You Can't Manage What you Can't Measure - Assessment of Learning in a Connected World

By Jonathan R. Cornwell and W. Reid Cornwell

"Teaching without learning is just talk." (*Unknown)*

Any discussion about educational reform must establish a working concept of knowledge and how education serves the acquisition of knowledge.

Master said, "Yu I shall I teach you what knowledge is? When you know a thing, to recognize that you know it, and when you do not know a thing, to recognize that you do not know it. That is knowledge" (Confucius 551-479 BC)

Compare this to a modern concept of knowledge acquisition:

1. Unconscious incompetence: I am ignorant and therefore without any ability.
2. Conscious incompetence - I am acquainted with the knowledge but my abilities require great conscious effort with many mistakes.
3. Conscious competence - I am familiar with the knowledge and my abilities are natural but still deliberate.
4. Unconscious competence - I have mastered the knowledge, my abilities are fully developed and require little or no conscious management (Unknown).

In the first example, Confucius contemplates knowledge as a duality of knowing verses ignorance mediated by awareness of one's internal state in relation to this duality. The use of knowledge is implied but not central. In the second example, knowledge contains both information and action. The usefulness of knowledge is assumed and central to the concept. In contrast to Confucius, the modern concept contains a transitory conscious awareness of knowing and focuses on the processual nature of learning; true knowledge is mastery beyond deliberate thought. "Certainty" is an unspoken element of both concepts of knowledge and a key element of facilitating the acquisition of knowledge within the Connected Learning framework. Technology applied to the acquisition of knowledge is referred to as Knowledge Technology (KT).

Certainty/Confidence

Knowledge provides order to our lives. On a personal and social plain, the things that we hold to be correct with certainty provide a framework on which we build predictability and develop action. Knowledge predicates control or, in the absence of control, adaptation. The notion of certainty, as a component of knowledge, has been discussed by philosophers and scientist for centuries. (e.g. Confucius, Aristotle, Auden, Russel, and McCombs). Despite the ubiquitous understanding that certainty plays a role in knowledge, it is only recently that science has attempted to investigate this domain. In the realm of education, certainty has been ignored because the tools to assess this domain did not exist. Current computer technology remedies this by providing robust data management, analysis and storage. These are the necessary elements of Certainty Based Assessment (CBA). This will be elaborated later.

True or Correct

Entire sections of libraries are devoted to discussions of truth and correctness. This epistemological discussion is not a subject we need to address. In the context of educational reform, "No Child Left Behind" (NCLB) has provided an operational framework that defines what is true or correct implicitly within the structure of "National Standards", "Expected Outcomes" and, "Units of Practice". Each subject area is defined in detail. To meet or exceed these standards is the goal of all educational reform. Connected Learning is no exception.

Unconscious Competence and Retrieval

Imagine a student in the 5th grade who must count on his or her fingers and toes to answer the question, What is the sum of 9 + 9?? A priori, we know that mastery has not been achieved in basic addition since one of the goals of knowledge is to quickly render fundamental facts and skills accessible. In the language of the second example of knowledge at the beginning of this section, this student has not acquired "unconscious competence". Standardized tests, such as SAT, GRE, MAT, ACT and a host of state-specific assessments, attempt to evaluate the accessibility of knowledge as a dimension of mastery by placing time restraints on the test. Unfortunately, these tests are administered at the end of a learning cycle, or outside the context of learning entirely, and the opportunity for improve ones mastery has expired.

To facilitate mastery, latency of response time is a measurable phenomenon with the aid of computer technology. There is an axiom in business that says, "You cannot manage, what you can't or don't measure". In knowledge acquisition, this axiom not only applies but is crucial. Our project contemplates and addresses the fundamental issues in knowledge acquisition by proposing technology-based methods for measuring, storing and analyzing the learning process in real time.

Assessment of Knowledge Acquisition

At the beginning of this section we quoted the statement, "Teaching without learning is just talk." The way we assess whether learning is in fact taking place is some sort of "testing". In the best of all possible worlds, testing should result in a realistic measurement of both teaching and learning. However, the current educational environment tends to skew measurement with the weight being placed on the learner side of the equation.

Before addressing the different types of assessment, it is instructive to delineate assessment's purposes. Kellough & Kellough (1999) state that the purpose of assessment is:

1. To assist student learning.
2. To identify student's strengths and weaknesses.
3. To assess the effectiveness of a particular instructional strategy.
4. To assess and improve the effectiveness of curriculum programs.
5. To assess and improve teaching effectiveness.
6. To provide data that assist in decision making
7. To communicate with and involve parents.

Please note that students are mentioned only twice. Five of the seven purposes of assessment are related to the system of education that supports student learning. As

a whole, this description of the purpose of assessment is another base from which to approach learner-centered principles in education as developed by Dr. Barbara McCombs and others.

While in principle there is general agreement in the education community that the focus of the education system is the academic development of individual students, actually applying learner-center principles is difficult if not impossible within the current model of education. How, for example, do you develop and execute individualized learning plans for thirty students in a classroom? The simple answer is that, in the absence of new technologies and methodologies to improve both teaching and learning, you cannot. While many other tools are required to support a learner-centered education environment, dramatically improved assessment methods are the keystone. Sections 1.5 through 1.7 describe the three domains of assessment within the Connected Learning framework.

Diagnostic Assessment

A great many tools exist to map individual abilities and differences but diagnostic test are typically used episodically such as when a teacher suspects that a student has a learning disability. Episodic diagnosis has at least three consequences. First, without baseline diagnostic information, it is not possible to empirically construct individualized learning plans. Second, problems with student learning are typically not caught until the problem becomes acute and remediation to catch up with the class becomes extraordinarily difficult. Third, marginal problems in student performance remain undiagnosed and often are explained away as a normal variation in performance that fits within the standard distribution of a class.

Diagnosis of the aptitudes, abilities and progress of each student must be unambiguous, anticipatory, timely, frequent and dynamic. Just as content changes, learners change as the result of physical, social, personal, familial and a myriad of other circumstances. Diagnostic assessment enables the education system and the student to anticipate and solve problems.

Formative Assessment

There are several extant definitions of formative assessment. Operationally, formative assessment is periodic evaluation of student performance for the purpose of improving instruction. The outcome of formative assessment can be summarized in two extremes: continuing the lesson using alternative instruction methods, examples, etc. until each student masters the material or bypassing a portion of the lesson plan if student mastery is quickly achieved. In combination with diagnostic assessment, whole lesson units may be skipped entirely or reviewed only in brief if a student already demonstrates mastery.

While it is common in many classrooms to use low-stakes assessments such as quizzes to broadly assess relative performance, low-stakes assessments mainly serve as an early warning system for later high-stakes tests such as mid-term and final examinations. Dylan William, director of the Learning and Teaching Research Center at the Educational Testing Service, states "What I mean by formative assessment is not assessment that takes place every five to six weeks, but assessment that takes place every 10 seconds." (Sausner, 2006) Rick Stiggins at the Assessment Training Institute argues that improving scores on standardized tests and learning overall depends on instructing teachers how to conduct assessments on a daily basis (Chappuis, Stiggins, Arther, & Chappuis, 2001). However, William and Black (1998) note that formative assessment is not something that can be "tacked

on" to current teaching methods; ideally, teaching, learning and assessment are an integrated whole.

The power of diagnostic and formative assessment is seen in the results from North Topsail Elementary in Hampstead, North Carolina. Principal Sylvia Lewis (Sausner, 2006) was quoted as saying, "Formative assessments catapulted the Title 1 school's proficiency rating from just below 80 percent to 98.4 in a handful of years." The means to integrate formative assessment into teaching methodology currently exists. The missing pieces are user-friendly tools to easily create assessments, and capture, store and analyze the data. This is one of the goals embodied in Connected Learning.

Summative Assessment

Summative assessment is the process of evaluating student achievement after a unit of learning, e.g. quizzes, examinations, writing assignments, etc., and represents the primary way that student achievement is evaluated today. Summative assessments typically rely upon objective testing methods such as true/false, multiple-choice and matching. Other methods such as fill-in-the-blank, short answer and essay are used with less frequency. While objective testing methods are convenient and often benefit from the use of technology such as Scantron response card readers, they are subject to test taking strategies which compromise the validity of the assessment and are severely limited in what dimension of learning they evaluate. In fact, human learning is far too complex to be relegated to simple forced choice examinations. Funderstanding (n.d.) developed an alternative to current summative assessment strategies within a concept referred to as Authentic Assessment. Funderstanding claims that Authentic Assessment accomplishes the following goals:

> Requires students to develop responses rather than select from predetermined options.
> Elicits higher order thinking in addition to basic skills.
> Directly evaluates holistic projects
> Synthesizes with classroom instruction.
> Use samples of student work (portfolios) collected over an extended time period.
> Stems from clear criteria made known to students.
> Allows for the possibility of multiple human judgments.
> Relates more closely to classroom learning Teaches students to evaluate their own work? (Funderstanding, n.d.).

TCFIR has not had an opportunity to evaluate these claims. However, the goals of Authentic Assessment illustrate a set of alternatives to current summative assessment methods that should be considered if a full understanding of student development is to be accomplished. Not all of the assessment methods embodied in Authentic Assessment are amenable to automation. However, by utilizing technology where applicable, more time can be liberated to conduct assessments that require human intervention.

Certainty-Based Assessment (CBA)

In the start of Appendix B, we present confidence or certainty as a key element of knowledge. The pioneering work of A.R Garner-Medwin at University College London forms the basis for our proposals in the domain of certainty-based assessment (CBA). Simply described, CBA adds a ?certainty? dimension to technology-mediated assessment, i.e. the student is required to place a value on how confident they are in

their answers. Correct answers receive extra credit based on the certainty value given to the answer. Incorrect answers are penalized based on the certainty value.

"Automated assessment suffers from two problems that are considered here. Firstly, it seldom makes use of information about how confident a student is in the answer given, which is part of what we take into account in assessing students person-to-person. Secondly, it often involves the construction of complex questions to ensure that students cannot get good marks by a combination of partial knowledge and guesswork." (Gardner-Medwin, A.R., 1995)

Through the coordinated efforts of several medical schools in London, a project was initiated under the name of the London Agreed Protocol for Teaching (LAPT) in 1994. The resulting CBA system currently contains over 10,000 questions and processes more than one million graded questions per year at the University College London alone. TCFIR offers a demonstration of CBA technology at http://tcfir.org/lapt/sys/options.cfm based on the UCL site at http://www.ucl.ac.uk/lapt. Gardner-Medwin (2006) states that CBA provides several distinct advantages over current assessment methods, including:

Encouraging students to think more carefully about questions in objective tests.
Improving student involvement and interest in the process of self-assessment.
Helping students identify uncertainties and misconceptions in their understanding.
Mitigating the effects of guessing on standardized tests, improving test reliability.
Facilitating the setting of questions for students with a large range of abilities.
Providing immediate feedback of results for teachers and learners.
Helping students who lack self-confidence, or are overly confident.
Facilitating remediation.

It should be noted that certainty assessment is not an assessment domain like the diagnostic, formative or summative methods. Rather, CBA is a methodology of assessment design that can be incorporated in each of the domains as appropriate.

Bringing the Elements of Assessment Together

By enriching summative assessment with certainty data, and adding diagnostic and formative assessment methods, the Connected Learning framework proposes a ?learning lifecycle? approach to assessment that addresses all seven purposes for assessment. Starting conditions in the learning ecology are revealed through diagnostic assessment, enabling the development of lesson plans targeted to individual students. Formative assessment provides ongoing feedback during the learning process to both teacher and student, making it possible to alter the lesson plan based on student performance. Summative assessment is then transformed into the role of a confirmation of the success of the lesson plan rather than the primary means of assessing the learning process. All assessment methods provide additional diagnostic data for the next lesson plan. Testing the certainty with which students hold knowledge provides a qualitative, diagnostic dimension to that enables even more precise fine-tuning of teaching and learning. With additional technology relating student performance to particular teaching methods and materials, parents, teachers

and administrators alike can evaluate the effectiveness of the learning ecology with more and better data than is produced currently. Assessment becomes the means not only for enabling significantly higher levels of student achievement but is also instrumental in the Connected Learning principle of developing a self-reforming, self-improving and self-documenting system of education.

Applying Technology-Mediated Assessment

High standards for assessment are critical in the Connected Learning framework. Based on what we have developed in this section about knowledge, and diagnostic, formative and summative assessments, let us speculate for a moment about where technology-mediated techniques can apply.

Knowledge Acquisition

1. Present standards-based content to know
    a. Computer generated multi-media presentation in all subject areas.
    b. Repetitive practice in all subject areas.
    c. Consistent high quality.
2. Assess certainty in answers
    a. Most easily implemented in fill-in-the-blank, matching, multiple-choice and true/false.
    b. Possible to implement with other test types.
3. Measure actionable mastery
    a. Possible for all of the above by measuring and recording answer latency.
    b. More subtle measures of performance may be developed, e.g. relating time spent reading a section of material to formative assessment results thereby deriving overall reading traits like speed and comprehension.

General Purposes of Testing

1. To assist student learning.
    a. Provide immediate feedback in real time
    b. Provide graphic representation of progress
    c. Record question level granularity
    d. Make it fun
2. To identify students? strengths and weaknesses.
    a. Subject level granularity will clearly reveal level of success and failure
3. To assess the effectiveness of a particular instructional strategy.
    a. Question level granularity will reveal teaching effectiveness
    b. Agglomeration of class data will discriminate individual or systemic failures
4. To assess and improve the effectiveness of curriculum programs.
    a. Question level granularity will reveal teaching effectiveness
    b. Agglomeration of class data will discriminate individual or systemic failures
5. To assess and improve teaching effectiveness.
    a. Granularity will reveal instructional quality
    b. Granularity will reveal teacher preparation
6. To provide data that assist in decision making
    a. Provide classroom management with intricate detail

7. To communicate with and involve parents.
    a. Graphic presentations will aid in communicating with parents
    b. Details based on small-scale evaluations will help parents understand the unique characteristic of their child
    c. Progress measurement will aid in earlier intervention when there is a problem

Summative / Authentic Assessment

1. Requires students to develop responses rather than select from predetermined options
    a. It is possible to merge content presentation with assessment
    b. It is also possible to offer multiple presentation and assessment options.
2. Elicits higher-order thinking in addition to basic skills
    a. Simulations can be created that stimulate both deductive and inductive reasoning.
    b. Connecting subject areas for a more elaborate and enriched understanding, e.g. how developments in timekeeping devices,
3. Directly evaluates holistic projects
4. Synthesizes with classroom instruction
    a. Tools can be created to accomplish this integration
5. Uses samples of student work (portfolios) collected over an extended time period
    a. Student data of an objective nature can be store easily and can be used to form both a quantitative and qualitative map
6. Stems from clear criteria made known to students
    a. Based on empirically derived data individual learning plans can be created and modified as needed.
    b. Progress against those plans can be accessed in real time
7. Allows for the possibility of multiple human judgments
    a. Clever assessments can be created that test judgment
8. Relates more closely to classroom learning Teaches students to evaluate their own work
    a. Certainty assessment require the student to self-assess

Curriculum Design

1. Analysis of patterns revealed by associating assessment data with lesson plans, learning/teaching materials, teaching methodologies, etc. coupled with Internet-enabled pooling of data create opportunities to identify effective education elements.
    a. Proven lesson plans, i.e. those associated with high assessment results, could be accessed and used by teachers as a whole or in parts.
    b. Diagnostic data could be used to form ?student types?; lesson plans and curricula could then be associated as effective with certain student types.
2. Data-enabled curriculum design; assessment and standards integration via technology would allow for more complex approaches to subjects while still meeting education standards.
    a. Integrated, multi-subject designs.

       i.   Theory-to-application, e.g. geometry/trigonometry and applications in navigation, construction, engineering.
      ii.   Connections designs, e.g. how a watch, sextant and GPS receiver can be the basis for exploring history, physics, geography, mathematics, etc.

3. Bounded self-study designs.
4. Collaborative designs.

Bibliography

Black, P., William, D. (1998). Inside the Black Box: Raising Standards Through Classroom Assessment. *Phi Delta Kappa International*. Retrieved June 20, 2006, from http://www.pdkintl.org/kappan/kbla9810.htm.

Chappuis, S., Stiggins, R., Arther, J., & Chappuis, J. (2001). *Assessment for learning: an action guide for school leaders*, (2nd ed.). Portland, OR: Assessment Training Institute.

Funderstanding (n.d.). *Authentic assessment*. Retrieved June 28, 2006, from http://www.funderstanding.com/authentic_assessment.cfm.

Gardner-Medwwin A.R. (2006) Confidence-Based Marking - towards deeper learning and better exams. In Innovative Assessment in Higher Education, Bryan C and Clegg K (eds) Taylor and Francis, London

Gardner-Medwin, A.R. (1995). Confidence assessment in the teaching of basic science. *Association for Learning Technology Journal,* 3, 80-85. Retrieved June 13, 2006, from http://www.ucl.ac.uk/%7Eucgbarg/tea/altj.htm

Kellough, R.D. & N.G. Kellough (1999). *Middle school teaching: A guide to methods and resources* (3rd ed.). Upper Saddle River, NJ: Merrill/Prentice Hall.

Sausner, R. (2006). Making assessment work. *District Administration*. Retrieved June 27, 2006, from http://www.districtadministration.com/page.cfm?p=1188.

University College London (2005). *Certainty-based, or confidence-based marking (CBM)*. Retrieved June 3, 2006, from http://www.ucl.ac.uk/lapt/index.htm.