

UNIV 100 Report: Comparison of Pretest/Posttest Results from 2005 and 2008

By Katherine S. Dabbour and Susanna Eng-Ziskin
Reference & Instructional Services Dept., Oviatt Library
Spring 2012

Method

For the purposes of the current study, the authors focused on data from 2005, which was the year prior to the hiring of the FYE librarian, and the latest year that the data was collected (2008) to see if there were differences that could be attributed to changes in pedagogy. The 2008 data was the latest available due to the unexpected demise of the Flashlight Online system previously available to outside institutions from the CTLT at Washington State University, which left CSUN with only partial data and not enough time to switch to SurveyMonkey™ for the fall 2009 cohort. While frustrating, it also coincided with the library assessment coordinator's beginning work on testing information literacy rubrics for assessing student work, which is still in progress.

In the fall of 2005 and fall of 2008, U100 students took the pretest at the beginning of the semester, prior to "Library Week" and the posttest at the end of the semester, after attending all of the library instruction sessions and completing an annotated bibliography or other library research assignment. U100 instructors distributed the pre-and posttest URLs to their students through various means, including personal email, WebCT™ pages, and the official U100 course web site. As an incentive, instructors gave students extra credit to complete the surveys. Students provided their names in the surveys to get credit, which in turn, allowed the pairing of pretests and posttests for the data analysis. However, the librarian did not share individual test results with the instructors and later redacted the students' names to keep the results anonymous.ⁱ

Results and Discussion

Since U100 is not required of first year students at CSUN, one cannot generalize about all such students from this study; therefore, the analysis of the results is limited to students enrolled in the course. In the fall of 2005, there were 3,720 first-time freshmen, out of which 625 (16.8%) were enrolled in Univeristy100. Out of 4,625 first-time freshmen enrolled in the fall of 2008, 852 (18.4%) students took U100.

In 2005, of the 247 paired respondents who took the pretest and posttest, 61.2% or ($n=150$) indicated they had formal library instruction prior to starting CSUN and 17.5% ($n=43$) had instruction at CSUN, before their first U100 library instruction session. Out of the 273 paired respondents who completed the pretest and posttest in 2008, 58.5% or ($n=158$) indicated they had formal library instruction prior to starting CSUN, similar to 2005. However, in 2008, 31.5% ($n=86$) had instruction at CSUN, before their U100 library week, almost twice the rate of students in 2005's cohort (Figure1). While the timing of the administration of the pretest and U100 library instruction remained constant from 2005 to 2008, students in 2008 also took a pre-library instruction online tutorial, created by the FYE librarian, which demonstrated much of the material that used to be covered by the librarian during the first session. Therefore, this most likely explains the difference.

Figure 1. Library Instruction Experience, 2005 and 2008

| | Prior to CSUN | | At CSUN | |
|------|---------------|------|----------|------|
| | <i>N</i> | %* | <i>N</i> | %* |
| 2005 | 150 | 61.2 | 43 | 17.5 |
| 2008 | 158 | 58.5 | 86 | 31.5 |

***Percentages do not add up to 100% due to non-response.**

Students also rated their library and Internet research skills in both the pretest and posttest. Since some might question the usefulness of a self-rating of skills, the authors were simply looking for descriptive data in support of the instruction program. As one would hope, in both years, the number of students who rated their skills as “very good” to “excellent” went up proportionally to those who rated their skills as “fair” to “poor” went down from pretest to posttest. In particular, for both years, the percentage of students rating their skills as “very good” after library instruction went up at the same time as those rating their skills as “fair” went down (Figures 2 and 3).

Figure 2. Self-Rating of Research Skills, Pretest to Posttest, 2005

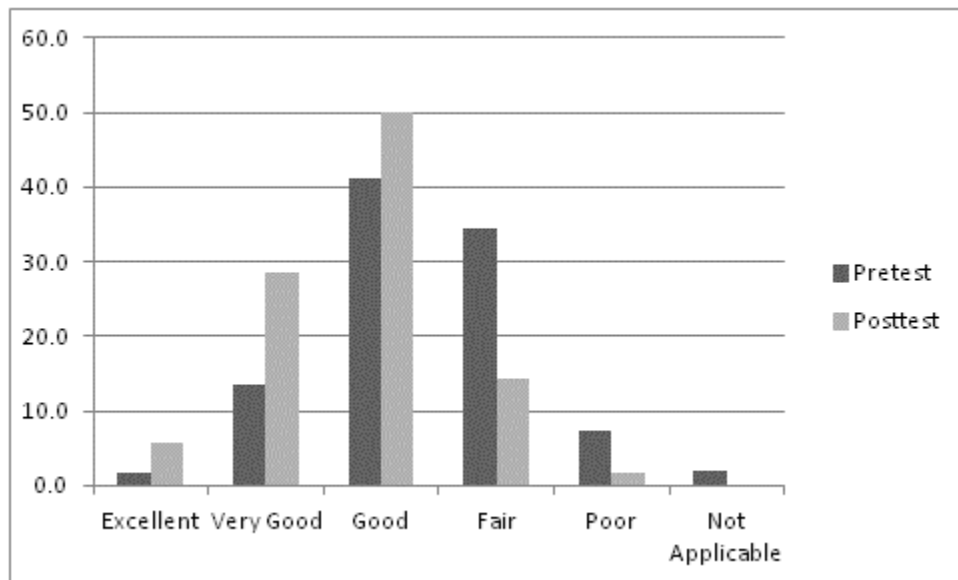
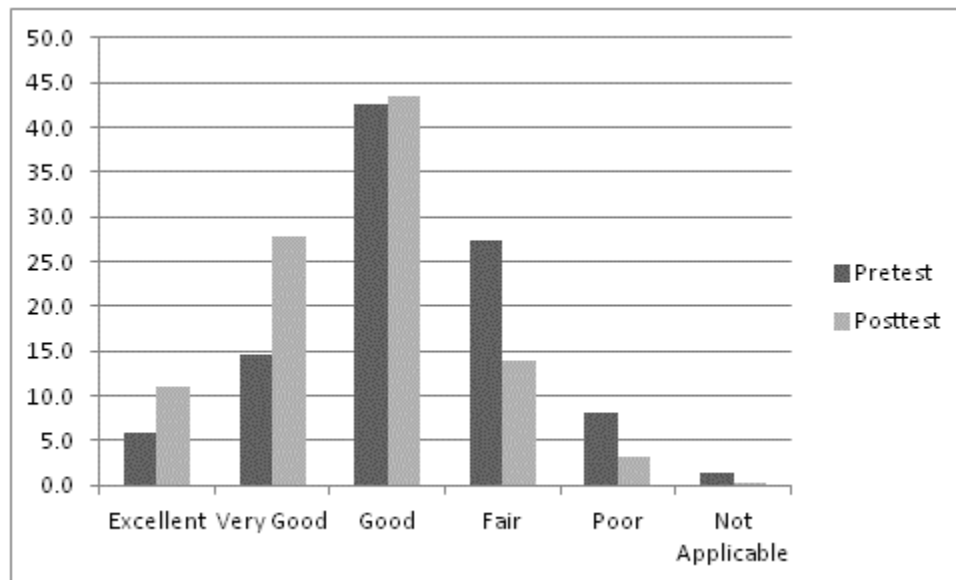


Figure 3. Self-Rating of Research Skills, Pretest to Posttest, 2008



As previously mentioned in the methodology section, while there were differences in some of the test questions in 2005 and 2008, the concepts tested remained the same. As previously stated, therefore, from a purely scientific research standpoint, one cannot draw definitive conclusions about the relative performance of U100 students in 2005 versus 2008, but in practical terms, the data can be useful for assessing the impact of changes in pedagogy.

Figure 4. Question-by-Question Summary of 2005 and 2008 Pretest-Posttest Results

| Question | 2005 Percent Correct | | | 2008 Percent Correct | | |
|--|----------------------|----------|--------|----------------------|----------|--------|
| | Pretest | Posttest | Change | Pretest | Posttest | Change |
| #1: Identify Keywords | 85.8 | 91.5 | 5.7 | 48.0 | 49.5 | 1.5 |
| #2: Recognize Journal as Most Scholarly | 29.1 | 57.1 | 28 | 46.2 | 79.1 | 32.9 |
| #3: Recognize Today's News Source | 60.3 | 50.2 | -0.1 | 71.8 | 61.2 | -10.6 |
| #4: Define Scholarly Journal | 33.2 | 45.3 | 12.1 | 38.1 | 46.5 | 8.4 |
| #5: Recognize Book Chapter Citation* | 20.3 | 23.9 | 3.6 | * | * | * |
| #6: Recognize an Article Citation | 29.6 | 34 | 4.4 | 33.3 | 35.2 | 1.9 |
| #7: Recognize a Book Citation | 49.4 | 59.7 | 10.3 | 52.7 | 50.2 | -2.5 |
| #8: Know Use of Library Catalog | 90.3 | 91.9 | 1.6 | 87.2 | 92.7 | 5.5 |
| #9: Know Use of Periodical Index | 28.7 | 34.8 | 6.1 | 37.4 | 42.1 | 4.7 |
| #10: Identify Boolean "Or" Logic* | 28.3 | 35.2 | 6.9 | * | * | * |
| #11: Identify Boolean "And" Logic | 26.3 | 34.4 | 8.1 | 60.4 | 70.7 | 10.3 |
| #12: Recognize Source Bias | 45.7 | 45.7 | 0 | 57.1 | 63.7 | 6.6 |
| #13: Critically Thinking About Internet | 25.1 | 45.3 | 20.2 | 77.7 | 78.8 | 1.1 |
| #14: Define Plagiarism | 36.8 | 49.8 | 13 | 89.0 | 89.0 | 0.0 |
| #15: Recognize Citation Elements | 61.1 | 70 | 8.9 | 50.5 | 67.4 | 16.9 |
| #16: Understand Appropriate Topic Choice** | ** | ** | ** | 59.0 | 61.9 | 2.9 |

*Not asked in 2008

**Not asked in 2005

As shown in Figure 4, the percent correct on the pretests for both years for Question #1, which was on keyword choice, were drastically different. In 2005, 85.8% were correct in the pretest, which went up to 91.5% in the posttest, a change of 2.7%. However, in 2008, students scored much lower in the pretest (48% correct) and stayed low in the posttest (49.5%), a change of 1.5%. At first glance, it appears that students in 2005 came in knowing much more about keyword searching than students in 2008, and in both cases, the pedagogy had a minimal effect. Upon careful consideration of the characteristics of the students and the pedagogy in this regard, it appears neither were the case; rather, the change in the question text and response choices most likely created confusion. In 2005, Question #1 read:

Choose the best set of keywords that you would use to search for recent research on the possible link between cell phone use and cancer.

Cell phone, use

Cell phone, cancer (correct)

Research, cell phone

Research, cancer

Don't know

In 2008, the question was revised to update it to a more newsworthy topic as follows:

Choose the best set of keywords that you would use to search for recent research on the possible impact of global warming on weather patterns.

Global warming, impact

Research, weather patterns

Global warming, weather (correct)

Weather patterns, impact

Don't know

In analyzing the raw data, the authors found that while 48% answered correctly (Global warming, weather), 43% chose, "Global warming, impact." Upon further analysis of the text of the test question, it seems the most likely cause of this difference was the deletion of the word "patterns" from the phrase "weather patterns" in the list of possible answers. This highlights one of the lessons learned when revising test questions: in an online test environment, it is quite simple to cut corners and make changes without careful item analysis; however, this can lead to sloppy test question construction, thus rendering the data useless.

The results from Question #2, which was about recognizing “journal” as the most scholarly source in the choices presented, reveals that in 2008, students had higher scores in the pretest (46.2% correct in 2008 versus 29.1% correct in 2005). However, in the posttest, scores for both years went up at a similar rate: 28% in 2005 and 32.9% in 2008. Again, wording changes might have contributed to the lower scores. In 2005’s question text, students did not have a topical example that perhaps might have helped further define the concept of scholarly source, whereas in 2008, students were given such an example in the question text. Therefore, comparisons between the two years in regards to this question are not possible.

Respondents’ choices for Question #3, which asked students to choose “the fastest, most accurate source” of today’s news, were “a book,” “today’s newspaper,” “this week’s *Time* magazine,” “CNN.com or another news website,” or “don’t know.” In 2005, 60% correctly answered “CNN.com or another news website” in the pretest, but in the posttest, correct responses went down to 50.2%, a difference of -10.1%, despite identical wording of the question in both the pretest and posttest. Similarly, in 2008, 71.8% of the students were correct in the pretest, but in the posttest, scores fell to 61.2%, a difference of -10.6%. To what can one attribute this negative change in scores after receiving library instruction? Perhaps students were off put by the usual warnings about the credibility of Internet sources compared to more traditional library sources, and therefore, may have reinterpreted the question text regarding “most accurate” to mean “today’s newspaper,” which would be the closest correct response. Furthermore, the assessment coordinator discovered too late that the question text broke one of the cardinal rules of survey design: never address more than one concept in a single question. Ideally, the question should have only focused on the “fastest” resource choice, leaving out the credibility judgment

of “most accurate” source. This is borne out by the literature cited by Manuel related to students’ difficulties learning new materials that do not conform to prior knowledge (Halttunen 2003).

Question #4 attempted to see if students could identify what constituted a scholarly journal so that they not only could recognize journals as more scholarly than other sources (See Question #2 discussion above) but also demonstrate an understanding of the concept of scholarliness.

Among the librarians in the Reference and Instructional Services department at CSUN, there were differences of opinion as to how to teach freshmen this concept. In 2005, some librarians instructed students to simply look for visual clues in the article, such as the presence of a references cited list and the lack of advertising. Other librarians emphasized the peer-review process and the differences between scholarly and journalistic treatment of a topic. In 2008, the peer review process was explained, but not all students were actually required to locate scholarly articles for their assignments. Instructors were encouraged to let their students use popular sources instead. While the question text for both years was changed to reflect changes in pedagogy and philosophy, it is telling that for both years’ respondents, posttest scores were still only 45-46% correct, a dismal performance, and perhaps indicative of how changes to the assessment instrument and/or pedagogy were still needed.

Recognition of citations for various sources was considered important enough to put on both years’ tests (Questions #5, #6, and #7), but the book chapter citation question was dropped after 2005 for being too advanced for freshmen. In any case, as indicated in Figure 4, the respondents’ ability to recognize an article citation was low in the pretest in both years, and remained low in the corresponding posttests. Furthermore, while book citations were more recognizable in both years, as shown in the pretest results, in the posttest results, the 2008 respondents’ number of correct responses went down slightly while the 2005 respondents went up by 10.3%. While

librarians emphasized searching the book catalog and article databases in both years, clearly a different approach was needed.

Students in both years knew what the library catalog was used for (Question #8) but were quite unclear on what was found in a periodical index (Question #9), even after the aforementioned emphasis on periodical databases during the lectures. In 2005, respondents were expected to understand the definition of “periodical” in order to correctly select “newspaper, journal, and magazine articles” as the three items out of a list of six that were searchable in a periodical index. Because of the low performance pretest to posttest in 2005, the question was revamped in 2008. Instead of expecting students to know the three types of periodicals, the question was revised to include a topic and emphasize the source of a periodical article rather than the types: “To search for reviews about the new Apple iPhone, of the following, the BEST source to use is a . . . ,” and the options were, “library catalog, almanac, periodical database, directory, and don’t know.” The correct response, “periodical database,” was chosen by 37.4% in the 2008 pretest and 42.1% in the posttest, a negligible difference. Again, one would expect that the emphasis on periodical databases during instruction would influence posttest performance, but clearly, this was not the case. More than likely, there was confusion over the use of the word “reviews” in the 2008 question text; in retrospect, it would have been better to use the word “articles” in the question text. Therefore, it is difficult to say definitively if the continued low performance was due to flaws in pedagogy or test question wording in both years. The term “periodical database” may have further confused students since the FYE librarian emphasized “databases” and rarely used the jargon term “periodicals,” favoring the more user friendly term, “articles.”

Concomitant to knowing what a periodical database and library catalog are for, one should understand something of Boolean logic. In 2005, the test asked students to identify which search

would retrieve the most results in a database search (Question #10), and the correct response was the “or” statement: “grammar or English.” However, as demonstrated in Figure 4, respondents did poorly in the pretest, and while the percentage of correct scores went up 6.9%, the overall scores were still low. “And” logic had similar results in 2005 (Question #11). In this case, pedagogy did not exactly match what was being tested; because of this, the Boolean “or” question was dropped in the 2008 test due to this mismatch. However, the “and” logic question was left in the 2008 test, and students not only started out much higher in the percentage of correct responses in the pretest (60.4%), but their scores went up 10.3% in the posttest. Most likely, this discrepancy in percentages correct as compared to the 2005 test was another change in the wording of the question. In 2005, students were asked to identify which search statement would retrieve the “least number of results”; in 2008, the question was “which search would retrieve the most specific (fewest)” results. While 2008’s wording was not only awkward but also poor response construction since it combined two answers in one (“most specific” and “fewest”), it also did not require students to understand the mathematical (and seemingly contradictory) concept of the Boolean “and” retrieving less rather than more. Instead, the question was worded as such that students would understand that “and” lead to greater specificity in addition to the “fewest” results.

Questions #12 and #13 attempted to test in an objective, multiple-choice format, the complex concept of critically thinking about resource use. Question #12 remained the same in both 2005 and 2008:

If you had to give an informative speech on the benefits and limitations of electric-powered cars, which of the following resources should provide the most balanced or unbiased information?

A report issued by the American Petroleum Institute

An article from the Electric Car Retailers Association Newsletter

An article in Consumer Reports magazine (correct)

Don't know.

The goal of Question #12 was to make respondents think about competing information resources and potential bias. In the pretest for both years, the percentage of correct responses was low and in the posttest, it either stayed the same (2005) or only went up slightly (2008).

Question #13 attempted to ascertain if respondents understood how to critically evaluate information found on the Internet. While the question text changed from 2005 to 2008, essentially respondents were asked what was most important to consider before citing an Internet resource:

The authority of the person or organization responsible for the information.

The information is up-to-date enough for your research topic.

The information is based on cited facts, or if it is opinion, that is made clear.

Your professor allows you to use the Internet as a source.

Obviously, in this list of possible responses, all would be correct. In 2005, students were instructed to “click all that apply,” and in 2008, they were given an “all of the above” as a possible (and correct) response. As shown in Figure 4, in 2005, the percentage of correct responses went up dramatically, by 20.2%, which is significant, although the overall percentage correct at the posttest was still only 45.3%. On the other hand, 77.7% of 2008’s respondents were correct in the pretest, but it only went up to 78.8%, a negligible difference. One could speculate that the inclusion of an “all of the above” response made the question too easy for students in

2008; whereas the 2005 cohort had to closely read the possible responses as well as follow the instruction to “click all that apply.”

Understanding what constitutes plagiarism (Question #14) is another problematic concept to test with an objective format, which was further confounded by the option of an “all of the above” response included with 2008’s choices and a “click all that apply” instruction in the 2005 question. Again, while the question wording changed over the years (when must you cite the source), the responses stayed constant:

When you copy a whole paragraph

When you summarize or write a paragraph over in your own words

When you quote one sentence

In 2005, the percentage correct went from 36.8% to 49.8%, a difference of 13%. In contrast, 89% were correct in the 2008 pretest, which did not change at all in the posttest. Again, the 2008 respondents had the luxury of an “all of the above” option whereas those in 2005 had to read and select all that applied to get it right. Furthermore, by 2008, the campus had ratcheted up its anti-plagiarism campaign by requiring all faculty to include the campus’ policy on academic dishonesty in their syllabi, and subscribing to Turnitin.com™, which might also account for the higher scores.

Related to recognizing plagiarism, it is essential that students understand citation elements. As was previously discussed, the percent correct response to the questions on book, chapter, and article citation recognition were equally disappointing in 2005 and 2008. However, in both years, respondents did well recognizing which group of record elements from a database search would be needed to create a citation, regardless of style. The percent correct responses to Question #15

in both years were not only high to begin with (61.1% in 2005 and 50.5% in 2008), but increased in the posttests (70% in 2005 and 67.4% in 2008).

Finally, as shown in Figure 4, respondents were to identify from a list of possible topics which would be the most appropriate to research for a particular assignment (Question #16). Due to poor results attributed to the wording, the question was revised more than once over the years, and was left off the 2005 version. In 2008, the percent correct from pretest to posttest rose almost imperceptibly from 59% to 61.9%, although it still fared better than other seemingly less ambiguous questions did. The 2008 question read as follows:

Which of the following topics is the MOST GENERAL for a 5-minute informative speech:

Image of women in violent video games

Women in popular culture

Popular culture (correct)

Image of women in mass media

Don't know

In retrospect, it probably would have been better to word the question less ambiguously: “Which of the following topics is TOO GENERAL for a 5-minute informative speech?” However, would the recognition that “popular culture” is more general than “image of women in violent video games” also translate into measuring the ACRL Information Literacy Outcome, “The information literate student determines the nature and extent of the information needed?” Most likely, more questions needed to come to any conclusions about U100 students’ competency in this area.

ⁱ The pretests for 2005 and 2008 are available at http://library.csun.edu/kdabbour/2005_U100_Pretest.pdf and http://library.csun.edu/kdabbour/2008_U100_Pretest.pdf. The posttests for both years are identical to the pretests, except for the exclusion of redundant demographic questions, and as such are not provided. Note that the order and numbering of the questions on the actual pretests do not always conform to how they are reported in the results section of this report.