Using Calibrated Peer Review™ to Mediate Writing and to Assess Instructional Outcomes

Patricia A. Carlson, Frederick C. Berry, and David Voltmer
Department of Humanities and Social Sciences / Department of Electrical and Computer Engineering
Rose-Hulman Institute of Technology
Terre Haute, Indiana 47803

Writing in Engineering Education

The written word is crucial to engineering for at least two compelling reasons. First, the texts of engineering – publications that report findings or describe/advocate “products” – are the lifeblood of the profession. Project reports, feasibility studies, proposals, user/client support manuals, libraries of design documents, product specifications, risk communication, and a myriad of memoranda and business letters – this list is but a part of the many text forms required of a practicing engineer.

Second, writing is important to engineering – as it is to any knowledge worker – because the act of placing ideas into language mediates higher-order intellectual activities that are foundational to critical thinking and multi-staged problem solving. Though other symbol systems – notably mathematics – play a major role in engineering reasoning, language fosters mental manipulations such as synthesis, analysis, classification, inferring, definition, hierarchical order, comparison/contrast, and elaboration/extension. Some of the most respected of twentieth-century educational theorists have endorsed this notion of writing as a medium for learning and for understanding. Vygotsky, Luria, and Bruner, to name only a few, have pointed out that higher cognitive functions seem to develop most fully only with the support system of verbal language – particularly, of written language.

Several contemporary researchers have demonstrated that carefully designed and well-integrated writing assignments improve concept learning in content areas courses. However, until recently, most traditional engineering courses made use only of rather pro forma written genre, such as lab and project reports.

But, change is at hand. Certainly, recent journal articles and conference papers provide accounts of many solid programs where communication faculty work with technical faculty to create rich and challenging courses. Yet, we posit that these collaborations are the exception rather than the rule. Specifically, we address our presentation to engineering faculty members who do not
have access to the resources and expertise of showcase programs, who may be feeling pressured to include writing without being given the needed guidance or support to do so, and who may be feeling reluctant, dubious, or anxious about these new expectations.

Our experience with Calibrated Peer Review™ (CPR™) in several courses at Rose-Hulman Institute of Technology suggested that this robust instructional technology which partners both with the instructor and with the student has the potential to –

- Increase competence, creativity, and confidence in exploratory inquiry and reasoning.
- Promote and sustain interest in engineering practice.
- Engage the user and transfer powerful strategies for problem-solving.
- Bridge the gap between process (thinking) and product (writing).
- Improve the quality of writing by improving quality of thinking.

What is Calibrated Peer Review™?

CPR™ is a component of a large-scale, National Science Foundation-supported project led by a team of educators at UCLA to develop a completely digitized, network-delivered Molecular Science Curriculum. The fully integrated CPR™ contains an assignment authoring tool for custom crafting of writing tasks and a library of edited assignments contributed by instructors from varied institutions. Currently hosted at UCLA, the system draws from the model of manuscript submission and peer-review in the conduct of scientific inquiry.11

Components that Enable Learning

Four structured workspaces perform in tandem to create a rich series of activities that reflect modern pedagogical strategies for using writing in the learning process. Table A summarizes these stages in a typical CPR™ session.

Table A: Four Structured Workspaces of CPR™

<table>
<thead>
<tr>
<th>SEGMENT</th>
<th>ACTIVITY</th>
<th>COMMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Assignment and Text Entry: Students are presented with a challenging writing task, with guiding questions to act as scaffolding for the demanding cognitive activities.</td>
<td>Instructors work with the authoring interface and are guided through the construction of a task that elicits active learning. Students compose using a word processor, but upload the finished text as an HTML file. Some graphics and all tables are supported in the upload.</td>
</tr>
</tbody>
</table>
| 2       | Calibrations: After electronic submission of their texts, students read through three “benchmark” samples and assign each a score based on a series of evaluative questions | Modeled on the same process used in large scale writing evaluation projects, this segment mitigates the common objection to peer review in the undergraduate classroom: that the experience reduces itself to the-blind-
(a rubric). Students are then given a “reliability index” from 1 to 6, based on their demonstrated competency in these exercises.

| 3 | Reviews of Texts Submitted by Classmates: After becoming a “trained-reader” – and being assigned a credibility weighting – students read and provide written feedback on three anonymous peer essays using the same rubric as embedded in the calibrations. They also assign each essay a holistic score from 1 to 10. | Many years of classroom observations augmented by highly structured research on writing as used in a broad spectrum of learning situations indicate the power of peer review. As early as the 1970s, Ken Bruffee and his colleagues demonstrated that students paid more attention to critiques of their writing when done by peers than when done by an instructor. Looking at the other side of the coin, providing commentary on texts submitted by classmates also helps novice writers to sharpen their abilities to recognize aspects of writing that meet the performance standards of the assignment. |

| 4 | Self-Assessment: As a final activity, students evaluate their own essay using the same learning template instantiated in the calibration and in the peer review segments. | As with segments 2 and 3, students use the same “rubric” (set of performance standards for the task). Only this time, they apply the standards to their own text. Having “trained” on benchmark samples, and then applied their expertise in evaluating peer text, students now engage in a reflective, final activity by assessing their own submission. Students are encouraged at this time to make comments to themselves that capture the evolving insights they have gained in the previous two segments. They may also be invited to reflect on whether they have gained a deeper level of understanding for the assignment and its outcomes. |

**Components that Facilitate Usage**

The CPR™ system contains several components that facilitate both usage and learning. Though the multiple features make the system seem complex, following a typical session path...

*Proceedings of the 2003 American Society for Engineering Education Annual Conference & Exposition Copyright ©2003, American Society for Engineering Education*
demonstrates both CPR™’s power and its ease-of-use. Figure 1 – a conceptual overview – guides the discussion for the features and functions of CPR™.

Figure 1: Conceptual Overview of CPR™

Item 1  The student interface guides the learner through all four segments in a seamless fashion. All components – from assignment and attendant resources to peer feedback and consolidation of performance indicators – are viewed in a unified framework.

Item 2  The instructor interface contains a number of powerful features. An "authoring space" guides the construction of assignments. A “progress” function gives real-time reports on the ongoing activities of a class or a single student. A number of “tools” handle anomalies – such as accommodating a student who misses a deadline for valid reasons.

Item 3  The “library” contains a number of “edited” assignments contributed by a network of CPR™ users throughout the country. Individuals or institutions can also maintain a “private use” library.

Item 4  A sophisticated set of class and student accounts are maintained on the main server, currently hosted at UCLA. Quantitative reports are calculated through a well-designed set of algorithms and returned to both
student and instructor. Qualitative feedback is also available to both student and teacher.

**Item 5** The results (comments and text) may be viewed as finished, or they may be used in further iterations, or as the foundation for another “sequenced” assignment.

**How Does CPR™ Improve Student Learning?**

Not so long ago, engineering was viewed as applied science. The curriculum was filled with exercises in scientific fundamentals, applied to what might be termed classical engineering situations. Much time was spent in laboratories, working through standard exercises in data collection and analysis. Unlike the teaching that now characterizes medicine or the law, engineers were taught in an environment curiously devoid of a sense of practice. Over the years, this “decontextualization” caused distortions in attitudes and values, along with graduating engineers who had little exposure to the types of professional maturity and judgment they would need from the very first day they stepped into the work world.¹⁶

Most would agree that expectations for today’s engineering education incorporates richer notions of how people learn and of how to nurture the higher-order skills necessary for a knowledge worker in the 21st century. The rich dynamics of the four workspaces make CPR™ an excellent tool for modeling learning tasks that promote both cognitive and affective development.

**Facilitation for Cognitive Development**

Based on the ideas of learning psychology theorists, CPR™ instantiates the “writing as a way of learning” approach to pedagogy.¹⁷ Practitioners who have pursued the notion that writing is a heuristic for higher-order cognition report their students to be more actively engaged in learning and also find improvements in critical meta-cognitive abilities (or thinking about one’s own thinking). An annotated bibliography, prepared at Brigham Young University, indicates the richness of implementation associated with the “writing as thinking” approach to teaching in content courses.¹⁸ In our presentation, we provide examples of well-designed CPR™ sessions that foster student learning as defined by Bloom’s Taxonomy.

**Facilitation for Affective Development**

A second important model recurrent in today’s engineering education literature – the schema codified by William Perry – focuses on a student’s development of mature decision making and ethical sensibility.¹⁹ Essentially an epistemological model, the Perry Scheme traces the development of mature thought processes through nine stages in which the student progressively moves from depending on external, “teacher-centered” authority to a more self-assured ability to reconcile multiple perspectives, to tolerate ambiguity, and to reflect on the process of knowing (meta-cognition). In our presentation, we provide examples of CPR™ sessions which foster student learning as defined by the Perry Scheme.
How Does CPR™ Help to Construct Writing Assignments?

ABET’s EC2000 requirements, changing realities of the workplace, and the growing awareness of language in the learning process all place added emphasis on writing in today’s engineering curricula. However, merely providing more exposure to the traditional forms of technical communication instruction prevalent in academia today does little to prepare students either to become accomplished thinkers or to be skillful at the practice of engineering.

Using writing as a heuristic for learning significantly enriches the instructional system. The approach to writing as a cognitive tool is cogently articulated by. A product of the cognitive science research of the1990s, this approach takes into account modern discourse analysis and modern communication theory. In this model, the learning being enabled is an enactment of the cognitive skills that underlie multi-staged problem-solving, applied creativity, and complex decision analysis.

In a well-designed “writing as learning” activity, the instructor creates an open-ended, problem- or project-oriented assignment. Students are provided with resources that serve as “thought mediators” for the particular type of activity being modeled. Students work through these mediators and create a written artifact that moves beyond the mere transcribing of pre-existing knowledge. To use the words of noted researchers Carl Bereiter and Marlene Scardamalia, students are encouraged to abandon their less-robust methods of “knowledge-telling” in favor of more powerful strategies for “knowledge-transformation.”

Crafting a writing assignment that guides students through a series of higher-order mental manipulations is not an easy task. However, the authoring functions of CPR™ and the inter-dynamics of the four structured workspaces provide an instructor with a mental model for the writing process. This framework for the entire writing process aids instructors in developing assignments that both model behavior as well as foster in-depth processing of course content. As Figure 2 indicates how CPR™ helps instructors to develop integrated writing tasks that move students purposefully through the cognitive transformations need to turn observations into data, data into information, and information into knowledge.

Rather than focusing on a re-telling of the course “content” or – alternatively – emphasizing the surface features of writing as the “container” for pre-fabricated thought, a well-formulated CPR™ session encourages students to develop awareness of higher-level reasoning behaviors such as

- developing strategies for information gathering,
- identifying patterns of meaning and recognizing relationships,
- practicing processes of inquiry,
- making inferences from observation.

Also indicated in Figure 2, student learning gains in one area are consolidated and carried forward to the next workspace so that students are continuously challenged but never overwhelmed. In short, a good CPR™ session uses writing as an analog for thinking and encourages students to develop strategies that will endure even in the absence of the computer.
mediated learning environment. Our presentation includes examples of how to construct assignments that fully exploit these CPR™ capabilities.

Figure 2: Conceptual View of the “Writing/Revision Process” Instantiated by CPR™

How Does CPR™ Measure Learning Outcomes?

Good instructional design depends on having both objectives (desired outcomes) and a plan (a means of accomplishment). As illustrated in the seven levels of Figure 2, the instructor sets the goals in the writing prompt and thus initiates a guided inductive-path for attaining them by setting expectations (performance standards) that are reverberated throughout the next three levels of activity (calibration, peer review, self-assessment).

We claim that the multi-staged workspaces in a typical CPR™ session encourage students to develop higher-order reasoning processes such as discerning patterns of meaning, practicing processes of inquiry, and drawing inferences from observation. Such goals are easy to enumerate, but more difficult to substantiate using the assessment methods currently applied to writing outcomes. Thus, the second part of our claim is that CPR™’s built-in, comprehensive data collect provides a range of in-situ observations from which significant learning outcomes can be measured through standard and specialized data reduction methods. These learning outcomes can be represented as growth for individual students or aggregates over time or they can be given as synoptic overviews resulting from statistical methods such as ANOVAs or MANOVAs, to name just two forms of analysis that can be applied to the data routinely collected during a CPR™ session.
Interpretive inferences (assessments) can be drawn from the comprehensive features of the system. Each of the four structured workspaces returns data that indicates various aspects of student performance. For example, as shown in Figure 3, the calibration sequence tracks four specific indicators of student performance.

**Assignment:**
The guided-inductive instructions serve as performance requirements that anticipate the standards set forth in the rubric.

**Rubric:** Consists of a set of custom-designed questions that guide students in their evaluation of the 3 sample responses.

<table>
<thead>
<tr>
<th>Student</th>
<th>% Style</th>
<th>% Content</th>
<th>Avg. Dev.</th>
<th>Score</th>
<th>RCI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smith</td>
<td>55.55</td>
<td>77.78</td>
<td>1.33</td>
<td>1.67</td>
<td>4</td>
</tr>
<tr>
<td>Jones</td>
<td>88.89</td>
<td>100.00</td>
<td>2.00</td>
<td>10.00</td>
<td>5</td>
</tr>
</tbody>
</table>

**Figure 3: Indicators Tracked in the Calibration Workspace**

The results of the calibration exercise are reported in five separate numbers, as indicated above. (Note that “Score” is not an separate measure; rather, it is a composite result derived by CPR™’s algorithms and used as an element in determining the students overall “grade” for all four workspaces.) Both “Style” and “Content” column numbers indicate the percentage of questions about the calibration texts for which the student’s answer matched that of the “answer” determined by the instructor.

Each benchmark text is also given a holistic rating from 1 to 10 by the assignment author. The “Average Deviation” (Avg. Dev.) number indicates how far away from the rating standard the student was for all three of the calibration texts. The “Score” is the number of points (a portion of the potential total of 100) the student received for the calibration section of the session. (The instructor sets the point distribution between text, calibration, peer review, and self-assessment.) The “RCI” (Reviewer Competency Index) – an integer from 1 to 6 – indicates how well the student “trained” during the calibration. This score is used later in the session – in the peer-review workspace – to determine a “weighting” for this student’s rating of the three peer texts she reviews. (The algorithms embedded in CPR™ are beyond the scope of this paper. Suffice it to say that the calculations are very robust and that the instructor can set the level of tolerance for many of the indicators.)

In a well-crafted assignment, the calibration questions contained in the rubric are indicative of the learning objectives implicit in the assignment, which have been indicated in the guiding questions of the writing prompt. Thus, in performing the calibration, the student demonstrates the degree to which she can enact the knowledge learned in the assignment. Aggregating the...
scores for the entire class (or multiple sections of classes) provides insight into the learning outcomes for groups and sub-groups, helps the instructor to adjust the assignment, or aids in modifying upcoming activities. At a minimum, the data serve as a snapshot of how well the students understood the bundled concepts presented as “Style” and as “Content” questions.

A simple case example demonstrates how the data may be used. A strong negative correlation between either “Style” or “Content” (or their average) and “Avg. Dev.” indicates that the rubric has taught the student well. In other words, the higher the percentage of questions the student can answer correctly either in “Style” or in “Content,” the better she should be able to apply this knowledge in rating the sample texts. The correlation is negative since the better the student does in applying the rubric, the lower her deviation from the norm on the holistic number for the text. In other words, as one number increases, the other should decrease, if learning is taking place.

For measures of interactions, consider that both the peer review workspace and the self-assessment workspace also return data. More complex statistical manipulations may provide deeper interpretation of the learning outcomes. For example, multivariate analysis (MANOVA) potentially will provide insights into patterns that would not be readily apparent if only the attendant artifacts of the exercise (texts, peer-review sheets, and self-assessment reflective statements) were evaluated by more standard means. Additionally, as noted in the literature, assessment accomplished by drawing inferences from the final product (the written text) is both labor-intensive and subject to the criticism of subjectivity.25

As a second, more sophisticated, example of how CPR™ data may be used to provide robust indications of learning outcomes, consider the question of inferring intellectual growth through the analysis of a CPR™ session. Mapping gains based on observations of text samples has a fairly long history in the study of the writing process. Kenneth Pike makes a case for specific linguistic markers as evidence that certain mental operations have taken place.26 Lee Odell extends the idea that identifiable discourse features are evidence that corollary types of mental operations have occurred.27 More recently, Tamor and Bond have developed a text-analysis system that links “. . . task, cognitive system, and performance” (p. 113).28

Using this approach, a CPR™ session can be built around four features extracted from an examination of the differences between expert and novice performance in a writing task. Table B elaborates on the four “markers” used to track growth. Designing a rubric (set of performance standards) that asks students to identify facets of these “markers” (a) within the benchmark texts, (b) in their peer review samples, and (c) within their own submitted text generates data from which a measure of development can be inferred.

<table>
<thead>
<tr>
<th>COGNITIVE MANIPULATION</th>
<th>COMMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstraction</td>
<td>Experienced thinkers cluster specifics within a propositional frame-of-reference based on similarities, differences, likely outcomes, or proximity in time and space. Likewise, much of the reasoning process requires the thinker to move from parts</td>
</tr>
</tbody>
</table>
with whole, or to draw from a body of evidence some inference or enumerative generalization about the body of observations.

<table>
<thead>
<tr>
<th>Elaboration</th>
<th>Experienced thinkers have a variety of strategies for discovering appropriate content for a written artifact. They also have filtering mechanisms for selecting only those ideas appropriate to the audience and purpose of the rhetorical situation.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focus</td>
<td>Advanced thinkers are adept at establishing a purpose while at the same time providing a rich texture of detail. They do this by appropriately foregrounding ideas of interest while backgrounding other contextual elements.</td>
</tr>
<tr>
<td>Sequence</td>
<td>Expert thinkers formulate relationships of ideas in their minds and then convey these connections over large blocks of text through devices for coherence and cohesion that may be as simple as using transition words or as complex as interweaving a &quot;given&quot; in one sentence with the &quot;new&quot; in a subsequent sentence.</td>
</tr>
</tbody>
</table>

With experience, instructors become proficient at constructing assignments that elicit the desired learning outcomes. However, an even more exciting potential of the software is to develop a series of interlinked assignments so that gains in one CPR™ session can be consolidated and provide the foundation for the next assignment. As illustrated in Figure 4, sequential CPR™ sessions might reflect the same design coordination and instructional intentionality as would, for example, cascading mathematical problem sets that teach progressively more difficult domain concepts and cognitive manipulations.

![Course Learning Objectives](image)

Figure 4: “Iterative” CPR™ Sessions
Conclusion

The last decade has seen a national consensus emerge that engineering education must be renewed and improved. In particular, advocates for change call for innovations that:

- Incorporate findings from basic cognitive research to build a "modern" pedagogy of learning.
- Integrate the various performative skills foundational to engineering, rather than isolating them in artificial subsets.
- Teach the process as much as the product of engineering.
- Include advanced computer-mediated educational technologies for instruction.
- Accommodate individual differences in learning styles, gender, and background.
- Make the concepts of science and mathematics accessible through a naturalistic curriculum and authentic exercises, guided-inductive learning, and cognitive apprenticeship models of instruction.

In our presentation, we will demonstrate how Calibrated Peer Review™ may help to achieve some of these lofty goals.

Bibliography


Acknowledgements
Patricia Carlson’s participation in this research was funded by the National Science Foundation, Grant No. CCLI-9980867. We express our appreciation for this support.

PATRICIA A. CARLSON
Pat Carlson is Director of Writing Programs at Rose-Hulman Institute of Technology. She has been a National Research Council Senior Fellow, as well as having had several research fellowships with NASA (Langley and Goddard) and the Army’s Aberdeen Proving Ground. She has also been a research fellow at NASA’s Classroom of the Future located in Wheeling, WVA.

FREDERICK C. BERRY
Fred Berry is the Head of the Electrical and Computer Engineering Department at Rose-Hulman Institute. Long an advocate of engineering educational reform, he has been active in such innovations as project learning, entrepreneurial studies, and advanced IT in the classroom. He also participates in many outreach programs to interest middle school children in engineering.

DAVID VOLTMER
Dave Voltmer, Professor of Electrical and Computer Engineering at Rose-Hulman Institute of Technology, is currently focused on creativity and student design projects. He excels in the area of
electromagnetic fields, microwaves, and antennas. Spare moments are spent on long distance cycling, clawhammer banjo, or with his first grandson.