Paper ID #19321

Computer-Mediated Peer Review: A Comparison of Calibrated Peer Review and Moodle's Workshop

Dr. Patricia Carlson, Rose-Hulman Institute of Technology

Dr. Patricia "Pat" A. Carlson is a transplanted middle westerner, having spent her childhood in Norfolk, Va. She came to Rose-Hulman Institute of Technology early in her teaching career and has taught a variety of courses over the past three decades. Dr. Carlson has held a number of American Society for Engineering Education summer fellowships that have taken her to NASA-Goddard, NASA-Langley, the Army Research Laboratory in Aberdeen, Maryland, and NASA's Classroom of the Future in Wheeling, W.Va. She was on loan to the Air Force Human Resources Laboratory from 1989 to 1995, managing a project to transition advanced instructional technologies to ten different middle schools located in five states. She is on the editorial board of three professional publications and has served as National Research Council Senior Fellow assigned to the Air Force Human Resources Laboratory. In her spare time, Pat enjoys reading and gardening.

Ryan Smith, Rose-Hulman Institute of Technology

Ryan Smith has served as webmaster and system administrator of the PRISM Project for the past 14 years. He is a 2002 computer engineering graduate of Rose-Hulman Institute of Technology. As part of his duties, Smith manages the PRISM server infrastructure, performs database administration, manages systems security, and coordinates distance education online courses on the PRISM website. In addition, Smith researches new web technologies to implement for the PRISM Project. In 2008, Smith received the Hoosier Educational Computer Coordinators HERO award for his contributions to integrating technology in Indiana education. In 2011, the PRISM team received the Rose-Hulman Institute of Technology Excellence in Service Award.

Computer-Mediated Peer Review: A Comparison of *Calibrated Peer Review™* and Moodle's *Workshop*

Abstract

Based on our extensive experience with Computer-Mediated Peer Review (CMPR) software in engineering education, we compare *Calibrated Peer Review* (developed at UCLA) and Moodle's *Workshop*. Our examination includes four criteria for each platform's effectiveness, both for students and for instructors. First, does the software include a cohesive mental model to facilitate a complex task? Second, does the scaffolding elicit / improve reliable student response? Third, are students encouraged to use peer commentary in the learning process? Fourth, does the platform collect and return empirical results that can be used as measures of course learning outcomes? We end with suggestions for improving return-on-investment for instructors and with advice for dealing with student receptivity.

1. Peer Review in Student Learning

Peer review has become an important pedagogical strategy in higher education. Based upon a well-known process in academic and business communities, the process of colleagues providing advice to colleagues has migrated to the classroom. Topping provides an overview both of the gains and of the costs of implementing peer review in various discipline-specific classes.[1] Furthermore, research on collaborative learning has established credibility for students in giving feedback to their colleagues. The notion of students helping other students in reading and writing was propagated by early advocates, such as Kenneth Bruffee.[2]

Computer-mediated peer review (CMPR) offers advantages on several dimensions: instructor return on investment, convenience of use, uniformity of delivery, data collection, and gains in student learning outcomes. Today, emerging technologies make asynchronous peer assessment more available and engaging. Additionally, significant research studies of fielded CMPRs add to the knowledge base for designing and implementing this category of educational technology.

The authors have over a decade of classroom experience using computer-mediated peer review software. Carlson has held three NSF grants to study *Calibrated Peer Review™* (*CPR*) in engineering education. She is not, however, a member of the platform's design / development team. Neither author is a developer for Moodle, although we do run a Moodle website for K-12 STEM educators (http://rose-prism.org). Neither of the authors has a financial interest in the commercial version of *CPR*. Elsewhere we have examined how *CPR* provides evidence for learning outcomes.[3-6] Here, we present a comparison / contrast between *Calibrated Peer Review* and Moodle's activity tool, *Workshop*. Our observations are informed by the central questions for peer review as a pedagogical device in engineering education suggested by Gehringer in papers given at ASEE and FIE.[7-8]

Both platforms can be used to facilitate peer-review for a range of assignment types: text documents, engineering drawings and visual representations, slide decks to accompany talks, even

videos of class presentations. In this review, we focus on written products and their composition process in the context of the writing-as-a-way-of-learning movement.

2. Evolution of Computer-Mediated Peer Review (CMPR) Systems

Despite proven benefits, integrating effective peer-review into a course requires much effort. Using computers to facilitate the process of peer-review was a logical progression. Early peer response systems made use of email exchanges among student reviewers. With digital advances in the 1990s, CMPR systems – such as MUCH (Many Using and Creating Hypermedia, 1994) – automated allocation of files for review, stored responses, calculated results, and gave access to peer feedback.[9] Also, Eschenbach exploited web-enabled software to integrate e-assessment in an engineering design course.[10]

In the 2000s, increasing use of computers in education, more robust internet connectivity, and advances in peer-to-peer sharing software resulted in improved CMPR systems. Gehringer's group at the University of North Carolina developed ***PG***, "a portable, Web-based … system written in Java. . . " (p. F1B-3).[7] Of note, ***PG*** may well have been the first CMPR to include a grade adjustment algorithm that rewarded students for giving quality feedback.

Other richly featured systems moved the CMPR paradigm forward. SWoRD (Scaffolded Writing and Rewriting in the Disciplines), developed at the University of Pittsburgh by Christian D. Schunn, added layers of mediation to the basic design and facilitated a richer e-assessment and feedback analysis.[11] Another system, SPARK (and SPARK^{PLUS}) – developed at the University of Technology, Sydney, Australia – also moved CMPRs closer to becoming highly interactive learning environment rather than automated systems for overcoming logistical hurtles. SPARK^{PLUS} focuses on mediating group-work by including various peer- and self-assessment workspaces. Like ***PG***, the grading algorithms in SPARK take into account the quality of feedback contributions and the level of engagement from all members of a team.[12]

In summary, today's CMPR systems make use of the improved technical capacity of web-enabled platforms. Thus, many of the burdensome aspects for face-to-face peer review have been alleviated. As we have noted elsewhere, "today's CMPR systems have the characteristics of what an educational technologist would call a 'cognitive tool,' a mental-modeling device aiding learners to enact more powerful strategies for problem-solving than possible without the scaffolding, heuristics, and visualization embedded in the device itself."[5]

3. Background for UCLA's *Calibrated Peer Review* and Moodle's *Workshop*

CPR is one of the earliest, sophisticate CMPR systems still widely available.[13] (See http://cpr.molsci.ucla.edu/.) *Workshop,* on the other hand, mirrors a more typical Open Source development history.

Originally fielded in 1995, major *CPR* refinements have been funded through several NSF grants, and the literature contains a rich body of studies on the system's efficacy. Having gone through six releases, the current version – *CPR6* – has some barriers for usage. First, this version is no longer free. Second, the purchasing institution must run the core package locally and absorb

the overhead for software / network maintenance and for the storage of student submissions. However, UCLA continues to offer access to a free, older version, which we here designated as ***CPR1***. This release is fully maintained, supported, and periodically updated. Like most legacy software, it contains some idiosyncrasies that require getting used to for a smooth operation. Because of its wide accessibility, we used *CPR1* in constructing the current comparison / contrast.

Workshop is available to those who have access to the Learning Management System Moodle. *Workshop* was originally a 3rd-party plug-in created by UK developer Ray Kingdon in 2003 for Moodle 1.0. The module was integrated into the core software with Moodle 1.1 (released May 28, 2003). The component had a slow start; the interface wasn't easy to use; additionally, the setup process, interaction protocol, and scoring scheme were opaque both to students and to instructors.

By 2009, the module was so rarely used that it was slated to be removed from the 2.0 Moodle release.[14] However, the software was given new life through a redesign by developer David Mudrák in 2009 - 2010 and was elevated to a core activity in Moodle 2.0 release. Improvements were made to the interface, particularly the instructor setup phase. Also, the scoring algorithms were reworked to return results that were more understandable to the end-users.

Today, *Workshop* reviewers still underscore the module's complexity, recommending it as an advanced tool to be used after an instructor is thoroughly familiar with Moodle's *Assignment* activity. Vogel notes that even with the latest *Workshop* version, continued low usage stems from the "daunting number and nature of the settings . . . . [T]here are several pitfall to avoid which aren't obvious on the first pass" (para. 4).[15]

4. Criteria for an Effective CMPR System

CMPR platforms contain a suite of components. As Figure 1 shows, systems require an interface that accommodates the instructor's need to author lessons and to access student submissions and reviews. The system also needs another interface that facilitates the students' authoring submissions, creating insightful evaluations, and assimilating the constructive criticism given by the classmates in the e-assessments.

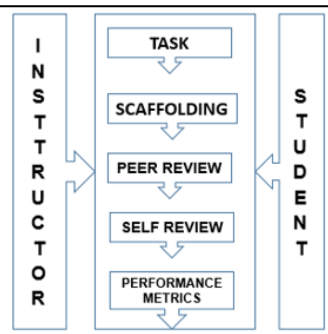


Figure 1: Fundamental Components of a CMPR Platform

Session components (the middle panel on the diagram) might include five segments. The first gives the assignment, and may include exercises to aid both in understanding and in performance. Next, embedded scaffolding supports the students' engagement with the task and models robust strategies for completing the assignment. The third and fourth segments (Peer Review and Self-Review) accelerate learning through application of a rubric and reflection on self-performance. In these segments, students demonstrate an understanding of task requirements by critiquing the work of others. Then, students consolidate their learning gains by reviewing their own submission and reflecting on

ways to improve. The fifth and last segment collects performance data /peer commentary and displays results both to instructors and students. This last segment reinforces learning by giving a composite grade, rewarding achievement, and helping students to see missed opportunities.

From this overview of a basic CMPR system (Figure 1), we extract four criteria for examining *CPR1* and *Workshop*

1. Does the system include a cohesive mental model that deconstructs the process and demonstrates staged problem-solving?
2. Does the system include scaffolding to move students forward, both in task accomplishment and in enhanced independent learning?
3. Does the system encourage the student to learn from peer-feedback?
4. Does the system provide data / outcomes for instructors to assess both assignment-specific and programmatic gains for individuals and for larger aggregates?

Segments 4.1 through 4.4 examine how the two platforms address these four criteria.

4.1 Includes a Cohesive Mental Model for a Complex Task

Like other digital devices used in staged-problem solving, a well-design CMPR helps the student to partition a complex cognitive task into smaller, more manageable units. As such, a CMPR operates as a process-modeler similar to software used in teaching engineering systems design. Ideally, a model of the workflow helps students to engage in sequence of the writing / reviewing process without being overwhelmed by the intricacy of the task.

Both *CPR1* and *Workshop* meet this requirement, but in different ways. Neither platform has a synoptic graphical representation for students to track progress or to visualize the entire process. Yet, both systems help students to deconstruct the work into meaningful, smaller units and to deploy problem-solving techniques appropriate for the situation.

CPR1 is rigorous in its workflow enforcement. The instructor sets the deadlines – the openings and closings for each of four stages. Typically, the session takes place over several days. Four stages are enacted:

1. **Submit** -- Read the prompt, examine explanatory / supplementary materials, and upload completed work as a file.
2. **Calibrate** – Train on using an instructor-authored rubric by applying it to an instructor-created set of samples. Correctness is determined by proximity to the instructor's application of the rubric. Trained students are given a Reader Competency Index (RCI) score from 1 to 6, which adjusts how much their rating weights in evaluating the work of their classmates.
3. **Peer Review** – Each student applies the same Calibration rubric to the submissions of three classmates.
4. **Self Review –** Each student now applies the rubric for the seventh application, this time to her own submission. Students are encourages to use reflection in this stage and to construct a plan for improvement.

*CPR*1 workflow instantiates the educational goal of teaching students to mirror expert behavior by learning to apply performance standards, as demonstrated by assessments done by the instructor. Basically, the scoring for each phase is a function of how far the student's assessments are from the instructor's in the Calibration, and how far from the group's norm in assessing a particular colleague's work.

Both *CPR1* and *Workshop* are built upon a similar, intrinsically obvious pattern for enacting structured peer review. However, *Workshop* consists of five workspaces: (1) Setup (instructor only), (2) Prompt and Submission, (3) Peer Assessment, (4) Calculation of Scores, and (5) Closing, or display of results. Note that this sequence is slightly less illustrative of the writing / reviewing process because it more closely follows what the software does than modeling the complex process of guiding the student through a multi-stage problem. For example, Setup is a dashboard available only to the instructor. Similarly, Calculation of Scores isn't an interactive workspace for students, nor does it demonstrate to the student how her actions resulted in specific scores. Also, Closing is a relatively static display of two quantitative results: a score for peer assessments and a score for text submission.

In summary, *CPR1* enacts a sophisticated, four-stage process that directly mirrors the cognitive tasks the students will accomplish in a complete assignment. On the other hand, *Workshop*'s sequence of activities is a linear, dashboard representation more useful as an aid to instructors in setting lessons than it is as a mental model for students working through the assignment.

4.2. Effective Scaffolding Enhances Reliability in Student Responses

Modern CMPRs offer more than just digital convenience for automating the review process. In 4.1, we examined how workflow representation in a CMPR might because a mental model for students to manage a complex task. We now turn to the second, equally critical function of a CMPR: scaffolding novice behavior by embedded coaching mechanisms that elicit quality feedback when reviewing a peer's submission.

Both *CPR1* and *Workshop* have the characteristics of what an educational technologist would call a "cognitive apprenticeship model."[16] This term comes from Vygotsky and is used to explain the mentoring dynamics that occur between master and apprentice and also between peers during collaboration.[17] To be effective, a CMPR system should scaffold the sophisticated judgments necessary for students to give helpful commentary. To accomplish this, both *CPR1* and *Workshop* require that instructors author a structured evaluation guide or rubric, suitable for the specific assignment.

CPR1's rubric asks students to provide two levels of peer evaluation: analytical (a list of orthogonal, atomistic criteria) and holistic (a single score based on the overall success of the submission). Within the *CPR1* authoring template, the instructor composes the analytical questions, using as many individual items as need. The system automatically includes a last item asking the student reviewer to rate the whole piece on a scale of 1 to 10. As indicated below, three types of rating scales are available for analytical items and one for the holistic rating.

Response Type	Scoring Method	Display
Analytical questions	Binary – Yes / No	Selection displayed as a radio button. Text box for written commentary – can be required or optional
	Gradient – None / Some / Many	
	Multiple Choice – A / B / C defined choices	
Holistic rating	Fixed scale from 1 (low) to 10 (high)	Text box for written commentary – can be required or optional

Workshop also has a guided workspace for instructors to author an evaluation rubric to scaffold student responses. However, *Workshop* offers the instructor four different options for scaffolding student reviews of peer work. These response styles are listed below:

Response Type	Scoring Method	Display
Accumulative	Evaluation guide consists of a set of analytical, orthogonal criteria. Peers score each by a number (out of 100) or on another scale. Each criteria can be weighted for impact on the final tally	Question followed by a blank for the score. Text box for reviewer's comments on each criteria.
Comments Only	Returns only qualitative results. I.e., reviewers provide written feedback for a specific prompt. The criteria are written by the instructor and can be very specific, nuanced, and task-specific.	Question and text box. Displays only qualitative feedback (comments). No quantitative scoring is used
Number of Errors	Items elicit a binary response. The criterion is either present or not. Various words can be set to express the pass or failure state – e.g. Yes/No, Present/Missing, Good/Poor. Each response item can be given its own weight, to reflect importance or added emphasis.	Question followed by a radio button. Text box can be present to gather qualitative comments, e.g., explanation for why a judgment was given.
Rubric	Similar to a survey, criteria are listed by the instructor and the student indicates the level of quality in the submission.	Question followed by radio buttons indicating level of performance. Text box can be present to gather qualitative comments.

Karen Spear, among others, makes a case for addressing reliability in student-contributed feedback in peer review.[18] Scaffolding – in the form of structured review – mitigates the common objection to peer review in the undergraduate classroom: that the experience reduces itself to the-blind-leading-the-blind. Both *CPR1* and *Workshop* improve reliability in peer comment by giving explicit task requirements and by asking student reviewers for targeted advice on each designated element of the assignment.

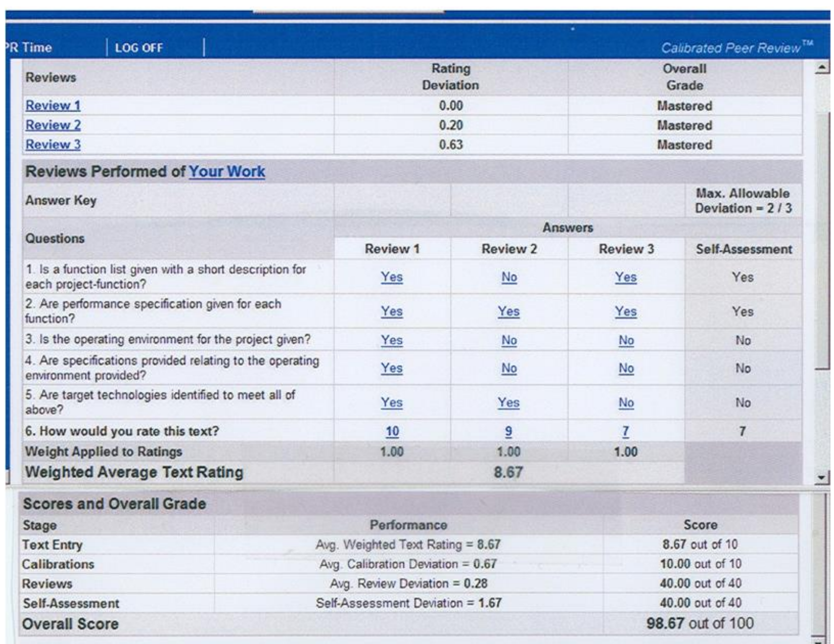
4.3. Students are Encouraged to Use Peer-feedback in the Learning Process

The third criteria – at the heart of collaborative learning – examines how these CMPRs encourages students to value peer feedback as an enabler for learning. The concept that students refine and mature in their socio-cognitive behaviors by interacting with other students echoes Bruffee's theory that communication and interaction are critical to enabling authentic achievement.[19]

Both *CPR1* and *Workshop* elicit quantitative and qualitative results as students assess the submissions of their classmates. Although similar in intent, both platforms implement this function in different ways. Two differences are (1) calculations for final scores and (2) methods for displaying feedback to the student. We do not contend that grades are the sole motivator for students. We do, however, submit that a composite performance score – with clear links to praise and targeted advice for improvement from peers – will motivate students to contemplate more deeply the evaluations made by their colleagues.

Both *CPR1* and *Workshop* tally results from all phases in the session and report a score, based on a 100 point scale. Both platforms allow instructors to set weights for computing each element in the total score. Both also display the sub-score for different phases of the session, so that a student's learning performance is available for each identified stage.

However, the quantitative results given in *CPR1* are more fine-grained than those available in *Workshop*. As indicated in Figure 2, for *CPR1* many different elements go into the calculation of the final score for each student. Also, the composite score sheet displayed to the student is both synoptic and easy to explore. A matrix shows all four stages of the session and the quantitative results for each: (1) ability to apply the assessment guide during the calibrations, (2) quality of the submitted text, (3) quality of the reviews given, and (4) consistency of self-assessment. The single interface uses hyperlinks to access the peer qualitative comment and justification associated with each quantitative judgment.


PR Time | LOG OFF | Calibrated Peer Review™

Reviews	Rating Deviation	Overall Grade
Review 1	0.00	Mastered
Review 2	0.20	Mastered
Review 3	0.63	Mastered

Reviews Performed of Your Work

Answer Key				Max. Allowable Deviation = 2 / 3
Questions	Answers			
	Review 1	Review 2	Review 3	Self-Assessment
1. Is a function list given with a short description for each project-function?	Yes	No	Yes	Yes
2. Are performance specification given for each function?	Yes	Yes	Yes	Yes
3. Is the operating environment for the project given?	Yes	No	No	No
4. Are specifications provided relating to the operating environment provided?	Yes	No	No	No
5. Are target technologies identified to meet all of above?	Yes	Yes	No	No
6. How would you rate this text?	10	9	7	7
Weight Applied to Ratings	1.00	1.00	1.00	
Weighted Average Text Rating		8.67		

Scores and Overall Grade

Stage	Performance	Score
Text Entry	Avg. Weighted Text Rating = 8.67	8.67 out of 10
Calibrations	Avg. Calibration Deviation = 0.67	10.00 out of 10
Reviews	Avg. Review Deviation = 0.28	40.00 out of 40
Self-Assessment	Self-Assessment Deviation = 1.67	40.00 out of 40
Overall Score		98.67 out of 100

Figure 2: Composite results screen for a student session in *CPR1*

Workshop, on the other hand, calculates final scores based on only two components: (1) grade for submission (the average score from all the peer reviews) and grade for assessments (in essence, this is a number derived from the notion of inter-rater reliability). The instructor can give a percentage for each of the two components in the final grade. For example, the

submission grade may be set at 80%, and assessments are set at 20%. In *Workshop*, the same notion of distance from the norm determines the assessment ratings. The implication is the same for both platforms: student group performance in applying the assessment criteria determines their grades. This notion shares similarities with the more familiar idea of grading on a curve.

We observe from our experience using *CPR1*, that the composite display – with the ability to drill down to each associated comment – motivates students to spend time absorbing qualitative advice and to view the comments as elaborations on the quantitative numbers.

4.4 CMPR Results Can be Used as a Measure of Course Learning Outcomes

At a minimum, CMPRs help instructors by providing a reasonably accurate grade for a relatively sophisticated experience. As an extension, the data collected by the system can also be used to measure learning outcomes in a number of other ways. Elsewhere, we have elaborated on using a CMPR's results to capture learning outcomes appropriate for an ABET-style assessment.[5] Here we suggest some ways in which instructors can realize a return on investment by using CMPR system data as indicators of student progress. For example, if more than one peer review session occurs, scores for individuals can be examined longitudinally to track learning across the entire course. Alternatively, aggregates of student results can be examined to determine if results correlate with other identified variables.

Both *Workshop* and *CPR1* collect and return results to the instructor during the session. Both systems store results, and instructors can return to results for review long after the assignments have closed. In each, the instructor can watch – in real time – as the students work within the different stages of the exercise. Also, while reviewers are anonymous to the student, the instructor can see names and can monitor commentary for appropriateness. In both systems, the instructor can batch download the submissions and upload that file to a plagiarism checker, such as *TurnItIn*. However, both systems are not equal in what level of data they allow the instructor to access easily.

In *Workshop* instructors have many options for constructing the rubric and for setting the weights in the final calculations, yet the system returns only two composite grades (submission and reviews). These scores may also be transferred to the Moodle Gradebook. As indicated in Figure 3, the instructor sees the two grades at the top level, and can click on each reviewer to move to the qualitative and quantitative input for each. Unfortunately, the numerical elements of the composite scores are not downloadable to a spreadsheet, meaning that only limited statistical analysis can be done.

First name / Last name	Submission / Last modified	Grades received	Grade for submission (of 80)	Grades given	Grade for assessment (of 20)
Five Student	amtmoore5 modified on Saturday, March 22, 2014, 6:41 PM	80 (20)< One Student 80 (20)< Three Student	80	72 (20)> One Student 60 (20)> Three Student	20
Four Student	amtmoore4 modified on Saturday, March 22, 2014, 6:39 PM	40 (20)< Two Student - (-) @ 0< Six Student	40	40 (20)> Six Student 40 (20)> Two Student	20
One Student	amtmoore1 modified on Saturday, March 22, 2014, 6:40 PM	72 (20)< Five Student 80 (20)< Three Student	76	80 (20)> Five Student 44 (20)> Three Student	20
Six Student	amtmoore6 modified on Saturday, March 22, 2014, 6:41 PM	40 (20)< Four Student 76 (20)< Two Student	58	- (-) @ 0> Four Student - (-) @ 0> Two Student	-
Three Student	amtmoore3 modified on Saturday, March 22, 2014, 6:39 PM	60 (20)< Five Student 44 (20)< One Student	52	80 (20)> Five Student 80 (20)> One Student	20
Two Student	amtmoore2 modified on Saturday, March 22, 2014, 6:40 PM	40 (20)< Four Student - (-) @ 0< Six Student	40	40 (20)> Four Student 76 (20)> Six Student	20

Figure 3: Instructor's view of student scores for a completed Workshop session.
Note: students' names in this example have been replaced by placeholders, e.g., One Student.

This sparse return of data is a lost opportunity when compared with the richness of data available in *CPR1*. As Figure 4 shows, instructors can see a numerical profile of every student's progress through the four stages of the exercise. The disaggregated overall score consists of 12 data points, each measuring performance on a specific task during the overall session.

S1.	99.33	9.33	9.33	100.00	79.17	0.33	10.00	6	0.37	40.00	0.33	40.00
S2.	95.34	8.67	8.67	66.67	66.67	1.00	6.67	4	1.57	40.00	0.33	40.00
S3.	98.01	8.01	8.01	100.00	66.67	0.00	10.00	6	0.72	40.00	0.01	40.00
S4.	90.35	3.68	3.68	100.00	70.83	2.33	6.67	4	1.18	40.00	1.32	40.00
S5.	97.01	7.01	7.01	100.00	70.83	1.33	10.00	6	1.15	40.00	0.99	40.00
S6.	95.66	8.99	8.99	66.67	70.83	1.67	6.67	4	0.73	40.00	0.99	40.00
Class Averages												
Categories	Overall Grade	Text		Calibrations					Reviews		Self-Assessment	
		Rating	Score	% Style	% Content	Avg. Dev.	Score	RCI	Avg. Dev.	Score	Dev.	Score
Class Averages	87.46	7.83	7.83	87.04	68.13	1.22	7.90	4.87	0.86	37.95	0.96	38.78

Figure 4: Sample of Instructor's Report from a CPR1 Session

All 12 distinct scores are downloadable to an EXCEL spreadsheet, where various forms of statistical analysis can be performed. Instructors can track progress of sub-groups within the class and – based on numerical analyses – can modify the assignment for later use. More specifically, because each of the 12 data points (in Figure 4) measures an identified activity, instructors can determine which aspect of the assignment was weaker than others. If students as a group have low Reader Credibility Index (RCI) numbers, the calibration samples and / or the rubric need further refinement from the instructor. Additionally, empirical assessment is especially useful for overall programmatic evaluation.

5. Discussion

Both *CPR1* and *Workshop* are fully-developed CMPR systems. Each is delivered as an "empty engine," and both are remarkably powerful cognitive tools for integrating collaborative learning exercises within a content course. Both systems guide the instructor in authoring assignments and constructing review rubrics to match the needs of the course's pedagogical purposes. (*CPR1* also has a system library where instructors can use / adapt previously submitted lesson plans.) On balance, both systems have positive and negative attributes; thus, we will not make a definitive recommendation. Instead, we end by commenting on two important facets of implementation and efficacy.

5.1 Return on Investment for Instructor: Using transformative pedagogies such as writing-for-learning requires substantive reworking of course materials. Indeed, as those teaching in engineering programs can attest, integrating writing assignments to meet current ABET criteria requires coordination and ingenuity.

Using a CMPR system certainly provides an advantage over using the paper process. However, using either *CPR1* or *Workshop* is labor intensive. First, instructors will find a steep learning curve for setting up either system because of the range of options available. Second, constructing the training segment for either system requires that the instructor find benchmark samples and evaluate them based on performance specifications used as the assessment rubric.

This can be challenging: for example, *Calibrated Peer Review* uses the convention of having a low, middle, and high quality example in the training suite. Finding authentic samples that fit the assignment can be problematic, unless one has access to previously submitted student contributions that can be used without violation of FERPA guidance. Also, to harvest the benefits of modeling mature commentary for students, faculty must create richly contextualized examples of written reviews for each element on the rubric and for each sample.

Less the above caveats seem daunting, we attest that using either system becomes easier with time. We also advise setting up collaborations with other instructors of the same or similar courses and sharing the design, insights, and authoring work. Despite the upfront investment, we note these significant likely payoffs delivered by either system.

- **Improve the use of writing in the learning process:** The close sequencing of a lesson necessitated by both *CPR1* and *Workshop* helps faculty better to understand the connection between writing tasks and learning. The authoring functions of both platforms help faculty to visualize a complex, multi-staged process. Within the authoring environment, instructors set learning goals and devise scaffolding by which novice learners can accomplish the target performance levels. "In essence, instructors unpack an assignment through a cognitive task analysis and accompanying outcomes, and then ask students to rate how well their colleagues have met structured performance standards."[8] All of this requires more profound scrutiny of desired learning and assignment enactment. Though labor intensive, the process helps instructors more fully appreciate incorporating writing within a content course.

- **Improve student engagement in the learning process***:* In both systems, the composition/review/revision process – perceived as opaque and arbitrary by many students – is reified into a set of understandable "state transitions" through the robust "thinking frames" used in each design. Students are trained on what to look for in evaluating the specific artifact, and this – in turn – improves the quality of student feedback, enhances learner confidence, and promotes deep engagement in the process. True to the tenets of collaborative learning, after evaluating peer contributions, students can consolidate and apply their gains by reflecting on their own submission (in light of gains made from examining materials from fellow students).

- **Qualitative / Quantitative Measures Verifying Student Learning Outcomes***:* Transformative engineering education moves beyond a surface grasp of course content. As we have noted elsewhere, engineering education now encompasses (1) mature inquiry as a precursor to knowledge, (2) reflective judgment as a foundation for decision making, and (3) meta-cognition as the initiator of self-awareness.[6] Finding evidence in student outcomes that such lofty goals have been met proves problematic. Both *CPR1* and *Workshop* collect evidence of student performance, which can then be analyzed to support student gains in identified learning goals.

5.2 Student Receptivity*:* Results from studies on student receptivity to CMPR are mixed. However, several cautions must be given because of the many variables involved in any such empirical research. Earlier implementations tended to meet student resistance potentially because web-delivered applications and classroom innovations such as blended learning

environments were still very new. Thus, students had to overcome the learning curve of the technology as well as become comfortable with the notion of a classroom without walls.

An insightful recent study on using *Workshop* suggests that students' negativity may be based on more nuanced and ingrained student attitudes.[20] In this study, students' concerns arose from issues of trust and of professionalism. The first is a perennial issue in peer-review: to what extent is a classmate capable of determining a grade for a fellow student? The second, also a common complaint, centers on a traditional notion of the responsibility and authority of the teacher. Some students see using a CMPR as shifting the workload off the instructor and onto the student. A third criticism raised by students – in a different study – concerns collaborative versus competitive environments, with concerns that a CMPR system rewards lazy students who are given access to several peer submissions, and then merely imitate what the better students have worked hard to achieve.[21]

After years of experience with a CMPR systems, we suggest these strategies for helping students see the value of the process:

- Use the CMPR session as part of a review and re-write process, never as a final grade for the artifact. Allow the students to use the peer-provided feedback as part of an improvement phase. Allow students to resubmit the end-product. This approach best brings out the collaborative learning potentials in the software.
- Alternatively, use the CMPR session to critique a specific small component of paper or project. Enact an iterative process for students to work on a final, larger composite where they harvest the good advice provided by colleagues.
- Avoid using the assessment from CMPR as the only grade for an entire assignment. Keep in mind that both systems are intended to provide formative, not summative, evaluation.
- Since CMPR sessions can be done as out-of-class work, use several, shorter exercises within the expanse of an entire course. Explore ways in which these assignments can take the place of more traditional homework exercises.

As a final word, we underscore that using any CMPR platform to its full advantage requires much work and might best be done as a multi-semester effort accomplished by a team of instructors. However, the rewards are many. We close with an observation by Shoshana Zuboff (*In the Age of the Smart Machine*) cautioning against false expectations from digital applications.[22] Simple automation – she suggests – takes away value from a task for humans. As an alternative, Zuboff argues for using Information Technology (IT) to "informate" complex tasks – that is collect and analyze data embedded within the process of accomplishing the work for the purpose of deeper engagement in the task. We believe the two CMPR systems reviewed here "informate" rather than automate. They orchestrate a sophisticated group activity and provide an effective feedback loop on performance; thus, they serve both as a learning tool and as an assessment tool.

References

1. Topping, K.J. & Ehly, S. W. (2001). Peer assisted learning: A framework for consultation. *Journal of Educational and Psychological Consultation*, *12*(2), 113-132.

2. Bruffee, K. A. (1978, Dec.). The Brooklyn Plan: Attaining intellectual growth through peer-group tutoring, *Liberal Education*, 64, 447-468.
3. Carlson, P., Russell, A. A., Waggenspack, W. N.,Wilmot, C. G., Bowles, D., Voltmer, D. R., Monroe, W. T., Hull, W. R., & Raubenheimer, D. (2012). Improving engineering education with enhanced Calibrated Peer Review: Assessment of a collaborative research project," *Proceedings*, ASEE Annual Conference, San Antonio, TX, 10 – 13 June.
4. Berry, F. C. & Carlson, P. (2010, Dec.). Assessing engineering design experiences using Calibrated Peer Review, *International Journal of Engineering Education*, *26*(6), 1503-1507.
5. Carlson, P. & Berry, F.C. (2008). Using computer-mediated peer review in an engineering design course, *IEEE Transactions on Professional Communication*, *51*(3), 264-279.
6. Carlson, P. & Berry, F. C. (2003). Calibrated Peer Review™ and assessing learning outcomes. *Proceedings*, ASEE/IEEE Frontiers in Education Conference, Westminster, CO.
7. Gehringer, E. F. (2000). Strategies and mechanism for electronic peer review. *Proceedings* of the 30th ASEE/IEEE Frontiers in Education Conference, Kansas City, MO.
8. Gehringer, E. F. (2012). Applications for supporting collaboration in the classroom. *Proceedings*, ASEE Annual Conference, San Antonio, TX, June 10–13, #2012-4769.
9. ACM (1994). The many using & creating a hypermedia (MUCH) system, *SIGBIO Newsletter*, *14*(3), 12.
10. Eschenback, E. A. (2001). Improving technical writing via web-based peer review of final reports, *Proceedings*, ASEE/IEEE Frontiers in Education Conference, F3A1-5. Reno, Nevada.
11. Cho, K. & Schunn, C.D. (2005). Scaffolded writing and rewriting in the disciplines: A web-based reciprocal peer review system. *Computers and Education*, *48*(3), 409-426.
12. van Zundert, M., Sluijsmans, D. & van Merrienboer, J. (2010), Effective peer assessment processes: Research findings and future directions. *Learning and Instruction*, *20*(4), 270-279.
13. Chapman, O.L. & Fiore, M.A. (2000, Winter). Calibrated Peer Review, *Journal of Interactive Instruction Development*, 11-15.
14. Figueira, A. & Cunha, E. (2009), Work in progress: An easy-to-use Workshop module. *Proceedings*, ASEE/IEEE Frontiers in Education Conference, W2J-12.
15. Vogel, M. (2015, March). A good peer review experience with Moodle Workshop. Available at: https://www.ucl.ac.uk/teaching-learning/case-studies/good-peer-review-experience-moodle-workshop.
16. Brown, J. S., Collins, A., & Duguid, P. (1989). Situation cognition and the culture of learning, *Educational Researcher*, *18*, 32 – 42.
17. Vygotsky, L. (1962). *Thought and Language*. Eugenia Hanfmann & Gertrude Vakar (Trans.). Cambridge, MA: MIT Press.
18. Spear, K.I. (1988). *Sharing writing: Peer response groups in English classes.* Portsmouth, NH: Boynton / Cook.
19. Kelly, J. (2002). Collabortive learning: Higher education, interdependence, and the authority of knowledge by Kenneth Bruffee: A critical study. *Journal of the National Collegiate Honors Council*. Available at: http://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1090&context=nchcjournal
20. Wilson, M.J., Diao, M. M., & Huang, L. (2015). 'I'm not here to learn how to mark someone else's stuff': An investigation of an online peer-to-peer review Workshop tool. *Assessment and Evaluation in Higher Education*, *40*(1), 15-32.
21. Mostert, M. & Snowball, J. D. (2013). Where angels fear to tread: Online peer-assessment in a large first-year class. *Assessment & Evaluation in Higher Education*, *38*(6), 674-686.
22. Zuboff, S. (1988). *In the Age of the Smart Machine: The Future of Work and Power.* New York: Basic Books.