

CILT99 Assessments for Learning Abstracts

Presentations

Using the Scaffolded Knowledge Integration framework to implement new assessment practices in the Web-based Integrated Science Environment

Britte Cheng

<http://wise.berkeley.edu>

This talk will describe the on-line assessment approach of the Web-based Integrated Science Environment (WISE) and findings from a unit on water quality based on a theoretical framework (Scaffolded Knowledge Integration or SKI) which derives from findings of the Computer as Learning Partner project, and technology principles that derive from our Knowledge Integration Environment (KIE) Project (Linn and Songer, 1992; Linn, Bell and Davis, 1995; Slotta and Linn, in press). WISE allows students to access the Internet in an environment which provides technology supports informed by the four principles of SKI. Our environment is available to new collaborators at: <http://wise.berkeley.edu>. I will show how the Scaffolded Knowledge Integration framework informs assessment design by illustrating the four principles: a) Incorporate personally relevant context and accessible levels of instruction; b) Develop lifelong learning skills; c) Provide social supports for learning; and d) Make student thinking visible.

Using the SKI framework, our group concurrently designed new instruction and assessments for a two-week water-quality curriculum called Strawberry Creek. Challenging assessments were seamlessly incorporated into classroom activities to engage students' knowledge, observe conceptual development, and support further knowledge integration. Our group took advantage of several forms of on-line assessment including: 1) construction of causal models; 2) discussion; 3) note-taking; and 4) data analysis and added new approaches.

I will conclude with analysis of how the broadening range of assessment forms available through new and future technologies can support new assessment practices.

The Native American Distance Education Community: On the Road to Forming an Indigenous Distance Education Institute.

Evans Craig

<http://www.arc.unm.edu/Alliance/Tribal>

The Albuquerque High Performance Computing Center (AHPCC), a Center of the University of New Mexico, is supporting building infrastructure for the US Native American (American Indian) communities as part of a National Computational Science Alliance (NCSA) effort in Education, Outreach, and Training (EOT) through the Women and Minority Programs. AHPCC also participates and supports the American Indian Higher Education Consortium (AIHEC Computational Science) activity to plan, to build

consensus, and to seek funding for significant technology development and deployment at AIHEC Tribal Colleges. This is being accomplished in two areas, Community Outreach and K-12 Educational Curriculum Development, through two programs; Tribal Computational Science Program, <http://www.arc.unm.edu/Alliance/Tribal/>

The Albuquerque High Performance Computing Center is developing programs to reach Native American students in supercomputing and computational science, with emphasis on Internet technologies and other related distance learning technologies, in approximately 50 Native American serving schools. Working in concert with the American Indian Higher Education Consortium (AIHEC - 31 Tribal Colleges in the US), the North Dakota Association of Tribal Colleges (NDATC - 5 Tribal Colleges in North Dakota), the Montana Consortium (6 Tribal Colleges in Montana), the College of Rural Alaska (7 Native American serving Community Colleges in Alaska), the Hawaiian Community College System (7 Native Polynesian serving Community Colleges in Hawaii) and several other Native American serving Schools & Institutions, pilot groups are being identified to build the models needed to serve various Indigenous Communities (Schools, Tribes, and Native organizations).

This will be accomplished by providing additional support for national/regional Native American programs and specifically additional Alliance coordination for the AIHEC effort.

If you or your group is interested in participating in creating another possible future for our children, contact: Evans Craig, EOT Mgr., ecraig@arc.unm.edu, Albuquerque High Performance Computing Center. Albuquerque, NM 87131, 505-277-8249

Integrating Assessment Into the User Experience for Web-based Programs

Alex Cuthbert

<http://www.worldtrek.org>

The Odyssey is a free, non-profit, online educational project launched on January 15, 1999, already used by more than 700 teachers throughout the US and the world. Via The Odyssey website. K-12 students follow along and interact with a Team of adult volunteers doing a trek around the world. The Team documents the lives and cultures of local people, updating The Odyssey website twice/week, and supports live interactions, often with prominent figures. The Odyssey leverages the Web to:

- Engage students more directly in subjects seemingly far-removed - Make available to students information not typically available in school materials
- Reach students with diverse learning styles through the multimedia and interactive capabilities.

The Odyssey seeks to formalize an existing collaboration between:

- Dr. Bonnie Scott, an experienced educational program evaluator with WestEd
- Alex Cuthbert, a doctoral student at the Graduate School of Education at UC Berkeley and a computer programmer
- Jeff Golden, the Founder of The Odyssey and an experienced elementary and high school teacher

The goal is to design an assessment program that:

- Is integrated into the website as part of the users' normal experience following the World Trek storyline (to make gathering assessment information easy and immediate)
 - Stores this information uniquely for every user and makes it available on demand (to allow users and The Odyssey to assess progress)
 - Provides positive reinforcement for a user's success/progress (to support intrinsic motivations for learning)
 - Documents affective learning and the use of higher-order thinking skills. (to allow documentation in this key area among a large number of users spread over vast distances)
- Very initial conversations have considered a guided use of discussion boards, entry and exit quizzes, and surveys.

Assessing Knowledge Construction Processes in On-Line Learning Communities

Sharon Derry

We are developing methods for automated and partially-automated assessment of collaborative work within on-line learning communities. The testbed for this approach is STEP, the secondary teacher education project at the University of Wisconsin-Madison. The goal of the STEP project is to create a model for teacher education in which pre-service teachers learn about teaching as they collaborate with cooperating teachers and faculty mentors (including scientists and mathematicians) to design and evaluate learning environments in secondary schools. Because such project teams are difficult to coordinate and members are geographically dispersed, by the spring 2000 semester, teams will be expected to conduct a significant amount of their work on line. Our assessment approach involves taking advantage of the on-line environment's ability to collect evidence that can be used to draw inferences about collaborative process.

To help determine standards of evidence that are appropriate for judging on-line work groups as knowledge-construction entities, we articulated a theoretical framework based on four highly regarded views of social knowledge construction: the situative, sociocognitive, argumentation, and group information processing theories. Assessment was conceptualized as a problem of abductive inference, or case identification given particular theories and relevant evidence pertaining to them. The theories predict indicators that can be represented as nodes that are linked by subjective probabilities within a Bayesian network. Observable features of on-line interactions provide input to the net. Outputs are probability distributions indicating the degree to which a target group meets the normative standards prescribed by each theory. Such networks can produce general assessments and meaningful diagnostic profiles for groups and subgroups within large communities.

This assessment approach should enable early identification and intervention for individuals and groups experiencing difficulties within large communities. Such networks can produce general assessments and meaningful diagnostic profiles for groups and subgroups within large communities. This assessment approach should enable early identification and intervention for individuals and groups experiencing difficulties within large communities.

Utilizing Brain Research and Technology to Support Instructional Change

Patrick Faverty, Ed.D.

<http://www.acms.nvusd.k12.ca.us>

At American Canyon Middle School (Napa Valley Unified School District) we have used the current developmental and brain research along with the availability of current technology to build a middle school for the 21st Century. As an Apple Distinguished School, we are utilizing the current Apple products and services to support the shift from a focus on teaching to a focus on learning. We use project/problem/scenario based applications to motivate student learning. Technology is used as a tool, embedded in the program, not taught. Students are treated as learners, not as compliant subjects. This school represents the application of theory. It was designed and built not with incremental change but with radical departure from the current middle school model.

Partnerships for Post-secondary Programs

Robert Holloway

<http://www.kern.org/edoutreach/>

The mission of the Education Outreach Department is to help educators and community members reach their educational goals through distance learning and other technology applications designed to minimize disruption to their daily routine. Distance learning is a rapidly growing field that offers scheduling flexibility to individuals who need to expand their education but have little time to attend the traditional classes.

Selection of technology for development, establishing working relationships with other agencies, and finance issues are critical process steps. So far the initial commitments have been made, with some promising projects and some mistakes. The purpose of this session is to share this initial process.

Aligning TIMSS with Math Standards

Ken Koedinger

The poor performance of US students on the Third International Mathematics and Science Study (TIMSS) has received wide public attention. If technology innovations can lead to dramatically improved performance on such items, this should provide a clear message that such innovations do work and are worth pursuing. However, a focus on specific items has a serious risk. It can lead to "teaching to the test" in a way that the underlying general cognitive strategies and concepts the item is intended to test may not in fact be achieved. Thus, item selection must be accompanied with a clear articulation of the targeted cognitive processes. Furthermore, such processes should be important to students' readiness for the work place as well as future academics. National standards efforts provide a way to assess what processes have been judged by education experts to be important. Thus, it is critical to not only select TIMSS items that show where US students are lagging behind, but furthermore, to align these items with national standards to be clear about the underlying cognitive processes that are the true focus of instruction. This mini project seeks to identify key items for the TIMSS and align them with NCTM

standards and instructional materials associated with the cognitive tutors developed at CMU (e.g. Pump Algebra Tutor (PAT)).

Applications of the Intelligent Essay Assessor

Darrell Laham

<http://www.knowledge-technologies.com/>

The Intelligent Essay Assessor (IEA) is a set of software tools for automatically scoring the quality of expository essay content. The IEA uses Latent Semantic Analysis (LSA), which is both a computational model of human knowledge representation (Landauer & Dumais, 1997) and a method for extracting semantic similarity of words and passages from text. Simulations of psycholinguistic phenomena show that LSA reflects similarities of human meaning effectively (Landauer, Foltz, & Laham, 1998).

To assess essay quality, LSA is first trained on domain-representative text. Then student essays are characterized by LSA representations of the meaning of their contained paragraphs and compared with essays of known quality on degree of conceptual relevance and amount of relevant content. Over many diverse topics, the IEA scores agreed with human experts as accurately as expert scores agreed with each other.

At a minimal level, the IEA can be applied as a consistency checker for teacher scoring. Because the IEA is not influenced by fatigue, deadlines, or biases, it can provide a consistent and objective view of the amount and quality of relevant essay content. The IEA can further be used in summative testing for large classes or standardized tests, by either providing consistency checks or serving as an automatic grader.

At a more interactive level, the IEA can help students improve their writing by providing formative evaluations of the quality and scope of the conceptual content in their essays, and by automatically referring students to sources for missing knowledge. The IEA permits students to receive additional practice in written expression of knowledge without requiring all essays to be evaluated by the teachers. Because the IEA's evaluations are immediate, students can receive feedback and learn by making multiple successive revisions. This approach is consistent with the goals of the Writing across the Curriculum movement.

Laham, D. (1997). Automated holistic scoring of the quality of content in directed student essays through Latent Semantic Analysis. Unpublished master's thesis, University of Colorado, Boulder, Colorado.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review*, 104, 211-240.

Landauer, T. K., Foltz, P. W. & Laham, D. (1998) An introduction to Latent Semantic Analysis. *Discourse Processes*, 25, 2&3, 259-284.

Landauer, T. K., Laham, D., Rehder, B., & Schreiner, M. E., (1997). How well can passage meaning be derived without using word order? A comparison of Latent Semantic Analysis and humans. In M. G. Shafto & P. Langley (Eds.), *Proceedings of the 19th annual meeting of the Cognitive Science Society* (pp. 412-417). Mahwah, NJ: Erlbaum.

Using the Progress Portfolio as a tool for assessment: Supporting reflection and the construction of artifacts to document thinking

Sue Marshall

<http://www.ls.sesp.nwu.edu/sible/>

Computer-based learning environments provide unprecedented opportunities for scientific inquiry using large databases and sophisticated simulation and analytical tools. But these complex environments also create new challenges for students, who often become performance-oriented, lost in the activities of doing inquiry. This problem is compounded by the addition of computer technologies that encourage browsing. Rather than blindly forging ahead in their investigations, students need to be reflective inquirers, and to periodically step back to document and monitor their progress, review their understanding and conclusions, and communicate their understanding to others.

We have designed software, called the Progress Portfolio, to help students reflect on the inquiry process as they construct artifacts that represent the progress of their investigations. It provides tools to document these otherwise invisible processes: capturing states of work, documenting thoughts, observations, direction and purpose with annotation tools, organizing work through data management tools, and communicating the products of an investigation through presentation tools. These inscriptions of the work process provide tangible artifacts for learning about the process of inquiry through self-reflection and social discourse. Additionally, teachers can customize the Progress Portfolio with structured workspaces and prompts to support their own ideas about what is important for inquiry.

In collaboration with CILT partners, we are interested in exploring the Progress Portfolio's potential for supporting assessment activities. Although our current research efforts have focused more on the utility of the Progress Portfolio as a performance support system, some recent classroom pilots have pointed to interesting ways in which teachers used the tool in authentic assessment activities. The main mode of use wherein students capture and document their thinking is particularly suited for both student self-assessment and teacher assessment; it embeds the creation of portfolio artifacts into the process of an investigation.

A Web environment that assesses student preparedness to learn mathematics

Joyce Moore

This project's objective is to create an intelligent web-based environment in which students invent mathematical formalisms that simultaneously develop and assess their preparedness to learn from other forms of instruction. Most classroom assessments are not designed to be sensitive to whether students are prepared to learn, and thus can mischaracterize student understanding and lead to ineffective instructional choices. Our environment will include automated assessments that not only indicate what kinds of problems a student can or cannot solve, but also whether a student is ready to learn.

To test these ideas, we conducted a study in which undergraduates completed a lesson about a measure of variability and its notation. Invention students created procedures for capturing the variability of contrasting distributions of numbers; procedural students

practiced a procedure for measuring variability; and baseline students had completed an introductory statistics course. Students then evaluated nonstandard procedures for measuring variance. Invention students learned to reflect on the quantitative properties of distributions, and to evaluate statistical procedures in terms of their ability to differentiate those properties. Procedural and baseline students tended to evaluate a formula by whether it was the correct formula. We believe that assessments in which students evaluate unusual formulae are an excellent way to determine preparation for future learning. If we had assessed the students' ability to compute the standard deviation, procedural and baseline students' performance would have been superior. However, we do not believe that computational fluency adequately indicates how well students are prepared to learn. Given instruction on variability, we believe invention students would have learned much more.

We are designing a web-based environment to allow students to solve problems by entering invented notations. The system evaluates a notation to determine which relevant quantitative properties are captured, then presents additional problems that require identification of other quantitative properties. This prepares students to understand the point of the conventional notation during classroom instruction. Teachers can determine whether students need more opportunities to complete invention activities or are ready for instruction before the invention sequence is completed. Under this model, assessments become truly dynamic by affecting the course of learning.

Digital Portfolios

David Niguidula

The Digital Portfolio project of the Rhode Island Skills Commission stems from research at the Coalition of Essential Schools. The Skills Commission is a consortium of eight school districts, higher education, and business leaders, which will award a Certificate of Initial Mastery to high school students who demonstrate that they have achieved a set of standards. The Digital Portfolio project provides a web-based tool for students to collect evidence that they have met those standards. (The evidence will include student work samples, on-demand tasks, such as the New Standards Reference Examinations, and extended tasks drawn from an item bank developed by the consortium.) Current work has focused on what it takes to help schools put such a system in place; besides the technology, schools have worked to develop common understanding of the standards, and development of appropriate assessment tasks that will allow students to demonstrate those standards.

During the next year, plans are for teachers and other members of the consortium to develop rubrics and other assessment protocols to score such portfolios to ensure a high degree of reliability across the districts, ultimately leading to a statewide system for assessing portfolios. The increased reliability (plus the involvement of all members of the Skills Commission's consortium) will also help to position the Certificate of Initial Mastery as a useful indicator to business leaders, higher education faculty, and policy makers of a student's abilities, and, indirectly, improve faith in teachers' abilities to assess students against appropriate standards and goals.

21st Century Assessment: Using Technology To Support Student Science Assessment

Edys Quellmalz

Practitioners involved in educational reforms are citing the need for credible and feasible methods to assess student growth and to document accomplishments of innovative programs. It has become increasingly evident that technology can help by supporting the assessment of many student performances that have not been readily accessible through traditional testing methods. Technology-based assessments can represent natural or man-made phenomena, systems, substances, or tools that are too large, too small, too dynamic, too complex, or too dangerous to be adequately represented using a paper-pencil test or a performance test format. This presentation will describe a number of technology applications that support science learning that could be generalized and repurposed or redesigned for assessment purposes. In our view, many of these technologies can be used or adapted to elicit, collect, document, analyze, appraise and display kinds of student performances that have not been readily accessible through traditional testing methods. Furthermore, these technologies open the possibilities of ongoing, formative assessment of investigations-in-progress, in addition to the design of summative, end-of-project evaluation. We will present a conceptual model that portrays technologies that have been developed for seven very general components that occur in many project-based science inquiry curricula: (1) rich environments with authentic problems, (2) collaboration, (3) planning, (4) investigating, (5) analyzing and interpreting, (6) communicating and presenting, and (7) monitoring, evaluation, reflection, and extension. Technology applications have also been developed for assembling electronic notebooks or digital portfolios, and resource libraries. We argue a number of the technology applications that support science learning could be extracted, tuned, generalized, and re-purposed or re-designed for assessment purposes. These technology-based approaches could be made more widely available to assessment development teams of curriculum, assessment, and technology developers, and teachers. In addition, the affordances of some of these technologies could be used to support more explicit, systematic student assessment within their existing programming environments. . By identifying and adapting promising technology supports for scientific learning to improve the assessment of student educators can enhance further assessment development, encourage the integration of curriculum and assessment, index technological innovations to their potential assessment roles, and speed their implementation in other projects.

A Video Exploration of Classroom Assessment (CD-ROM)

Tina Syer

<http://www.irl.org/assess/assess.html>

A Video Exploration of Classroom Assessment (CD-ROM) highlights that no one assessment system is the right system for all teachers in all classrooms. Therefore, this CD-ROM puts forth examples of many techniques that teachers can try out and see if they work for themselves and their students. The disk's content evolved out of five years of discussions with 15 teachers who focused on classroom assessment.

Video examples give teachers a window into a real classroom, so that they can see what the proposed techniques look like in practice. The real, unstaged nature of the video sets this CD-ROM apart from other assessment tools. Teachers also appreciate the interview

videos because they offer a variety of perspectives on the different techniques. All assessment materials seen on the disk can be printed directly off of the disk, allowing teachers to implement new techniques immediately.

Teachers can use this CD-ROM in small groups or individually to informally learn about classroom assessment. Moving through the material at their own pace, teachers can self-organize their own small groups to explore the disk when the time is right for them. Working with the disk in this way allows teachers to share their own stories about assessment and discuss whether or not the techniques they see in the videos will work in their own classrooms. They can also brainstorm about changes they may need to make so that the methods will be successful. After small groups have worked with the disk, individual teachers often like to take the disk away and work with it on their own. The CD-ROM is useful both as an assessment reference and a discussion starter. Teachers K-12, in all subjects, find A Video Exploration of Classroom Assessment helpful.

Measures of Readiness to Participate in Technology-Infused Constructivist Reform Projects

Jason Ravitz & Hank Becker

There is considerable demand for an instrument that can assess teacher and school readiness for participation in technology-related projects, particularly innovative projects that are consistent with constructivist-based reforms. This presentation provides results from several studies that provide an empirical basis for predicting how teachers use computers, and, by inference, how they might respond to invitations or encouragement to participate in design experiments and other pioneering technology-based reform programs. The study of teachers in the National School Network shows correlations between a set of environmental conditions and teachers' extent of use of Internet technologies in their teaching. The Validation Study for Teaching, Learning, and Computing: 1998 (TLC) provides a concurrently validated index of constructivist philosophy and teaching practice. And the TLC national survey shows how a range of personal background variables, teaching philosophy, general pedagogy, and school environmental factors all combine to successfully predict how teachers use technology resources.

School Professional Cultures and Constructivist-Compatible Uses of Technology

Margaret Riel & Hank Becker

Using information from Teaching, Learning, and Computing: 1998, a national survey of 4,100 teachers across 1,100 schools including schools involved in major reform programs, this paper examines the hypothesis that in schools where professional cultures have merged—where teachers are at the center of the process for improving teaching practice—teachers' use of computers is more likely to reflect a constructivist perspective on organizing learning tasks for students. We also explore a second hypothesis: that under these same conditions, teachers are more likely to report that computers are responsible for changing their practice in more constructivist directions. To examine these hypotheses, information is used about teachers' informal contacts with other teachers, their leadership roles in professional development activities, pressures that do or do not exist at their

school for teachers to follow more "traditional" practices, and other related elements of professional culture.

Assessing children's oral reading via speech recognition technology

Susan Williams

Watch Me! Read software uses speech recognition technology to help an emerging reader (ages 4-8) read an on-screen book and create a multimedia performance of the book featuring the child's voice and video commentary. Although the recognition accuracy of this technology in actual classrooms is still being evaluated, it offers great promise for providing frequent formative assessments of children's reading.

Speech recognition for adults has been widely available for some time, but similar, commercially available systems based on children's voices and vocabulary did not exist. This was due, in part, to the lack of appropriate acoustic models for children's voices. Recently, researchers at IBM's T.J. Watson Research Center have developed the necessary models based on speech samples collected from over 1800 linguistically diverse children at 20 sites across the U.S. Despite this breakthrough, there are still numerous challenges for software developers using speech recognition for reading instruction. These challenges include: -- creation of an interface simple enough to be navigated by a young children without diminishing the experience of reading -- design of technology that tracks a child's progress through the text despite numerous pauses, omissions, repetitions, insertions, and mispronunciations -- improved recognition of alternate forms of pronunciation common in regional dialects and of children's speech at varying stages of maturation -- creation of appropriate feedback for situations in which the system can't be certain of the correctness of the child's performance.

Watch Me! Read is designed to overcome these problems in order to provide independent oral reading practice opportunities for young children. Development and refinement of the system has been guided by testing in several public school classrooms. These pilots have resulted in improved recognition of commonly mistaken words and the creation of a diagnostic tool that provides formative assessments of individual student's reading performance. Members of Vanderbilt's Learning Technology are currently collaborating with IBM's Watson Research Center to evaluate a system-wide implementation of Watch Me! Read in the Houston public school system.

MathLab: Technology assessment in M.S. mathematics

Andrew Zucker

This Theme Team has observed that it is well documented that the nature and content of the assessments used in classrooms have a strong influence on what and how teachers teach. This is clearly a barrier to the use of technology. Fortunately, the situation is changing. E.g., Virginia now requires that all its students learn to use computer spreadsheets. State-mandated assessments reflect this priority. Still, there is a long way to go to align assessment with new curricula that effectively use technological tools.

The MathLab project at SRI is developing software for teaching and learning middle school mathematics. One goal is to teach students to use a suite of computer-based tools

to solve mathematics problems. We address the issue of technology support for assessment in several ways. We believe the combination of these approaches will help provide powerful incentives to teachers and students to use technology in mathematics.

First, students' responses to problems will be in the form of open-ended, text-based communications in a real-life context. They may write a letter to Uncle Max or a note to a peer's mother. These responses will need to incorporate students' mathematical thinking. We believe this format will be appealing and, of course, the approach builds on the NCTM standards for Reasoning and Communication.

Second, the problems are such that students' responses will be assessed in part by how well they use the technology tools. This means that goals such as those adopted in Virginia will be explicitly integrated into the assessment of students' work with the software. The uses of technology will become part of students' constructed responses.

Finally, we are embedding a type of rubric for assessing the students' work, based on Polya's general scheme for solving mathematics problems. As an example, students will be expected to restate problems in their own words. They will know that this is one dimension on which their work will be assessed.

We hope that MathLab will contribute to a better alignment between the new goals for including technology in mathematics and effective classroom assessment.

Posters

Software Design Apprenticeships in Elementary Science Classrooms: Development of Evaluative Standards by Newcomers and Oldtimers

Yasmin B. Kafai, Cynthia Carter Ching UCLA Graduate School of Education & Information Studies & Cathleen Galas Corinne Seeds University Elementary School

Project-based learning approaches have become an umbrella term for reform-minded efforts in mathematics and science education. While many studies have demonstrated the academic learning benefits, another aspect of project learning which we call mindful practices has been more difficult to document. Mindful practices describe a rich set of practices such as students' evaluative standards, project management skills and tool competencies that develop through participation in communal activities. In our poster presentation, we will focus on students' development of evaluative standards within the context of a project-based software design activity. A class of 31 students composed of seven teams of fourth and fifth graders participated with their teacher in a three month-long science project in which they designed and implemented educational software simulations about neuroscience for younger students, third graders, in their school. Each team was composed of "oldtimer" students who had participated in a previous design project and "newcomer" students who were new and apprenticed by the oldtimers into the design project. In addition, eleven newcomers had served as software users when in third grade. Debriefing interviews were conducted with all students at the end about their project learning experiences. The answers are compared across the groups of users,

newcomers and oldtimers. Results indicate that oldtimer students have developed an expanded repertoire of evaluation standards when compared to newcomers. In the discussion we evaluate what we have learned about the nature of students' mindful practices and, in particular, their development of evaluative standards, and the design of a classroom learning environment to sustain the development of mindful practices through the apprenticeship model. We will also consider implications for the design of technology-based evaluation support.

Automated text analysis to assess cognitive outcomes of collaborative learning

Chris Teplovs, Darrell Laham, Carl Bereiter, Marlene Scardamalia, Peter Foltz and Tom Landauer.

Semantic analysis of texts based on word co-occurrences is decades old. Until recently, however, this kind of analysis was used almost entirely for descriptive and indexing purposes. Armed with both statistical models and cognitive models more powerful than those available in the past, researchers have demonstrated an amazing range of scientific and practical applications of Latent Semantic Analysis (LSA), among which is assessment of individual levels of content knowledge (Deerwester et al. 1990; Landauer & Dumais, 1997; Kintsch, 1998; Landauer et al. 1998). Knowledge assessment begins by analyzing a body of texts representing known levels of knowledge in a domain (for instance, textbooks at elementary to advanced levels). This produces a many-dimensional semantic space, and subsequent analyses can situate students' texts in this space and provide a measure of semantic distance between a student's text and the various reference texts. Learning can be demonstrated by showing that texts produced later in a learning sequence are closer to more advanced parts of the semantic space.

Whereas this kind of assessment has been applied to essay exams, texts derived from online discussions, problem solving, and the like represent additional challenges which telelearning researchers have recently begun to investigate. Among the problems are the quoting or paraphrasing of source material, which can produce spuriously high evaluations, and off-topic social and procedural discourse that might produce misleadingly low ones. The general strategy suggested by Landauer et al. for dealing with such problems is to do supplementary analyses aimed at measuring the problematic behaviors. In this way LSA can expand to yield additional variables of potential interest in their own right. In this poster, we present some early results on the use of LSA in assessing cognitive outcomes of online discussions.

References

- Deerwester, S. Dumais, S. T., Furnas, G. W., Landauer, T. K., and Harshman, R. (1990). Indexing by Latent Semantic Analysis. *Journal of the American Society for Information Science*, 41: 391-407.
- Kintsch, W. (1998). *Comprehension: a paradigm for cognition*. New York: Cambridge University Press.
- Landauer, T. K. and Dumais, S. T. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review*, 104, 211-240.
- Landauer, T. K., Foltz, P. W. and Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes* 25: 259-284.

Demonstrations

Measuring Students' Environmental Awareness and Skill in Data Analysis and Interpretation: A Web-Based Assessment for the GLOBE Program

Bill Penuel, Elaine Coleman

<http://www.wise.sri.com/servlet/globe.gat/GetFile/index1.html>

In this session, we will describe and demonstrate a web-based assessment task designed to help evaluate the Global Learning and Observations to Benefit the Environment (GLOBE) program, an international environmental science research and environmental education program. The GLOBE involves students, teachers, and scientists from across Earth in collecting, sharing, and analyzing data about Earth's dynamic land, air, water, and biology systems.

The assessment is designed to measure middle- and high-school students' qualitative understanding of ecology or environmental awareness (i.e., the recognition of the environment as a coherent set of interdependent and interconnected adaptive elements), and to measure students' skill in data analysis and interpretation using data of the kind typically collected by students in GLOBE. Because GLOBE students interact with the broader GLOBE community, which includes scientists, other GLOBE teachers and students, and they submit their data via the Internet, we designed a web-based task to measure student performance.

The web-based task requires students to (a) identify patterns and sets of relationships between ecological variables within images of a natural scene, and (b) to analyze climate data to decide where to host the next Winter Olympics, based on guidelines given to students for the best conditions under which the games might be hosted. In part b, students must develop an argument in favor of choosing a particular city and use the data to present evidence that supports their choice of site. A pilot conducted at a Bay Area high school implementing GLOBE indicates the assessment's viability for assessing student skill in data analysis and interpretation and for measuring student environmental awareness.

Performance Assessment Links in Science(PALS)

Edys Quellmalz

<http://pals.sri.com>

SRI International has received funding from the National Science Foundation to develop a greatly needed, specialized type of electronic library—an on-line, standards-based, interactive resource bank of science performance assessments, Performance Assessment Links in Science (PALS). States, districts, and teachers are seeking alternative assessments appropriate for measuring achievement of their science reform goals. Experience indicates, however, that the level of effort and costs of developing performance assessment tasks, scoring rubrics, and rate training materials, and then ensuring their technical quality are very high. The PALS Web site provides an innovative approach for sharing exemplary assessment resources, collaborating on the development

of new ones, and understanding how the use of standards-based performance assessment can advance educational reform at all levels of the educational system.

PALS on-line assessment resources are designed to serve two purposes and user groups: (1) the professional development needs of classroom teachers, and (2) the accountability requirements of state education agencies, districts, and specially-funded programs. Our partners include the Council of Chief State School Officers (CCSSO), three states (Connecticut, Illinois, and Kentucky) and two assessment consortia (the CCSSO State Collaborative on Assessment of Students and Standards (SCASS) for science and the California Systemic Initiative Assessment Consortium (CSIAC). The PALS project has two primary goals. (1) To develop a two-tiered on-line performance assessment resource library composed of performance assessment tasks for elementary, middle and secondary levels from multiple sources such as state assessment programs and consortia and national reference exams (NAEP, TIMSS, New Standards). One tier will be for use by teachers and professional development organizations. The second tier will be a password-protected, secure Accountability Pool of science performance assessments for use by state assessment programs and systemic reform programs (e.g., Systemic Initiatives). The Professional Development Pool will provide performance-based science assessments that have been used successfully in large scale (state or national) assessment programs and that have been released. The Accountability Pool will be composed of science assessments that have undergone a test development process and with data documenting their technical quality.

(2) To identify, study, and evaluate the effectiveness of policies, implementation models, and technical quality requirements for the use of the two tiers of PALS. The resources included on the PALS Web site consist of science performance tasks, student directions and response forms, administration procedures, scoring rubrics, examples of student work, and technical quality data calculated from field testing, and training packets for conducting on-line rater training. Performance tasks are compared for their alignment with the National Science Education Standards (NSES). In the second and third years of the project, SRI will embed on-line collaborations on science assessment within its virtual professional development center, TAPPED IN.

We encourage organizations to use any of the PALS resources in a variety of ways and share their experiences with us. You can visit the PALS Web site at: <http://pals.sri.com>

Technology Transfer To the Community

Ron Stevens

<http://www.immex.ucla.edu>

Over the last decade, the IMMEX Project at the UCLA School of Medicine has moved from software designed exclusively for future doctors to become a flexible instructional and assessment system for both secondary and college level courses. To date, the system has been used in over 70 schools in the Los Angeles area and involved over 300 teachers in training opportunities. Over 100 problem sets have been created by teams of high school, middle school, or elementary school teachers along with academic partners; many of these have been validated by teachers, university faculty, and now are used extensively in the classroom.

IMMEX can best be described as software that allows teachers to create "snapshots" of complex real-life situations that require students to make important investigative decisions. IMMEX also gathers information on the processes students use in making those decisions, giving teachers insight into the more subtle aspects of student thinking and learning. These insights help individualize education and provide valuable information on how the curriculum can be reformed.

Three important principles have guided the evolution of the IMMEX Project over the past decade. First, systemic educational reform - and the true realization of the potential of technology in the classroom - will not occur until teachers are confident and committed to the use of these technologies to enhance classroom practices. Second, educational technologies and software cannot be crafted for a single class, a single grade level, or a single school but must be part of a broader design that articulates curricular content and processes across disciplines and levels of ability. Third, the need for in-depth assessment of student performance, curriculum effectiveness, and teaching practices require that evaluation is a central component of all educational software development and implementation.

In this demonstration, Dr. Ron Stevens will present selected problems, provide details of the professional development activities and outcomes, and discuss how teachers are using the evaluation components of IMMEX in the classroom.