Digital Commons @ LMU and LLS

Digital Commons@
Loyola Marymount University
and Loyola Law School

LMU Librarian Publications

William H. Hannon Library

1-1-2011

Library Instruction for Freshman English: A Multi-Year Assessment of Student Learning

Susan Gardner Archambault
Loyola Marymount University, susan.gardner@lmu.edu

Repository Citation
Archambault, Susan Gardner, "Library Instruction for Freshman English: A Multi-Year Assessment of Student Learning" (2011). *LMU Librarian Publications.* Paper 6.
http://digitalcommons.lmu.edu/librarian_pubs/6

This Article is brought to you for free and open access by the William H. Hannon Library at Digital Commons@
Loyola Marymount University
and Loyola Law School. It has been accepted for inclusion in LMU Librarian Publications by an authorized administrator of Digital Commons@
Loyola Marymount University
and Loyola Law School. For more information, please contact digitalcommons@lmu.edu.

Evidence Based Library and Information Practice

Article

Library Instruction for Freshman English: A Multi-Year Assessment of Student Learning

Susan Gardner Archambault
Head, Reference & Instructional Services
Loyola Marymount University
Los Angeles, California, United States of America
Email: susan.gardner@lmu.edu

Received: 16 Aug. 2011 **Accepted:** 28 Oct. 2011

© 2011 Gardner Archambault. This is an Open Access article distributed under the terms of the Creative Commons-Attribution-Noncommercial-Share Alike License 2.5 Canada (http://creativecommons.org/licenses/by-nc-sa/2.5/ca/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly attributed, not used for commercial purposes, and, if transformed, the resulting work is redistributed under the same or similar license to this one.

Abstract

Objective – The objective of this study was twofold: 1) to assess the effectiveness of curriculum changes made from the 2009 freshman English library instruction curriculum to the 2010 curriculum at Loyola Marymount University (LMU); and 2) to evaluate the effectiveness of library instruction delivered via a "blended" combination of face-to-face and online instruction versus online instruction alone.

Methods – An experimental design compared random samples of student scores from 2009 and 2010 worksheets to determine the effects of a new curriculum on student learning. A second experiment examined the effect of delivery method on student learning by comparing scores from a group of students receiving only online instruction against a group receiving blended instruction.

Results – The first component of the study, which compared scores between 2009 and 2010 to examine the effects of the curriculum revisions, had mixed results. Students scored a significantly higher mean in 2010 on completing and correctly listing book citation components than in 2009, but a significantly lower mean on constructing a research question. There was a significant difference in the distribution of scores for understanding differences between information found on the Internet versus through the Library that was better in 2010 than 2009, but worse for narrowing a broad research topic. For the study that examined computer aided instruction, the group of students receiving only computer-assisted instruction did significantly better overall than the group receiving blended instruction. When separate tests were run for each

skill, two particular skills, generating keywords and completing book citation and location elements, resulted in a significantly higher mean.

Conclusions – The comparison of scores between 2009 and 2010 were mixed, but the evaluation process helped us identify continued problems in the teaching materials to address in the next cycle of revisions. The second part of the study supports the idea that computer-assisted instruction is equally or more effective than blended instruction.

Introduction

Information literacy in higher education is defined as a set of abilities requiring individuals to "recognize when information is needed and have the ability to locate, evaluate, and use effectively the needed information" (ACRL 2000). Lawson (1999) wrote "although bibliographic instruction has been ongoing in libraries for many years… no single method has been established as the best" (p. 77). The Association of College & Research Libraries' professional standards on information literacy, introduced in 2000, have helped guide the practice of library instruction, but there is still no established "best" method. Anderson & May (2010) state "essentially, one class period is inadequate to provide the necessary information to gain the IL skills set forth by the ACRL" (p. 496). Houlson (2007) notes "one-shot workshops and orientations typically cover too much information and rely on passive learning" (p. 104). According to Dewald (1999) the traditional "best practices" of good class-based library instruction dictate that it be assignment-related; include active learning exercises; accommodate more than one learning style (auditory and visual); and have clear educational objectives (pp. 26-27). She advocates that these same components be used as a guide in the design of web based instruction. Dewald cautioned that online tutorials "cannot completely substitute for a human connection in learning" and they are "best used in connection with academic classes rather than in isolation" (p. 31).

Online or computer-assisted library instruction has several benefits over face-to-face instruction, including greater immediacy, greater flexibility and convenience, instant feedback, self-pacing, and greater consistency, since every student receives the same information (Holman, 2000, p.54). In most cases, the instruction remains accessible for repeat learning, if desired. However questions remain regarding the effectiveness of computer-assisted library instruction compared to face-to-face instruction at teaching undergraduates basic library skills.

There is significant pressure for libraries to adopt more evidence-based practices and to measure the ways those practices are contributing to student learning within the framework of the university. With regard to library instruction, the library "must move from a content view (books, subject knowledge) to a competency view (what students will be able to do)" (Smith, 2001, p. 32). Oakleaf (2009) describes an Information Literacy Instruction Assessment Cycle (ILIAC) framework for libraries to follow that consists of seven steps: identifying learning goals and their outcomes; creating and enacting learning activities; gathering and interpreting data to measure learning; and enacting decisions and changes based on that data. After step seven, the cycle starts again with new objectives. This assessment cycle provides feedback that librarians can use to improve their own skills, reflect on their teaching, and examine their assumptions about learning (Oakleaf, 2009, p.541). Samson notes "an assessment is only valuable when the analyses are used to augment or change the program being assessed" (p. 341). Oakleaf (2006) adds "to close the loop, educators finish the assessment cycle by using data to improve teaching and learning programs (p. 51).

Librarians at Loyola Marymount University (LMU) have been examining the ways they teach information literacy to undergraduate students. Following revisions to the curriculum, they were keen to establish whether the revisions made a difference as well as the methods in which the instruction was provided. This paper describes a study comprising two experiments that examines these issues.

Literature Review

Using Assessment to Improve Teaching

Walsh (2009) discusses the following nine methods used in case studies to assess student learning and information literacy skills following library instruction: multiple choice questionnaires, quizzes/tests, bibliographies, essays, portfolios, self-assessment, observation, simulation, and final grades. Knight (2006) advocates using "authentic assessment" methods that measure how students apply their knowledge to real-time tasks and incorporate that knowledge into academic work, rather than the artificiality of traditional standardized tests that fail to measure higher order thinking (45). Although many case studies describe the results of testing information literacy skills during one year, not many discuss how they used the data to improve their instruction programs or compare the results of using the same assessment technique across multiple years. An exception is Scharf, Elliot, Huey, Briller, & Joshi (2007), who assessed information literacy through the writing portfolios of seniors in humanities classes. The authors used previous writing portfolios to develop information literacy variables, and then analyzed the results of that assessment to address instructional issues raised by the assessment and ways to improve the teaching of those skills. Warner (2003) piloted formal assessment in 2002 on a group of pre-freshmen doing three assignments and then changed and improved teaching methods to address the learning problems she discovered for the following year. Fain (2011) examined the results of using a pre-/post-test library skills assessment over a five-year period to evaluate the effectiveness of an information literacy program and make changes. Burkhardt (2007) discussed results of pre-/post-tests over a five-year period on a three-credit undergraduate class and how improvements would be made to the teaching of low-scoring areas. Finally, Oakleaf (2009) discusses two rounds of an assessment cycle using a rubric and how the results were used to improve an online information literacy tutorial for undergraduates taking a required first-year writing course.

Face-to-Face Versus Online Learning

Many articles have compared the effects of face-to-face instructional delivery versus online instructional delivery on student learning, but the aggregate results are inconclusive. Some researchers found no significant difference in student scores for the same teaching material delivered in traditional face-to-face format versus a computer-assisted format (Germain et al., 2000; Kaplowitz & Contini, 1998; Vander Meer & Rike, 1996; Zhang, Watson, & Banfield, 2007; Holman, 2000; Koufogiannakis and Wiebe, 2006). Contrary to this, Lawson (1989) found students using a tutorial performed better than face-to-face instruction, while Anderson & May (2010) found online instruction yielded higher scores than blended or face-to-face instruction. Madland & Smith (1988) and Churkovich & Oughtred (2002) reported that students receiving face-to-face library instruction performed slightly better than computer-assisted learners; and Kraemer, Lombardo, & Lepkowski (2007) reported students receiving online-only instruction scored lower than both face-to-face or blended instruction.

Context

Loyola Marymount University (LMU) is a private, Jesuit university in Los Angeles, California. The LMU Freshman English Program (consisting of English 110: Introduction to College Writing classes, required of all freshmen) is the primary

avenue through which students are introduced to the library. The Director of the Freshman English Program requires all English 110 instructors to bring their classes to the library once during the Fall semester for a one-shot instruction sessions lasting 50 or 75 minutes. The Reference Department designed an "English 110 Library Worksheet" and an accompanying online five-module introduction to the research process promoted via a LibGuides-based guide. The first two modules of the worksheet are completed prior to the face-to-face library session as an independent homework assignment and using the LibGuide as a support resource. During the face-to-face session, students progress through the last three modules while completing the related sections of the worksheet.

The original five student learning outcomes in the LMU freshman English library instruction program developed for 2009 are:

1. Given a broad research topic, use the 4W questions (who, what, where, when) to write a research question (ability to define a research question that is significantly narrower than the original topic).

2. Given a research topic, pick out the key concepts and compile a list of search terms or keywords (ability to compile five relevant keywords for the research topic that would yield useful results if typed into the library catalog or an article index).

3. Given background information about Google and the Library, list two differences between the two related to content, organization, quality, or access (specific teaching points are authority; "invisible web;" free versus fee-based; quality control; personal assistance; Pagerank technology; and scholarly versus popular).

4. Given a research topic and access to the library's catalog, find 2 relevant books on the topic and record all relevant citation information (citation elements are title, author, location, call number, subject heading, availability, and floor).

5. Given a research topic and access to a general article index database, find 2 relevant articles on the topic and record all relevant citation information (citation elements are author, title, publication name, volume/issue, date, and pages).

After scoring the worksheets from 2009, librarians made the following changes to the curriculum to address lower-scoring areas:

- In an effort to help students learn how to define a research question that is narrow, staff inserted five additional topic examples and resulting research questions into the LibGuide (Module 1).
- To address difficulties in selecting key concepts and generating keywords, we added five additional examples to the LibGuide. We also added a video showing a student brainstorming for keywords, and two interactive concept mapping tools. Finally, we changed the worksheet to allow the topic and keywords to be updated at the end of the worksheet as they evolved (Module 2).
- Due to large numbers of blank answers in the section on articulating key differences between information found on the Internet versus the Library, our librarians reduced the number of teaching points to give students more time. To address answers showing misunderstandings, a visual image on the LibGuide shows a person "fishing" for information in an ocean with information from both the visible and the invisible web (Module 3).
- In response to a large number of blank answers for book floor location and book availability, our staff modified the worksheet to include designated

boxes for those answers. Because we could still assess what we needed by decreasing the quantity, we reduced the number of required books from two to one, and reduced the required citation elements to conform more strictly to MLA style citation components (we removed subject headings and availability status, but added location-city, publisher, year, and medium) (Module 4).

- Since we could still assess what we needed by decreasing the quantity of articles discovered, we reduced the required number of articles from two to one. This gave students more time to focus on relevancy. We modified the required citation elements to conform to MLA style (we added elements for database, medium, and date of access) (Module 5).

The literature provides few examples of the use of assessment to improve the curriculum at a basic undergraduate level, and provides conflicting evidence on face-to-face versus online learning. The purpose of this study was to analyze the effect of the revisions made to freshman English library instructional materials on student learning. Specifically, to compare student scores between 2009 and 2010 and to compare blended versus computer-aided learning. The final objective was to identify ongoing or new problem areas to feed into the next cycle of revisions.

Methods

Assessing the Effectiveness of Revisions to the Curriculum Between 2009 and 2010 (Blended Learning Instruction Models)

Librarians made revisions, described above, to the 2009 curriculum to develop the 2010 curriculum. More specifically, librarians targeted the lower-scoring modules from 2009 and made the changes outlined in the "Context" section. To determine the effects of a new curriculum on student learning, staff used an experimental design to compare student scores from 2009 and 2010, and used a random number table to sample 100 worksheets from the total number collected from both the Fall 2009 (755) and the Fall 2010 (587) cohorts. All eight members of the Reference & Instruction Department, consisting of seven librarians and one library assistant, graded the worksheets using the scoring rubric described below. Each grader scored their assigned worksheets and recorded the scores in a Google form using an analytic scoring rubric (see Appendix A). Staff then calculated averages for all modules and subsections for the Fall 2009 and Fall 2010 cohorts. A t-test was used to determine significant differences in means for each subsection in 2009 and 2010, along with a chi-square test of independence to see if the number of scores in each scoring group in 2009 and 2010 was different or equal.

Comparing Blended Learning to Online Learning

A second experiment examined the effect of delivery method on student learning by comparing scores from a group of students in 2010 receiving only online instruction through the LibGuide against the 2010 group receiving blended instruction. The students receiving online-only instruction completed all five modules online; the students receiving blended instruction completed the first two modules online and the final three modules in class during face-to-face instruction. In total, 46 worksheets (the entire sample available) from the online-only group were single-graded by library staff, who independently recorded the scores in a Google form using the same analytic scoring rubric mentioned above. Staff calculated averages for all modules and compared the averages against the 100 sample worksheet scores from 2010's blended instruction group in Part 1. The t-test was used to determine significant differences in means for each subsection.

Scoring Rubric

Librarians employed an analytic scoring rubric as the primary means of assessment to measure student learning in both experimental studies. An analytic rubric divides performance into "separate facets and each

facet is evaluated using a separate scale" (Moskal, 2003), but each facet can also be summed to form a total score. Our rubric (see Appendix A) separates each of the five learning modules into subsections with stated student learning outcomes and corresponding ACRL Information Literacy Standards, then lists overall evaluation criteria for each subsection as well as specific evaluation criteria for scoring each task. The rubric ranks each student on a 1-3 point scale of beginning, developing, or proficient for all areas. The rubric was developed in-house; for more on that development process see Gardner and Acosta (2010). Librarians established inter-rater reliability of the rubric prior to the study through a twenty worksheet sample in which the percent-agreement across multiple judges was at least 92% for each worksheet.

Results

Assessing the Effectiveness of Changes to the Curriculum Between 2009 and 2010 (Blended Learning Instruction Models)

To assess the effectiveness of the curriculum changes made between 2009 and 2010, librarians compared student scores for both years to see if there were significant improvement in scores for each module in 2010. Overall, the results were mixed. When averaging scores across all five modules, student worksheet totals were relatively the same in 2010 (2.49) compared to 2009 (2.48). When considering the mean aggregates for each Module separately, students did worse on Module 1 (narrowing a topic and defining a research question) and Module 2 (picking out key concepts and listing keywords) in 2010 than 2009, but they did better on Module 3 (listing two differences between information found on Google versus through the Library), Module 4 (finding books), and Module 5 (finding articles). When each sub-module was tested for statistical significance, students did significantly better in 2010 on Module 3 (listing two differences between information found on Google versus through the Library) and parts of Module 4 (completion and accuracy of book citation elements). However, they did significantly worse on Module 1 (completion of "who, what, where, when" questions about a topic and developing a research question from the topic).

Module 1: Narrowing a Topic

Module 1A measures completion of answering four "who, what, where, when" questions about the topic on the worksheet. Students receive a proficient (3) score for answering all four questions; a developing (2) score for answering two or three of the questions; and a beginning (1) score for answering one or none of the questions. The mean score for the 2009 students (M=2.75, SD=.54) was not significantly larger than the scores for the 2010 students (M=2.68, SD=.70) using the two-sample t-test for equal variances, t(198)=.80, p=(.43). The chi-square test for independence indicates a difference between the distribution of scores for 2009 and 2010, $X^2(2)=7.42$, p=.02 (Table 1). The higher number of "beginning" scores in 2010 reflects a higher number of students who left that section blank. More than half of the "beginning" scores from 2010 did not do their assigned homework and left all sections of modules 1 and 2 blank on the worksheet. Reasons for this difference are

Table 1
Module 1A: Distribution of Student Scores

Evaluation Criteria	Beginning (1)		Developing (2)		Proficient (3)	
Narrows topic (completion)	2009	2010	2009	2010	2009	2010
	5	13	15	6	80	81

Table 2
Module 1B: Distribution of Student Scores

Evaluation Criteria	Beginning (1)		Developing (2)		Proficient (3)	
Defines research question	2009	2010	2009	2010	2009	2010
	7	12	31	42	62	46

unknown, but might include less conscientious students in 2010, or a higher rate of student absences in 2010. Since the scoring rubric did not have a separate category for blank answers, staff re-examined the worksheets to discover this problem. Students receiving a "developing "score in both groups did not answer all four questions about their topic, and it appears to be because they did not think all the questions applied to their topic.

Module 1B measures the ability to construct a research question. Students receive a proficient (3) score for a question that is narrower than the original topic and specific enough for a research question; a developing (2) score for a question that is still too broad for a research question; and a beginning (1) score for a question that is not much narrower than the original topic. In the final overall worksheet score this section counts double since it is considered a higher-level skill. The mean score for the 2009 students (M=2.55, SD=.63) was significantly larger than the scores for the 2010 students (M=2.34, SD=.68) using the two-sample t-test for equal variances, t(198)= 2.26, p=(.02). The chi-square test for independence indicates no significant difference between the distribution of scores for 2009 and 2010, $X^2(2)=5.34$, p=.07). The main reason for the difference in "beginning" scores between 2009 and 2010 is because seven students from 2010 left it blank, failing to do this part of their homework. Students also lost points for failing to narrow the topic at all (such as "who is Tom Cruise?") or choosing an unrealistic research question (such as "how many writings does Robert Frost have?").

More students from 2009 received a "proficient" score, while more students from 2010 received a "developing" score for questions that were still too broad for a research question. A typical example of a question receiving a "developing" score is "what characteristics determine whether or not students will succeed in college?" This question is narrower than a question about general college success, but still not specific about what characteristics to determine success they want to research. To receive a "proficient" score, the question could be modified to "how good of a predictor is standardized testing for college success?" Because the scores in this section significantly decreased in 2010, we continue to seek ways to improve the teaching material. We will take the six research question examples on the LibGuide and illustrate what would constitute a "beginning" question, a "developing" question, and a "proficient" question to further illustrate the refining process. Also, we can encourage faculty to assign a general research topic to the class.

Module 2: Key Concepts

Module 2A measures the ability to list three key concepts within a research question. Students receive a proficient (3) score for listing all three of the most important concepts from their research question; a developing (2) score for listing two of the three; and a beginning (1) score for listing one or none. The mean score for the 2009 students (M=2.47, SD=.67) was not significantly larger than the scores for the 2010 students (M=2.41, SD=.78) using the two-sample t-test for equal variances, t(198)=.58, p=(.56). The chi-square test for independence indicates no significant difference between the distribution of scores for 2009 and 2010, $X^2(2)=4.11$, p=.13 (Table 3). In the 2010 group, twelve students left this

Table 3
Module 2A: Distribution of Student Scores

Evaluation Criteria	Beginning (1)		Developing (2)		Proficient (3)	
Lists key concepts	2009	2010	2009	2010	2009	2010
	10	18	33	23	57	59

section blank compared to only four in the 2009 group. Only six students in each group received a "beginning" score because they tried and failed to list more than one of the key concepts from their research question. Common mistakes for students in each group receiving a "developing" score were adding an extra term that was related to but not part of the original research question; or forgetting to include the "who" part of the research question or the "what" part in favor of a less important adjective (for a topic about "modern undergraduates," including both 'modern' and 'undergraduates' as key terms in lieu of the main verb).

Module 2B measures the ability to list keywords related to a research topic that would lead to search results in a library catalogue or article index. Students receive a proficient (3) score for listing a total of five or more relevant keywords related to their research question; a developing (2) score for listing three or four keywords; and a beginning (1) score for listing two or less keywords. The mean score for the 2009 students (M=2.47, SD=.78) was not significantly larger than the scores for the 2010 students (M=2.32, SD=.84) using the two-sample t-test for equal variances, t(198)=1.31, p=(.19). The chi-square test for independence indicates no significant difference between the distribution of scores for 2009 and 2010, $X^2(2)=1.77$, p=.41 (Table 4). Students who received a "beginning" or a "developing" score often had trouble with the concept of a synonym, either conceptually or technically (picking a word not recognized as a linguistic unit, such as "automotist" as a synonym for driver). It was also common for students to select descriptive phrases rather than a singular synonym or keyword (e.g, using "family falls apart" instead of "divorce"). As the additional examples and tools added to the LibGuide in 2010 for this section did not increase scores, and the number of hits for this portion of the LibGuide was only 632, we will add a mandatory interactive keywords exercise to help students understand the concept of a keyword and gain more practice.

Module 3: Information Found on the Internet versus the Library

Module 3 measures the ability to list differences between information found on the Internet and the Library related to our teaching points. Students receive a proficient (3) score for listing a total of two or more differences; a developing (2) score for listing one difference; and a beginning (1) score for listing no differences. The mean score for the 2009 students (M=2.06, SD=.92) was not significantly smaller than the scores for the 2010 students (M=2.19, SD=.83) using the two-sample t-test for equal variances, t(198)=–1.05, p= .29). The chi-square test for independence indicates a significant difference between the distribution of scores for 2009 and 2010, $X^2(2)=6.36$, p=.04 (Figure 5). There were less "beginning" scores in 2010, and a primary reason was because fewer students in 2010 left this section blank than in 2009 (18 compared to 35) We shortened the module in 2010 to contain fewer and more standardized teaching points, which may have allowed more time for

Table 4
Module 2B: Distribution of Student Scores

Evaluation Criteria	Beginning (1)		Developing (2)		Proficient (3)	
Compiles keywords	2009	2010	2009	2010	2009	2010
	18	24	17	20	65	56

Table 5
Module 3: Distribution of Student Scores

Evaluation Criteria	Beginning (1)		Developing (2)		Proficient (3)	
Lists differences or characteristics: Lib & Internet	2009	2010	2009	2010	2009	2010
	39	26	16	29	45	45

completion. We also added a visual image to the LibGuide of a person "fishing" for information in an ocean containing information from both the visible and the invisible web in an attempt to help illustrate differences in information types. Since more students received a "developing" score in 2010 than in 2009, and since more students were able to list one difference related to our teaching points rather than no differences, the image might have been partially effective. The same number of students received a "proficient" score both years, however, so we cannot be sure of the extent of our changes. We will independently test Module 3 on a group of students both with and without the visual image. Also, we will do a peer review observation of Module 3 to check for standardization. Students lost points for factually incorrect answers (such as "the Internet gives you web sites with false information") or answers that were not one of our teaching points (such as "libraries can give you what you need"). Module 3 was consistently the lowest scoring module among students from both years.

Module 4: Finding Books

Module 4 measures the ability to find one book in the library catalogue (4A: quantity), with separate scores for completeness (4B) and the accuracy (4C) of the following citation elements: location, subject heading, availability, call number, floor, author, title, location/city, publisher, year, and medium. The relevancy of the book to the research topic (4D) illustrates a more sophisticated level of development, therefore this score counts double in the final overall calculated worksheet score. The mean scores for 2009 and 2010 were not significantly different for the section relating to quantity (4A) and the section relating to relevancy (4D); almost all students in both groups could find one relevant book using the library catalogue.

For the section relating to completeness (4B), the mean score for the 2009 students (M=2.43, SD=.57) was significantly smaller than the scores for the 2010 students (M=2.71, SD=.54) using the two-sample t-test for equal variances, $t(198)=-3.56$, $p=(.00)$. The chi-square

Table 6
Module 4: Distribution of Student Scores

Module	Evaluation Criteria	Beginning (1)		Developing (2)		Proficient (3)	
4A	(Quantity) Lists 1 book	2009	2010	2009	2010	2009	2010
		1	2			99	98
4B	(Completion) Citation elements	4	4	49	21	47	75
4C	(Accuracy) Citation elements	5	4	51	27	44	69
4D	(Relevancy) Links book to topic	11	10			89	90

test for independence indicates a difference between the distribution of scores for 2009 and 2010, $X^2(2)=17.63$, p=.00. The main reason for the increase in scores between 2009 and 2010 is likely due to a lack of designated answer boxes on the 2009 worksheet for the book floor and availability questions. In 2009, students left blank the element for "floor" 38 times, followed by "availability" 20 times, resulting in lower scores. After we modified the worksheet to have separate boxes for each individual citation element, the 2010 students only left the "floor" element blank eight times. Also, the changes made to citation elements collected makes this section now match its learning outcome of recording "all pertinent" citation information needed for Modern Language Association (MLA) citation style.

For the section on accuracy (4C), the mean score for the 2009 students (M=2.39, SD=.58) was significantly smaller than the scores for the 2010 students (M=2.65, SD=.56) using the two-sample t-test for equal variances, t(198)= -3.22, p= (.00). The chi-square test for independence indicates a difference between the distribution of scores for 2009 and 2010, $X^2(2)=13.03$, p=.00. In 2009, a lot of answers counted as incorrect were blank for "floor" and "availability." When looking to see how many of the answers were incorrect rather than merely blank, the results are "subject" (5 times) and "floor" (5 times). In 2010, the only element that was attempted (not blank) but incorrect was "floor" (15 times). This shows that "floor" is still problematic, and students either fail to locate the floor directory or do not know how to interpret it. We have requested that the floor location of a book appear as a popup over the call number in the next upgrade of our catalogue.

Module 5: Finding Articles

Module 5 measures the ability to find one article in the article index (5A: quantity), with separate scores for completeness (5B) and the accuracy (5C) of the following citation elements: author, article title, source, volume, issue, date, pages, database, medium, and date of access. There is also a separate score for relevancy to the research topic (5D), which counts double on the overall worksheet score since it illustrates a higher lever of development. The mean scores for 2009 and 2010 were not significantly different for any of the sections in Module 5; almost all students in both years could find one article using an article index and complete the citation elements correctly. Despite the fact that the curriculum was changed in 2010 to require only one article to give students more time to focus on relevancy (section 5D), students did not do significantly better on that section. The fact that students in both years scored low on finding keywords (Module 2B) also probably influenced the relevancy score for articles, since they were instructed to type these same keywords into the article index.

Table 7
Module 5: Distribution of Student Scores

Module	Evaluation Criteria	Beginning (1)		Developing (2)		Proficient (3)	
5A	(Quantity) Lists 1 article	2009	2010	2009	2010	2009	2010
		4	7			96	93
5B	(Completion) Citation elements	7	7	18	18	75	75
5C	(Accuracy) Citation elements	8	8	20	18	72	74
5D	(Relevancy) Links article to topic	12	12	21	15	67	73

Overall, the average student scores were better in 2010 than 2009 on Module 3 (listing differences between sources found on Google versus the Library), Module 4 (finding books), and Module 5 (finding articles). When each sub-module was tested for statistical significance, though, the significant results were Module 3 (listing two differences between information found on Google versus through the Library) using the chi-square test for independence; and parts of Module 4 (completion and accuracy of book citation elements) using the two-sample t-test for equal variances. Unfortunately, the average student scores were worse in 2010 than 2009 on Module 1 (narrowing a topic and defining a research question) and Module 2 (picking out key concepts and listing keywords). When sub-modules were tested for statistical significance, Module 1A (narrowing a topic by completing "who, what, where, and when" questions) was significant using the chi-square test for independence, and Module 1B (developing a research question for the topic) was significant using the two-sample t-test for equal variances.

Comparing Blended Learning to Online Learning

The second part of the study involved assessing the effectiveness of delivery method on students receiving the 2010 curriculum. Staff compared scores from students receiving a blended combination of face-to-face and online instruction through the LibGuide against the scores of students receiving only online instruction. The mean worksheet scores across all five modules for the blended instruction group of students (M=2.49, SD=.39) was significantly smaller than the scores for the students receiving only online instruction (M=2.63, SD=.25) using the two-sample t-test for equal variances, t(198)= 2.11, p=(.04). Looking at each module individually, students receiving online-only instruction did slightly better on Module 1 (defining a research question), Module 2 (picking out the key concepts of the research question and compiling keywords), Module 3 (listing two differences between information found on Google versus through the Library), and Module 4 (finding books). They scored exactly the same on Module 5 (finding articles). However, when each subsection was tested for statistical significance only two of them were statistically significant: Module 2B (listing keywords) and Module 4B (completion of book citation elements).

Module 2B: Listing Keywords

The mean score for the 2010 online-only group on Module 2B, listing keywords (M=2.67, SD= .52) was significantly larger than the scores for the 2010 blended group (M=2.32, SD=.84) using the t-test for equal variances, t(144) =– 2.64, p=(.01). The 2010 blended group was negatively affected by seventeen blank answers. The group doing the entire worksheet as a homework assignment using only the LibGuide had almost no blank answers. It is possible that the online-only group paid more attention to the supplemental materials listed on the LibGuide, such as the interactive concept map tools and the brainstorming video, leading to this difference. These students may have read more carefully in the absence of any personal assistance. The blended group had the benefit of seeing the keywords module demonstrated by a librarian during the face-to-face session, which could have led to higher scores but did not. A higher number of students in the online-only group indicated at the end of the worksheet that they went back and added keywords to Module 2B, which probably increased their scores. It is also possible that students who completed the worksheet during a face-to-face instruction session had limited time and may not have seen the prompt at the end of the worksheet to add more keywords.

Module 4B: Completing Citations

The mean score for the online-only group on Module 4B, completing book citation elements, (M=2.91, SD=.28) was significantly larger than the scores for the 2010 students receiving blended instruction (M=2.71, SD=.54) using the t-test for equal variances, t(144)=– 2.41, p=(.02). Key differences centered on the "floor" element and "medium" element.

"Floor" was left blank eight times for the blended group but only two times for the online-only group. The floor element was not on the book record in the library catalogue, but could only be learned by consulting a separate directory. The "medium" element was left blank twelve times by the blended group but only one time by the online-only group. Again, time constraints may have played a factor for the blended instruction group, since they had to complete Module 4B during a 50-minute class period and the online-only group had time for repeated readings or more reflection.

Discussion

This two-part study has examined the effects of curriculum changes on student learning by comparing differences in student scores in 2009 and 2010. Furthermore, it has compared the scores of students who received instruction by blended learning against those who received an online tutorial.

The average worksheet scores across all five modules were almost the same when comparing the effects of changes to the curriculum in two different years (2.49 in 2009 versus 2.48 in 2010), but there were significant differences across individual subsections. Students in 2010 showed significant improvement in Module 3, listing two differences between information found on the Internet versus through the Library. Fewer students left this section blank in 2010, probably because the module contained fewer and more standardized teaching points, and students had more time to complete that section of the worksheet in class. The impact of adding a visual image of a person "fishing" for information to illustrate the visible web versus invisible web is unknown; we plan to test the image against a control group. Since this was the lowest scoring module for both years, we will also undergo a peer review check for consistency of the teaching points amongst all librarians. The other result implying positive and effective changes made to the curriculum in 2010 was Module 4 (completion and accuracy of book citation elements). The improvement in student scores was largely due to less blank answers, and therefore having separate, designated fields appears to encourage responses to individual components of a question on a worksheet.

It was disappointing that students scored significantly worse in 2010 on Module 1, narrowing a broad research topic and constructing a research question. More students in 2010 left this section blank, effectively failing to do their homework. Since this segment was completed independently as a homework assignment, we can only speculate on the cause but will work with instructors to ensure greater accountability. Because of the lower scores due to blank answers, we will differentiate between a "blank" answer and an "incorrect" answer in future rubrics to separate "time" versus "learning" problems. For Module 1B (developing a research question), the six topic examples on the LibGuide will extend to illustrate for each a "beginning" research question, a "developing" question, and a "proficient" question. Despite these mixed results, the feedback loop continues to give us valuable information to improve the teaching materials for the next cycle. For a complete list of changes we will make next year, see Appendix B.

The results for the part of the study that compared blended learning against online learning showed that online learning students scored significantly better on average across all five modules than the group receiving blended instruction (2.63 versus 2.49). When subsections were broken down and tested for statistical significance, listing keywords (2B) and completion of book citation elements (4B) tested positive using the t-test for equal variances. For both sections, more students in the blended instruction model had blank answers. This suggests that the time constraints of a 50-minute class period played a negative role, but it also highlights the fact that research process is cyclical and it is important to simulate this by allowing students the flexibility to go back and change their topic or to add keywords as the research

evolves. The self-paced nature and flexibility of online instruction better accommodates this. This part of the study adds to the research on the effects of instructional delivery on student learning. Because the online-only group scored significantly higher than the blended instruction group overall, and two individual skill elements also scored significantly higher when isolated, it supports the findings of Anderson & May (2010), who found that students receiving online instruction scored better when tested on their ability to find sources than students receiving blended instruction. It contrasts with the findings of Kraemer, Lombardo, & Lepkowski (2007), who found that students receiving online-only instruction scored lower than those receiving blended instruction. It indirectly supports Lawson (1989), who found an online tutorial more effective than face-to-face instruction. More research needs to be conducted comparing blended instruction to other models. As highlighted in the literature review, the majority of studies that compare online instruction to face-to-face instruction find no significant differences in the two delivery methods. In this study, the blended instruction group did not appear to benefit significantly from the face-to-face demonstrations of each module during class time. The ability to work at their own pace and complete the worksheet when they wanted, rather than during a compressed class period, appears to have a greater positive impact than live modeling from a librarian, leading to higher scores for the online-only group.

The major limitation of this study is its methodology. Since the design was a field experiment, many different classes were included in the random sample. Further, the attitude of the instructor in each class may have impacted how seriously students took the worksheet assignment. Similarly, seven different librarians taught the face-to-face segments, and while standardization was attempted, there may have been some inconsistencies in delivery. No control group was used to test how well a student could complete the worksheet skills on their own without any aid, either from the LibGuide or blended instruction. Another limitation was that there was no pre-test, and so it is unknown whether the students in each year started from a different baseline of knowledge. Also, no student demographic data was collected to check whether the students in each group were similar. Finally, the in-house rubric used to score student worksheets only has three possible scores, so parts of the high inter-rater reliability achieved at calibration might have been reached by chance and not because the judges truly scored the same way. Nevertheless, the study highlights issues involved in conducting experiments in practice as well as the limitations of using experimental methods for changing practice.

Conclusion

The purpose of this study was to assess the effectiveness of curriculum changes made from the 2009 LMU freshman English library instruction curriculum to the 2010 curriculum, and also to compare the effectiveness of library instruction delivered online-only versus a "blended" combination of face-to-face and online instruction. Results were mixed; students scored significantly higher in 2010 on comparing Internet and library resources and accurately completing book citation components, but scored lower on narrowing a topic and defining a research question than students taking the course in 2009. Despite these mixed results, the evaluation process helped us identify continued problems in the teaching materials for the next revision cycle. Results of the instructional delivery method comparison revealed students receiving only computer-assisted instruction did better overall, and significantly better on the individual skill elements of generating keywords and completing book citation and location elements than students receiving blended instruction. This suggests that computer-assisted instruction through a LibGuide holds promise as an alternative to face-to-face or blended instruction. The teaching materials from this study can be adapted for a variety of instructional settings using any combination of delivery method.

Acknowledgements

A version of this paper was presented at the 6th International Evidence Based Library and Information Practice (EBLIP) Conference, held in Salford, UK (June 2011). The author would like to thank the LMU Instruction Coordinator, Elisa Slater Acosta; all members of the LMU Reference & Instruction Department; Laura Massa in the LMU Office of Assessment; and the William H. Hannon Library for a 2010/11 Research Incentive Travel Grant.

References

Anderson, K. & May, F.A. (2010). Does the method of instruction matter? An experimental examination of information literacy instruction in the online, blended, and face-to-face classrooms. *Journal of Academic Librarianship, 36*(6), 495-500.

Association of College and Research Libraries. (2000). *Information literacy competency standards for higher education.* Retrieved from http://www.ala.org/ala/mgrps/divs/acrl/standards/informationliteracycompetency.cfm

Burkhardt, J.M. (2007). Assessing library skills: A first step to information literacy. *portal: Libraries and the Academy, 7*(1), 25-49.

Churkovich, M. & Oughtred, C. (2002). Can an online tutorial pass the test for library instruction? An evaluation and comparison of library skills instruction methods for first year students at Deakin University. *Australian Academic and Research Libraries, 33*(1), 25–38.

Dewald, N.H. (1999). Transporting good library instruction practices into the web environment: An analysis of online tutorials. *Journal of Academic Librarianship, 25*(1), 26-31.

Fain, M. (2011). Assessing information literacy skills development in first year students: A multi-year study. *Journal of Academic Librarianship, 37*(2), 109-119.

Gardner, S. & Acosta, E.S. (2010). Using a rubric to assess freshman English library instruction. In S. Hiller, K. Justh, M. Kyrillidou, & J. Self (Eds.), *Proceedings of the 2010 Library Assessment Conference: Building effective, sustainable, practical assessment* (pp. 159-173). Washington, DC: Association of Research Libraries.

Germain, C.A., Jacobson, T.E., & Kaczor, S.A. (2000). A comparison of the effectiveness of presentation formats for instruction: Teaching first-year students. *College & Research Libraries, 61*(1), 65-72.

Holman, L. (2000). A comparison of computer-assisted instruction and classroom bibliographic instruction. *Reference & User Services Quarterly, 40*(1), 53-60.

Houlson, V. (2007) Getting results from one-shot instruction. *College & Undergraduate Libraries, 14*(1), 89-108.

Kaplowitz, J. & Contini, J. (1998). Computer-assisted instruction: Is it an option for bibliographic instruction in large undergraduate survey classes? *College & Research Libraries, 59*(1), 19-27.

Knight, L.A. (2006). Using rubrics to assess information literacy. *Reference Services Review, 34*(1), 43-55.

Koufogiannakis, D. & Wiebe, N. (2006). Effective methods for teaching information literacy skills to undergraduate students: A systematic review and meta-analysis. *Evidence Based Library and Information Practice,*

1(3). Retrieved from http://ejournals.library.ualberta.ca/index.php/EBLIP/article/view/76/153.

Kraemer, E.W., Lombardo, S.V., & Lepkowski, F.J. (2007). The librarian, the machine, or a little of both: A comparative study of three information literacy pedagogies at Oakland University. *College & Research Libraries, 68*(4), 330-342

Lawson, M.D. (1999). Assessment of a college freshman course in information resources. *Library Review, 48*(2), 73-78.

Lawson, V.L. (1989). Using a computer-assisted instruction program to replace the traditional library tour: An experimental study. *RQ, 29*(1), 71-79.

Madland, D., & Smith, M.A. (1988). Computer-assisted instruction for teaching conceptual library skills to remedial students. *Research Strategies, 6*(2), 52-64.

Moskal, B.M. (2003). Recommendations for developing classroom performance assessments and scoring rubrics. *Practical Assessment, Research & Evaluation, 8*(14). Retrieved from http://PAREonline.net/getvn.asp?v=8&n=14

Oakleaf, M. (2006). Assessing information literacy skills: A rubric approach. (Doctoral dissertation). Retrieved from Proquest Dissertations and Theses database. (UMI No. 3207346)

Oakleaf, M. (2009). The information literacy instruction assessment cycle: A guide for increasing student learning and improving librarian instructional skills. *Journal of Documentation, 65*(4), 539-560.

Samson, S. (2000). What and when do they know? Web based assessment. *Reference Services Review, 28*(4), 335-342.

Scharf, D., Elliot, N., Huey, H., Briller, V., & Joshi, K. (1997). Direct assessment of information literacy using writing portfolios. *Journal of Academic Librarianship, 33*(4), 462-477.

Smith, K.R. (2001). New roles and responsibilities for the university library: Advancing student research through outcomes assessment. *Journal of Library Administration, 35*(4), 29-36.

Vander Meer, P.F., & Rike, G.E. (1996) Multimedia: Meeting the demand for user education with a self-instructional tutorial. *Research Strategies, 14*, 145-158.

Walsh, A. (2009). Information literacy assessment: Where do we start? *Journal of Librarianship and Information Science, 41*(1), 19-28.

Warner, D.A. (2003). Programmatic assessment: Turning process into practice by teaching for learning. *Journal of Academic Librarianship, 29*(3): 169-176.

Zhang, L., Watson, E.M., & Banfield, L. (2007). The efficacy of computer-assisted instruction versus face-to-face instruction in academic libraries: A systematic review. *Journal of Academic Librarianship, 33*(4), 478-484.

Appendix A
Title here beneath the label

	Student Learning Outcomes	Evaluation Criteria	Beginning = 1	Developing = 2	Proficient = 3	
Module 1						
1.a.	Defines or modifies information need to achieve manageable focus through the dissection of a broad topic [ACRL Standard 1, indicator 1.d]	Narrows topic (completion)	Answers zero or 1 of the "when, where, who, what" questions about the topic in the box	Answers 2 or 3 of the "when, where, who, what" questions about the topic in the box	Answers all 4 of the "when, where, who, what" questions about the topic in the box	(x1)
1.b.	Defines or modifies information need to achieve manageable focus through the construction of a specific research question	Defines research question	Constructs no question or a question that is not much narrower than the original topic	Constructs a question that is narrower than the original topic, but still too broad for a research question	Constructs a question that is narrower than the original topic and specific enough for a research question	(x2)
Module 2	Student Learning Outcomes	Evaluation Criteria	Beginning = 1	Developing = 2	Proficient = 3	
2.a.	Identifies key concepts and terms that describe the information needed [ACRL Standard 1, indicator 1.e]	Lists key concepts	Lists 1 or less of the most important concepts from the research question	Lists 2 out of the 3 most important concepts from the research question	Lists all 3 of the most important concepts from the research question	(x1)
2.b.	Identifies keywords, synonyms and related terms for the information needed [ACRL Standard 2, indicator 2.b]	Compiles keywords	Lists a total of 2 or less relevant keywords for the research topic. If in doubt, type the keywords into the library catalog or article index to test for relevancy	Lists a total of 3-4 relevant keywords for the research topic. If in doubt, type the keywords into the library catalog or article index to test for relevancy	Lists a total of 5 or more relevant keywords for the research topic. If in doubt, type the keywords into the library catalog or article index to test for relevancy	(x1)
Module 3	Student Learning Outcomes	Evaluation Criteria	Beginning = 1	Developing = 2	Proficient = 3	
	Investigates the scope, content, or organization of two information retrieval systems [ACRL Standard 2, indicator 1.c]	Lists differences or characteristics	Lists no differences between or characteristics of Google, the Internet, or the library related to authority; invisible web; free versus fee-based; quality control; personal assistance; Pagerank technology; popularity; or scholarly	Lists 1 difference between or characteristic of Google, the Internet, or the library related to authority; invisible web; free versus fee-based; quality control; personal assistance; Pagerank technology; popularity; or scholarly	Lists 2 or more differences between or characteristics of Google, the Internet, or the library related to authority; invisible web; free versus fee-based; quality control; personal assistance; Pagerank technology; popularity; or scholarly	(x1)
Module 4	Student Learning Outcomes	Evaluation Criteria	Beginning = 1	Developing = 2	Proficient = 3	
4.a.	Recognizes relevant information sources using the library catalog and records all pertinent citation information for future reference [ACRL Standard 2, indicators 5.c and 5.d]	a. (Quantity) Locates 1 book	Finds no books in the library catalog and doesn't write down any citation information	No "2" value for this x	Finds 1 book in the library catalog and writes down the citation information	(x1)

4.b.		b. (Completion) Lists location and citation components	Location (1 point) Call Number (1point) Floor (1 point) Author (1 point) Title (1 point) Location/city (1 point) Publisher (1 point) Year (1 point) Medium (1 point)	Location (1 point) Call Number (1 point) Floor (1 point) Author (1 point) Title (1 point) Location/city (1 point) Publisher (1 point) Year (1 point) Medium (1 point)	Location (1 point) Call Number (1 point) Floor (1 point) Author (1 point) Title (1 point) Location/city (1 point) Publisher (1 point) Year (1 point) Medium (1 point)	(x1)
			Total is 4 or less T	Total is 5-8	Total is 9	
4.c.		c. (Accuracy) Lists location and citation elements in correct fields	Location (1 point) Call Number (1 point) Floor (2 points) Author (1 point) Title (1 point) Location/city (1 point) Publisher (1 point) Year (1 point) Medium (1 point)	Location (1 point) Call Number (1 point) Floor (2 points) Author (1 point) Title (1 point) Location/city (1 point) Publisher (1 point) Year (1 point) Medium (1 point)	Location (1 point) Call Number (1 point) Floor (2 points) Author (1 point) Title (1 point) Location/city (1 point) Publisher (1 point) Year (1 point) Medium (1 point)	(x1)
			Total is 3 or less	Total is 4-8	Total is 9-10	
4.d.		d. (Relevancy) Links book to research topic	Finds a book that is not relevant to the research topic. If in doubt, look at the subject field, title field, call number field, or table of contents to make the determination	No "2" value for this x	Finds a book that is relevant to the research topic. If in doubt, look at the subject field, title field, call number field, or table of contents to make the determination	(x2)
Module 5	Student Learning Outcomes	Evaluation Criteria	Beginning = 1	Developing = 2	Proficient = 3	
5.a.	Recognizes relevant information sources using an article index; Understands the elements and correct syntax of an article citation; and Records complete citation information from the index for future reference [ACRL Standard 2, indicators 5.c and 5.d]	a. (Quantity) Locates 1 article	Finds no articles in the article index and doesn't write down any citation information	No "2" value for this x	Finds 1 article in the article index and writes down the citation information	(x1)
5.b.		b. (Completion) Lists citation components	Author (1 point) Article Title (1 point) Source/Pub (1 point) *Volume (1 point) *Issue (1 point) Date (1 point) Pages (1 point) Database (1 point) Medium (1 point) Date/Access (1 point) *if newspaper, no vol/issue so give 1 point	Author (1 point) Article Title (1 point) Source/Pub (1 point) *Volume (1 point) *Issue (1 point) Date (1 point) Pages (1 point) Database (1 point) Medium (1 point) Date/Access (1 point) *if newspaper, no vol/issue so give 1 point	Author (1 point) Article Title (1 point) Source/Pub (1 point) *Volume (1 point) *Issue (1 point) Date (1 point) Pages (1 point) Database (1 point) Medium (1 point) Date/Access (1 point) *if newspaper, no vol/issue so give 1 point	(x1)
			Total is 5 or less	Total is 6-9	Total is 10	

5.c.		c. (Accuracy) Lists citation elements in correct fields	Author (1 point) Article Title (1 point) Source/Pub (2 points) *Volume (1 point) *Issue (1 point) Date (1 point) Pages (2 point) Database (1 point) Medium (1 point) Date/Access (1 point) *if newspaper, no vol/issue so give 1 point	Author (1 point) Article Title (1 point) Source/Pub (2 points) *Volume (1 point) *Issue (1 point) Date (1 point) Pages (2 point) Database (1 point) Medium (1 point) Date/Access (1 point) *if newspaper, no vol/issue so give 1 point	Author (1 point) Article Title (1 point) Source/Pub (2 points) *Volume (1 point) *Issue (1 point) Date (1 point) Pages (2 point) Database (1 point) Medium (1 point) Date/Access (1 point) *if newspaper, no vol/issue so give 1 point	(x1)
			Total is 4 or less	Total is 5-10	Total is 11-12	
5.d.		d. (Relevancy) Links article to research topic	Finds an article that is not relevant to the research topic. If in doubt, look at the abstract or full text of the article	Finds an article that is somewhat relevant to the research topic, but is general or broad. If in doubt, look at the abstract or full text of the article	Finds an article that is relevant to the research topic. If in doubt, look at the abstract or full text of the article	(x2)

Appendix B
Future Changes to Teaching Materials

- For Module 1B (developing a research question), the six topic examples on the LibGuide will extend to illustrate for each a "beginning" research question, a "developing" question, and a "proficient" question. A dartboard analogy getting closer to its target will be visually created. Also, we will push for the instructor to assign a research topic.

- For Module 2B (keywords) we will add a mandatory keywords exercise for extra practice.

- For Module 3 (Internet resource versus Library resource), we will test the visual "fishing" image used and also do a peer review check for consistency of the teaching points.

- The rubric will be modified across all modules to include a "0" score for blank answers, to differentiate a blank answer from an incorrect one.

- The entire worksheet with all five modules will be assigned as homework using only the LibGuide prior to the face-to-face library session rather than only modules 1 and 2.