Calibrated Peer Review: A New Tool for Integrating Information Literacy Skills in Writing-Intensive Large Classroom Settings.

Michael Fosmire

Purdue University, fosmire@purdue.edu

Follow this and additional works at: http://docs.lib.purdue.edu/lib_fsdocs

Recommended Citation

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.
Calibrated Peer Review: A New Tool for Integrating Information Literacy Skills in Writing-Intensive Large Classroom Settings

Michael Fosmire

portal: Libraries and the Academy, Volume 10, Number 2, April 2010, pp. 147-163 (Article)

Published by The Johns Hopkins University Press
DOI: 10.1353/pla.0.0099

For additional information about this article
http://muse.jhu.edu/journals/pla/summary/v010/10.2.fosmire.html
Calibrated Peer Review: A New Tool for Integrating Information Literacy Skills in Writing-Intensive Large Classroom Settings

Michael Fosmire

abstract: Calibrated Peer Review™ (CPR) is a program that can significantly enhance the ability to integrate intensive information literacy exercises into large classroom settings. CPR is founded on a solid pedagogic base for learning, and it is formulated in such a way that information skills can easily be inserted. However, there is no mention of its application for information literacy in the library literature. A sample implementation of CPR in a course co-taught by science disciplinary faculty and librarians at Purdue University is presented with recommendations for optimal use of this resource. CPR is a valuable new tool that librarians can use to further their collaborations with disciplinary faculty.

Introduction

As librarians seek to increase the number of active-learning exercises in their instruction, writing assignments that incorporate peer review can be a valuable addition. Calibrated Peer Review™ (CPR) is a program that can significantly enhance the ability to integrate intensive information literacy exercises into large classroom settings. CPR is founded on a solid pedagogical base for learning, and it is formulated in such a way that information skills can easily be inserted. Despite the inherent possibilities of CPR, the author has been unable to find any mention of Calibrated Peer Review™ in the library literature. This paper, which describes an effort at Purdue University, will hopefully raise awareness of the technique in the library community.
Calibrated Peer Review: A New Tool for Integrating Information Literacy

Nancy Falchikov and Keith Topping have documented the benefits of peer review as a learning technique. Well-crafted writing assignments probe all levels of Benjamin Bloom’s taxonomy of educational objectives, from simple knowledge and comprehension, to the synthesis of ideas, and, ultimately, to the highest level of the taxonomy—evaluation. The peer-review process appears almost exclusively in the evaluation level of Bloom’s taxonomy, as students critique, recommend, and evaluate each other’s (and their own) work. However, in large classroom settings, a substantial barrier to peer review is the sheer magnitude of the effort to collect and distribute writing samples, in addition to the challenge of quality control for grading. With those issues in mind, Orville Chapman, a chemistry professor at UCLA, developed the program Calibrated Peer Review™ (cpr.molsci.ucla.edu). He wanted to scale writing to a large classroom setting and introduce the scientific process of peer review to students. The CPR software program automates the entire process of submitting, distributing, and compiling grades for an assignment. Instructors at over 700 institutions, including faculty in statistics, biology, and communications at Purdue University, have adopted CPR.

The process is described more fully by Orville Chapman and Michael Fiore and Ralph Robinson, so only a brief outline is given here. In essence, CPR consists of the following steps:

- Students are given a writing assignment, often based on a reading selected by the instructor.
- Students compose and submit their essay by a certain deadline to the CPR software server.
- The CPR system then provides students with three instructor-created “calibration” essays to grade according to a provided rubric.
- After “passing” the calibration essays, to ensure they understand the grading criteria sufficiently, students receive three of their peers’ essays to grade against the same rubric.
- Students then evaluate their own essays, and those scores are compared to a weighted average of peer evaluations to determine if the students accurately evaluated their own work.
- Instructors may review any essays and change scores when students feel they were unfairly graded or when the instructor notices potentially anomalous grades.

CPR: How It Works

For instructors, CPR provides a fairly straightforward and robust way to design and grade assignments. Certainly, more time goes into the conceptual construction of an as-
ignment than in the mechanics of implementing it in the software. Everything is done in a Web interface that allows cutting and pasting of text, and the software program leads one sequentially through the steps of assigning a writing prompt, creating guiding questions, developing a grading scale for the assignment, and setting deadlines for each step of the CPR assignment. Instructors can save their progress at any time, and the assignments can be shared with other instructors through an assignment library. The system has some default point distributions for the reviews and thresholds for success, but instructors can easily modify them to meet the needs of the assignment. Instructors allocate a percentage of points in four areas: (1) the main essay itself (the grade for that is determined by the peer evaluation scores given by other students), (2) evaluation of the calibration essays (the only standard grading component because the instructors create the grading key for those essays), (3) evaluation of other students’ essays (the grade is determined by how well the student’s evaluation matches the grades given by other peer evaluators), and (4) self-evaluation of their own essay (the grade is determined by how well self-evaluation scores match grades assigned by peer evaluators).

Thus, if the instructor wanted to focus on the content of a student’s essay, more points would be allocated to the first area; whereas, if the focus is to be on evaluation skills, more points would be allocated to the latter categories. By requiring three peer reviews for each essay, the CPR system can accurately automate the grading of student papers. With two peer reviews, the computer would not be able to determine which review was more accurate in the case of substantially different grades. However, with three grades, it is more likely that two of the scores would be close together, and the computer can thus determine which score is the outlier. The CPR software also takes into account how well student reviewers performed on the calibration essays and gives more weight to the evaluations of high-performing reviewers.

There are two ways instructors can determine proficiency in the calibration phase of the CPR process—by the overall score for an essay or by correct answers to specific questions about the essay (such as components of the grading rubric). For example, the instructor might require a student’s overall grade for the calibration essay to be within one point of the instructor’s assigned value in order for the student to receive credit or that a student must answer at least 75 percent of the individual questions in the rubric correctly, or both. Furthermore, one can also designate a question in the rubric as “content” or “style,” and separate thresholds for each can be imposed on the calibrations (for example, 75 percent of the content questions must be answered correctly and 50 percent of the style questions). For the peer and self-evaluations, one can also set two thresholds of success, one level for full credit and another for half credit. In this case study, the author’s assignments, and indeed the course itself, focused on content and critical thinking, so stylistic concerns were of secondary importance, and the style category was not used.

In terms of monitoring grades, the CPR system provides a spreadsheet of student performance, including a real-time indication of how many steps (submission of essay, calibration, peer evaluation, and self-evaluation) have been completed by the student. One can click on a student’s record to see how they performed in each section and, from there, access their actual submissions and evaluations. The instructors can easily review work and reassign grades as appropriate, either in the final score or within each section. Changes to the grades may then change grades for peer and self-evaluations,
which the CPR program automatically generates and updates. Both the instructors and the students can see grade changes, and the instructor can leave notes for the students indicating why a grade was changed. The CPR program allows students to see their own performance on each section of the process as well, including the scores and comments provided by their anonymous peer evaluators. The CPR program also has some internal mechanisms to facilitate the grading process for instructors, such as flagging essays that might have been incorrectly graded (for example, where peer evaluations differed substantially). Overall, although the interface can be a little unsophisticated in places (most notably, it takes several clicks to access student comments on the evaluations), it is quite robust in terms of streamlining the workflow of facilitating writing assignments in large classroom settings and organizing the results of student work.

CPR has been evaluated and tested, mainly in the science education community, and found to be a valuable tool. For example, Arlene Russell found a 10 percent increase in mid-term test scores in a course using the CPR method compared to lectures and textbooks alone.5 R. Dean Gerdeman, Arlene Russell, and Kelly Worden found scores increased most for low-performing writers, and self- and peer-assessment deviations decreased over the course of a semester.6 Susan Plutsky and Barbara Wilson, in a non-science class, found no significant difference between writing assignments graded by faculty, group review, or CPR, indicating that using this process did not have a negative effect compared to the more labor-intensive method of instructor-graded writing assignments.7 Keith Topping and Nancy Falchikov and Judy Goldfinch also found that peer review was no worse than faculty review of student work, and Lawrence Margerum et al. and Teresita McCarty et al. found positive results for students using CPR.8 On the other hand, Julie Reynolds and Cary Moskovitz analyzed the content of an assignment library housed on the CPR main server and concluded that many assignments do not probe higher level skills but rather are too concrete and fact based.9 Mark Walvoord et al. observed other difficulties with the system and recommended changes.10

Russell summarizes the concept behind CPR, “CPR empowers students to write to learn rather than learn to write. …When students write, they are required to organize their thoughts, make decisions about what is relevant, convey their thoughts, and arrive at conclusions. This means students are active in constructing their understanding of the material.”11 Indeed, this constructivist basis for CPR is its greatest selling point. Since CPR assignments almost always contain readings, Russell’s assessment comes tantalizingly close to describing information literacy skills. Other than the “searching” component, Russell is describing the core information literacy competencies of organization, evaluation, utilization, and integration of information, as detailed in the next section.
Information Literacy Tie-Ins of CPR

Before discussing the information literacy skills related to CPR, a brief review of those skills is in order. The Association of College and Research Libraries’ (ACRL) information literacy competency standards for higher education define six major information competencies. According to the standards, students need to be able to:

- Determine the extent of information needed
- Access the needed information effectively and efficiently
- Evaluate information and its sources critically
- Incorporate selected information into one’s knowledge base
- Use information effectively to accomplish a specific purpose
- Understand the economic, legal, and social issues surrounding the use of information and access and use information ethically and legally

In addition, the ACRL Science and Technology Section’s (STS) version of information literacy standards also maintains that a student “understands that information literacy is an ongoing process and an important component of lifelong learning and recognizes the need to keep current regarding new developments in his or her field,” which the author considers an unofficial “7th” standard. Indeed, this last standard also resonates with the ABET accreditation criterion (3.i), which requires that undergraduate engineering students develop “a recognition of the need for, and an ability to engage in, life-long learning.”

It is not difficult to see how these standards can be applied to a typical CPR assignment. For example, when given parameters of an essay assignment, students need to determine what information they need to glean from that reading (standard 1). Although traditionally CPR assignments only ask students to analyze a provided reading, assignments can also require students to gather and incorporate external sources of information into their essays (standard 2). Students must evaluate their readings to determine their relevancy (standard 3) and extract and synthesize the information from those articles (standard 4) to accomplish their task of writing an essay addressing their guiding questions (standard 5). By requiring students to appropriately cite or refer to their sources, standard 6 can easily be probed; and “standard 7” can be reinforced by using current readings for the assignments, supplementing textbook content. Laura Pence and Harry Pence, for example, introduced students to using RSS feeds to identify articles to read for their assignments.

Several authors have discussed the need for discipline-specific information literacy instruction programs, and integrating writing exercises into information literacy instruction has received substantial interest from the science library community. Deborah Huerta and Victoria McMillan describe a course co-taught by a librarian and a biologist that requires students to write review papers and popular science articles, incorporating peer and self-review processes; although the instructors still “each read every assignment and grade them together.” Brian Winterman describes a course co-developed with a biology faculty member whose focus was the creation of a research proposal. Winterman found significant gains in confidence by students in reading scientific journals, expressing scientific ideas in writing, and forming questions for scientific research.
describes a quite extensive semester-long process for constructing a literature review of a topic for an advanced course on cell biology, which he believes leads to superior student projects. Margie Krest and Daria Carle developed a course, Introduction to Scientific Writing, which addressed their concern that “when students do learn to write, research, and think critically, they often learn them in isolation and rarely learn how to integrate them.”

When size is mentioned, the authors in this literature review reported class sizes of 30 or fewer students. Porter, for example, states that “the entire group of assignments as described here is probably most appropriate for relatively small, upper-level courses.” Thus, developing scientific writing and research skills in large classroom settings is a challenge and one that Calibrated Peer Review™ can address. At a university of the size of Purdue, with approximately 3,000 science undergraduates, scalability of instruction is a significant concern.

CPR Implementation at Purdue

In 2007, the Purdue University College of Science implemented a new core curriculum for its students, which incorporates not only disciplinary competencies but also “professional skills,” such as effective communication and the ability to work in teams. One of the new requirements was that students take a “great issues” course that “addresses the impact of science on society and the ramifications of scientific advances.” The author participated in a previous version of a science and society course (not required by the college) and saw an opportunity to rekindle collaboration and to integrate information literacy skills throughout the class. The end result was that the author, Professor Andrew Hirsch from physics, and Jane Yatcilla, the mathematical sciences librarian, co-taught the only college-sponsored course to date, SCI 490: Great Issues in Science and Society. The course was targeted to upper-level students because they have sufficient subject background to apply their disciplinary knowledge to the course.

The first implementation of the course simulated the workings of a think tank. The 56 students were broken up into teams of three or four students and assigned a semester project. The project was to develop a white paper and an accompanying short video proposing a policy that analyzed both scientific and social aspects of a potential solution to or amelioration of an important challenge facing society, such as energy, water, population, or the environment. Since great issues, by definition, are topics that are of current interest in the public discourse as well as in scientific research, there is a strong case for the need for information literacy skills. Textbooks cannot keep up with discoveries and the latest public perceptions of energy or the environment, for example. Textbooks provide some core background to the issues, so students have a basis from which to judge new publications; but they are insufficient for students to remain on top of these rapidly evolving fields.

As such, the instructors incorporated recent publications throughout the class assignments and asked students to demonstrate proficiency in information literacy skills. Calibrated Peer Review™ was seen as a potentially valuable tool for probing and evaluating these skills. The instructors created Calibrated Peer Review™ assignments for each of the great issues explored in the class, so students had to read, understand,
extract, and synthesize information from readings in those areas. The instructors also incorporated non-content objectives for the Calibrated Peer Review™ assignments. These included the ability to critically evaluate the content of papers by asking students when a model or worldview applies and when it does not; comparing information between two sources and resolving differences in the conclusions; citing information appropriately; generating research questions based on knowledge gained from a source (and information missing from that source); understanding the role of data and simulations in the information gathering/knowledge creation process; developing tools to start to evaluate validity of simulation data; learning to read a scientific paper effectively and efficiently; and proposing the use of experiments or gathering of data as ways to help answer research questions.

Thus, in the implementation for this course, the Calibrated Peer Review™ assignments probed all the information literacy standards except the second one—effectively and efficiently locating information. That skill was evaluated and reinforced in the course in an information portal assignment\textsuperscript{24} and in the students’ final projects. The Calibrated Peer Review™ assignments served to introduce students to several different kinds of information as a way of helping the students “learn by doing”—understanding the purpose, value, and limitations of different types of information so that they could determine what kinds of information were appropriate for their final projects.

Since the Calibrated Peer Review™ process has several steps, the instructors limited themselves to five assignments during the semester. This allowed students a week to read the background information and compose their essays and a week to complete the review portion of the assignment. Colleagues who had used CPR had advised the instructors that more than five assignments would be too overwhelming for students.

**Sample Assignment**

To illustrate how CPR works in practice, what follows is an assignment used in the course. This assignment asked students to compare the information presented in two non-technical reviews of the current and future state of American and global energy. Students read “Electricity Without Carbon” and “What You Need to Know About Energy.”\textsuperscript{25} These articles had different scopes and purposes. One dealt exclusively with carbon neutral sources of energy from a global perspective. The other article analyzed the current state of energy in the United States, including for example fossil fuel usage, and made predictions for likely future U.S. energy needs and adoption rates of new energy sources.

The instructors communicated to the students the following goals for the assignment:

- Determine the scope and purpose of a review article
- Extract and synthesize relevant information from multiple sources
- Generate new research questions based on initial readings

The instructors then provided guiding statements to shape the essay:

1. This essay should summarize the main purpose and scope of each article, noting the differences between the articles.
2. Also, pick one energy source where the two articles differ in their assessment for the future prospects of the source. Describe, quoting appropriate sections of the text or referencing specific pieces of information, the fundamental differences between the two articles’ assessment.

3. What questions about the energy source were raised by your readings? How would you go about finding information to answer those questions?

The formal writing prompt states:

Write an essay of 300–500 words that summarizes and compares the two aforementioned review articles about the current and future state of energy in the United States and the world. Included in the essay should be an analysis of an area in which the two articles reach different conclusions about the prospects of one energy source, with an analysis of (1) why they reach different conclusions, and (2) what questions for further research would help resolve those different conclusions.

For the review portion of the assignment, students used the following grading scale:

1. Does the essay correctly describe the scope of the “Energy Without Carbon” article? It should include, for example, that the article looks at energy from a global standpoint, only looks at carbon-neutral energy sources, assigns potential capacity of each source, looks at environmental and economic issues of the source. (Essays can be given 0–2 points, depending on how well they describe the article.)

2. Does the essay correctly describe the scope of the “What You Need to Know about Energy” article? It should indicate that, for example, it looks primarily at energy from a U.S. standpoint, analyzes not only energy sources but also transmission and conservation of energy as part of the overall energy discussion, provides an overview of all major energy sources, their supply and the demand for them, not just carbon-neutral sources as in the prior article. (Again, 0–2 points, depending on how well the essays describe the scope)

3. Does the essay identify one energy source to compare between the articles?

4. Does the essay use quotes or specific data from the articles to justify the differences between the articles?

5. Are the quotes or specific data relevant to the discussion?

6. Does the essay include at least one question that the author thinks is unanswered by the article or for which there is conflicting information presented by the articles?

7. Does the essay suggest where or how to find the missing information? Does it suggest specific papers mentioned in the articles, or types of resources that might have that kind of information (for example, government information, scientific papers, industry reports).

8. Does the essay contain correct spelling and grammar?

9. How would you rate this text? 1–10. (Rating should bear some correlation to the answers of the previous questions.)

By using the scale, the main areas of emphasis of the assignment were reinforced for the students. Each time they went through the review process, they engaged again
the question of whether data were quoted in the essay and if that data were relevant to the argument. The students evaluated calibration essays, so instructors could easily tell whether a student had correctly used the scale. Each student essay was reviewed by three peers, so the instructor could see which students did not meet the requirements of the essay and, thus, might need some remedial help with the concept; and the instructor could see whether reviewers disagreed with each other on the rating for a student’s essay. This would enable the instructor to go in and rectify the disagreement. Finally, the student’s self-evaluation showed the instructor whether he or she had internalized the lifelong and self-directed learning skill of self-correction.\(^\text{26}\) One of the fundamental skills for self-directed learning is for students to know whether they know something (as opposed to “thinking” that they know something) and to be able to accurately review their own work, which is not a trivial process. The instructors found that many students could accurately determine other’s level of success on an assignment but had more difficulty evaluating their own work objectively.

**Results**

Overall, students were able to use the Calibrated Peer Review™ system without too much technical difficulty. One of the instructors, Jane Yatcilla, created a non-threatening first Calibrated Peer Review™ assignment, asking students to take a brief personality test and read a little bit about their personality type and how it might affect their interactions on teams. Students were asked in the assignment to report their type, the strengths and weaknesses of the type, and, in retrospect, how they felt their personality type may have influenced their work in groups in the past. Thus, this assignment, with very little chance for “wrong answers,” allowed students to focus on and master the Calibrated Peer Review™ tool itself. The instructors found, to their surprise, that about half the students had already experienced CPR in a previous communications class.

The instructors asked students to write weekly reflections about their experiences in the class, how the experience reinforced or changed their perspectives on an issue, or if they wanted to share personal experiences they had related to course topics. This encouraged students to integrate information they received into their existing knowledge bases (information literacy standard 4), and we saw ample evidence in their reflections that students considered what they discussed in class, what they read for the CPR, and did in fact draw connections between those topics and their previous experiences and observations. Since this was the first offering of the course, we also encouraged input on how to improve the class as well.

In this way, the instructors gathered significant input from the students about how they regarded Calibrated Peer Review™. In general, students who had prior experience had a negative attitude about CPR, but several student comments indicated that they appreciated our implementation. There may be a few reasons for this. First, many
students expressed displeasure with having their peers evaluate their work instead of the instructor, which is common for peer review processes in general. They felt their peers were not qualified to rate their papers. The instructors remedied this by stressing that any student could ask the instructor to re-grade an assignment, and the instructors also monitored anomalous looking grades and proactively re-graded papers that might have been poorly evaluated. Consequently, only a handful of students asked that their assignments be re-graded. In light of the preexisting negative attitude of students, the instructors were pleasantly surprised that the students’ peer reviews were generally more accurate than their self-evaluations. The instructors also emphasized that only a minority of the grade was based on the peer’s evaluation of a paper. If a student did well in his or her reviews, he or she would do well on the assignment. This was borne out in the final grades, as typical average scores on the essay portion of the assignments varied between 7.5 to 8.5 out of 10 points, while the average overall grade for the assignments hovered around 90 percent. Students also seemed to engage more with the subject matter of these CPR assignments. We received many comments that the readings chosen for the CPR assignments were quite interesting and relevant to the class and their major. We believe that framing the context of these assignments as having students “write to learn” rather than “learn to write” helped motivate them to accept using CPR.

Some students did complain that they spent too much effort reviewing with CPR. The student has to review three calibration essays, three peer essays, and then evaluate their own essay. A couple of students reported “review fatigue” by the end of an assignment. This potentially can lead to a degradation of the quality of reviewing as students rush through the grading of “non-calibrated” essays. The instructors sympathized with this situation, especially when it came time to grade final projects. However, we feel there is pedagogic value in the peer and self-review process. Perhaps the easiest way to decrease the workload for students would be to decrease the number of calibration reviews. Having fewer than three peer reviews makes it much easier for inaccurate peer review grades to be assigned to an essay, so it would be difficult to reduce the number of those a student must do. The base CPR software does not allow for such modifications, but assignments can “fool” the system to reduce the number of student reviews.

Finally, the instructors experimented with the best format for the guiding questions to attach to the assignments. Some assignments emphasized open-ended questions, and other assignments emphasized more specific questions. We received about the same number of comments from students who preferred the open-ended questions and from those who preferred the more specific questions. Students who favored open-ended questions felt motivated by the ability to be creative in writing their essays. Those who favored more specific questions either felt they were easier to grade or were more comfortable finding specific answers rather than tackling less well-defined questions.
In terms of student performance, the authors analyzed the assignment reported above (the third assignment of the five given in the course). We focused on the three information literacy questions in this assignment,

1. Does the essay use quotes or specific data from the articles to justify the differences between the articles?
2. Are the quotes or specific data relevant to the discussion?
3. Does the essay suggest where or how to find the missing information? Does it suggest specific papers mentioned in the articles or types of resources that might have that kind of information (for example, government information, scientific papers, industry reports)?

Results for this assignment are summarized in table 1 and are illustrative of the kinds of issues that arise in the CPR process. In general, students have a tendency to overestimate their achievement, as witnessed by the difference between self- and peer evaluation of success, in which fully 10–14 percent of students thought they had answered the “data” questions correctly when they had not. There were a few students who did not think they had answered those questions (false negatives) that actually had done so; but, as shown in the table, 20 percent of students thought that their quotations and data were relevant when they actually were not (and instructor review of papers confirmed the results).

Also, the instructors were concerned that more than half of the students did not or were not able to make a suggestion for finding more information to answer questions raised by the articles. The threshold for success for this question was pretty low. We did not ask students to actually find information, just to identify a likely source of relevant information. The texts students read for the assignment even included references to papers that contained source material for some of the conclusions reported, but almost no students identified those as potential sources of information. Some students indicated in their essays that they did not have any questions and, thus, did not address the need for more information. Other students just did not know how to answer the question. This reinforced the instructors’ impression that the need to teach information literacy skills is of fundamental importance and that students’ skill levels are not as high as we sometimes like to believe.

We did not give the students any preparation for finding external information prior to this assignment, treating this as a diagnostic question of the readiness of students for their research projects. Since this assignment demonstrated students’ inability to articulate where to find information, they were more receptive to learning about this concept. As has been noted by Kate Manuel, students’ self-confidence with using information technologies can lead to overconfidence in their ability to find, evaluate, and use information effectively. The instructors treated the results as a teachable moment not only for introducing the searching process but also to illustrate that any paper one reads should raise questions for further consideration. Since a scientific paper only probes a certain, well-defined question, it necessarily is not the end of the discussion but rather just the beginning. This issue resurfaced in the final CPR assignment, in which students were asked to identify a research question unanswered by the scientific paper they read and to formulate an experiment or method of collecting data that could help answer that
research question. Almost all students proposed experiments, including several who suggested that the data might already exist and where one might find it.

The instructors were encouraged that, especially with an initial negative attitude toward CPR, in a final survey of the students, comments were evenly divided positively and negatively; and only one or two reviews were completely negative. Most indicated that there were strengths and weaknesses with the system. Although this was not an overwhelming endorsement of the tool from the students, the instructors were encouraged by the level of support from students in our first attempt at using CPR. Selected, illustrative comments from the final survey have been compiled in the appendix.

As mentioned previously, the instructors received many comments from students in their weekly reflections regarding their preference for open-ended versus more specific guiding questions. Several students also commented that the CPR system was a good vehicle for learning content but less effective at improving writing skills (mainly because students did not believe their peers were providing good feedback). Another recurring issue for the students was with the grading process. Mainly, students thought that the grading scales appeared to expect specific answers to open-ended questions. A few instances were the error of the instructor, but many grading guidelines indicated that answers “could include” some specific information, not that they had to include it. Sometimes student interpretation of the guidelines led to confusion. Students also came to realize that, even though some questions are open ended, there are better and worse answers; and successful essays need to address some key concepts or facts.

**Recommendations**

Overall, the Calibrated Peer Review™ tool shows substantial promise in support of integrating information literacy competencies in a writing intensive, large-classroom environment. Instructors can easily probe information literacy skills with this tool; and the automated, student-driven process for grading assignments makes it manageable to incorporate writing extensively in larger courses. The instructor still must devote time to quality control of the grading, but it is a much smaller investment for the instructor than grading all the papers in their entirety. The instructors believe that teaching students the peer and self-reviewing process and providing them with grading guidelines also give them more tools with which to evaluate their own learning, although we did not probe that research question specifically.

The calibration essays that the instructor prepares, as well as the grading scales, also provide opportunities to teach students skills by modeling proper (and improper) behavior. For example, in the assignment that asks students to find other sources of information, the calibration essays contained specific examples of appropriate sources of information. Such “stealth teaching” reinforces the concepts the instructors want to get across—namely, evaluating whether sources are valid in the context of a concrete application. This is more effective than merely lecturing in the abstract about information sources.

As noted above, instruction can be embedded in the assignments themselves so that librarians can collaborate on the construction of the assignment in lieu of providing in-person instruction. This can assist with load balancing of effort because assignments can
be created in non-peak times of the semester. The assignments can also be coupled with librarian-led instruction in order to reinforce the importance of the librarian’s presentations with a follow-up, concrete assignment in line with best practices for instruction.

Although the CPR assignments in the pilot offering of the Great Issues in Science and Society course did not ask students to find their own information resources, the next iteration of the class will require students to incorporate additional resources to supplement those given in the assignments. The instructors anticipate that this reinforcement for finding appropriate information will improve student outcomes in that area. We anticipate that it will make peer review of essays more interesting for students as well because the essays will incorporate different sources of information, which will yield greater variety in the end products.

Conclusion

Calibrated Peer Review™ is not a panacea for all things concerning information literacy, however. As with any teaching tool, just putting something in a CPR format does not make it a good assignment. Instructors need to create clear, focused, and, if possible, student-tested assignments in order to make sure students and instructors have a common understanding of the assignment expectations. For this implementation, the instructors reviewed each other’s assignments. Additionally, some student library workers, as a convenient sample, provided input on the clarity of the assignments.

At this stage of development, Calibrated Peer Review™ remains largely unexamined by the library community. Thus, many opportunities exist to more rigorously evaluate the effectiveness of the technique compared to other methods of teaching information literacy competencies in different instructional settings. As indicated in this paper, prob-

<table>
<thead>
<tr>
<th>Table 1. Student performance on information literacy questions.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>N=50 (completed assignments)</strong></td>
</tr>
<tr>
<td>---------------------------------</td>
</tr>
<tr>
<td>Data used to justify differences</td>
</tr>
<tr>
<td>Data used were relevant</td>
</tr>
<tr>
<td>Made suggestions for finding more information</td>
</tr>
</tbody>
</table>

Each column indicates whether the student (Self) or peer reviewer (Peer) believed that the student answered that question correctly. False positives indicate the number of times a student’s self-evaluation was “Yes” and their peer evaluation was “No.” A false negative was a situation when the self-evaluation was “No” and the peer evaluation was “Yes.”
ing and developing many of the “non-searching” information literacy competencies—such as topic or problem articulation, evaluation and extraction of information from resources, and integration of that information to solve a problem—came most naturally to the author as uses for CPR. Whether this technique is the most appropriate or the best method for teaching or assessing search skills, however, is certainly an open question and one that would benefit from study and the creativity of the library community in devising assignments.

Other librarians should consider adding CPR to their toolkit of instructional techniques. The level and depth of interaction and engagement with information literacy competencies that can be facilitated are worthwhile investments of time. The CPR Web site lists institutional contacts in order to assist in identifying a local community of users. For institutions not yet using Calibrated Peer Review™, it is possible to investigate with the information technology professionals how to implement CPR locally to support innovation in instruction. The CPR system can be run from the servers at UCLA; or, as Purdue has just done this year, it can be hosted at the local institution level. As Alexius Smith Macklin and Michael Fosmire noted, offering a new technology or pedagogic approach to teaching faculty can be an effective catalyst to develop good collaborative relationships and a springboard for in-depth integration of information literacy into their courses.

Acknowledgements

The author wishes to thank Professors Andrew Hirsch and Jane Kinkus for their collaboration on the pilot implementation of the SCI 490 class. Also, conversations with Ellen Gundlach and Nancy Pelaez about their use of Calibrated Peer Review™ in statistics and biology classes, respectively, allowed us to make many fewer mistakes than we otherwise would have in our first attempt at using CPR. The author also wishes to thank the anonymous peer reviewer for many substantive and thoughtful comments that substantially improved this paper.

Michael Fosmire is head of the Physical Sciences, Engineering and Technology Division, Purdue University Libraries, Purdue University, West Lafayette, IN; he may be contacted via e-mail at: fosmire@purdue.edu.

Appendix

Student Comments on Calibrated Peer Review Assignments

“For next semester’s class, I wouldn’t change much concerning CPR.”

“About the CPR, I think it’s a good thing. Like every other system it might have a few drawbacks, but I think it should be used in more classes.”

“I think the CPR system does a good job of forcing students to actually read the articles. Even if a person just goes through the list of prompts, they still have to read the article.”
“The CPR assignments were pretty informative and were very pertinent to our in class discussions and topics. They all seemed to be fair, although the last one [reading a scientific paper] was a little more challenging.”

“CPR is a good tool because it can help us know what to write and focus in our papers. Also, it can help improve my writing skills.”

“I felt like the information we were reviewing for the CPR’s were very interesting, and I enjoyed reading most of the articles for them...[however], I would have enjoyed doing a CPR on the essays we were reading for our guest lectures. I felt that would have tied the course in more closely with the activities we were doing. ...I feel the CPR could be a very integrated part of the next classes of this course.”

“I think that these assignments should stay in the curriculum for this course. I truly did learn a great deal about great issues during these assignments.”

“The CPR assignments...were kind of a pain. I don’t feel I learned a lot from them, and it was difficult to grade them when we’re not the ones giving the assignment.”

“Open ended questions are fine for writing the essays, as long as the rubric is more open as well. Problems arise when there is an open-ended question,...but the rubric asks for specific answers. As long as the material is relevant to what is going on in class, the CPR’s are a good tool.”

“I found the whole CPR process to be tedious and somewhat annoying some weeks, and then other weeks I thought that it was a good learning experience that was well within my academic ability to complete.”

Notes


27. Walvoord et al.


29. Reynolds and Moscovitz.