In the spring of 1978, the individualized Latin program at the Ohio State University underwent an evaluation to determine the effectiveness of the experiment and to consider making the program a permanent part of the curriculum. Although the competency-based learning (CBL) component had worked satisfactorily for a majority of the students, about 27% of the students had received failing, incomplete, or inflated grades as a result of the testing system. In adjusting the program, it was felt that by focusing on the testing system, the greatest amount of improvement could be achieved with the least amount of damage to the concept of a variable paced variable credit program. Problems included the tendency of CBL to siphon off "under-achievers" to the classroom track, the questioned applicability of CBL to foreign language study, and the "frustration factor" in repeated examination. Also, it was determined that 80% passing levels gave both students and instructors a false sense of security. A new testing system was devised that defined true mastery as 90%. Students are limited to two repetitions of testing. Petroactive testing is an experimental aspect of the new system. Finally, quizzes are now being graded and have become part of the overall unit grade. (JB)
In the autumn of 1975, the Classics Department of The Ohio State University began an experiment in individualized Latin. This experiment later became a part of a grant from the National Endowment for the Humanities for individualized instruction in six foreign languages: Arabic, French, German, Latin, Russian, and Spanish. In the original proposal, the individualized projects were defined by three components: variable pacing, variable credit, and competency-based learning. The individualized track was to run as an alternative to the lock-step classroom in the elementary language programs and to provide a place in these programs for intensive students, slower students, and students whose work or study schedules would not allow time for a five hour per week classroom meeting. The purpose of this report is to survey the competency-based learning component of the Latin program and to relate the results of this part of the experiment and the subsequent modifications.

When the competency-based learning (CBL) component was first defined for the program, 80% was set as the lowest passing grade on all quizzes and tests. According to the original plans, there would be one quiz for each lesson, a practice test for each unit,1 and an achievement test for each unit. The quizzes and practice tests were to be graded on a pass/not pass basis (with 80% as the minimum for a pass), while the achievement tests were to carry letter grades (with a B− as the minimum for a pass). Students who could not achieve the 80% minimum on any given quiz or exam would be required to restudy the material and take alternate forms of the test until the minimum level of proficiency was attained. The assumption was that since the students were no longer hampered by requirements of time and deadlines, they would be expected to use the program not simply to complete the
language requirement, but to gain some proficiency in the subject as well.

In the spring of 1978, the Latin program was scheduled to undergo a departmental evaluation to determine the value and effectiveness of the experiment and to consider making the program a permanent part of the Latin curriculum. In this evaluation process, the CBL component of the program was reviewed carefully; as a result of this review, it was recommended that CBL undergo substantial modifications, so that it might be more effective and beneficial to the program as a whole. Several problems, it seems, had developed as a direct result of the CBL component; these problems had surfaced in a thorough analysis of the data for the project.

Although the CBL component had worked satisfactorily for a majority (about 65%) of the students, about 27% of the students over a three year period had received failing grades, incomplete grades, or inflated grades as a direct result of the testing system. Since the program had had an enrollment of well over 1000 students in that three year period, the number of problems was significant. A case study of the testing procedure revealed the following:

1. Because the program had been described as one in which the minimum passing grade was a B-, a number of students (as stated in discussions with them and on evaluation forms) felt that the program guaranteed a B-. They placed the burden on the instructional staff and the materials to provide a relatively high grade in a difficult subject; their own efforts in the learning process had, consequently, suffered.

2. Approximately 12% of the students had required more than two alternate forms on quizzes and tests in the early stages of the course. Given the limited amount of material covered in the early phases of a language course, it was often difficult, if not impossible, to come up with valid alternate forms beyond the second alternate. This rate of failure continued throughout the period of the project, despite the many and varied attempts at providing remedial drills and prescribing ways by which the students might
absorb the necessary material.

3. At the later levels of the program, the number of students requiring more than two alternate forms increased to approximately 21%. While, at this point, alternate tests were easier to make up, prescriptive assistance was much more difficult because of the vastly increased amount of material to be covered.

4. At the end of the quarter, when the traditional amount of procrastination during the quarter on the part of the students begins to take its toll, there were always significant numbers of students who found themselves without sufficient time to achieve "mastery" by means of alternate tests. The instructional staff was frequently left with three unsatisfactory alternatives. They could assign incomplete grades to students who could not make the 80% minimum on their last test. An incomplete grade, however, should not, by definition, exist in a program in which variable credit is available, unless, of course, a valid excuse is presented to the instructor. As a rule, incomplete grades led to even greater difficulties in the following quarter. They could assign a failing grade (E), but this often seemed quite unfair for students who fell into the 70% range. As a last resort, a barely passing grade (B-) could be assigned to a test that was close, but not actually passing. This last resort frequently led to grade inflation of the worst kind.

Although many of the problems described above could be remedied by other changes in the program, it was felt that by focusing on the testing system, the greatest amount of improvement could be achieved with the least amount of damage to the basic concept of a variable paced and variable credit program. During the evaluation process, certain questions arose regarding the validity of CBL in a foreign language; these questions could not be answered sufficiently enough to merit the retention of that component in its original state if the project was to succeed.
First, one of the natural effects of CBL in the individualized program was to siphon off "under-achievers" to the classroom track. Although frequently the discipline of the classroom was good for the students who had not succeeded in the individualized program, significant numbers of students found themselves in even worse circumstances when faced with a five day per week class that moved at a relatively demanding pace. They seemed to be students who could clearly benefit from the less competitive atmosphere of the individualized track, but who, because, of the alleged demands of the testing system, had found themselves with no alternative but the classroom. The classroom Latin courses thus became strained by the sharp divisions between the well-prepared students and the slower students who had not succeeded in the individualized program.

A second question dealt with the applicability of CBL to foreign language study. It was acknowledged that most of the material in the first three quarters of Latin concerned the basic skills. Basic skills in a foreign language, however, should not always be tested in a vacuum, that is, without a context that demands thorough retention of previous material in a cumulative manner. To learn forms and rules is one thing; but to apply those forms and rules to new and old contexts is something which re-study and re-testing cannot necessarily achieve. Thus, the notion of competency, or mastery, becomes extremely complex beyond the first few stages of a foreign language. Several students who had failed to achieve 80% on a given test had done so not because they did not know the current material, but because their retention of previous material was weak, or because they had great trouble applying concrete rules to the abstract thought process that the idiom of language is meant to convey.

A third question concerned what can be termed the "frustration effect." Students who are made to retake examinations frequently become frustrated after their second try. The frustration begins to take its toll by the third try (second alternate), where, in 37% of the cases the grades actually began to decline. In some instances, where more alternates were not available, a student would be
given as the third alternate a test form which he or she had already seen. In every such case that this was tried (to our knowledge), the grade on the second effort at the same form was even lower than the first try. In some of these cases, students were allowed to bend the rules and move ahead with less than passing grades. These students, as part of an experiment, were told to come back to the weaker tests later. In this instance, every student was able to pass exams that were previously "unpassable." The experimentation with retroactive testing led to some of the modifications that were eventually implemented into the individualized program.

Finally, it was determined that 80% passing levels gave both students and instructors false senses of security. It was originally felt that since knowledge was cumulative, passing knowledge in one lesson would lead to passing knowledge in the next. Lack of knowledge, however, is also cumulative. A student who has passed three tests at 80% (given, for example, 100 items on each test) has missed 60 items out of 300. If, in a random test on the fourth unit, a disproportionate amount of those 60 items reappears in a new format, that student may find that he or she has suddenly dropped to a 60% or 70% level. Despite the valid reasons for occurrences of this nature, such a drop can only serve to frustrate the student and lead to claims of deception. Even with intensive remedial work, the potential for ever-increasing gaps in knowledge grows with each test, even though "mastery" may have been achieved. Again, remedial prescriptions become increasingly difficult over a long period of time, since the amount of material covered increases with each unit. Efforts made at constructive grading of the quizzes and tests did little to remedy the deterioration effect, which, though it occurs in all courses, is harder for students to comprehend and deal with in a CBL situation.

The results of the three year experiment led the department to consider major modifications in the testing system of the individualized program. The academic year 1978-1979 was given over to, among other things, the development of an improved testing program. The new program, begun in the summer of 1979, is a combination of
aspects of CBL, aspects of the traditional classroom system, and some altogether new policies.

Basically, the new testing system defines true mastery, especially in the early stages of the language, as 90%. Even though that is still an imperfect definition and can lead to results similar to those of the 80% level, it is a more realistic definition of what is necessary in the beginning of a language to succeed with high grades throughout the four quarter requirement. The newest feature of this system, however, is that it makes "mastery" an option open to the students to select or reject as they desire. Thus, all grades (A, B, C, D, and E) are now possible in the individualized program, just as they are in the classroom.

Students are told that if they pass a test with a 90% or better, they will automatically be sent ahead to the next step. If they receive below a 90%, they are given the option of proceeding or retaking the exam up to two additional times. Thus, the student is restricted to two alternates, and the last grade received, not the highest, is the one that counts. The burden for improvement and mastery is now placed on the shoulders of the student. Success is well defined for the students, and they are warned about the pitfalls of early low grades. Nevertheless, weaker students are not automatically relegated to the classroom.

At the end of the quarter, if certain students have procrastinated too long to allow themselves time for retakes, every grade can count, and they can progress to the end of what they have agreed to complete.

Retroactive testing is an experimental aspect of the new testing system. Within a given unit, students may retake quizzes in any order, provided they have not attempted the unit test. This aspect is being monitored to determine if, in fact, students can consistently improve their grade during the quarter by being allowed to proceed and return to problem areas. If it does appear to be helpful, this option will be augmented to include some form of retroactive testing at the unit test level. Plans for that, however, will have to be thought out care-
fully, lest chaos result. 8

Finally, quizzes are now being graded and have become a part of the overall unit grade. It was felt that a pass/not pass system did not provide the incentive for good performance on the quizzes that a grading system does. Likewise, the practice tests are no longer administered in the individualized center, but have become a part of the individualized workbook, which contains model tests with answer keys. This change alleviates some of the pressures of the testing process late in the quarter.

It is far too early to give definitive results; nevertheless, the new program has been in effect for two quarters, and some 120 students have now been involved in it. The preliminary figures are encouraging. It should be noted that changes in the materials and other aspects of the program have contributed to some of the improvement; largely, however, the immediate improvement in the learning process and in the administration of a sometimes unwieldy program has come as a result of the new testing system.

1. The quarterly rate of progress has been raised from an average of 2.5 hours of credit per student to 4.1 hours per student (the classroom is a 5 hour course).
2. The number of incompletes and failures has dropped from 19% to 5%.
3. Grades have not changed dramatically; but a small number of students have received grades below the B- level of the past and have continued to progress through the second quarter of the program. The inflated B- has disappeared, since it is no longer a necessary grade for passing a quiz, test, or the course.
4. The retroactive testing option appears to be benefitting almost 75% of the students who attempt to use it.
5. Student attitudes toward the program in general have changed in positive ways. They are working more throughout the quarter, and they are taking
tests more frequently. Even those who procrastinate have been able to complete their work, by and large, by the end of the quarter.

Most importantly, the new testing system recognizes an important factor in a language program in a university which maintains a four quarter language requirement. Although high grades are certainly desirable for all students, there has to be a distinction between true mastery of the elementary and intermediate phases of a language and passing knowledge. In a sense, it has restored some respectability to the C grade and put the B and A grades back into the perspectives in which they properly belong: above average work and superior work. Likewise, the students for whom the individualized option was intended, namely, slower students, intensive students, and students with difficult schedules, can now be kept in the individualized program as long as it is the better, or perhaps only way by which they can complete the language requirement. Good performance is encouraged, but is not used as the only tool for determining the track in which the student is better placed.

Notes:

1. The individualized program was divided into 15 units, called modules, each worth one hour of credit. This was the equivalent of the three basic elementary Latin courses, worth five hours each. Within a module, each lesson was to be tested by means of a quiz. From 2 to 5 lessons were included in each module.

2. The remaining 8% had had grade problems not as a result of the testing system, but as a result of other factors, such as non-attendance.

3. Initially, the first two quarters (10 hours) were devoted to the basic skills, while the third quarter was primarily a reading course. The third quarter has now been changed to include a continuation of the basic skills, with a much more limited amount of reading to be done.
4. A total of 25 students were used in this experiment, during the 1977-1978 academic year.

5. This did, in fact, happen with frequency on the eighth, ninth, and tenth units, which dealt with the more complicated aspects of Latin grammar and syntax.

6. Findings indicate that there are two types of "weak" students in the individualized track: those who are weak because of lack of motivation and self-discipline, and those who are weak because of genuine learning problems in Latin. The former group is, as a rule, better off in the structured environment of the classroom, while the latter group is much better off remaining in the individualized track. Under the new testing system, there is the ability to analyze the problem or combination of problems for these students and to evaluate which option is actually better for each person. Under the old system, they would all be handed over to the classroom.

7. Change of credit contracts are signed in the seventh week of the quarter, at which time the student may lower the number of hours, raise the number of hours, or maintain the original number of hours of credit. What the students sign up for in the seventh week is what they are held responsible for at the end of the quarter. Students are given a great deal of advice during the period of contracts, to avoid over-estimates of possible achievement in the last three weeks of the quarter.

8. Some time requirements, for example, may be enforced if and when retroactive testing is expanded.

5. To date, approximately 32% of the students appear to be trying this option.

Douglas N. Lacey
The Ohio State University
2 November 1979