

An Experimental Testing Program in Elementary Chemistry: A Preliminary Report

BILLIE G. BROACH and HOWARD L. HODGES
Department of Chemistry, University of Arkansas at Little Rock,
Little Rock, Arkansas 72204

ABSTRACT

An experimental testing program is described which utilizes questions that are partly computer composed, in addition to a section composed by the instructor, and a retesting option to the student. Results from a trial of the program for one term indicate that (1) course grades were improved, (2) the student withdrawal failure rate was unaffected, and (3) the employed students took greater advantage of the retest than did the unemployed students.

INTRODUCTION

In the fall of 1975 the authors introduced an experimental testing program at the University of Arkansas at Little Rock in the course titled, "Elementary Chemistry I." The course is designed for several categories of students: those who have had no high school chemistry; those who feel their chemistry background is weak as a result of inadequate high school preparation or an interruption of several years in their college career; those whose background in mathematics is weak; those who are pursuing professional or preprofessional careers in nursing, home economics, or agriculture; and those who need one semester of an introductory laboratory physical science course to fulfill graduation requirements. The student population in this particular class included 13 declared science majors in chemistry, physics, biology, or engineering; 32 students in health-related sciences such as nursing, premedicine, prepharmacy, medical technology, radiology technology, dental hygiene, physical therapy, and respiratory therapy; 6 students in such fields as law enforcement, psychology, sociology, and physical education; and 4 students who had not declared a major.

Experience had shown that students who have little or no confidence in their ability to succeed in chemistry, for whatever the reason, pose a challenge. It was also apparent from experience that these students were the ones who ventured timidly and reluctantly into Elementary Chemistry I. Various approaches to testing had been taken previously, and late in the fall of 1974 a grant to implement a new method of testing was applied for and received from the Donaghey Foundation through the Innovative Teaching Committee at UALR. The proposed testing program was as follows. An extensive pool of multiple choice questions would be compiled by the authors and computerized by topics; at test time multiple tests would be generated by the computer over selected topics from the pool of questions; a second section written by the instructor, including problem solving and discussion questions, would be duplicated and would complete the test packet. Students would retain the computerized part of the test and turn in only the standard answer sheet form and the duplicated part. The answer sheets would be processed by the computer and a printout of the results of the first section posted; the second section would be graded by the instructor and returned to the student. An optional discussion period would be scheduled at a time other than regular class time for questions concerning the test and the material covered on it. Shortly thereafter a three-hour period of time would be set aside to enable a student to take a retest over the same material. The three-hour period was chosen arbitrarily as a compromise between a full day for retesting which seemed highly desirable and the block of time that could be worked into the instructor's schedule.

The authors' ambition was to collect definitive data on several unanswered questions.

1. Would students achieve at a higher level in a test-retest situation than they would without the retesting?
2. Could the high percentage of withdrawals and failures that plague this course be reduced?
3. Could some of the trauma experienced by many students when confronted by a test be reduced?

4. Was there a pattern of achievement and/or retesting of the employed versus the unemployed student?

PROCEDURE

Mechanically, it was decided to use three questions per topic from the pool of questions for the computerized part of the test. For a 20-question test this meant a subset of 60 questions from the large pool. This number was used to ensure that the questions acquired by the students would be limited and that the subset could be replaced in the pool with relative ease. Questions from the first test did not reappear on the retest or the final examination although similar ones replaced them. Students were required to turn in both sections of the retest and the computerized part was not returned to them. However, any student could obtain the computerized test for study in the instructor's office at any time after the results of the retest were posted. For both the initial test and the retest three parallel tests for the noncomputerized part were prepared and identification of the three was simplified by duplicating each test on paper of a different color (test A on blue paper, test B on yellow, and test C on white, for example). This practice seems to be common among instructors of large classes. A room for the retest period was designated and students were allowed to come in at any time during a three-hour time period to be retested. It was possible for a student to spend the entire period on the test and the 50-minute classroom limit was not observed.

The method of testing was well received by the students. The grading scale was fixed and the scores were not curved. Retesting was optional and the higher of the two scores achieved by the student was recorded. There were several problem areas. It was necessary to schedule computer time for test generation rather far in advance because the computer was new to the campus and subject to being shut down frequently. There was often a somewhat long time lag between testing and receiving test results from the computer. It was difficult at first for the students to read computer "handwriting" and to fill in the ID field on the answer sheets properly. The method of testing also is very time consuming from the instructor's standpoint. However, by the end of the semester most of these problems had smoothed out. The wide margin on the right side of each computer test was especially useful to the student and the ability to keep the computerized part of the first test seemed helpful.

RESULTS

The preliminary results of the testing were not as decisive as had been anticipated. Table I summarizes the results of the test-retesting. Column I shows a breakdown by letter grade of the total number of students who took the initial test for each of the three lecture tests given during the semester. Column II shows the number who attempted to raise their score by retesting. The percentages shown in parentheses were obtained from the numbers in Columns I and II. Column III shows the number of students attempting to raise their grades who were successful. The percentages in this column were obtained from the numbers in Columns II and III. A larger

percentage of students had been expected to take advantage of the retest opportunity than actually participated. Numerically, the totals are about the same for all three tests. However, 49% of the students elected to retest on test 3, and the percentage of those who raised their scores was encouraging. As anticipated, the students who took the retest were predominantly in the C, D, and F range.

Statistically, 34 students, about 62%, took advantage of the retesting sometime during the semester; 19 students retested only once, 16 retested twice, and 3 students retested on all three tests. The benefits of the retesting are summarized in Table 2. Retesting on the final examination was not permitted.

One of the most striking observations about the testing results is that not even one student who withdrew from the course took advantage of the retesting program. It is especially surprising because many of these students did not withdraw until after the third test. Further study of this finding is indicated. It is also of interest to note that a change from a score that would place a student, for instance, in the low "C" range without retesting but in the high "C" range by retesting would not be reflected in the "Change after 3 tests" column, but was reflected in the "Change in final grade" column where the score was averaged with those for the remaining tests of the semester.

In the final analysis 18 of the 55 students in this study who began the course, or 32.7%, were able to achieve a higher final grade through retesting than would have been possible without retesting. If this figure is based on the 40 who completed the course, the percentage rises to 45.0%.

In an effort to determine whether or not this method of testing was discriminatory to the employed student, a study was made of the retesting pattern of the employed versus the nonemployed student. The questionnaires filled out by the students attending classes during the first week of the semester were examined carefully and the findings are presented in Table III.

Contrary to what might have been expected, a substantially higher percentage of employed students participated in the retesting program than did the nonemployed students. The withdrawal rate contrast is not as dramatic but the rate is still noticeably higher for the nonemployed. Any change in the employment status of the students during the semester was not reported to the authors, but might have some bearing on the statistics. However, conclusions that can be drawn from the figures in the table are that the employed student seems to have been able to find time for retesting and that the withdrawal rate is lower for these students.

Table I. Summary of Tests and Retests by Student Populations

Grade	Column I No. who earned this grade without retest			Column II No. attempting score rise by retest			Column III No. raising score by retest		
	Test			Test			Test		
	1	2	3	1	2	3	1	2	3
A	12	3	4	0(0%)	1(33%)	1(25%)	0(0%)	0(0%)	0(0%)
B	11	2	5	3(27%)	0(0%)	0(0%)	0(0%)	0(0%)	0(0%)
C	14	12	11	6(43%)	5(42%)	5(45%)	2(33%)	1(20%)	5(100%)
D	4	10	6	1(25%)	6(60%)	4(67%)	1(100%)	4(66%)	3(75%)
F	14	18	15	8(57%)	9(50%)	6(40%)	7(88%)	5(56%)	4(75%)
Total	55	45	41	18(33%)	21(47%)	20(49%)	10(56%)	10(48%)	12(60%)

Table II. Testing Results

Course grade after 3 tests	By grade			By degree of change	
	Students using no retesting	Students raising grade by retesting	Student not raising by retesting	Change after 3 tests	Change in final grade
A	1	2	1	1 C A	1 D A
B	4	7	1	1 B A	2 B A
C	3	3	4	6 C B	2 C A
D	2	1	2	1 D B	1 D B
F	4	0	3	3 D C	6 C B
W	15	0	0	1 F D	4 D C
Subtotals	29	13	12		1 F C
Total			55	13	2 F D
					18

Table III. Record of Employed Versus Nonemployed Students

	Initial enrollment	Participants in retesting	Withdrawals	Completing course
Employed	32	23 (71.8%)	8 (25.0%)	24 (75.0%)
Nonemployed	23	12 (52.2%)	7 (30.4%)	16 (69.6%)
Total	55	35 (63.3%)	15 (27.2%)	40(72.7%)

DISCUSSION

A comparison was made of the withdrawal-failure percentages in the fall semesters of 1973, 1974, and 1975. In the fall of 1973, the withdrawal-failure percentage for Elementary Chemistry I was 34.7%. In the same semester of 1974, it had dropped to 20.0%. This dramatic reduction is believed to be the result of an exclusive computer testing-retesting program in which the retest questions were taken from the same pool as the test questions. The authors recognized that students could easily increase their test scores by acquiring a pool of questions and memorizing the answers without understanding the material. It is believed that this held students in class but was not a desirable method of testing. The percentage rose to 38.1% in the 1975 fall semester; this increase is believed to reflect the change in the testing procedure. Certainly the 38.1% rate of fall 1975 indicates a failure to retain or even approach the low of the previous year. Little comfort can be drawn from the fact that not one of the students who withdrew from the course elected to take even one retest. Without exception these students were unattracted to the program.

The following comments can be made concerning the questions the authors are seeking to answer.

1. A substantial number of students, 32.7% of the total beginning students and 45.0% of those who completed the course, were able to achieve a higher grade against a fixed, uncurved grading scale with the test-retest program. No

attempt was made in this study to determine whether or not students gained a better understanding of the material covered except from the information gleaned from the comparison of the withdrawal-failure rates.

2. Disappointingly, the test-retest program was not as successful as the previous year's program in reducing the withdrawal-failure rate. However, it is believed to be a more valid program. Further study will be made of succeeding classes.
3. The question of the reduction of testing trauma through the test-retest program has yet to be resolved by data. It was evident from informal student response that a large number of students would have taken advantage of a retest on the final at any hour of the day or night had it been available. This indication does not constitute hard evidence, however, and a questionnaire on this topic will be administered as the study proceeds.
4. The data collected for this study show that the employed student takes advantage of the test-retest opportunity and is more successful than the nonemployed student in raising his score and grade.

Additional data will be collected and a comparison made of the achievement of two parallel classes differing only in the method of testing.