

Gibson, D. (2003). Network-based assessment in education. *Contemporary Issues in Technology and Teacher Education*, 3(3), 310-323.

# Network-Based Assessment in Education

[David Gibson](#)

*National Institute for Community Innovations, Vermont Institutes, USA*

---

Assessment, for both the improvement of performance and evaluating learners, is most effective when it reflects learning as "multidimensional, integrated, and revealed in performance over time" (Walvoord & Anderson, 1998). With that in mind, what do networks and new media have to offer that can assist and improve educational assessment? This paper asserts that network-based assessment offers fundamentally new possibilities for knowing what students know.

Networks, as used here, are an integration of the Internet, computers, intranets and humans offering new forms of instruction and assessment. Network-based assessment is emerging within educational testing and measurement, as well as online teaching and learning (Bennet, 1999; Mislavy, Steinberg, & Almond, 2000). In educational testing and measurement, there are examples of large-scale online testing and scoring, such as the online SAT and GRE. In online teaching and learning, there are numerous reflective writings, small sample studies of classes, and innovative experiments documented in the Society for Information Technology in Teacher Education (SITE) conference proceedings and journals over the last several years.

A great deal of the literature on online assessment is concerned with design and delivery to students in online courses. These studies primarily offer advice on ways to reproduce face-to-face methods and standards of quality, with some suggestions about ways to use standard telecommunications tools such as email and discussion threads to determine what students know and can do (e.g., Perrin & Mayhew, 2000; Robles & Braathen, 2002; Roblyer, & Ekhaml, 2000). Some policy groups underscore the same view of technology as "almost like" face-to-face settings.

For example, the first assumption of The American Distance Education Consortium's guiding principles for distance teaching and learning is that "the principles that lend themselves to quality face-to-face learning environments are often similar to those found in web-based environments." (*Editor's note:* The URL for this web site and others are located in the [Resources](#) section at the end of this article.) The definition of good teaching articulated by the American Association for Higher Education's "Seven Principles of Good Practice in Undergraduate Education" remained the same after being revised for online teaching. I am, nevertheless, convinced that the new media means more than "almost face-to-face"; new media has changed the landscape of teaching, learning and assessment.

Researchers who agree that the landscape has changed are interested in the unique affordances of network-based teaching and learning and have begun to articulate a general framework for assessment. For example, some have outlined a broad framework for assessment, from which we can build a new architecture for network-based assessment (Almond, Steinberg, & Mislavy, 2002;

Pellegrino, Chudowsky, & Glaser, 2001). Others have begun to outline techniques for estimating the best problem or resource to present to a learner given a set of problems already completed (Almond & Mislevy, 1999; Hawkes & Derry, 1989; Steinberg & Gitomer, 1996).

Innovators in the field of latent semantic analysis and applications of Bayesian theory are beginning to show essay scoring results that rival human scoring (McCallum & Nigam, 1998; Rudner & Liang, 2002). Others, using neural net analysis, can categorize the problem-solving approach of learners in a web-based environment (Stevens, Lopo, & Wang, 1996). These examples begin to point to a qualitatively new role for Internet-based technologies in assessment — one that is rich with multimedia, responsive to learners, flexible over many situations, unobtrusive to the natural actions of learning, and assisted by artificial and network intelligence.

Network-based assessment methods and media have the potential to transform how assessments help us know what students know. The new technology enhanced conception of assessment stands in contrast to the traditional view of assessments as "tests" of knowledge remembered. Instead, the new perspective on assessment seeks to create a body of "evidence" of usable and available knowledge observed in natural settings of the learner (Greeno, Collins, & Resnick, 1997; Mislevy et al., 2000). In contrast, some have argued that the fundamental adjustment needed in online assessment is primarily due to a lack of face-to-face contact (O'Malley & McCraw, 1999).

Others have pointed to the difficulties of preserving secrecy of items in traditional item response theory tests (Perrin & Mayhew, 2000). But the development of effective and reliable assessments for online students requires a great deal of innovation and departure from traditional practices (Ryan, 2000). Because technology mediates learning in new ways, it engenders new forms of knowledge, as well as possibilities for documentation and analysis (Bransford, Brown, & Cocking, 2001; Bruce & Levin, 1997; Greenfield & Cocking, 1996; Kafai, 1995) and should, therefore, focus our attention on expanding our conceptions of assessment.

What follows is a brief outline of the mediating role of technology and what those new affordances mean for teaching and learning. That discussion is followed by an introduction to the major elements in contemporary designs for assessment systems and ways network-based assessment processes can take advantage of these perspectives. Finally, the article uses a case study example to illustrate the new elements in use in a network-based assessment system.

## **Technology as Mediator in Teaching, Learning, and Assessment**

"When used to its full potential, the computer is more than a tool for efficiency and automation: it transforms thinking and creates new knowledge." (Kalik, 2001)

Technology mediates knowledge and thus fundamentally changes learning, teaching, and assessment — what we can know about what students know (Bruce & Levin, 1997). This view contrasts with traditional views of the computer as an automaton, a tool for efficiency in searching, organizing, and communicating knowledge and a place to store information. The mediating role of technology, extended to network-based assessment, also contrasts with the traditional view of an assessment as providing documentation of what has been learned. In place of these views, the computer combined with global networks is seen as an extension of thinking, inquiry, and expression that transforms the reach and power of the mind. In this section, the claim is made that a new landscape of learning has appeared with network-based technologies, and the changed environment is briefly outlined, with implications for network-based assessment.

We begin by considering "the effects of technologies as operating to a large extent through the ways that they alter the environments for thinking, communicating, and acting in the world.

Thus, they provide new media for learning, in the sense that one might say land provided new media for creatures to evolve" (Bruce & Levin, 1997). A partial listing of the fundamentally new affordances made possible through network-based technologies includes:

1. *Access to an abundant multimedia global knowledge storehouse.* Network-based resources include digital libraries such as the NICI Virtual Library, real-world data for analysis, and connections to other people who provide information, feedback, and inspiration, all of which can enhance learning and assessment. Furthermore, today's Internet, as vast as it seems, is just the beginning of network-based multimedia and represents a small fraction of global knowledge available now in digital form — with an even vaster array of information in non-digital form that is quickly finding its way onto the web. "Deep web" and "interoperability" methods will soon make available several orders of magnitude more information than is available today. (See the W3C web site for more information on these topics.) In addition, that access is multimedia—involving texts, images, sounds, digital video, and more—which evidence suggests is a rich and effective environment for learning (U.S. Department of Education, 2000). Design and delivery of multimedia assessment is in its infancy, as is the use of globally linked multimedia resources in network-based assessments.
2. *A vastly expanded range of tools for inquiry and expression.* New network-based media is more than a storage medium for information; it is a new environment for inquiry, expression, construction, and communication. The frontiers of science illustrate this, as they are dominated by new visualization, aural, and analytic capabilities that have only become available within the last few years (Novak, 2002). Yet, teaching, learning, and assessment have yet to take full advantage of these developments. Technologies can help users visualize difficult-to-understand concepts — a boon for learners as well as teachers. In assessment for example, that enhanced capability can help teachers see the conceptual growth of the learner or view the structural shape of performance of a group of learners (Stevens, 1991). Learners can work with modeling software similar to the tools used in scientific and work related environments, which can "increase their conceptual understanding and the likelihood of transfer from school to nonschool settings" (Bruce & Levin, 1997). Most important, the new range of inquiry and expression changes the nature and extent of knowledge and its acquisition.

For example, new forms of computational proof and demonstration have opened up branches of mathematics that were considered intractable in the past (Wolfram, 2002), and the role of computational simulation has taken on enlarged importance for the sciences, including cognitive science (Holland, Holyoak, Nisbett, & Thagard, 1986). Additional examples include the use of visualization, simulation, and network-based communities in the discovery of new chemical materials, the human genome project, astronomy, and physics. In network-based assessment, the techniques of remote sensing can lead to unobtrusive observations of learners who, rather than taking a test, are making decisions, constructing artifacts, and thinking aloud as they work in a naturally productive setting.

3. *More interactive and responsive applications.* "Because many new technologies are interactive, it is now easier to create environments in which students can learn by doing, receive feedback, and continually refine their understanding and build new knowledge" (Bruce & Levin, 1997). However, thus far much of the development of interactivity has taken place in home-based entertainment and educational games. Those applications are just beginning to tap the potential of network-based technologies, for example, in globally extended massively multi-player online role-playing games (MMORPG). In addition, as global network-based interoperability takes hold, new forms of responsive dissemination are emerging (Gibson, Knapp, & Kurowski, 2002), which are making it possible to envision learning environments in which the active status of the learner launches a

variety of software agents that search the global knowledge store. Agents can return with links to resources and people and present the next best item for consideration, study, or enjoyment. Thus, the creative impulses of the learner can be met by interactive multimedia technology, providing new avenues to draw upon a learner's strengths, interests, and aspirations.

4. *New social networks and schools of thought.* The Internet makes unrestricted social networks possible, as well as the possibility that new forms of school and other social organizations may arise in response to the thoughts and actions of groups who share common goals. As network-based technologies become embedded in daily social life, they tend to become invisible; "we focus less on the fact that they may be consciously employed as a tool to do a task, and come to see the task itself as central, with the technology as substrate" (Bruce & Levin, 1997). Today, for example, some people can contact nearly everyone they work with at anytime via an electronic message system. Yet for all the ways that technologies are becoming an invisible part of our lives, education is still largely organized around traditional face-to-face settings, except for a few "leading edge" projects.

Perhaps most importantly, the new social communications systems are interactive, and conducive to active, engaged learning. Network-based assessment systems, for example, are just now emerging that take advantage of social groups (Gibson, 2002). Students can choose what to see and do, and the media can unobtrusively record as well as extend what they learn. Learning can be, more than ever and in ways not possible without networks, driven by the individual needs and interests of the learner in balance with the social goals of education. (Bruce & Levin, 1997; Friedrichs & Gibson, 2001).

In addition to these new affordances, network-based assessment systems can also take advantage of recent advances in the science of assessing thinking and learning (Pellegrino et al., 2001) including the following:

1. *Complex performances can be supported and documented* in network-based assessments via multimedia, multileveled, and multiconnected bases of knowledge.
2. With many instances of the learner interacting with applications in different times, places and contexts, network-based assessments can build a long-term record of documentation, *showing how learners change over time.*
3. *Analysis of expert-novice differences* can be facilitated across groups, across space and time, drawing from an evolving common knowledge store.
4. The interactive potential of network-based assessment opens up new possibilities for *fostering and determining metacognitive skills* of the learner.
5. Emerging capabilities in metadata generation offer the potential for *identifying the problem-solving strategies* of learners.
6. Unobtrusive observation techniques combined with libraries of evidence and tasks can make possible *timely feedback to learners and teachers* and matching of current needs with best "next step" materials, tasks and challenges, including tasks that involve transfer of learning to new contexts.
7. Network-based assessments can include *statistical analysis and displays* of information to assist learners and teachers in making inferences about performance.

This section briefly outlined eleven ideas — four broad categories of the new mediating potential of network-based technologies and seven criteria for assessment systems grounded in recent research — that might form a set of criteria or an agenda for developing a network-based assessment system. The next section outlines two of the potential implications of these elements on teaching, learning and assessment — the potential for developing adaptive expertise in learners, and an expansion of methodologies for assessing the range of knowledge and skill of the learner.

## **Implications of the New Affordances on Teaching, Learning, and Assessment**

Teaching and learning supported by the elements outlined in the previous section can shift from an overdependence on short-term memory and using procedures to creative, interdependent, and iterative processes of knowledge construction. Such a shift is necessary in order to deal with massive access to information, essentially limitless bounds for social interactions, and completely novel ways of interacting with and expressing information and ideas. Some commentators have noted this shift as part of a larger movement from an industrially based economy to a knowledge-driven society, bringing with it new demands of flexible and adaptive responses by learners. As Kalik (2001) observed:

The chief implication of a shift to "knowledge work" is that knowledge workers adapt their responses to a given situation instead of carrying out standard operating procedures. They attempt to understand what would be an appropriate response to a situation, then marshal the necessary resources and capabilities to get it done. They are good problem solvers.

These new cognitive demands on learners are a sign of what cognitive scientists call "adaptive expertise" (Bransford et al., 2000), which network-based assessments can be designed to measure. In systems designed to develop and measure adaptive expertise, learners are viewed on a continuum with other knowledge workers, including their teachers. Teachers, in turn, who want to be flexible and adaptive themselves, must become curriculum designers who assist learners in planning, marshalling resources, and validating that learning has taken place. Assessment methods and reporting must follow these trends in order to stay well aligned and to measure what is important, as well as what is actually taught and learned. Assessment designers thus need to understand and begin with a model of cognition that includes problem solving, analysis skills, and varying degrees of expertise. (Pelligrino et al., 2001).

The cognitive model elements of problem solving, analysis, and adaptive expertise are measurable within a performance range that differs for various learners, identified by Vygotsky (1978) as the "zone of proximal development." The increased interactivity and responsiveness of network-based assessments improves the measurement of the top of the zone. The zone represents the difference between what learners can do with help and what they can do without guidance and, thus, has a minimum as well as a maximum that should be measurable by assessments. We can assume that what a learner can do without help or guidance, as is often the case in traditional "test" settings, measures near the bottom of the zone.

In summary, the previous two sections begin to show that there is a unique new potential for network-based assessments to measure what students and teachers know and can do. The forms of delivery and interactions are dramatically different from traditional assessments, giving rise to new possibilities for forms of collecting and analyzing information that are better aligned with what we know about how people learn. To take advantage of the new potential, researchers and developers can take advantage of a new model for the design and delivery of assessments that can be applied to network-based technologies, including those that combine computers and human expertise.

### **A New Model for Assessment Design and Delivery**

Recent work has led to a new model of assessment design. Pellegrino et al. (2001) showed that every assessment, regardless of its purpose, involves three fundamental components: "a model of how students represent knowledge and develop competence in the subject domain, tasks or situations that allow one to observe students' performance, and an interpretation method for drawing inferences from the performance evidence thus obtained." In addition to this triadic

internal structure, assessments only operate successfully in a context in which learners have been given an opportunity to learn, for example, through curriculum and instruction. The assessment tasks or situations must be aligned with actual opportunities to learn in order to provide good information to any intended audience (learner, teacher, public) for an assessment.

Deepening and extending the three-part model, Almond et al. (2002) outlined several submodels in the design as well as delivery of assessment systems (see Figure 1). Relating their core models to the above and including a brief description produces an architecture for building network-based assessment systems:

A model of how students represent knowledge and develop competence in the subject domain.

*Student model*— specifies the dependencies and statistical properties of relationships among variables that lead to claims about the knowledge, skills, and abilities of the learner. A scoring record holds the values of those variables at a point in time.

Tasks or situations that allow one to observe students' performance.

*Task model*— specifies variables used to describe key features of tasks (e.g., content, difficulty), the presentation format (e.g., directions, stimulus, prompts), and the work or response product (e.g., answers, work samples)

*Presentation model*— specifies how a task will be rendered (e.g., on screen, audio, on handheld)

An interpretation method for drawing inferences from the performance evidence thus obtained.

1. *Evidence model*— specifies how to identify and evaluate features of the work or response product and how to update the scoring record.

Almond et al., (2002) described two submodels that fall outside of the Pellegrino et al., (2001) model, since they deal more with construction and delivery than design.

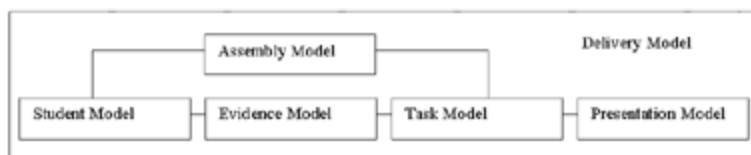
Methods for assembling and delivering assessments.

*Assembly model*— specifies how an assessment will be assembled (e.g., iterative and interactive as online, redundant and complete as on paper)

*Delivery model*— is a catch-all container for all of the above models and includes constraints that do not fit elsewhere (e.g., security, backup, administration control)

The network-based system illustrated by the case study in the following section extends the above four-system model of assessment to include a globally shared library of resources behind each submodel described above. As the new assessment architecture becomes a common vocabulary among assessment designers, the possibility increases for sharing that vocabulary and structure as a web-based ontology for searching and finding assessment objects. Utilizing XML and RDF schema, researchers are beginning to develop interoperable systems that allow the creation of a wide variety of locally relevant assessments from globally available resources.

The essential tools and approaches of the architecture have been developing within the global "World Wide Web Consortium" (W3C), which develops interoperable technologies, specifications, guidelines, software, and tools, to lead the World Wide Web to its full potential. W3C is a forum for information, commerce, communication, and collective understanding.



**Figure 1.** A model for design and delivery of assessment systems, taken from Almond, Steinberg, and Mislevy (2002).

With a general model of an assessment system available and the mediating potential outlined above, we next turn to case study examples to illustrate the new elements of network-based assessment.

### Case Study Examples

The Educational Technology Integration and Implementation Principles (eTIP) Cases project, funded by the Preparing Tomorrow's Teachers to Use Technology -Catalyst grant program has built a number of online simulations intended for preservice teacher education programs. The simulations are set in the context of imitation web sites for several different schools and provide online, multimedia case-based instruction and assessment that can help preservice teachers and teacher education faculty learn about effective integration and successful implementation of educational technology.

The content of the cases draws from the National Educational Technology Standards, the Interstate New Teacher Assessment and Support Consortium standards, and the National Staff Development Council standards for staff development programs, as well as the experience of the case writers. A matrix of "sim-schools" has been created, in which rural, suburban, and urban settings were crossed with high-performing, mid-performing, and low-performing student results and staff development data. This produced a rich simulation context of schools, in which questions of technology innovation, teacher preparation, and staff development can be raised.

Each question creates a new "case." Several cases are brought together into a "problem set," and several problem sets can exist within one over-arching "problem space" created by the matrix of school types and characteristics. The flexibility and reusability of the major elements — cases, sets, and spaces — form the heart of the task model of the network-based assessment system. Two items contribute to the definition of each case within a problem set: the prologue, which sets out the challenge or situation and requests a student work product or response, and a table of weights, which determines the relevancy of the description items for a particular prologue. The relevancy table is used in the analysis of the resources that learners use while constructing their responses and, thus, functions as an idealized student model, detailing how an expert learner would view the relevancy of the contents in the site concerning the question at hand.

The presentation model for each case includes a unique prologue that frames the challenge or situation and calls for learners to make a decision and produce a response. Then, through a menu of hyperlinks, learners explore the range of information available to use in developing their response to the challenge. Context-rich descriptions of classroom and school settings are presented in text, visual, and audio formats. The multimedia elements and descriptions, which are also items or data variables for the assessment analysis in the evidence model, can be selected in any sequence. The hyperlinked items, the scenario posed, and the case's weighted contents constitute a specific problem space context through which learners navigate as they construct their response. Responses can be either machine or human scored, including remote scoring by

social networks of peers and experts. While the overall approach is constructivist, each case is not so open-ended and complex as to overwhelm users (Mayer, 1997).

To illustrate, "H. Usher Elementary School" is one of the sim-schools set in an urban location. It is a medium size school, with about 700 students. Although the learners do not know it when first encountering the simulation, H. Usher is a high performing school. According to the prologue, its faculty and administration perceive that the school has a problem with student results. The prologue to the simulation states that the second-grade students are not meeting the district goals and need to advance their reading comprehension at a faster pace. The learner is challenged to explore the school context to understand more about the learning environment in which this situation has occurred, decide what went wrong, and write a response explaining what to do differently as a second-grade teacher, given the resources that are available.

In this case, as learners try to figure out how this school works, they find evidence of a high performing staff development program and a school that outperforms the district and state. How will inexperienced future teachers view this situation calling for a complex performance (deciding what information is relevant and not, deciding what options might work in this setting, writing about their decision and justifying it)? The eTIP Cases project is designed to help preservice educators and future teachers find out.

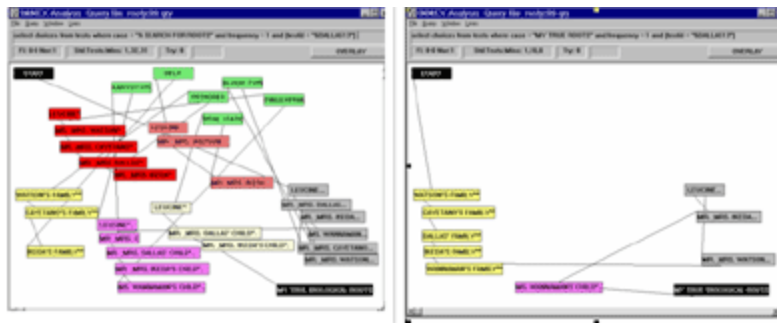
Problem spaces like H. Usher contain many potential challenges or situations and solution paths. The content of the school's web site contains an abundance of rich information that allows several prologues to be created. Each prologue can ask different meaningful questions, such as questions about technology integration in the fourth grade, the principal's attitude toward peer support systems, the state of professional development, the needs of students given their performance on state assessments, and so on. This allows a single problem space to function as a generic task and presentation model over many "cases" and "problem sets."

### **Visualizing and Analyzing Problem Solving**

As the learner navigates around the problem space, reading, watching, and listening to the items, the application tracks the sequence and timing of items used and collects the learner's response product in the form of essays, which can be scored by the teacher and others. By tracking the learner's use of items, the application creates a performance record as part of the evidence model that documents the development of learner-reasoned relationships among problem space variables. In addition to the performance record, which is captured as an unobtrusive observation, a work product in the form of an essay is gathered. The narrative of the essay stores information directly from learners concerning their decisions, rationale, and what was meaningful in their analysis.

The heart of the eTIPs application is the IMMEX system, developed first for chemistry and the physical sciences, and extended by the Vermont Institute of Science, Math, and Technology (VISMT) to include an essay scoring capability and an online campus to help introduce new teacher education faculty to the process of using the cases. IMMEX provides timely feedback to learners and teachers through a number of quantitative displays of the performance records of users — including visual displays called "search path maps" (Stevens, 1991). In these maps, each student action is represented by a rectangle that is colored to visually relate items closely linked by content, concepts, or type within the content domains in the problem space. These icons are organized in different configurations, and lines connect the sequences of items selected by the students while performing the case (Figure 2).

Teachers can use these maps in multiple ways. For teacher educators the maps provide a validity check on their classroom preparation and emphasis, as well as a source of information about student performance differences.



**Figure 2.** Sample student search path maps. The map on the left represents a student who explored many menu items, making a complete search of the problem space. The performance of the student at right shows that only two general areas of the problem space were explored, indicating a lack of grasp of the concepts underlying the problem. Taken from eTIPs documentation.

By comparing earlier to later maps, one can determine a learner's progress overtime through refinements of problem solving approaches. Providing students with their own maps encourages reflection, which can be combined with in-class discussion and writing. Search path maps are particularly important for examining and promoting the metacognitive aspects of problem solving, such as persistence, elimination of alternative hypotheses, efficiency, confidence, and certainty. The maps also supply artifacts for developing problem-solving scoring rubrics and for discovery of problem solving strategy patterns across groups of performances, including by artificial neural network analysis (Kanowith-Klein, Stave, Stevens, & Casillas, 2001), an approach that helps automate the interpretation process through pattern recognition.

VISMT enhancements to IMMEX add essay scoring to the feedback available to the learner. Essays offer a way to enhance the metacognitive skills of students. The application supports the creation of scoring rubrics (Figure 3), which have been used by the eTIPs project to create six rubrics, one for each eTIP. The rubrics are viewable and printable by teachers, and can be used to guide essay writing. An essay grading form is provided to record essay scores (Figure 4) and reports can be generated that compare performances across several essays in a problem set.

Criteria:	Level 1	Level 2	Level 3
1. Validation: Explains the <b>central classroom challenge</b> in the case	Does <b>not</b> present an explanation of instructional planning in relation to curriculum goals or standards	Presents a <b>limited</b> explanation of instructional planning in relation to stated curriculum goals or standards	Articulates a <b>clear</b> explanation of instructional planning in relation to stated curriculum goals or standards
2. Evidence: Identifies <b>case information</b> that must be considered in a decision about using technology aligned with curriculum goals or standards	Does <b>not</b> identify aspects of case information, including appropriate technology uses, that align with curriculum goals or standards	Identifies aspects of case information, including appropriate technology uses, <b>without</b> explanation or examples of how these align with curriculum goals or standards	Identifies aspects of case information, including appropriate technology uses, <b>with</b> explanation or examples of how these align with curriculum goals or standards
3. Decision: States a <b>justified recommendation</b> for implementing a viable classroom option to address the challenge	Does <b>not</b> state a recommendation for using, or not using, a particular technology in the curriculum	Presents a <b>limited</b> recommendation for using, or not using, a particular technology in the curriculum	Presents a <b>well-justified</b> recommendation for using, or not using, a particular technology in the curriculum

**Figure 3.** Sample Etip Rubric (Etip 1). The rubric maker can accommodate any number of criteria and score points.

In his/her own words (200-300), the credential candidate's essay clearly demonstrates his/her decision-making process about using technology to meet a range of learning needs of the diverse group of students in the case.			
Criteria:	Level 0	Level 1	Level 2
1. Validation: Explains the <b>central classroom challenge</b> in the case			X
2. Evidence: Identifies <b>case information</b> that must be considered in a decision about using technology to differentiate instruction to meet the diverse needs of learners		X	
3. Decision: States a <b>justified recommendation</b> for implementing a viable classroom option to address the challenge		X	
<b>Global Score:</b> Does Not Meet Expectation because Decision score above = 0 Does Not Meet Expectation because both Validation and Evidence score above = 0 Meets Expectation because scores in both Decision/Validation or Decision/Evidence = 2 Exemplary because scores for all three criteria (Validation/Evidence/Decision) = 2 ✓ <b>Somewhat Meets Expectation because your scores meet none of the above rules</b> Global score not available.			
<b>Narrative Evaluation:</b> Test of the system...with new rubric scorer.			

**Figure 4.** Sample essay score using the Essay Grading Tool. Total score and average score are computed based on the rubric scores for each criteria, as well as global score.

Essays scores can be compared with search path map information, for example, by comparing the justification of a decision with the knowledge domains visited during the search for information.

Also, with relevancy scores available for each item in the problem space, a score can be created that relates the efficiency of searches with the scores on the essay. An overall relevancy score is computed that relates the total items visited to the sum of the level of relevancy of the items. A high relevancy score with a ratio to all searches that approaches "2" (meaning that all items searched were highly relevant) might represent an expert score, which can be used in an analysis of expert-novice differences. Changes in performance over time can also be used to show those differences.

At present, the evidence model of the eTIP Cases is in an early stage; thus, there is still much to learn about computing relevancy, relating it to score profiles on essays, and comparing that relationship with search path map data. However, it is clear that there is potential for documenting complex performances that involve problem solving, analysis, and metacognition.

## The Future of Network-Based Assessment

The future of network-based assessment will take advantage of World Wide Web architecture - the Semantic Web - for interoperability of systems. The Semantic Web (Berners-Lee, Hendler, & Lassila, 2001) allows applications to share data, even if they were built independently and remotely from one another. For example, the eTIPs instruction and assessment application on IMMEX in California sends essays to Colorado, where they are picked up and scored by people in Vermont (using the VISMT essay scoring tool), and then returned to a classroom for display to the teacher, who may be in Minnesota. The future of network-based assessments seems headed toward such distributed systems.

Semantic Web applications will enable the building of digital catalogs of resources that take advantage of a decentralized network of experts, such as the scorers in Vermont adding information to a classroom in Minnesota. Intelligent routing of those resources can then respond to queries that express the essay score, a multidimensional score from a survey, and other profiles of a user's strengths, interests, and aspirations. Human advisors and teachers can utilize new forms of network-based assessment to provide guidance to learners and validation of learning, resulting in highly personalized instruction, guidance, and assessment applications.

As these systems are developing, they will be guided by new conceptions of teaching, learning, and assessment, where teaching is seen as a guiding activity for planning, marshaling resources, and validating learning, learning is seen as a process of developing patterns and procedures to acquire and use knowledge in social and technological settings, and assessment is seen as an unobtrusive network-based activity that produces a rich record for analysis and making inferences about learners.

## References

- Almond, R., & Mislevy, R. (1999). Graphical models and computerized adaptive testing. *Psychological Measurement, 23*, 223-237.
- Almond, R., Steinberg, L., & Mislevy, R. (2002). Enhancing the design and delivery of assessment systems: A four process architecture. *The Journal of Technology, Learning, and Assessment, 1*, 5.
- Bennet, R. (1999). Using new technology to improve assessment. *Educational Measurement: Issues and Practice, 18*, 5-12.
- Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The semantic web: A new form of web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific American, 284*, 5, 34-43.
- Bransford, J., Brown, A., & Cocking, R. (2000). *How people learn: Brain, mind, experience and school*. Washington: DC. National Academy Press.
- Bruce, B., & Levin, J. (1997). Educational technology: Media for inquiry, communication, construction, and expression. *Journal of Educational Computing Research, 17*, 1, 79-102.
- Friedrichs, A., & Gibson, D. (2001). Personalization and secondary school renewal. In A. DiMartino & J. Clarke (Eds.), *Personal learning: Preparing high school students to create their future* (pp. 41-68). Lanham, MD: Scarecrow Press.
- Gibson, D., Knapp, M., & Kurowski, B. (2002, October) *Building responsive dissemination systems for education with the semantic web: Using the new open-source "liber" application* A paper presented at the 2003 EdMedia conference, Montreal, Quebec.
- Greenfield, P., & Cocking, R. (Eds.). (1996). *Interacting with video*. Greenwich, CT: Ablex.
- Greeno, J., Collins, A. & Resnick, L (1996). Cognition and learning. In D. Berliner & R. Calfee, (Eds.), *Handbook of educational psychology* (pp. 15-46). New York: MacMillan.
- Hawkes, L., & Derry, S. (1989). Error diagnosis and fuzzy reasoning techniques for intelligent tutoring systems. *Journal of AI in Education, 1*, 43-56.
- Holland, J., Holyoak, K., Nisbett, R., & Thagard, P. (1986). *Induction: Processes of inference, learning, and discovery*. Cambridge, MA: MIT Press.
- Kafai, Y. (1995) *Minds in play: Computer game design as a context for children's learning*. Hillsdale, NJ: Lawrence Erlbaum Associates.

- Kallik, B. (2001) Teaching thinking through technology: Introduction to Chapter 10. In Arthur L. Costa (Ed.), *Developing minds: A resource book for teaching thinking* (3rd ed.). Retrieve September 24, 2003, from [http://www.ascd.org/publications/books/2001costa/kallick\\_chx.html](http://www.ascd.org/publications/books/2001costa/kallick_chx.html)
- Kanowith-Klein, S., Stave, M., Stevens, R., & Casillas, A. (2001.) Problem-solving skills among pre-college students in clinical immunology and microbiology: Classifying strategies with a rubric and artificial neural network technology. *Microbiology Education*, 2, 1, 25-33.
- Mayer, R.E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist*, 32,1-19.
- McCallum, A., & Nigam, K. (1998). *A comparison of event models for Naive Bayes text classification* In AAAI 1998 Workshop on Learning for Text Categorization (pp. 41-48). Retrieved September 24, 2003, from <http://citeseer.nj.nec.com/cache/page/489994/1>
- Mislevy, R., Steinberg, L., & Almond, R. (2000). *Leverage points for improving educational assessment*. A paper prepared for an invitational meeting, entitled The Effectiveness of Educational Technology: Research Design for the Next Decade, Menlo Park, CA, SRI International.
- Novak, M. (2002, November). Catalyst Leadership Retreat meeting notes. Washington, DC.
- O'Malley, J., & McCraw, H. (1999, Winter). Students' perceptions of distance learning, online learning and the traditional classroom. *Online Journal of Distance Learning Administration* [Online serial], 2(4). Retrieved April 21, 2003, from <http://www.westga.edu/~distance/omalley24.html>
- Pelligrino, J., Chudowsky, N., & Glaser, R. (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press.
- Perrin, K. M., & Mayhew, D. (2000). The reality of designing and implementing an Internet-based course. *Online Journal of Distance Learning Administration* [On-line serial], 3(4). Retrieved April 21, 2003, from <http://www.westga.edu/~distance/ojdl/winter34/mayhew34.html> .
- Robles, M., & Braathen, S. (2002, Winter). Online assessment techniques. *Delta Pi Epsilon Journal*, 44(1), 39-49.
- Roblyer, M. D., & Ekhaml, L. (2000). How interactive are YOUR distance courses? A rubric for assessing interaction in distance learning. *Online Journal of Distance Learning Administration* [On-line serial], 3(2). Retrieved April 21, 2003, from <http://www.westga.edu/~distance/roblyer32.html> .
- Rudner, L., & Liang, T. (2002, April). *Automated essay scoring using Baye's theorem*. Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans, LA.
- Ryan, R. C. (2000). Student assessment comparison of lecture and online construction equipment and methods classes. *THE Journal*, 27(6), 78-83.
- Steinberg, L., & Gitomer, D. (1996). Intelligent tutoring and assessment built on an understanding of a technical problem-solving task. *Instructional Science*, 24, 223-258.

Stevens, R. (1991). Search path mapping: A versatile approach for visualizing problem-solving behavior. *Academic Medicine*, 66, 9, S72-S75.

Stevens, R., Lopo, A., & Wang, P. (1996). Artificial neural networks can distinguish novice and expert strategies during complex problem solving. *Journal of the American Medical Informatics Association*, 3, 131-138.

U.S. Department of Education. (2000). *Two exemplary and five promising educational technology programs*. Retrieved September 24, 2003, from <http://www.ed.gov/pubs/edtechprograms/webproject.html>

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Walvoord, B.E., & Anderson, V. J. (1998). *Effective grading: A tool for learning and assessment*. San Francisco: Jossey-Bass, Inc.

Wolfram, S. (2002). *A new kind of science*. Champaign, IL: Wolfram Media, Inc.

## Resources

American Distance Education Consortium - <http://www.adec.edu/>

American Association for Higher Education's - <http://www.aahe.org/>

ETIPs - <http://www.etips.info/>

Interstate New Teacher Assessment and Support Consortium - <http://www.ccsso.org/intasc.html>

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

MORPG - <http://www.mmorpg.com/>

National Educational Technology Standards - <http://cnets.iste.org>

National Staff Development Council - <http://www.nsd.org/educatorindex.htm>

NICI Virtual Library - [www.vlibrary.org](http://www.vlibrary.org)

Vermont Institute of Science, Math, and Technology - <http://www.vismt.org/>

World Wide Web Consortium (W3C) - <http://www.w3.org/>

## Contact Information:

David Gibson  
National Institute for Community Innovations  
100 Notchbrook Road  
Stow, VT USA  
dgibson@vermontinstitutes.org

*Contemporary Issues in Technology and Teacher Education* is an online journal. All text, tables, and figures in the print version of this article are exact representations of the original. However, the original article may also include video and audio files, which can be accessed on the World Wide Web at <http://www.citejournal.org>