

# Evaluation of a Self-Paced Bibliographic Instruction Course

Maria R. Sugrañes and James A. Neal

*Procedures used to evaluate the instructional effectiveness of the materials and methods implemented in a self-paced credit course in library skills are discussed in this paper. The assessment strategies used include: multiple-choice assignments, an end-of-course test, and an attitudinal survey. The assessment was aided by computerized evaluation procedures available from California State University at Long Beach. Results seem to indicate that the instructional treatment contributed to (1) successful completion of the assignments after consultation with an instructional librarian, (2) a rise in test scores as measured by the administration of a pre- and posttest instrument, and (3) manifestation of positive student attitudes about the course.*



rogram evaluation is a key component of instructional design. Formative evaluation is the process of collecting and using data that will enable managers to make decisions for improvement of an educational program.<sup>1</sup> Evaluation is often the most neglected aspect of an instructional design because extensive planning is required and because evaluation techniques are perceived as complex, time-consuming, and generally difficult to implement. Sometimes evaluation is seen as a cumbersome exercise that culminates not in useful data but in voluminous reports few have time to read. There is also an inherent risk in evaluation since the results obtained may fail to prove conclusively the effectiveness of a program or study.

At the University Library, California State University, Long Beach, we were presented with the opportunity of designing a new library instruction course. At an early stage in program development, it was decided to include evaluation as a major component of the instructional plan. Our objectives were to continuously eval-

uate the instructional materials used, to assess patterns of students' library skills growth, and to quantify student attitudes about the library and the course. The ultimate aim of the process was to revise and update our materials, methods, and resources in order to promote program effectiveness. The following are the results of a study reviewing our first-year efforts in implementing this new course.

## DESCRIPTION OF COURSE

In the fall semester of 1981, the University Library began implementing the library skills component of University 100, a one unit course, entitled "The University in Your Future." University 100 was designed as a graduation requirement for all freshmen and transfer sophomore students entering the university in the fall of 1981 and in subsequent semesters. The course consists of three components: history and mission of the university, career planning, and library skills. The library component is designed as a self-paced course; the other components follow the traditional classroom lecture methodol-

---

Maria R. Sugrañes is assistant librarian and James A. Neal is user services manager, both of California State University-Long Beach.

ogy. Enrollment in the course numbered approximately twelve hundred students during the fall 1981 semester and fifteen hundred students in the spring 1982 semester.

To fulfill the requirements of the library component, students must read a library instruction workbook, complete four multiple-choice assignments that are designed to assess student competency in the instructional objectives of the course, and pass a thirty-question, multiple-choice test. Students are assigned to one of three five-week periods during the semester in which to complete their work. There are final deadlines for turning in all assignments and for testing, but there are no intermediate deadlines during the five-week period.

The workbook, which is available at the university bookstore, consists of ten chapters that include: (1) a tour of the library and information on locating materials in the library, (2) a review of basic reference sources, such as encyclopedias and the card catalog, (3) periodical and newspaper indexes, and (4) biographical and book review sources and an introduction to the development of a search strategy.

At the end of each chapter of the workbook, the student is instructed to complete a set of multiple-choice questions. The questions are printed on four optically scannable Scantron answer sheets that are custom designed to include the questions on the left side of the page. Each Scantron assignment contains the questions from as many as two or three workbook chapters. There are twenty alternate forms of each of the four assignments. This is to prevent many students from having to access the same reference sources used in the assignments and to discourage collaboration. Students complete an assignment and turn it in to be corrected at the Center for Bibliographic Instruction. There, two clerical assistants machine score and record the students' assignments within forty-eight hours. Students who miss more than a predetermined number of questions must correct their mistakes to receive credit for the assignment. Reference librarians are scheduled in the center forty-one hours per week to provide one-to-one

consultation and explain problem areas students missed on their first reading of the workbook.

After students have successfully completed the four assignments, they may sign up to take the end-of-course test, which is given at regularly scheduled times every day the center is open. To pass the test, which is available in two forms, students must answer correctly twenty of the thirty questions on the test. Students who fail the test attend a review session given by a librarian and later retake a different form of the test. Students who fail again must receive further remediation and are given a search strategy assignment that consists of applications of the skills on which they were previously tested. Thus, every student attempting the library component will, given time and remediation, pass the course.

### EVALUATION STRATEGIES

In order to validate the instructional materials used in the program and to begin to evaluate overall program effectiveness, three assessment strategies were devised:

1. Tabulation of the number of students receiving credit on the assignments on first attempt.
2. Assessment and monitoring of student performance on a criterion-referenced, end-of-course test.
3. Administration of a survey designed to measure students' attitudes toward the course.

#### *Assessment of Completion Rate of Assignments*

Each of the four research-skills assignments covering two or three chapters of the workbook contain up to twenty multiple-choice questions. Figure 1 contains a sample of typical questions used. Students generally receive credit for an assignment if they miss less than four or five questions. Some questions, however, test basic competencies that must be mastered before more advanced skills can be learned. Such questions, as one requiring students to identify subject tracings on a catalog card, are weighted, and students missing even one weighted "key" question are not given credit on the assign-

Use *Biography Index* to find a periodical article about ROGER LABRUSSE published between September 1952 and August 1955. What is the name of the periodical in which you can find an article about this person? Select the first if more than one is listed.

- a. *United States News and World Reports*
- b. *French Spy*
- c. *United Nations Bulletin*
- d. *People*

To locate information on the issues in a current State Assembly election in Los Angeles County you would consult

- a. periodical indexes
- b. biographical indexes
- c. newspaper indexes
- d. book review indexes

Look up REALISM in *Art Index* for November 1975–October 1976. What is the complete citation for the first article listed under this topic?

- a. Hillingford saga, F. G. Roe, *bibl f il Connoisseur* 190:50-5 S '75
- b. Photo Realism: post-modernist illusionism. L. Chase. *bibl f il por Art Int* 20:14-27 Mr-Ap '76
- c. Barkley Hendricks and his figurative drama. D. Mangan. *il (pt. col) Am Artist* 40:34-9+ J1 '76
- d. WPA & social realism. A. Werner. *il (inc. cover) Art & Artists* 10:24-31 O '75

If you have not found materials in the card catalog by using the search terms you have identified you should

- a. browse through the reference collection for additional sources
- b. check the *Library of Congress Subject Headings* for additional subject headings
- c. go back to the encyclopedia for ideas
- d. change your topic

#### FIGURE 1

#### Sample Research Skills Assignments

ment. Some students are initially resentful of the grading system; however, they soon understand the rationale of this rather strict grading policy. When queried, they agree that this policy and the resulting consultation with a librarian better prepare them for the rest of the assignments and the final test.

During the fall 1981 semester, 50 percent of the students received "no credit" on the first assignment (library tour, call numbers, locating materials in the library), 43 percent received "no credit" on the second assignment (encyclopedias, card catalog), 51 percent failed the third assignment (periodical and newspaper indexes), and 85 percent did not receive credit on the fourth assignment (biographical and book review sources, search strategy). Although the percentages differ a bit, the pattern of success and failure was repeated by the spring 1982 students (see figure 2).

In analyzing student performance, with the exception of the fourth assignment, it is important to note that more than 50 percent of the students could not meet the cri-

terion without the assistance of a librarian/instructor. The high failure rate could be attributed in part to the rigorous grading system. However, their apparent lack of mastery of the material should be a warning against total reliance on self-paced materials as a primary instructional strategy. A great deal of care should be taken so that, regardless of the possible effectiveness of the instructional materials, self-paced instruction will offer more than "correspondence course" methodology.

Librarian-student consultations yielded valuable information for revisions of the assignments and the workbook. A log-book was kept where observations and comments could be noted; review sessions and discussions among librarians often took place, particularly during the first semester of implementation.

Librarian and student feedback as well as assignment results showed that the fourth assignment covering biographical and book review sources and search strategy was very troublesome. The problem was isolated to several questions on search strategy. The high failure rate on

Assignments	Percent of students receiving No Credit	
	FALL 1981	SPRING 1982
1	50	44
2	43	34
3	51	45
4	85	79

Assignment 1: Tour of library and location information.

Assignment 2: Encyclopedias and card catalog.

Assignment 3: Periodical and newspaper indexes.

Assignment 4: Biographical and book review sources; search strategy.

**FIGURE 2**

Assessment of Student Success/Failure on First Attempt at Assignments

this assignment prompted the restructuring of the search strategy chapter in the next edition of the workbook and the revision of all the search strategy assignment questions.

#### *End-of-Course Test*

The major measure of student performance is a thirty-question, multiple-choice-test instrument. It is a criterion-referenced test containing questions designed to measure the specific instructional objectives of the course. The test items were created by a team of three librarians who are very active in the library's bibliographic instruction program. Two equivalent forms of the test were designed. Many items contain identical base questions that only differ by the inclusion of varying examples:

To locate critical evaluations of (book title) written by (author), you should consult a/an

- encyclopedia
- book review index
- periodical
- biographical index

This technique helps to prevent collaboration and at the same time helps to maintain equivalence of test forms.

The test items range from items requiring the use of basic skills to identify parts of a periodical index citation, to more complex items that require both recall and analysis. For example:

The French people elected François Mitterrand as their new president in May 1981; this resulted in various governmental changes. Where would you expect to find the most information on this topic?

- encyclopedias
- periodicals
- books
- almanacs

Functionally identical items appear in the same position on both test forms. Sets of three to ten items compose subtests of the test. These seven subtests—periodical indexes, function of reference sources, call numbers, card catalog, encyclopedias, evaluating sources, and search strategy—closely parallel the workbook chapters and the competency areas addressed in the course.

Scoring of the test results and evaluation of the instrument is facilitated by services provided by the Test Office and the Data Processing Department of the university. Once a week, our office submits a



batch of National Computer System (NCS) answer sheets to the Test Office. The answer sheets are scanned and the data are transmitted via tape to the university computer where a statistical program processes the information. The result is a complete testing report that includes:

1. Alphabetical listing by students' last names giving their scores.
2. List of wrong answers given by each student.
3. Scores listed by students' social security numbers.
4. Alphabetical list by student name of subtest scores and number of questions missed in specific subtests of the test: card catalog, search strategy, etc.
5. Item analysis of each question, including difficulty index and point biserial coefficient.
6. Histogram of total scores.
7. The mean, standard deviation, standard error of measurement, and the Kuder-Richardson (KR-20) for the test.

The detailed information obtained from the testing reports, allowed us to identify troublesome questions, assist students in remediation sessions by checking the items they missed and the subtests they were weak in, maintain accurate records of the students' course completion status, post student scores listed by social security number, and generally monitor students' performance.

One of the most frequently used reports was the printout describing each student's performance in the different subtests of the test. The subtest report allowed the librarian reviewing the test with students who did not initially pass it to diagnose more effectively the source of the student's difficulty and to prescribe ap-

propriate remedial instruction. In addition, the student performance data on the different subtests of the test enabled us to focus on possible areas of revision both in the test itself and in the assignments and workbook chapters. For example, the majority of the students did best on the periodical index subtest, which contains questions similar to ones previously asked in the assignments. Since 50 percent of the students had received "no credit" on the periodical index assignment, we could conclude that the librarian-student consultations had a positive effect. On the other hand, the search strategy subtest was the area where students consistently exhibited the most problems. Given that more than 80 percent of the students failed the search strategy assignment, the poor results were less than a surprise. However, librarians spent many consultation hours explaining the search strategy chapter, apparently with little success. It is this type of evaluative information that allowed us to make decisions for revision of our instructional materials. Although revisions were needed for the periodical indexes materials, a complete restructuring was necessary for the search strategy chapter.

Item analysis information consists of student-performance data on each of the test questions. Figure 3 illustrates a portion of a sample item analysis report. For each question, the report shows the percentage of students who answered the question correctly and scored in the bottom 27 percent or top 27 percent of the sample, and the discrimination index, which is the difference between the two percentages. The point biserial correlation, a statistic that measures the relationship of the question to the total score, is also computed.

ITEM NUMBER	LOW 27% % RIGHT	HI 27% % RIGHT	DISCRIM INDEX	POINT BISERIAL
6	55	99	44	.42
7	56	94	38	.37
8	15	23	8	.05

FIGURE 3

Excerpt from Sample Item Analysis Report

The item analysis reports are useful because troublesome questions can be isolated by reviewing the key statistics. One question, which we later discarded, required the identification of a type of catalog card, a sample of which was illustrated in the test. The question had a .05 point biserial and a discrimination index of 8. In reviewing the test with students, the source of the problem became clear. The illustration was a title card that contained a one-word title heading. Students could not match the one-word heading with the title and subtitle of the book as it appeared in the body of the card.

Although close attention should be paid to the discrimination index and point biserial statistics, their significance may be questioned when analyzing criterion-referenced testing instruments. In reviewing the item analysis reports, we noticed that there were quite a few questions with low discrimination indexes. A question that required students to identify the correct volume number of a journal article in a sample *Readers' Guide* entry had a discrimination index of 4: 96 percent of the bottom 27 percent of the students answered it correctly, and 100 percent of the top 27 percent of the students also answered it correctly. In other words, the question did not discriminate between the top and bottom students since their responses did not vary. What exactly then is the significance of this measurement? Either the question is too easy since nearly all the students answered it correctly, or practically every student understood the concept and met the criterion specified.

A criterion-referenced test by definition attempts to measure students' mastery of specified objectives.<sup>2</sup> The better the instructional treatment, the larger the number of students attaining mastery. According to W. James Popham, an authority on criterion-referenced tests, "The result of increased mastery, of course, is decreased response variance,"<sup>3</sup> thus, the low discrimination index.

In our program, we are committed to helping students master at least 60 percent of the instructional competencies specified. As such, the low discrimination indexes are valuable statistics, but not nec-

essarily for their usually intended purpose. However, this example points out a real concern in criterion-referenced test construction; namely, to what degree should questions be designed to measure mastery as opposed to discriminating between thoughtful and less-thoughtful students? The test to be used in the 1982-83 academic year will contain many of the same "nondiscriminating" questions since our goal is student mastery of specified competencies, and creating variance is not the chief concern of criterion-referenced testing<sup>4</sup> and competency-based instruction. However, a number of instructional objectives require analysis and synthesis. The questions that test students' mastery of those objectives have been sharpened to demand a higher level application of critical thinking skills. Needless to say, students' performance on these questions will be closely monitored and reviewed.

The test results have been, on the whole, quite gratifying. The criterion for passing was set at a minimum of twenty out of thirty possible questions. Ninety percent of the students passed on their first attempt in the fall 1981 and spring 1982 semesters. Statistical measures for both test administrations are outlined in figure 4.

The administration of the test during the two semesters yielded similar results. The mean or average scores show an inconsequential but significant difference; the standard deviations, measuring the dispersion of the scores, are comparable. The values of the standard error of measurement (SEOM) are similar. The SEOM indicates the range of scores that will include a student's true score with 68 percent certainty. For instance, given a mean of 24 and SEOM of 2, there is a 68 percent chance that a student's true score lies somewhere between 22 and 26. The KR-20 statistic, from which the SEOM is derived, measures the internal reliability of the test. It answers the question of how accurately the test measures whatever it is supposed to measure, how precise the scores are, and if the scores could be reproduced upon subsequent measurements.<sup>5</sup> A rule of thumb for measuring the reliability of

Statistical Measurements	Fall 1981		Spring 1982	
	A Form	B Form	A Form	B Form
Mean	24.1	23.5	23.5	24.1
Standard Deviation	3.0	2.9	3.3	3.3
KR-20	.56	.57	.63	.64
Standard Error of Measurement	2.0	1.9	2.0	2.0
Number of Cases	645	629	732	733

FIGURE 4  
Statistical Measurements for End-of-Course Test

teacher-prepared, as opposed to commercially prepared, tests is a coefficient of .70. Marshall and Hales state that "generally speaking, teacher made tests are infamous for their lack of reliability. Many classroom tests have coefficients of reliability approaching zero. Probably most fall in the range of reliability above .60."<sup>6</sup> Since our KR-20 values range in the upper .50s to low .60s, we are not far from the optimum level, but improvement is needed. It must be noted, however, that KR-20 is a measure of internal consistency weighing the interrelationships of questions. As illustrated in the previous discussion on discriminating and nondiscriminating questions, the questions for a criterion-referenced test are not designed for discrimination, thus the reliability of the measurement will probably not improve dramatically with revision.

Validity is the most important quality of any testing instrument. We have to know whether a test measures what we want measured: the precise competencies, skills, and behaviors addressed in the course. No one intentionally designs a testing instrument that does not attempt to measure what the students are to learn. However, errors and biases often intrude in the test-writing process and compromise the validity of the instrument. One unfortunate illustration of this was the inclusion of a test question that required examination of an excerpt from *Readers' Guide*. Although the copy was clear, some

students had difficulty reading the small print. To compound the problem, students also missed the question because when asked to identify the date of a particular periodical they, to our surprise, confused the abbreviations for January, June, and July.

Validity is difficult to establish. Good strategies for achieving a greater degree of content validity are constructing well-formed test specifications and field-testing questions with students and colleagues. Another strategy is to devise a matrix that cross-indexes instructional objectives, assignment questions, and test questions that assess student competency on that particular objective.<sup>7</sup> Figure 5 illustrates a sample of the matrix used to construct the test given in the 1982-83 academic year.

Since more than 90 percent of the students passed the test, it could be concluded that the test items are appropriately selected to test the skills and competencies addressed in the course. However, such a high success rate might also indicate that the instrument was so easy that students could pass it without the benefit of the instructional treatment.

In order to address the issue of alternative explanations for the increase in test scores, an evaluation design was created that would:

1. Compare pre- and posttest scores to assess gains in scores after instruction during a particular semester.

Chapter	Objectives	No. of Assignment Questions	No. of Test Questions
1 Library Tour	The student will 1. Take a self-guided tour of the university library and locate major resources (e.g., card catalog, serials record, general book collection, periodicals, government publications), major services (e.g., reference, interlibrary loan, circulation), and major equipment (e.g., photocopiers, microfilm readers) in the university library.	9	0
2 Locating Materials	1. Identify the alphabetic filing arrangement used in different reference sources. 2. Use a call number to locate a book on the shelves. 3. Identify the proper procedure for checking out books, audiovisual items, and reserve materials.	1 7 1	1 2 0
3 Basic Search Approaches	1. Identify the most effective method of searching for information, given a particular search problem. 2. Identify the type of reference source that will most effectively meet specific information needs, given a particular search problem.	2 3	1 3
4 Encyclopedias	1. Use a subject encyclopedia to locate an article and a bibliography on a given topic. 2. Read a short article in an encyclopedia to select key words to use as potential search terms.	1 1	0 1
5 Card Catalog	1. Locate the catalog card records of cataloged items in the university library by searching for the author or title or subject of the items. 2. Identify the notes and subject tracings on the catalog card records of a particular book or audiovisual item. 3. Identify and explain the usefulness of notes and subject tracings found on a catalog card record. 4. Identify the subject headings used in the card catalog for a particular topic by using <i>Library of Congress Subject Headings</i> .	4 2 2 2	1 2 2 1
6 Periodicals and Periodical Indexes	1. Locate an article in a periodical index and identify parts of the citation, given a particular subject. 2. Locate a volume of a specific periodical in the university library.	8 3	7 3
7 Newspaper Indexes	1. Identify the page, column, and title of a newspaper article using the <i>New York Times Index</i> and the <i>Newspaper Index: Los Angeles Times</i> , given a subject and date.	4	0
8 Evaluating Sources	1. Determine the potential usefulness of a book by applying indicators of relevance found on the text of a catalog card record. 2. Identify appropriate techniques to use in evaluating a particular book or author. 3. Use specialized biographical indexes to find articles and factual information about particular authors or individuals. 4. Use book reviewing indexes to find a book review in a periodical, given the author and title of the book.	2 3 2 2	2 1 0 0
9 Developing a Search Strategy	1. Identify the steps of a basic search strategy. 2. Select an example of a narrow topic for research.	5 1	2 1

FIGURE 5  
Evaluation Specifications Matrix 1982-83



2. Compare posttest scores of two samples, one pretested and the other posttested only, in order to assess the effect of pretesting on scores.
3. Compare the pre-post gains in scores of selected student groupings, such as class and rank in pretest scores.

In the fall 1981 semester, all students enrolled in University 100 were pretested with an instrument equivalent to the end-of-course test administered after instruction. The pretest contained thirty questions and was also administered in two forms. The pretest was not administered in the spring 1982 semester. Using the same passing score of twenty, only 30 percent of the students theoretically "passed" the pretest, whereas, more than 90 percent of the students passed the posttest in both the fall and spring semesters. Figure 6 lists pre- and posttest gain scores achieved by selected student groups within the total sample. As expected, there was a marked increase of scores in the posttest as compared to the pretest. There also seems to be very little difference between posttest scores taken by different student groups in the two semesters.

In order to evaluate statistically the significance of the increase of the student scores from pretest to posttest, a *t*-test was run using 1,182 pairs of scores (see figure 7). The evaluation was performed by creating a computer file containing the paired scores and subsequently using the Statistical Package for the Social Sciences (SPSS) software package.<sup>8</sup> After running

the program, the results indicate that the mean increase for the total sample was 6.4856 points; the *t*-value was 51.70, which given 1,181 degrees of freedom, indicates that the difference was significant at the .000 level of confidence. These results indicate that there is less than a 1 percent probability that the increase in scores was due to chance.

To determine the extent to which the gain in scores could be attributed to the practice provided by the pretest, a *t*-test was run comparing the differences in mean scores between pretested students taking the posttest in the fall 1981 and students taking only the posttest in the spring 1982. The results show that there is little difference in the mean scores: 23.8184 for spring 1982 and 23.7461 for fall 1981. The *t*-value is .59 with 2,596 degrees of freedom, which gives .556 probability that the difference between scores could be due to chance. In other words, the fact that one group of students was pretested did not measurably increase their posttest scores.

Another *t*-test was run comparing the difference of scores of students scoring in the bottom quartile of the pretest (16 cases) and the top quartile (175) cases. The mean increase for students scoring in the bottom quartile of the pretest was 16.5 points. The mean increase of students scoring in the top quartile was 1.8686. Using a pooled variance estimate, the *t*-value was 19.54 with 189 degrees of freedom and .000 probability that the increase in scores was due to chance. One interesting

Student Groupings	Mean Increase in Pre- and Post-test Scores
Upper Quartile	1.86
Bottom Quartile	16.50
Freshmen	6.59
Students other than Freshmen	5.33
All Students	6.49

FIGURE 6  
Pre- and Posttest Gain Scores of Selected Student Groups

	FALL 1981		FALL 1981	SPRING 1982
	PRE-TEST	POST-TEST	POST-TEST	POST-TEST
Mean Score	17.23	23.72	23.72	23.81
SD	4.05	2.90	2.90	3.30
N	1182	1182	1182	1377
Difference of Means	6.49		.09	
t	51.70		.59	
df	1,181		2596	
p	.000 significant		.556 not significant	

SD = standard deviation

N = number of cases

t = t-test value

df = degrees of freedom

p = probability

FIGURE 7

Comparison of Pre-Posttest Scores, Fall 1981, and Post-Only Scores, Spring 1982

note, this particular sampling is a good illustration of the "regression effect": in "virtually all test-retest situations, the bottom group . . . will on the average show some improvement on the second test—and the top group will on the average fall back."<sup>9</sup> This normal rise and fall is not caused by the course, it is merely due to the spread of scores in any given grouping or population.<sup>10</sup>

Lastly, an analysis was made to determine the mean increase of scores achieved by freshmen as opposed to sophomore and upper-division students enrolled in the course. The purpose of this analysis was to examine the possibility that non-course-related experience in the academic milieu could account for the increase in student scores. It is reasonable to assume that freshmen and nonfreshmen differ at minimum by one semester in overall academic experience and that both groups are gaining an additional semester's experi-

ence while taking the class. If general academic experience is a contributory factor, nonfreshmen students would be expected to achieve higher pretest scores and continue to demonstrate a higher learning rate (achieving higher gain scores) after the instructional treatment than the freshmen students in the sample.

The results of this analysis (see figure 8) suggest that general academic experience does not outstandingly contribute to gain scores:

1. Freshmen students' mean pretest score is 1.92 points lower than the mean score attained by the non-freshmen students. This may be due to the effect of academic experience such as test-taking and exposure to libraries.
2. Freshmen students' mean gain scores are higher than the more advanced students' scores. The initial "knowledge gap" of 1.92 points de-

		FALL 1981 PRE-TEST	FALL 81 POST-TEST	FALL 81 PRE-POST GAINS	SPRING 82 POST-ONLY
FRESHMEN STUDENTS	Mean	17.05	23.64	6.59	23.47
	SD	4.05	2.87		3.47
	N	1070	1070		618
NON- FRESHMEN STUDENTS	Mean	18.97	24.45	5.48	24.26
	SD	3.54	3.16		3.08
	N	112	112		759
DIFFERENCE IN MEAN SCORES		1.92	.81		.99
ALL STUDENTS	Mean	17.23	23.72	6.49	23.81
	SD	4.05	2.90		3.30
	N	1182	1182		1377

SD = standard deviation, N = number of cases

FIGURE 8  
Comparison of Mean Scores of Freshmen and Nonfreshmen Students

creased measurably to .81 points. There is again only a .99 point difference in mean scores of spring 1982 students who took the posttest only.

- There is a negligible difference in test sensitization that occurred in the sample of freshmen as opposed to nonfreshmen students. Pretested nonfreshmen received mean posttest scores that are .19 points higher than nonfreshmen taking the posttest only in the spring 1982 semester. Similarly, pretested freshmen received scores that are only .17 points higher than nonpretested freshmen.
- The percentage of freshmen students whose pretest scores were high enough to meet the posttest pass criterion was only 26 percent as contrasted to 50 percent for the pretested nonfreshmen. After instruction,

however, both groups had approximately the same pass rate with 94 percent of the freshmen passing the test on their first attempt compared to 97 percent of the nonfreshmen. Under the posttest-only condition, both groups performed at the same level with 91 percent of both groups passing.

Considering the initial difference in pre-post scores, the regression effect, and the possible one semester experience factor in the nonfreshmen students, the increase in scores is too high to be attributed to non-course-related academic experience alone and could therefore be attributed to the instructional treatment.

It needs to be emphasized that all of these statistical measurements were made possible by the computer programs and expertise available through the Test Office

and Data Processing Department of the university. If computerized statistical services are available they should be utilized. However, it is also important to note that the measurements can be done without the aid of a computer, albeit this is highly time-consuming and labor intensive.<sup>11</sup>

Although the statistical measurements ostensibly show an increase in scores after instructional treatment, it still cannot be concluded that the treatment was the only contributing cause for the rise in scores. There are many non-instructionally related background variables operating on the 1,182 students sampled, such as intelligence, study habits, practical experience in the library related to other classes during the instruction period, attitude toward the course, etc. Any one variable or combination of variables may account in part for gains in scores. It is extremely difficult, if not impossible, to conduct a study where change in behavior can be attributed to the instructional variable.<sup>12</sup> It might have been possible, however, to request the 1,182 students in the sample to provide us with measurement, such as SAT scores, or GPA in high school or in their current semester of enrollment. We then could have extracted the effect of that variable and obtained a more valuable conclusion. Unfortunately, although there was a commitment to evaluation in the planning stages of the course, at the start of the program's implementation, priority was given to preparing the workbook and assignments, hiring and training staff, establishing organizational procedures, not to creating a rigorous evaluation design. It is reasonable to assume, however, that the gain in scores is due to the effect of the instructional treatment.

#### *Student Attitude Survey*

To obtain a measure of student attitudes toward the course, a brief, ten-question survey was administered to all students who completed the library component of University 100. A Likert scale was not used because such an instrument was used to evaluate the entire course, and because we wanted unequivocal responses from our students.

The survey questions and the percent-

age of positive and negative responses are presented in figure 9. The results of the survey reveal an overall positive attitude toward the course and are consistent from semester to semester. One unpleasant finding was that a large percentage of students expressed no further need for library instruction. Since the University Library offers an extensive noncredit program of bibliographic lectures on special subject areas, which was attended by nearly six thousand students in the 1981-82 academic year, there is a definite need to review these students' observations. As one librarian suggested, perhaps the question implied to students that they might be expressing a desire for further *required* instruction. The question has been rephrased for the survey to be used next year, and students' responses will be reviewed with care.

#### CONCLUSIONS

The evaluation strategies used in this study have yielded important information that has been used to revise the instructional materials used in the library component of University 100. Additionally, the data gathered suggest that the program has had a positive effect on students in terms of their knowledge of library skills and in terms of general attitude toward the library:

1. Students successfully completed assignments assisted by librarians as necessary.
2. The high proportion of students passing the end-of-course test indicates that most students mastered the material to criterion. Evaluation of pre- and posttest scores shows a marked gain in scores.
3. Students' attitudes toward the course were significantly positive even though University 100 is a required course.

The evaluation techniques used for this research project are fairly standard and relatively simple to implement. They are not labor intensive due to the availability of computer programs to speed calculations and process data. Most importantly, our efforts at evaluation have yielded not only interesting research data but also



Percentage of Student Responses

	Fall 1981		Spring 1982	
	Yes	No	Yes	No
1. Was the library component difficult?	20	80	18	82
2. Will the library component be useful in your college career?	93	7	88	12
3. Would you be interested in further library instruction designed for your major?	14	86	38	62
4. Was consultation with librarians at the Center for Bibliographic Instruction (CBI) useful?	87	13	82	18
5. Were services efficient at the Center for Bibliographic Instruction (CBI)?	90	10	93	7
6. Did you receive adequate help on the assignments if you needed it?	89	11	93	7
7. Were the sources needed to complete assignments available to you when you wanted them?	88	12	89	11
8. Was the self-paced method a good feature of this course?	91	9	90	10
9. Does the <u>Library Instruction Workbook</u> have clear directions and explanations?	81	19	81	19
10. Did you receive library instruction in high school?	50	50	47	53

FIGURE 9  
Results of Student Attitude Survey

practical information that has been used to revise and refine instructional materials and strategies. It is our hope to expand the

evaluation process and thus promote the improvement of our program.

#### REFERENCES

1. Peter J. Taylor, "User Education and the Role of Evaluation," *Unesco Bulletin for Libraries* 32, no.4:254 (July-Aug. 1978).
2. William P. Gorth, Robert P. O'Reilly, and Paul D. Pinsky, *Comprehensive Achievement Monitoring: A Criterion-Referenced Evaluation System* (Englewood Cliffs, N.J.: Educational Technology Publications, 1975), p.50.
3. W. James Popham, *Criterion-Referenced Measurement* (Englewood Cliffs, N.J.: Prentice-Hall, 1978), p.106.
4. Robert L. Thorndike and Elizabeth Hagen, *Measurements and Evaluation in Psychology and Education* (4th ed.; New York: Wiley, 1977), p.214-15.
5. *Ibid.*, p.73.

6. Jon Clark Marshall and Loyde Wesley Hales, *Classroom Test Construction* (Reading, Mass.: Addison-Wesley, 1971), p.208.
7. Popham, *Criterion-Referenced Measurement*, p.156.
8. Norman Nie and others, *SPSS: Statistical Package for the Social Sciences* (2d ed.; New York: McGraw Hill, 1975).
9. David Freedman, Robert Pisani, and Robert Purves, *Statistics* (New York: Norton, 1978), p.159.
10. Ibid.
11. For a full discussion of the manual process, see James Rice, Jr., *Teaching Library Use: A Guide for Library Use Instruction* (Westport, Conn.: Greenwood Pr., 1981), p.118-25.
12. Larry Hardesty, Nicholas P. Lovrich, Jr., and James Mannon, "Library-Use Instruction: Assessment of Long Term Effects," *College & Research Libraries* 43:38-46 (Jan. 1982); Richard Hume Werking, letter to the editor, *College & Research Libraries* 43:353-54 (July 1982).