Potential Uses of On-Line Performance Assessments in Engineering Education: Measuring Complex Learning Outcomes and Processes

Gregory K.W.K. Chung[1], Thomas C. Harmon[2], Eva L. Baker[1]
[1]National Center for Research on Evaluation, Standards, and Student Testing, University of California, Los Angeles
[2]Department of Civil and Environmental Engineering, University of California, Los Angeles

Abstract

Modern engineering education is undergoing significant changes, particularly in the way engineering schools are adopting problem-based instruction to meet the changing demands of practice. Increasingly, engineering schools are requiring students to work on team projects that are open-ended with loosely specified requirements, produce professional-quality reports and presentations, consider ethics and the impact of their field on society, and develop lifelong learning practices. While there exist numerous implementations of courses adopting these methods to purportedly improve student learning, measuring the impact of problem-based instruction remains challenging. The existing evidence generally suffers from methodological shortcomings such as reliance on students' self-reported perceptions of learning (vs. more direct measures of learning), use of instructors' anecdotal reports of impact (vs. experimental manipulation), use of untested (vs. validated) measures, and small sample sizes.

This paper discusses the potential role of computer-based performance assessments in evaluating student learning on complex, open-ended problem-solving tasks. The discussion is framed within the research on computer-based performance assessments developed at the National Center for Research on Evaluation, Standards, and Student Testing (CRESST). Examples of three kinds of on-line performance assessments are presented as well as lessons learned.

I. Changes in Modern Engineering Education

Modern engineering education is undergoing significant changes, notably in the way engineering schools are adopting problem-based instruction to meet the changing demands of engineering practice[1-7]. Mastery of technical content is no longer sufficient. Increasingly, engineering schools are requiring students to work on team projects that are open-ended with loosely specified requirements, produce professional-quality reports and presentations, consider ethics and the impact of their field on society, and develop lifelong learning practices.

An implicit goal of this shift in curricula is to produce graduates who will be ready to assume engineering tasks upon graduation—that is, with the skills to develop solutions to problems under competing constraints of functionality, cost, reliability, maintainability, and safety[8]. Today's graduates are apprenticed for 1 to 2 years before they engage in meaningful engineering tasks. The half-life of engineering knowledge ranges from 2 to 7 years, yet the engineered systems are becoming more complex and multidisciplinary[9].

To encourage engineering schools to respond to these changes, the Accreditation Board for Engineering and Technology (ABET) has specified 11 criteria for graduating students in their Engineering Criteria 2000 (EC2000). Five of these criteria cover cognitive or teamwork skills. EC2000 specifies that graduating students should have (a) an ability to apply knowledge of mathematics, science, and engineering; (b) an ability to design and conduct experiments, as well as to analyze and interpret data; (c) an ability to function in multidisciplinary teams; (d) an ability to identify, formulate, and solve engineering problems, and (e) an ability to communicate effectively[1, 10].

II. Assessing Complex Learning

EC2000 specifies *cognitive skills* expected of graduating students. The criteria, while creating a difficult measurement problem, are far more consistent with what educators and industry value. That is, in many applied environments it is more informative to know whether students have learned conceptual knowledge, developed effective problem-solving skills, and developed the skills to work with others in a team. Of less interest is whether students have remembered definitions and formulas, or can solve problems "on paper." In environments that purport to promote the development of complex cognitive skills, evaluating the effect of instruction requires use of assessments that can provide evidence that students have (or have not) attained competency on these skills.

II.A) Simple Assessment Formats May Be Inadequate to Measure Complex Learning Outcomes and Processes. Two common test formats used to measure the impact of an instructional intervention are multiple choice tests and self-reports via a survey format. Typically, multiple choice tests are used to measure declarative knowledge of course content, and surveys are used to measure changes in attitudes, perceptions of learning, instructional impact, and frequency of use of particular behaviors or cognitive strategies. These formats are attractive because they do not require extensive classroom time, are easily administered, and are quickly scored. However, there are important disadvantages. It is difficult to measure processes and complex outcomes with multiple-choice formats, and surveys suffer from inaccuracies in recollection, are often biased toward socially desirable responses, and at best are indirect measures of the skills of interest. There may be quite a difference between students' perception of their task performance and the actual task performance.

II.B) Performance Assessments. While surveys and other simple response formats are convenient, they may not be the best format for measuring complex student learning. A more valid approach may be to use performance assessments—requiring students to demonstrate their competency on a task. Typically, the task is complex, requires an extended amount of testing time, is open-ended, and requires students to construct a response rather than select a response. Students are presented with a prompt and the resources needed to complete the task. The open-ended nature of the task requires students to frame the problem, create a plan to solve the problem, and then carry out the plan. Student responses are commonly products (often written) whose quality is determined by trained raters using a scoring rubric.

While performance assessments provide information about students' capability to deal with complex tasks, feasibility is the biggest obstacle to adoption. Performance assessments require extensive time to administer and are labor- and time-intensive to score (and thus expensive), and the results are unavailable for months. At CRESST, we have engaged in a program of research that has attempted to explore the extent to which technology can be used to deliver performance assessments on-line, simultaneously addressing issues of validity and feasibility.

II.C) CRESST On-Line Performance Assessments. All on-line assessment development has been guided by CRESST's model of learning. The model broadly characterizes learning as a function of content understanding, problem solving, self-regulation, collaboration, and communication. Table 1 lists brief definitions of each component[11-13].

Table 1
CRESST Model of Learning Components

Component	Definition
Content understanding	Understanding of subject matter content, which includes domain concepts, facts, principles, and procedures.
Problem solving	Activity directed at attaining a goal when the solution is not obvious. Problem solving involves content understanding, problem-solving strategies, metacognition, and motivation.
Collaboration/teamwork	Working with other members of a group or team to jointly complete a task.
Self-regulation	Includes metacognition, effort, self-efficacy.
Communication	The ability to express oneself clearly and effectively for various audiences and purposes.

The CRESST model has provided the framework for the design and development of several assessment technologies, and has led to the conceptualization of technology as an important component in the effort to measure complex student performance. Our assessment approach has been two-pronged. First, we employ a suite of assessment tools (rather than a single, monolithic tool) to measure specific components of the CRESST model. Second, we integrate the assessment tools with the task structure to provide a problem-based, relatively authentic context for students to work in.

Our view of computer-based assessments is that they provide new opportunities to measure student performance on complex, constructed-response tasks. Unlike traditional performance assessments, which are based on widely varying data sources and implementations, a computer-based task can provide a relatively stable measurement context (thus reducing methodological concerns), provide an open-ended environment that can vary in task complexity (thus providing opportunities for students to demonstrate a range of skills), and provide an extraordinarily rich environment to measure not only the products of student performance, but also the process of student learning. In the next section we briefly present examples of performance assessments developed at CRESST. For each product, we describe the task and related research findings.

III. Examples of CRESST On-Line Performance Assessments

III.A) Knowledge Mapping

III.A.1) Overview. Our use of knowledge mapping as a performance assessment requires students to demonstrate their understanding of a content area by creating a network diagram, where nodes represent concepts and labeled arcs describe how concepts are related. Individual concept-relation-concept tuples can be thought of as propositions. We have consistently found moderate, positive, and significant relationships between scores on students' knowledge maps (when scored against an expert's map) and other measures of content knowledge (e.g., essays[14]). Figure 1 shows a screen shot of our on-line knowledge mapper in human physiology.

III.A.2) Task. A typical knowledge-mapping task consists of providing the student with a fixed set of concepts and links, which allows for automated scoring of knowledge maps. A student is instructed to construct a map of his or her understanding of how the given concepts relate to each other. Students are free to configure their maps any way they choose, and they can add, delete, or move concepts and links at will. Students are given between 20 and 30 minutes depending on age and the number of concepts and links.

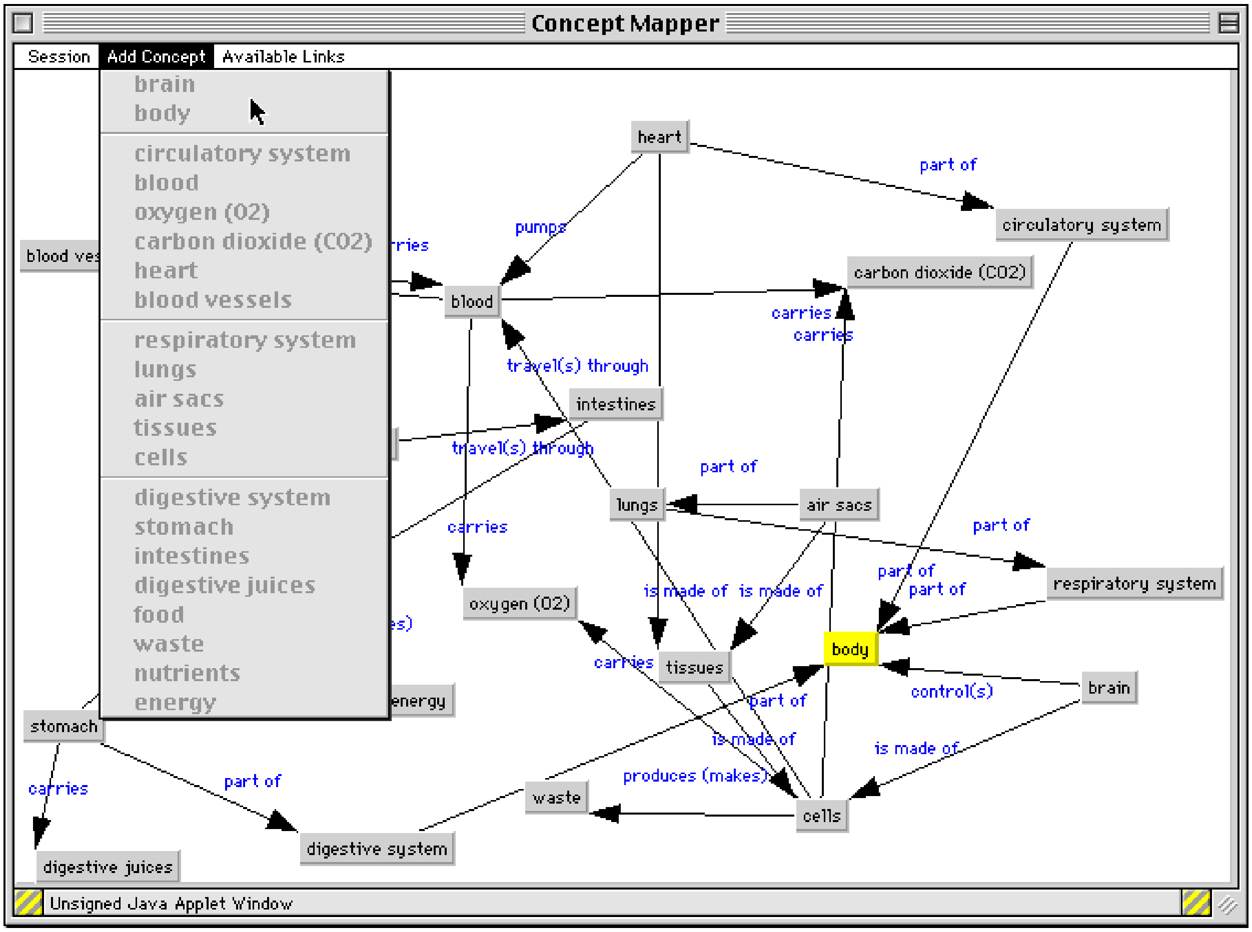


Figure 1. User interface of the on-line knowledge-mapping system, showing a knowledge map of digestive, respiratory, and circulatory systems.

<u>III.A.3) Research Findings.</u> We have tested knowledge mapping in elementary school (fourth/fifth grades), middle school (eighth grade), high school, college, and technical schools. In general, students quickly learn how to use the software with about 10 minutes of training.

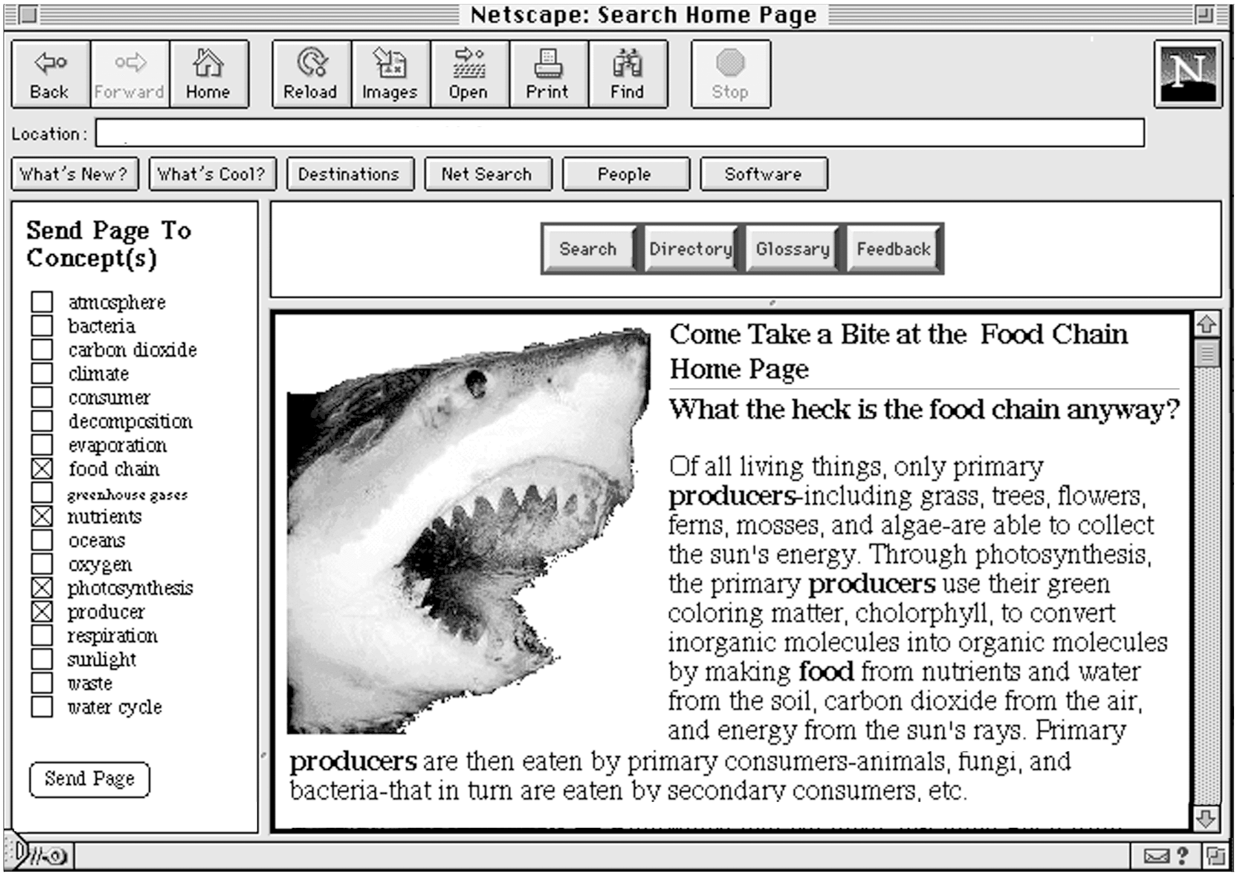
In our most recent use of knowledge mapping[15], we used knowledge mapping to assess student learning in a civil engineering capstone course. In the course, students were required to conduct a hazardous waste site investigation using simulation software designed specifically for the course. The software simulated physical processes as well as real-world engineering processes. A knowledge map task was administered before and after the site investigation project. We found students had significantly more deep propositions in their posttest knowledge map compared to their pretest knowledge map, suggesting instructional sensitivity [$M(SD)_{\text{posttest}}$ = 10.78 (2.37) vs. $M(SD)_{\text{pretest}}$ 8.47(3.87); $p < .05$]. Students also had more deep propositions than shallow ones in their posttest, although this difference was marginally significant ($p = .07$). Students' final course grade also was related to students' use of deep propositions ($r_{sp} = .49$, $p < .05$).

In an earlier study that investigated the effects of on-line knowledge mapping (in small groups) on learning of the respiratory, digestive, and circulatory systems of the human body, students who used the knowledge-mapping software over a period of 6 weeks gained a deeper understanding of the relationships between the scientific concepts, both within each system of the human body and between these systems. Students in the knowledge-mapping condition made more scientific [$M(SD)_{\text{mapping}} = 7.10(4.17)$ vs. $M(SD)_{\text{control}} = 4.77(3.28)$; $p < .05$] and also more highly principled propositions in their knowledge maps [$M(SD)_{\text{mapping}} = 1.15(1.35)$ vs. $M(SD)_{\text{control}} = 0.46(0.27)$; $p < .01$]. In addition, students' performance on the interconnection measures, which reflected linkages among the different systems, suggests that knowledge mapping helped students construct more interconnected understandings of the human body [$M(SD)_{\text{mapping}} = 4.40(0.50)$ vs. $M(SD)_{\text{control}} = 3.86(0.64)$; $p < .05$]. Students in the knowledge map condition had more interconnection propositions in their map compared to the control group. In addition, on an independent essay measure that asked students to explain the three systems and how they related to each other, students in the knowledge-mapping condition demonstrated a higher understanding of the interrelationships among the systems [$M(SD)_{\text{mapping}} = 2.48(1.24)$ vs. $M(SD)_{\text{control}} = 1.83(1.13)$; $p < .05$]. Positive changes in mapping scores for all students from pretest to posttest demonstrated the knowledge map task's sensitivity to instruction.

III.B) Integrated Simulation

III.B.1) Overview. In an attempt to increase the efficiency of the administration of an assessment, we created an *integrated simulation* assessment. By integrated we mean students are assessed on one or more components of the CRESST model one or more times as they go through the task. By simulation we mean approximating in a computer environment a real-world context for students.

The integrated simulation assessment was implemented with a combination of knowledge mapping and Web searching. For the Web site, we created an information space that was self-contained (i.e., students could not leave our Web site). This feature bounded the amount of information which students could access. Each page within the Web space was scored for relevance with respect to each concept in the knowledge map. Thus, we were able to link students' use of particular information to their knowledge map. Feedback on the quality of their knowledge map was available by a *Feedback* button on the Web page. The feedback was in the form of a report that was generated in real-time listing the concepts in a student's map that (a) needed "A lot" of improvement, (b) needed "Some" improvement, and (c) needed "A little" improvement.



Send Page To Concept(s)

☐ atmosphere
☐ bacteria
☐ carbon dioxide
☐ climate
☐ consumer
☐ decomposition
☐ evaporation
☒ food chain
☐ greenhouse gases
☒ nutrients
☐ oceans
☐ oxygen
☒ photosynthesis
☒ producer
☐ respiration
☐ sunlight
☐ waste
☐ water cycle

Come Take a Bite at the Food Chain Home Page

What the heck is the food chain anyway?

Of all living things, only primary **producers**-including grass, trees, flowers, ferns, mosses, and algae-are able to collect the sun's energy. Through photosynthesis, the primary **producers** use their green coloring matter, cholorphyll, to convert inorganic molecules into organic molecules by making **food** from nutrients and water from the soil, carbon dioxide from the air, and energy from the sun's rays. Primary **producers** are then eaten by primary consumers-animals, fungi, and bacteria-that in turn are eaten by secondary consumers, etc.

Figure 2. Example Web page that contains the concepts food chain, nutrients, photosynthesis, and producer. Bookmarking is handled by checking the appropriate concepts on the left side of the screen and clicking on the "Send Page" button.

III.B.2) Task. The integrated simulation task was made up of two stages. Students first created a knowledge map on environmental science. They were not given any supplementary information and thus their map was based on their existing prior knowledge of the subject. After completing their initial knowledge map, the maps were scored and general feedback returned to the students about which concepts "needed work." At that point, students were given access to the information space—Web pages on environmental science. Students could search for information, modify their knowledge maps, and request feedback. The one activity that was requested of students was for them to bookmark their concepts. That is, when students found information they believed relevant to a concept in their map, they were told explicitly to bookmark that page.

III.B.3) Research Findings. We tested the assessment system with eighth graders[16]. We were interested in measuring the extent to which students' information-seeking and problem-solving skills increased over the course of an academic year. Our study was a part of a larger program evaluation of the effectiveness of computer technology infusion into a school district.

Students were assessed during the fall and spring. Students were measured on the following information-seeking behaviors: (a) browsing, (b) focused browsing, (c) keyword searching, and (d) using feedback. Browsing was measured by counting the number of times a student clicked on a page from the directory, glossary, or a link within a page. Focused browsing was computed by counting the number of concepts that had three or more highly relevant (browsing) pages accessed. Keyword searching was computed by counting the number of times a student requested a keyword search. Use of feedback was computed by counting the number of times a student requested feedback. Table 2 shows the means and standard deviations for each measure.

Table 2
Descriptive Statistics for On-line Measures ($n = 20$)

	Pretest *M*(*SD*)	Posttest *M*(*SD*)	Difference between posttest and pretest
Knowledge map score	3.80(2.77)	10.40(4.45)	$p < .01$
Browsing	10.59(8.65)	15.33(5.88)	
Focused browsing	1.35(3.08)	8.61(4.29)	$p < .01$
Searching	1.53(2.24)	6.78(3.06)	$p < .01$
Using feedback	1.45(2.41)	5.00(2.79)	$p < .01$

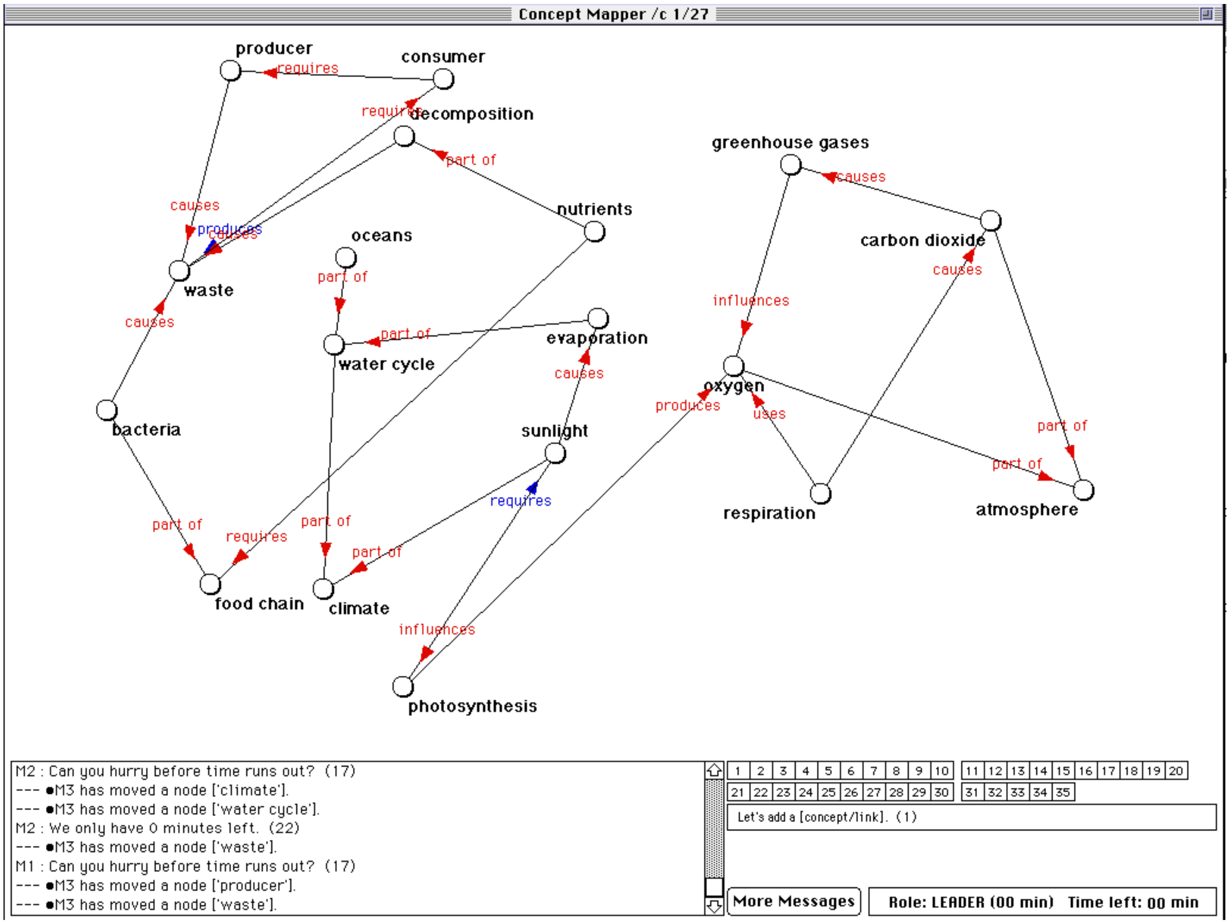
As Table 2 shows, there were significant differences between the pretest and posttest performances. In general, students improved on the measures at the posttest. We also found several significant correlations between search activity and students' knowledge map scores. The quality of a student's knowledge map score was related significantly to browsing ($r = .41$, $p < .01$), focused browsing ($r = .52$, $p < .001$), and using feedback ($r = .51$, $p < .001$). The more students engaged in any of these behaviors, the higher their knowledge map scores.

III.C) Collaborative Knowledge Mapping

III.C.1) Overview. Our attempts to assess team processes have centered around the use of networked computers to capture and measure—in real-time—team processes for individual students and teams. We use the taxonomy of 6 teamwork processes[17-19]: (a) adaptability—recognizing problems and responding appropriately, (b) communication—the exchange of clear and accurate information, (c) coordination—organizing team activities to complete a task on time, (d) decision making—using available information to make decisions, (e) interpersonal skills—interacting cooperatively with other team members, and (f) leadership—providing structure and direction for the team.

Existing approaches to measuring teamwork processes almost exclusively rely on observational methods (e.g., behavioral checklists, video- and audio-taped observation, and analysis of think-aloud protocols). These methods are labor-intensive and time-consuming and offer no opportunity for rapid analysis and reporting of team performance.

Our approach has been to provide team members with a set of predefined messages with which to communicate. These predefined messages are the sole means for members to communicate with each other, and the messages map on to the teamwork processes (adaptability, coordination, decision making, interpersonal skills, and leadership). As the team carries out the task, the software tracks which messages (and hence which categories) are used by each member. Thus, message usage provides an index of the kinds of teamwork processes members are using. Figure 3 shows a screen shot of the collaborative knowledge mapper user interface.



```
M2 : Can you hurry before time runs out?  (17)
--- ●M3 has moved a node ['climate'].
--- ●M3 has moved a node ['water cycle'].
M2 : We only have 0 minutes left.  (22)
--- ●M3 has moved a node ['waste'].
M1 : Can you hurry before time runs out?  (17)
--- ●M3 has moved a node ['producer'].
--- ●M3 has moved a node ['waste'].
```

Let's add a [concept/link]. (1)

Figure 3. Screen shot of the collaborative knowledge mapper (high-performing group).

<u>III.C.2) Task.</u> A typical task is to assign a group of three students to jointly construct a knowledge map. The members of the group are connected via a network and are assigned anonymous identifiers (i.e., "M1," "M2," or "M3"). One member of the group is initially assigned the role of the leader. Leadership rotates throughout the task, and only the leader can change the map. Non-leaders are instructed to advise the leader on a course of action. All computers are updated as changes occur (e.g., someone sending a message, or the leader making changes to the concept map); thus, the computers are synchronized with each other. Communicating between group members is done through the use of predefined messages. Members are given a list of 37 messages, and they send these messages to each other (e.g., "Let's link carbon dioxide to producer."). The rationale for using predefined messages is to provide a means to measure team

processes in real-time. All messages are coded *a priori* as reflecting a particular team process.

III.C.3) Research Findings. Our findings support the idea that students using our networked knowledge-mapping system were able to jointly construct a knowledge map. The number of concepts used by groups (across occasions) ranged from 7 to the maximum of 18, and the number of links ranged from 8 to 39. Thus, while some groups had sparse maps, others had fairly complex ones.

We found no significant correlations between most team processes and outcome measures; this was unexpected, especially for the use of decision-making messages. Our decision-making messages were designed to give participants the opportunity to consider alternatives among different links and concepts. We expected that the more teams used these kinds of messages the higher their team performance. We think the combination of a cognitively demanding user interface and the high knowledge demand of the task may explain these findings[18]. Exploratory analyses support this idea—low-performing teams, compared to high-performing teams, apparently spent more time focusing on the text box than on their knowledge map. Further, these participants were not able to profit as much from the discussion related to their knowledge maps, possibly because the large amount of messages they were sending was interfering with their ability to read and follow the discussion. The demands of the system may have induced too heavy a cognitive load on low-performing participants.

IV. Lessons Learned

Our approach to computer-based assessments is to use technology in ways that go beyond mimicking paper versions of assessments. We assumed that a far better use of technology would be to leverage the unique capabilities of technology to provide a clear advantage in terms of cost, utility, validity, reliability, access, or accommodation. Two major themes have emerged from our experience: (a) Technology affords unique measurement opportunities, and (b) measurement issues remain unchanged (e.g., need for reliability and validity).

Computer-based assessments provide the capability to measure complex learning. One of the most promising aspects of assessment technology is the capability to have students engage in constructed-response tasks and to measure both student performance and student learning processes. This capability is one of the most compelling reasons for using technology in assessment. Our experience to date points to the feasibility of developing powerful assessment environments that will provide authentic challenges to students. While this idea is not new and underlies many performance assessments, what is new—and only technology can feasibly provide this—is the capability to measure unobtrusively and more completely students' learning as they learn. The leverage computer-based assessments provide is the capability to design in (software) measurement points virtually at will and at any point in the interaction between student and computer. For example, in the CRESST integrated simulation students engage in a variety of activities—searching, browsing, and knowledge mapping. Embedding software measurement points to capture the events associated with these activities is a powerful way to measure the process underlying students' performance.

However, despite having the measurement capability, the placement of measurement points in the task must be driven by a cognitive model of student learning. Useful application of measurement points will occur only when the task is closely integrated with the user interface of the assessment system. The challenge is to design a task and interface that require students to interact with the computer. Ideally, this interaction reflects the results of students' cognition. We think that capturing behavior that reflects complex student thinking as students carry out the task will provide a far more complete picture of student performance.

Although much of our work has been devoted to issues of feasibility, utility, and cost, we recognize the importance of validating our systems not as technology systems, but as assessment systems that deliver high-quality measurement. Issues of validity and reliability become increasingly more important and more complex as new assessment formats go on-line. We have been generally successful with measuring individual performance and less so with measuring team performance. We are only beginning to understand the relationship between task, on-line behaviors and processes, student performance on complex tasks, and the usefulness of different measures toward characterizing student performance in an on-line environment. If on-line assessments are to be taken seriously as assessments of complex learning, future work must be directed at addressing the reliability and validity of these assessments.

Bibliography

1. Accreditation Board for Engineering and Technology, Inc. (2000). *Criteria for accrediting engineering programs*. Baltimore, MD: Author.
2. American Society for Engineering Education. (1996). *Assessment white paper: A framework for the assessment of engineering education*. Author.
3. American Society for Engineering Education. (1998). *How do you measure success? Designing effective processes for assessing engineering education*. Author.
4. Board on Engineering Education. (1995). *Engineering education: Designing an adaptive system*. Washington, DC: National Academy Press.
5. Coward, H. R., Ailes, C. P., & Bardon, R. (2000). *Progress of the engineering education coalitions*. Final report to the Engineering Education and Centers Division, NSF. Arlington, VA: SRI Internmational.
6. Dowell, E., Baum, E., & McTague, J. (1994). *Engineering education for a changing world*. Washington, DC: American Society for Engineering Education.
7. Meyers, C. W. (1995). *Restructuring engineering education: A focus on change*. Report of an NSF Workshop on Engineering Education. Arlington, CA: National Science Foundation.
8. Banios, E. W. (1991, September). Teaching engineering practices. In L. P. Grayson (Ed.), *Proceedings of the annual Frontiers in Engineering Education conference* (pp. 161–168). West Lafayette, IN.
9. Wulf, Wm. A. (2000). How shall we satisfy the long-term educational needs of engineers? *Proceedings of the IEEE, 88,* 593–596.
10. Besterfield-Sacre, M., Shuman, L. J., Wolfe, H., Atman, C. J., McGourty, J., Miller, R. L., Olds, B. M., & Rogers, G. M. (2000). Defining the outcomes: A framework for EC-2000. *IEEE Transactions on Education, 43,* 100–110.
11. Baker, E. L. (1995, September). *Finding our way*. Presentation at the annual conference of the National Center for Research on Evaluation, Standards, and Student Testing (CRESST), University of California, Los Angeles.
12. Baker, E. L., Abedi, J., Linn, R. L., & Niemi, D. (1996). Dimensionality and generalizability of domain-

independent performance assessments. *Journal of Educational Research, 89*, 197-205.
13. Klein, D. C. D., O'Neil, H. F., Jr., Dennis, R. A., & Baker, E. L. (1997, March). The five families of cognitive learning: A context in which to conduct cognitive demands analyses of innovative technologies. In H. F. O'Neil, Jr. (Chair), *An integrated simulation approach to assessmen*t. Symposium presented at the annual meeting of the American Educational Research Association, Chicago.
14. Herl, H. E., Baker, E. L., & Niemi, D. (1996). Construct validation of an approach to modeling cognitive structure of U.S. history knowledge. *Journal of Educational Research, 89*, 206-218.
15. Harmon, T. C., Burks, G. A., Chung, G.K.W.K., Baker, E. L. (in-press). *Evaluation of a Simulation and Problem-Based Learning Design Project Using Constructed Knowledge Mapping*. Proceedings of the 2001 American Society for Engineering Education Annual Conference and Exposition.
16. Schacter, J., Herl, H.E., Chung, G.K.W.K., Dennis, R.A., & O'Neil, H.F., Jr. (1999). Computer-based performance assessments: A solution to the narrow measurement and reporting of problem solving. *Computers in Human Behavior, 15,* 403–418.
17. O'Neil, H.F., Jr., Chung, G. K.W.K., & Brown, R. (1997). Use of networked simulations as a context to measure team competencies. In H.F. O'Neil, Jr. (Ed.), *Workforce readiness: Competencies and assessment* (pp. 411–452). Mahwah, NJ: Erlbaum.
18. Chung, G. K.W.K., O'Neil, H.F., Jr., & Herl, H.E. (1999). The use of computer-based collaborative knowledge mapping to measure team processes and team outcomes. *Computers in Human Behavior, 15,* 463–494.
19. O'Neil, H.F., Jr., Wang, S-L., Chung, G. K.W.K., & Herl, H.E. (2000). Assessment of teamwork skills using computer-based teamwork simulations. In H. F. O'Neil, Jr. & D. H. Andrews (Eds.), *Aircrew training and assessment* (pp. 245–276). Mahwah, NJ: Erlbaum..

Acknowledgments

The authors gratefully acknowledge the support of the National Science Foundation Combined Research-Curriculum Development Program, Award No. EEC-9700753 (Mary Poats, Director), and the patient perseverance of UCLA's C&EE 164 class members. In addition, the work reported herein was supported under the Educational Research and Development Centers Program, PR/Award Number R305B60002, as administered by the Office of Educational Research and Improvement, U.S. Department of Education.

GREGORY K. W. K. CHUNG

Gregory K. W. K. Chung is a Senior Research Associate at the National Center for Research on Evaluation, Standards, and Student Testing. His current work at CRESST involves developing problem-solving assessments for computer-based assessments, evaluating technology-based learning environments, conducting research on the measurement of team processes using networked simulation, and developing Internet-based assessment tools for diagnostic and embedded assessment purposes. Dr. Chung earned a Ph.D. in Educational Psychology from the University of California at Los Angeles, an M.S. degree in Educational Technology from Pepperdine University at Los Angeles, and a B.S. degree in Electrical Engineering from the University of Hawaii at Manoa.

THOMAS C. HARMON

Thomas C. Harmon is an Associate Professor in the Department of Civil & Environmental Engineering at the University of California, Los Angeles (UCLA). He received a B.S. in Civil Engineering from the Johns Hopkins University in Baltimore, MD, and an M.S. and Ph.D. in Civil Engineering from Stanford University. His research and teaching center around the fate and transport of contaminants in the environment. Dr. Harmon is the instructor for C&EE 164 Hazardous Waste Site Investigation and Remediation, and the lead content provider for ISIS program.

EVA L. BAKER

Eva L. Baker is co-director of the National Center for Research on Evaluation, Standards, and Student Testing and director of the UCLA Center for the Study of Evaluation. Her research is focused on the design and validation of technology-based learning and assessment systems and on new models to measure complex human performance in large-scale assessments. A professor in the Psychological Studies in Education Division at the UCLA Graduate School of Education & Information Studies, Dr. Baker also is co-chair of the Joint Committee on the Revision of the Standards for Educational and Psychological Testing. Professor Baker is involved in international, national, and state policy deliberations on assessment. She was recently selected to deliver the distinguished 1998 William Angoff Memorial Lecture at Educational Testing Service. Her publications include more than 450 chapters, books, and articles.