructors in this course (100 item, multiple-choice question tests). This change and its apparent effect on grades and attrition is the object of this study.
The change in evaluation was a very simple undertaking. The large lecture examinations were broken up into quizzes that culminated the study of single topics in biology. The topics selected were:

1. History of Biology
2. Chemistry of Life
3. The Cell
4. Nutrition
5. Transport and Gas Exchange
6. Regulation
7. Reproduction
8. Genetics
9. Development
10. Evolution
11. Ecology

Three or four one-hour lectures were given over these topics, and then a ten-point, ten-question quiz was given to the entire class. The lectures were highly structured and well organized, and a variety of audio-visual materials were used to support the lectures. The midterm examination included questions through the topic Regulation. The final examination included areas of Reproduction through Ecology only.

The main motivation behind this change in evaluation procedure was to determine if this change could improve grades and reduce attrition significantly. Two years ago the book, *A Modest Proposal: Students Can Learn*, by Roueche and Pitman, provided a real source of stimulation
to this author. In this book one statement actually provided the motivation for the entire course change.

"If clear objectives are presented and instructional sequences carefully designed to allow practice and feedback, the student knows where he is going and how he is doing throughout the learning sequence. If the sequence is then divided into short steps, each leading logically to the next step, the chance of a learner's successfully achieving the objectives is greatly increased. Under such a system the learner is constantly reinforced. Success breeds the desire for more success."

The key words seem to be "short steps" and success. In General Biology the first examination is usually given after about four weeks of instruction. It has been reported previously that a high percentage of attrition occurs just after the first examination. After that exodus, attrition levels out to a steady rate until about 25% to 30% have dropped by the end of the semester. If one could avoid giving that examination and give short quizzes that could positively reinforce the learning earlier, it is likely both grades and attrition could be affected positively.

A. Cohen (1969) suggests that student failure is really school and teacher failure in that failure is the result of vagueness in instructional goals rather than weakness in student ability.9 Hutchins (1968) speaks of the same problem as "educational shock."18 Perhaps the General Biology staff needs to reassess what that first major examination over three or four weeks of material does to students. It may be that the "educational shock" is too great for many and ultimately leads to a withdrawal or a poor grade in the course.
The "quiz system" in General Biology, as it will now be referred to, is really a take-off on part of the "Mink Model" for learning. The "Mink Model" is based on the principle of reinforcement learning. It includes four basic components of learning theory:

1. Predispositions for learning
2. Structure
3. Sequence
4. Reward

The "quiz system" seems to be supported by what B. F. Skinner says. "Man is a biological organism reacting to stimuli in his environment. He is a product of his experiences, his past learning or conditioning, and of potential reconditioning." Skinner sees man's behavior as predominantly influenced by feelings, frustrations or satisfaction. It follows that if a student can be positively reinforced early in a course of study, he will likely achieve better results and persist in the course.

P. Cross (1971) states that fear of failure will be the major impediment for the "new" students of the near future. The "new" students will be those who rank in the lowest third on tests of academic aptitude, and it is the group which constitutes the greatest reservoir of attenders. Students in this group have revealed an attitude of passivity toward learning and lack of interest in intellectual pursuits. It is very obvious that positive conditions for learning will assume increasing importance if we expect this group of students to learn.
The "quiz system" as discussed here adds even more structure to the evaluation system in General Biology throughout the semester. There are also more opportunities for reward and reinforcement along the way. Since the sequence of learning events is shorter and more clearly defined, it is more likely the student will learn and remember what is taught more readily. Bruner (1960) summaries succinctly the present state of knowledge about remembering and forgetting:

"Perhaps the most basic thing that can be said about human memory after a century of intensive research is that unless detail is placed in a structured pattern, it is rapidly forgotten. Detailed material is conserved in memory by the use of simplified ways of representing it. What learning general or fundamental principles does is to ensure that memory loss will not mean total loss, that what remains will permit us to reconstruct the details when needed."\(^7\)

Roueche and Pitman (1972) make the following observation:

"Learning in its natural state proceeds from a response. A small child hears sounds, makes a noise (response), and is reinforced. If he is punished, he is less likely to make further sounds. If he is ignored, he also begins to become more quiet. In short, a person does something and is then directed by the feedback from his action."\(^8\)

The change in the evaluation system means ultimately that the student gets more feedback and, most likely, positive feedback. If this occurs, he goes on and finally succeeds in finishing the course successfully.

Finally, the tests used in General Biology II A since the course became a large team-taught course in many instances seem to have one primary purpose: To sort students into letter grade categories.
Kryspin and Feldhusen (1974) suggest that tests have a multiple purpose:

1. Diagnostic - highlighting student strengths and weaknesses
2. Motivational - students are prompted to succeed on tests
3. Self-Evaluative - students learn to develop their own self-evaluation skills
4. Instructional - provides for a review of course material
5. Improving Teacher Effectiveness - indicate to the teacher how well material was covered
<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attrition Rate:</td>
<td>Percentage of students withdrawing from a college course during a semester.</td>
</tr>
<tr>
<td>Incomplete (I-grade):</td>
<td>Grade issued when a student cannot complete a course for reasons beyond his control and is doing passing work.</td>
</tr>
<tr>
<td>Non-Major:</td>
<td>Students with academic majors outside of the Life Sciences.</td>
</tr>
<tr>
<td>Traditional Course:</td>
<td>Three hours of lecture weekly with a three-hour laboratory weekly.</td>
</tr>
<tr>
<td>Traditional Evaluation:</td>
<td>Three or four 80-100 question multiple-choice examinations during the course of one semester.</td>
</tr>
<tr>
<td>Withdrawal (W-grade)</td>
<td>Grade issued when a student drops a course if the student is passing at the time of drop.</td>
</tr>
<tr>
<td>1-99 Courses:</td>
<td>University or senior college lower division courses.</td>
</tr>
</tbody>
</table>
LIMITATIONS OF STUDY

This study was confined to General Biology 11A taught at Pasadena City College during three semesters of the 1972, 1973, and 1974 school years. These classes are not true random samples of the school population.

Only final letter grades earned by students in General Biology 11A will be used to indicate student achievement.

No attempt was made to determine the reasons for student withdrawal in General Biology 11A during the study period.

No attempt will be made to compare student achievement in other General Biology classes taught by other instructors.

No analysis of the test questions used during the study period was made.
**BASIC ASSUMPTIONS**

Basic assumptions concerning this study are as follows:

The three classes used in this study are reasonably homogeneous.

The same teacher for these three classes provided equal learning opportunities for all students enrolled.

The three classes in the study each have a normal distribution of intelligence.

The test questions selected for the three classes were reasonably comparable in level of difficulty.
PROCEDURES FOR COLLECTING DATA

Three classes taught in three previous semesters were studied, and grade distributions and attrition rates were compared. Only grades A, B, C, and D were compared since there were no F grades given during the study period and only two E grades. (A grade of E is given only when a student has been doing satisfactory work but is unable to complete requirements of the course for reasons beyond his control.)

PROCEDURES FOR TREATING DATA

The comparisons of final grades (A, B, C, and D) achieved in the three classes were made using chi square as a test of significance.

The comparisons of attrition rates in the three classes were also completed using chi square analysis.
SIGNIFICANCE OF THE DATA

In the calculation of chi square for student attrition comparisons over the fall semesters of 1972 and 1973 and the spring semester of 1974, it was found that $X^2 = 5.35$ for two degrees of freedom, and this was approaching significance at the 5 per cent level. (Significance at the 5 per cent level is 5.99) (See Table 2) Since this was so close to significance, the two semesters (1972 and 1973) were averaged together and compared with 1974, and another $X^2$ analysis was made. This time $X^2 = 5.56$ for one degree of freedom, which is significant at the 2 per cent level. (See Table 3) Therefore, the null hypothesis is rejected that there is no significant difference in attrition rates in General Biology 11A where evaluation techniques have been changed to shorter and more frequent evaluations. Attrition percentage dropped from an average of 26% in the 1972 and 1973 semesters to approximately 12% during the spring 1974 semester. (See Figure 3)

Grades distribution in the two General Biology classes taught during the fall semesters of 1972 and 1973 compare favorably with the college grade distribution for 1-99 courses. Over half of the grades given in those two semesters were C grades. A and B grades constituted about 40% of the other grades. (See Figure 1 and Figure 2)

During the spring 1974 semester the most common grade was the B grade. The B and C grades reversed themselves compared to the 1972 and 1973 semesters. It was interesting that in the 1972 and 1973 semesters
A and B grades totaled 40%, and C and D grades totaled 60%. At the end of the 1974 semester A and B grades totaled 60%, and C and D grades totaled 40%. (See Figure 2, Figure 3, and Table 3)

In a calculation of chi square for final letter grade comparisons it was found $X^2 = 15.57$, which for six degrees of freedom is significant at the 2 per cent level. (See Table 4) Therefore, the null hypothesis for grades is rejected. The significance difference in grades was analyzed further, and a chi square of B and C grades was calculated comparing the 1972 and 1973 semesters with the 1974 semester. It was determined that $X^2 = 9.28$ with two degrees of freedom, which is significant at the one per cent level. (See Table 5) This was the most significant grade change resulting from the "quiz system" adopted in the spring 1974 semester.
CONCLUSIONS AND RECOMMENDATIONS

Conclusions:

This study was limited to two hundred and thirty students in three classes taught in three separate semesters by the same instructor. This study examined the effect of changing evaluation procedures, and the results must be limited to these parameters.

The findings indicate that student attrition can be significantly reduced when the evaluation is conducted over small segments of course material. The findings also indicate that final letter grades can be significantly effected when shorter tests are administered during the semester. More specifically, it would appear more students will perform above average and finally realize a grade of B or better. From these findings it seems that one of the crucial factors influencing final grades and attrition is the method used in student evaluation.

Recommendations:

(1) Investigate the attrition rates in all other General Biology 11A classes.

(2) Replicate the study. This would provide an opportunity to further validate the findings.

(3) Implement the "quiz system" in other classes.

(4) Use the "quiz system" with a more variable self-pacing individualized General Biology course. Students could take quizzes when they are "ready."
(5) If the "quiz system" is not adopted, at least give more examinations during the semester over smaller amounts of course material.

(6) Write definite course objectives that clearly identify the material that the instructor wants the students to learn.

(7) Determine students' attitudes toward evaluation methods used in General Biology 11A classes.

(8) Determine what the reasons are for student withdrawals.
SUMMARY

The problem of this study was to investigate and compare the effects of two methods of evaluation used in General Biology 11A. More specifically, two questions were asked:

(1) Could attrition be reduced with a change in evaluation procedures?
(2) Could grades be improved with a change in evaluation procedures?

These questions were answered affirmatively based on the results from this study. It is clear that the "new" students of the 1970's are here. Teachers must look for ways to motivate these students if these students are to learn.

There are many changes that can be made to make instructional programs more effective. If as much thought and planning went into the development of classroom tests as goes into the development of course content, the instructional program is very likely to improve in many ways.

This study provides a base for future studies and experimentation in the evaluation area. It is hoped that new imaginative instructional strategies will more perfectly meet the needs of students in the years ahead.
FIGURE 2
COMPARISON OF FINAL GRADES
(in Per Cent)

□ 1972 & 1973
□ 1974

BEST COPY AVAILABLE
FIGURE 3

COMPARISON OF GRADES 1972 - 1974
<table>
<thead>
<tr>
<th>Class</th>
<th>Number</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>W</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td>1972</td>
<td>81</td>
<td>3</td>
<td>15</td>
<td>32</td>
<td>7</td>
<td>23</td>
<td>1</td>
</tr>
<tr>
<td>1973</td>
<td>67</td>
<td>9</td>
<td>16</td>
<td>24</td>
<td>2</td>
<td>16</td>
<td>0</td>
</tr>
<tr>
<td>1974</td>
<td>82</td>
<td>10</td>
<td>33</td>
<td>21</td>
<td>7</td>
<td>10</td>
<td>1</td>
</tr>
</tbody>
</table>
TABLE 2

CALCULATION OF CHI SQUARE FOR ATTRITION COMPARISON

<table>
<thead>
<tr>
<th>Class</th>
<th>Withdrew</th>
<th>Persisted</th>
</tr>
</thead>
<tbody>
<tr>
<td>1972</td>
<td>23</td>
<td>67</td>
</tr>
<tr>
<td>1973</td>
<td>16</td>
<td>51</td>
</tr>
<tr>
<td>1974</td>
<td>10</td>
<td>72</td>
</tr>
</tbody>
</table>

1972 18.45 71.55 (Expected)
1973 13.74 53.26
1974 16.81 65.19

\[ X^2 = 5.35 \] which with df = 2 is approaching significance at the 5 percent level.

TABLE 3

CALCULATION OF CHI SQUARE FOR ATTRITION COMPARISON

<table>
<thead>
<tr>
<th></th>
<th>Withdrew</th>
<th>Persisted</th>
</tr>
</thead>
<tbody>
<tr>
<td>1972 &amp; 1973 (Averages)</td>
<td>20</td>
<td>54</td>
</tr>
<tr>
<td>1974</td>
<td>10</td>
<td>72</td>
</tr>
</tbody>
</table>

1972 & 1973 (Averages) 15.8 66.2 (Expected)
1974 14.2 59.8

\[ X^2 = 5.56 \] which with df = 1 is significant at the 2 percent level.
### TABLE 4

**CALCULATION OF CHI SQUARE FOR FINAL LETTER GRADE COMPARISON**

<table>
<thead>
<tr>
<th>Class</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>(Observed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1972</td>
<td>3</td>
<td>15</td>
<td>32</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>1973</td>
<td>9</td>
<td>16</td>
<td>24</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>1974</td>
<td>10</td>
<td>33</td>
<td>21</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>1972</td>
<td>7</td>
<td>20.4</td>
<td>24.5</td>
<td>5.1</td>
<td>(Expected)</td>
</tr>
<tr>
<td>1973</td>
<td>6.3</td>
<td>18.2</td>
<td>21.9</td>
<td>4.6</td>
<td></td>
</tr>
<tr>
<td>1974</td>
<td>8.7</td>
<td>25.4</td>
<td>30.6</td>
<td>6.3</td>
<td></td>
</tr>
</tbody>
</table>

\[ \chi^2 = 15.57 \] which with df = 6 is significant at the 2 percent level.
<table>
<thead>
<tr>
<th>Class</th>
<th>B (Observed)</th>
<th>C (Observed)</th>
<th>B (Expected)</th>
<th>C (Expected)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1972</td>
<td>15</td>
<td>32</td>
<td>21.3</td>
<td>25.7</td>
</tr>
<tr>
<td>1973</td>
<td>16</td>
<td>24</td>
<td>18.2</td>
<td>21.8</td>
</tr>
<tr>
<td>1974</td>
<td>33</td>
<td>21</td>
<td>24.5</td>
<td>29.5</td>
</tr>
</tbody>
</table>

\[ X^2 = 9.28 \] which with \( df = 2 \) is significant at the 1 percent level.
BIBLIOGRAPHY

1. Alter, M., "Retention As A Function of Length of Retention Interval, Intelligence and Training Time." Journal of Programmed Instruction, 2, 7-17, 1963.


