

EVALUATING A PRE-SESSION HOMEWORK EXERCISE IN A STANDALONE INFORMATION LITERACY CLASS

Joseph E. Goetz
Rice University

Catherine R. Barber
University of St. Thomas, Houston

In this study, researchers evaluate a homework exercise assigned before a standalone information literacy session. Students in a Master of Education program completed a worksheet using the ERIC database thesaurus. The researchers conducted pre- and posttests within a single library session to assess student learning, using a control group for comparison. The treatment group did not demonstrate better thesaurus skills than students who had regular library instruction alone, but results pointed the way to targeted improvements of pre-session learning materials. This approach could inform other information literacy homework applications such as flipping the classroom.

INTRODUCTION

Librarians teaching standalone “one-shot” instruction sessions may feel forced to choose among full content coverage, student practice, and assessment. Extending student learning time by assigning activities outside of class can lessen the constraints of that dilemma. But demonstrating the value of those learning activities can present instructional librarians with an additional challenge. In the traditional one-shot format, course instructors expect students to develop their skills independently after a self-contained library session. How should librarians make the case for assigning an activity before the session?

The argument to expand single-session information literacy instruction should rely on evidence of student learning. This article describes the evaluation of a pre-session homework activity’s impact on student learning compared with in-class instruction alone using pretest and posttest assessments administered within the standalone class. In keeping with its traditional one-shot context, this article does not assume that librarian instructors can reliably appropriate time outside the session to conduct assessments. The researchers—the information literacy librarian and the educational research program director—sought to answer this and other constraints with a site-specific approach to assessment. They found that despite limiting conditions, valid assessment data can point the way to iterative improvement of instructional practices and student learning.

LITERATURE REVIEW

In this study, assessments were used to

evaluate the impact on student learning of completing a pre-session exercise, assigned as independent homework, in addition to in-session instruction on database search skills and strategy. Given the approach and institutional context, the researchers saw the following constraints as imperative:

- The librarian would retain control over the administration and collection of assessment materials, including an in-class pretest and posttest.
- The assessment would compare the effects of pre-session homework on the treatment group versus the control group.
- Treatment and control groups would be established based on pre-existing course sections rather than individual random assignment.
- A posttest performance task would assess students’ skill application in an authentic database search scenario.

As with the authors of previous studies (Bryan & Karshmer, 2013; Hufford, 2010), the researchers had trouble finding cases in the information literacy instruction literature that fully applied to their own situation. In response, the researchers analyzed the literature for the ways librarians and their collaborators responded to methodological constraints: namely, the timing of pre- and posttests, the ethics and composition of control groups, and the use of performance assessments in information literacy. Overall, methodological strategy provided a useful framework for applying the arguments and lessons of previous researchers to a specially adapted assessment project.

Pretest and Posttest Timing

Pretests and posttests are commonly employed tools for assessing student learning from library instruction (Hufford, 2010, p. 140), but librarians face choices in when to assign these tests, and challenges in ensuring their completion. Assigning pretests and posttests outside the library session can both save time for instruction and affect the focus of the assessment. For example, Choinski and Emmanuel (2006), not wanting to lose “even a minute” of instructional time, had instructors assign one-minute paper assessments as extra-credit homework rather than an in-class activity (p. 151). In addition to making more time available, assigning pretests and posttests before and after the session may help assess students’ longer-term development, as discussed by Pierce and Fox (2012, p. 4). Carter (2002), despite having two sessions of a freshman seminar class available to teach research skills, arranged for a pretest to be administered during academic orientation and a posttest at the end of the semester (p. 38). Similarly, Swoger (2011) described pretests and posttests as part of a semester-length assessment cycle.

Conversely, conducting assessments outside the library session can lead to difficulties controlling the process. Brooks (2013) blamed low student motivation for poor response rates on voluntary posttests emailed to students. Portmann and Roush (2004) named “student apathy” as a “fatal flaw” to their research design (p. 464), and pointed to the need for grades to increase student motivation. Still, tests administered during instructors’ class time can also face interference from course priorities and other factors. Hsieh and Holden (2010) noted that “consistent and persistent” communication

efforts by the librarians were required to prevent drop-offs in the teaching faculty’s administration of tests and surveys (p. 466). Bryan and Karshmer (2013) also experienced low participation by instructional faculty; in response, the authors visited each class in person to administer pretests and posttests (p. 580). The need for this recourse points to the use of instructors’ class time as a challenge in terms of both logistics and scalability.

Control Groups: Whether and How

Librarians have not always seen control groups as necessary or desirable for assessment. Barclay (1993) pointed out the practical and ethical difficulties of creating control groups in library research, and argued that it is better to dispense with them than not to do research at all. On a practical level, different library instruction sessions are frequently too dissimilar to each other to serve as valid control and experimental elements in a single research project; Carter (2002), responding to practical difficulties, used pretests and posttests without control groups to measure student learning and improve the efficiency of class time. Ethically, control groups may seem to call for one group of students to be taught less well than another. Bryan and Karshmer (2013) addressed this dilemma by teaching both groups the same content with only differing methods of instruction; using a control group allowed them to compare outcomes. When the superiority of either instructional method is still unclear, the potential of long-term benefit can outweigh the risk of using multiple approaches. Nevertheless, librarians might hesitate to use techniques to benefit future students at the possible expense of those present.

Whatever the benefits of using a control group, educational researchers have long recognized the frequent impracticality for their field of individual random group assignment. Campbell and Stanley (1963) evaluated a range of models for performing quasi-experimental research, in which individual random assignment or some other requirement for true experimental research does not take place. Among those models, the nonequivalent control group design allows for division of treatment and control subjects by pre-existing groups, such as course sections. However, the model calls for a pretest to strengthen the evidence that these groups are not significantly different from each other in the area being measured (p. 47-48). Campbell and Stanley argued that the nonequivalent control group design, while not truly experimental, is “well worth using in many instances” and more secure (all else being equal) from threats to internal and external validity than pretest-posttest arrangements that forgo using a control group (p. 47).

Performance Assessment

Performance assessment, as advocated by Wiggins and McTighe (2005), gives students a chance to demonstrate a transferable understanding of skills that goes beyond recall-based knowledge (p. 153-155). In the context of library instruction, Oakleaf (2008) contrasted performance assessments with fixed-choice tests, pointing out the advantages of assessments that “reinforce the concept that what students learn in class should be usable outside the classroom” (p. 239). As one example of performance assessment in a standalone library session, Bluemle et al. (2013) described a “Source Evaluation Worksheet” that could be used in classes

across different departments, and either assigned within the session or as homework. In the worksheet, students were asked to find a source, write a citation for it and write several sentences evaluating its appropriateness. This task directly connected the instruction session’s learning objectives with students’ research work, calling for open-ended responses that allowed students to demonstrate applied understanding. Teaching a semester-long science information literacy course, Johnson et al. (2011) assessed students’ growth with a variety of performance-based homework assignments and exam questions together with surveys and citation analysis. Such approaches may seem best suited to extended course formats, but as Bluemle et al. (2013) pointed out, “carefully designed” assessments can elicit performances of higher-order tasks within short time frames (p. 300), meaningfully adding to the assessment picture for an instruction program based on standalone sessions.

METHOD

Participants

Study participants ($N = 138$) were graduate students in a Master of Education (M.Ed.) program that provided additional training in counseling, curriculum and instruction, dual language/bilingual education, educational diagnostics, educational leadership, exceptionality/special education, or reading. Regardless of their area of specialty, students completed two action research seminars that prepared them to analyze, plan, and conduct educational research. The study took place in the context of the first action research seminar, which involved a single, librarian-led, in-person group session on information literacy for educational

research. Students were informed on all written materials that their responses (if handed in) would be analyzed confidentially to improve library instruction; additionally, students in the experimental group were asked to give explicit consent for their work to be analyzed.

Measures

All participants completed three measures: pretest, posttest, and post-session survey (see Appendix A). The pretest was a two-part measure that rated students' self-assessed *familiarity* with library research processes (six items, using a 4-point rating scale with 1 = Not at all familiar; 4 = Very familiar) and their *knowledge* of specific resources (six items, using a multiple-choice format). This provided both subjective and objective information about students' information literacy baseline skills. A total score for familiarity was calculated for each participant, with possible scores ranging from 6 to 24, and for knowledge with possible scores ranging from 0 to 6. Reliability for each measure was calculated using Cronbach's alpha; the familiarity measure had acceptable internal consistency (Cronbach's $\alpha = .77$), while the knowledge measure's internal consistency was much lower than expected (Cronbach's $\alpha = .38$), suggesting that knowledge of these resources was not a unitary construct. Thus, for pretest knowledge, both the total score and the individual item scores were examined.

After receiving instruction about search strategies and the use of the ERIC thesaurus, students completed a posttest on those topics. This assessment included six multiple-choice items and a performance activity, both developed by the researchers

to assess key learning outcomes in a brief timespan. For the performance activity, participants read a research question scenario and a sample thesaurus entry; they were then asked to advise an imaginary peer on developing a search strategy with selected thesaurus terms. For the *searching exercise multiple choice* items, a total score was calculated, with possible scores ranging from 0 to 6. Cronbach's alpha was calculated as a measure of internal consistency; for this six-item posttest measure, reliability was again lower than desired ($\alpha = .37$). Therefore, in addition to looking at the six-item total score, student performance on individual items was also evaluated. For the performance-based *searching exercise activity*, the authors independently coded participants' open-ended responses according to a four-point rubric, with 4 indicating the highest level of proficiency with the thesaurus and 1 indicating the lowest level of proficiency (see Appendix B). Inter-rater agreement was lower than expected (Kappa = .42), though better than chance (60% agreement across four categories). Discrepancies among codes were discussed and resolved, resulting in a final set of codes used in the data analysis.

The final measure that all participants completed was a seven-item post-session survey that assessed participants' opinions about the relevance, value, and convenience of the library session. The post-session survey included three open-ended items about the aspects of the session that participants found most valuable, the topics that participants still had questions about, and any suggestions participants had for the librarian. Each item was examined independently.

In addition, treatment group participants completed a pre-class homework exercise (see Design and Procedure section below). The librarian scored successful completion of this pre-class homework exercise on a simple three-point scale with 3=successful completion, 2=partially successful completion and 1=unsuccessful completion.

Design and Procedure

A quasi-experimental (non-equivalent control group) design was used to determine whether exposure to a pre-session homework activity on the ERIC thesaurus would be associated with better performance on an in-session research activity. Participants were non-randomly assigned to one of two groups: the treatment group (which received the pre-session homework activity) and the control group (which did not receive the pre-session homework activity). Research instructors were invited to incorporate the pre-session homework activity into the lesson plan during the class prior to the library session. Those instructors who volunteered provided the pre-session homework activity handout, including instructions ([see Appendix C](#)), to their students, who constituted the treatment group; students whose instructors did not volunteer constituted the control group. All other aspects of the two groups' library instruction were identical, with the two library instructors having carefully coordinated lesson plans and presentations; however, differences between groups due to differences in instructors, location, etc., cannot be ruled out.

Participants attended one of five 90-minute library sessions as part of their class requirements; for most students, these sessions fell outside their regularly-

scheduled class times. Treatment group participants turned in their completed homework activity handout to their instructor, who submitted all handouts to the librarian conducting the session. Students in the session received a handout containing the pretest, posttest and post-session survey. Then the librarian administered the pretest. The fact that students in the treatment group had already completed the pre-session homework exercise before taking the pretest may seem a threat to the pretest's validity; however, the homework exercise was carefully designed so that students would learn to use the ERIC thesaurus nearly exclusively of other library skills. The researchers intended to thereby leave unaffected the general library familiarity and knowledge that the pretest measured in order to assess the groups' similarity.

Instruction focused on how to access, search and manage library resources in education. After hearing about the library's resources in education and how to find them, students viewed a demonstration of keyword searching. Then the librarian lectured on the purposes and structure of the ERIC thesaurus, including the elements of a thesaurus entry and the differences between searching with subject terms and with general keywords. Participants then completed both the *searching exercise multiple choice items* and the *searching exercise activity* in 10 minutes. In the last section of instruction, the librarian gave students methods for managing research information, including note-taking strategies and a demonstration of bibliographic management software. Participants then offered feedback through a post-session survey. As they handed in their written work, students received a handout with

descriptions and locations of key resources discussed during the session.

RESULTS

Data analysis involved three steps. First, means and standard deviations were calculated for each group (treatment and control) and the entire sample for these variables: pretest familiarity, pretest knowledge, posttest searching exercise multiple choice, and posttest searching exercise activity. In addition, frequencies were calculated for correct vs. incorrect responses to each pretest knowledge question, each posttest multiple choice question, each level of proficiency demonstrated in participants' responses to the posttest searching exercise activity, and degree of successful completion of the pre-session homework exercise (among

members of the treatment group only). Second, independent groups *t* tests compared the treatment and control group means on the main outcomes of interest (posttest searching exercise multiple choice and posttest searching exercise activity). Given the low reliability of the posttest multiple choice score, chi-square analyses were also performed on the correct vs. incorrect response frequencies for each item. Finally, participants' responses to the post-session survey were summarized.

A scoring of the pre-session homework exercise on a 3-point scale showed high rates of successful or partially successful completion among members of the treatment group, with 61% achieving full success and 20% achieving partial success.

Independent *t* tests revealed no difference between groups in terms of their pre-session

TABLE 1—ASSESSMENT MEANS AND STANDARD DEVIATIONS

Item	Treatment (N)	Control (N)	All (N)
Pretest Familiarity			
Mean	12.92 (42)	12.75 (93)	12.80 (135)
Stand. Dev.	3.89	3.65	3.56
Pretest Knowledge			
Mean	3.33 (43)	3.09 (92)	3.16 (135)
Stand. Dev.	1.08	1.48	1.37
Posttest Searching Exercise Multiple Choice			
Mean	3.83 (42)	2.93 (92)	3.22 (134)
Stand. Dev.	1.41	1.45	1.49
Posttest Searching Exercise Activity			
Mean	2.29 (31)	2.30 (77)	2.30 (108)
Stand. Dev.	1.10	0.78	0.88

familiarity, $t(133) = 0.26$, $p = .79$, or knowledge, $t(133) = 0.94$, $p = .34$ (see Table 1). However, given the low reliability of the knowledge measure ($\alpha = .38$), correct vs. incorrect responses to individual knowledge items were examined with chi-square analyses. The only significant pretest difference observed between the two groups was on the first knowledge question: “the list of subject terms is called...,” with more treatment group participants (93%) than control group participants (61%) responding correctly: $\chi^2(1) = 11.72$, $p < .001$. (Note that Yates’ correction for low cell size was used for this analysis.) Since the first knowledge question was the only one to deal directly with the content of the pre-session homework exercise and questions on other aspects of library use showed no significant difference, these findings were helpful for demonstrating that the treatment and control groups did not differ in their general familiarity with library and research procedures in ways that could influence their performance on the searching exercise measures.

Independent t tests revealed a significant difference between treatment and control groups in terms of their posttest searching exercise multiple choice scores, $t(132) = 3.35$, $p = .001$ (see Table 1). However, as noted above, the low internal consistency of this measure made it an unstable estimate of student performance. An analysis of individual posttest questions revealed the main source of this difference to be question 1, regarding keyword selection. Specifically, of six chi-square analyses comparing the percentages of correct responses, only the analysis of question 1 (about keyword selection) showed a significant difference between the treatment

group (86% correct) and the control group (54% correct): $\chi^2(1) = 13.04$, $p < .001$.

Posttest searching exercise activity responses showed no difference, $t(106) = -0.04$, $p = .96$ (see Table 1). Thus, the hypothesis that the treatment group would demonstrate better performance on a searching exercise was partially supported.

The results of the post-session survey indicated that participants generally perceived the session to be very relevant, that they felt moderately prepared to perform research, and that attending the session was moderately to very worthwhile. Database search skills and citation management software usage were frequently mentioned as valuable elements of the session. Common suggestions included handing out an outline during the session and being able to follow along on a computer.

CONCLUSIONS

Assessment results demonstrated that while students in the treatment group had a high rate of success completing the pre-session homework activity, that advantage did not translate to overall better performance on the in-class searching activity compared with students in the control group. The researchers can make reasoned guesses as to why this occurred. Perhaps the homework exercise’s mainly procedural activities did not lead to transferable understanding of thesaurus structure as called for in the in-class searching activity. Thus, students completing the pre-session homework exercise might have located a narrower or related term in a thesaurus entry, for example, without understanding what those

structural elements meant in context. Acting on this hypothesis, the researchers could develop more robust explanations of thesaurus structure for the pre-session homework exercise, perhaps including taxonomic diagrams or sample thesaurus entries with readily-understood terms from everyday life, to better build on students' previous learning and lead to greater gains in understanding.

It is difficult to make a case for improved student skills in keyword selection based on responses to a single multiple choice question in the posttest. Nevertheless, the researchers could consider what features of the keyword selection part of the homework exercise might have allowed students to build on their prior knowledge to develop a transferable skill, and how such features could be used in other parts of the exercise.

Students' low performance in the posttest assessment across both treatment and control groups might indicate that this assessment should be revised to better measure and reinforce student learning. The researchers could develop more multiple choice questions and gather feedback on the questions' clarity and perceived difficulty. The searching activity could be revised in light of demonstrated student difficulties such as not knowing the meanings of terms in the sample thesaurus entry or not being willing to engage with the assigned "email to a friend" genre. Such a revised posttest assessment could lead to even more targeted improvements of the pre-session homework exercise.

DISCUSSION

The insights gained through this study in

such a limited assessment time frame point to the potential of quasi-experimental approaches in evaluating pre-session library instruction tools. The use of a control group made it possible for the researchers to evaluate the pre-session homework exercise independently of the impact of classroom instruction, which was identical for both groups. One application of this study's approach would be to help develop the assessment of "flipped" information literacy classrooms, which assign homework before a session in order to focus class time on active learning experiences. Researchers assessing student learning in flipped classrooms have compared the outcomes of students in classes that flip with students in classes that do not (Arnold-Garza, 2014, p.19). While these results may point to the benefit of active learning in the classroom, they cannot indicate which elements of the flipped learning experience (including pre-session instructional videos, for example) had the most benefit. By iteratively building on the use of quasi-experimental methods to evaluate pre-session exercises, a flipped classroom project could better its position to increase student learning and impact academic culture.

REFERENCES

- Arnold-Garza, S. (2014). The flipped classroom teaching model and its use for information literacy instruction. *Communications in Information Literacy*, 8 (1), 7–22.
- Barclay, D. (1993). Evaluating library instruction: Doing the best you can with what you have. *RQ*, 33(2), 195–198.
- Bluemle, S. R., Makula, A. Y., & Rogal, M.

- W. (2013). Learning by doing: Performance assessment of information literacy across the first-year curriculum. *College & Undergraduate Libraries*, 20(3-4), 298–313. [doi:10.1080/10691316.2013.829368](https://doi.org/10.1080/10691316.2013.829368)
- Brooks, A. (2013). Maximizing one-shot impact: Using pre-test responses in the information literacy classroom. *The Southeastern Librarian*, 61(1), 41-43. Retrieved from <http://digitalcommons.kennesaw.edu/seln/vol61/iss1/6>
- Bryan, J. E., & Karshmer, E. (2013). Assessment in the one-shot session: Using pre- and post-tests to measure innovative instructional strategies among first-year students. *College & Research Libraries*, 74(6), 574–586. [doi:10.5860/crl12-369](https://doi.org/10.5860/crl12-369)
- Campbell, D. T., & Stanley, J.C. (1966). *Experimental and quasi-experimental designs for research*. Chicago: R. McNally.
- Carter, E. W. (2002). “Doing the best you can with what you have”: Lessons learned from outcomes assessment. *Journal of Academic Librarianship*, 28(1), 36–41.
- Choinski, E., & Emanuel, M. (2006). The one-minute paper and the one-hour class: Outcomes assessment for one-shot library instruction. *Reference Services Review*, 34(1), 148–155. [doi:10.1108/00907320610648824](https://doi.org/10.1108/00907320610648824)
- Hsieh, M. L., & Holden, H. A. (2010). The effectiveness of a university’s single-session information literacy instruction. *Reference Services Review*, 38(3), 458–473. [doi:10.1108/00907321011070937](https://doi.org/10.1108/00907321011070937)
- Hufford, J. R. (2010). What are they learning? Pre- and post-assessment surveys for LIBR 1100, Introduction to Library Research. *College & Research Libraries*, 71(2), 139–158. [doi:10.5860/0710139](https://doi.org/10.5860/0710139)
- Johnson, C. M., Anelli, C. M., Galbraith, B. J., & Green, K. A. (2011). Information literacy instruction and assessment in an honors college science fundamentals course. *College & Research Libraries*, 72(6), 533–547. [doi:10.5860/crl-166](https://doi.org/10.5860/crl-166)
- Oakleaf, M. (2008). Dangers and opportunities: A conceptual map of information literacy assessment approaches. *portal: Libraries and the Academy*, 8(3), 233–253. [doi:10.1353/pla.0.0011](https://doi.org/10.1353/pla.0.0011)
- Pierce, R., & Fox, J. (2012). Vodcasts and active-learning exercises in a “flipped classroom” model of a renal pharmacotherapy module. *American Journal of Pharmaceutical Education*, 76(10), 1-10. [doi:10.5688/ajpe7610196](https://doi.org/10.5688/ajpe7610196)
- Portmann, C. A., & Roush, A. J. (2004). Assessing the effects of library instruction. *Journal of Academic Librarianship*, 30(6), 461–465. [doi:10.1016/j.acalib.2004.07.004](https://doi.org/10.1016/j.acalib.2004.07.004)
- Swoger, B. J. M. (2011). Closing the assessment loop using pre- and post-assessment. *Reference Services Review*, 39(2), 244–259. [doi:10.1108/00907321111135475](https://doi.org/10.1108/00907321111135475)
- Wiggins, G. P., & McTighe, J. (2005). *Understanding by design*. Alexandria, VA: Association for Supervision and Curriculum Development.