Assessment in Learning Communities
Thursday, May 24, 2001
Breakout Session B

We began by considering the following definitions.

Assessment is the process of collecting targeted information. Assessment *data* can be collected informally or formally using various assessment instruments and assessment methods. Assessment *instruments* can include surveys, tests, completed written and oral assignments, journal entries, portfolio entries, videotapes, audiotapes, photos, web pages, and discussion forum entries. Assessment *methods* include a variety of approaches that are usually matched to a particular purpose for the assessment. For example, pre-post comparison of matched assessment data would depict students' gains (or losses) in performance on an assessment of knowledge before and after a selected amount of instruction or a particular learning activity. A descriptive case study would detail how a particular cohort of students approached a new learning experience.

Formative assessments are conducted as a process is forming and provide input into that process. **Summative** assessments are conducted at the end of a process and provide a measurement of the student's sum performance. A particular assessment can be formative or summative depending on its purpose. For example, the test at the end of Unit 1 is summative if it completes that portion of the course and you are not using the results of the test to inform your teaching. The test would be formative if it diagnosis weaknesses or strengths in the class that help you decide how to cover Unit 2.

A **test** is a comparison of a student's performance on a particular task or set of tasks to the performance of an expert on the same task(s). Often tests are given to determine whether or not students have met a particular set of learning objectives.

Students' performance on **criterion-referenced** tests are compared to criteria for those items (must make 70% to pass). Performance on **norm-referenced** tests are compared to the performance of other students through a process of establishing the norm among a sample of similar students(must score at the 70th percentile to be admitted).

Goals state intentions for specified learning events and are worded in terms of overall outcomes. **Objectives** state acceptable levels of student performance on specific behavioral task under controlled conditions using terms that clarify the assessment needed.

Valid assessments measure what they are intended to measure. Validity can be examined from the perspective of the content of the assessment, the construct of the assessment instrument/method, and the way in which the assessment data are analyzed. Test items that include wording which does not discriminate between those who guess well and those who have the knowledge tested are not clearly measuring what was intended. Assessment of group work that does not document the contributions of each group member does not accurately measure the group's work. What is assessed and how it is assessed affect validity.

Reliable assessments consistently yield the same information under the same conditions. The method of assessment contributes to reliability. For example, the conditions under which a standardized test are administered are noted each time a test is given and efforts are made to control those conditions so that it is reasonable to assume students had an equal chance to perform.

Evaluation is the process of selecting specific data sources, instruments, and method(s) of assessment and data analyses in order to determine (evaluate) the extent to which a particular intent (goal and/or objective) was achieved. The process of evaluation determines the value of the assessment data.

Grades are a form of evaluation in which a set of assessment information is compared to a standard in order to determine a student's success in a particular lesson, assignment, unit, course, or program.

Many assessment experts state that grades should only be based on students' achievement in the subject matter rather than effort or creative approach. They state that this is the only way for grades to have consistent meaning in deciding whether or not students have acquired the knowledge that is identified with that course. Most educators have a sense that the grades they have determined don't accurately assess everything the student has learned and account for this possibility by adjusting grades. Some adjust grades by comparing the distribution of grades to a norm such as a bell curve so that a given percentage of the students in a particular class receive each of the grades in the distribution. Others adjust grades by including extra credit options, credit for attendance and/or participation, or assignments that require minimal knowledge of the subject. Some drop the lowest test grade. Many write test items that include silly distracters which are clearly not correct, or give clues to help students guess at correct answers.

Fair assessments measure students' performance with an equal chance for success. The debate about fair assessments has been fueled by the desire to obtain objective measures of students' knowledge without the bias of subjective interpretation. Many educators are troubled by the task of explaining how subjective assessments can be fair and therefore avoid them, preferring the perceived safety of objective measures such as multiple choice tests. Other educators find that there is significant bias in all forms of assessment. Some researchers (Lincoln & Guba, *Naturalistic Inquiry*) respond to this by stating that the only fair approach is to acknowledge the bias inherent in all assessment so that it is available to all interpreting the assessment. Acknowledging subjectivity need not imply a lack of accuracy in measurement. Accuracy in measurement can be addressed by examining the validity and reliability of the assessments.

The ability to accurately measure a student's performance or knowledge is often exacerbated by the detail of reporting of grades. Explaining why an average of 89.4 is justifiably different than an average of 90.0 is complicated when students (or their parents) understand the use of significant digits, that the number of decimal places indicates the specificity of the measurement. The validity and reliability of the assessments are impacted by the analyses of the data, as a discussion of descriptive and inferential statistical analyses

would explicate, so it is important to report grades at the level of specificity of the assessments.

Quantitative assessment refers to assessment data that are primarily numerical and **qualitative** assessment refers to primarily descriptive assessment data. Most researchers agree that the type of assessment data collected should be determined by the purpose of the assessment. Most educators have experience collecting and reporting quantitative data. Qualitative data are sometimes coded so they are reduced to numerical data, however the value of the descriptions are reduced. The increase in portfolio work has prompted creating and refining of grading rubrics as valid means of grading portfolio work.

Discussion

Forcing ourselves to collect information about everything is forcing things and changing things, you change things just by taking the measurement. Always assume that there is some objective. Can we assess the creative process which is more subjective than objective?

Assessment processes in the end are disastrous, there should be quality control at each step (manufacturing example). One of the problems with assessment, once you have set your goals and objectives, you tend to work toward those, teaching to the test. Need some standard.

Should the objective(s) drive the activity and the assessment?

We don't like to lose control. There is this tendency to tame assessment, bring it under control. Need to simplify the goals.

Need to build into the assessment to include a measurement for creativity.

This isn't close-ended as it is listed here.

There is a difference between behavioral and qualitative models and they are somewhat at odds so we have an inclusive approach so that our definitions make room for both.

In college and university courses, many things are left unspoken (objectives), and it's important to bring those up.

Students need to be able to assess themselves.

How do they learn responsibility?

There is a way to create objectives without completely going with an instructional design model.

How do you defend qualitative results?

If we spell out every detail then we give too much information, you have to trust me as your professor.

You would not become a professor if you didn't have this type of judgment. Its part of the accountability.

You are no longer the sage on the stage. There are a lot of authority issues.

One bad apple spoils the barrel [referring to instructors who misuse authority].

They are arbitrary about their judgments

Authority point is interesting. I don't think that we can expect students to know things that we can't view them doing. If you don't have a creative professor, then you can't expect them to see creativity in their students. The fountain of assessment is how we think things are assessed. Sometimes things seem unreasonable when you put them on paper, and we need to look at those things. All of this is related to subjective evaluation.

We as professors have earned some of that authority.

But how did we get that, we have done that assistantship. They have assessed you, they have shown you the gate. Is that similar to what we do with our students?

Has presumably become part of a community of practice, but when it comes to assessment we don't seem to be part of the community. We have always looked at assessment etc. like multiple choice questions.

…And you can do the correlations, if you disagree…

If you think about it, that is still a tacit understanding of assessment.

We start talking…

I think that process would be greatly simplified if we limited the number of gradations. What it really comes down to can you do it or can't you, thumbs up or thumbs down. This is clearly passing or not passing.

I am forced to grade them every single semester. It is an evil, corrupt thing. I do it fairly, I think it's corrupt. There is set of published criteria, this last year it has been eating at me, I don't want to throw away these people. I tell a story about assessing, make grades formative instead of summative. We are playing the game just as much as they are playing the game.

You could actually do formative assessments within the 15 week semester.

That is what I do . . . you have to do 10 things and complete them adequately.

Are you offering formative assessments along the way?

Yes, I present the information and ask and answer questions [further description of discussion].

So discussion is a type of formative assessment?

Yes, students can talk to each other and share their work, and students can talk about each other's work.

Opportunity costs everywhere, alternatives have costs also.
We can send information about students into the future. It's a constraint that we live under.
We have been very successful in the US as flawed as it is.
You have come up with ways around the constraints.
Constraints… pressure to use online assessments, if you can't squeeze this assessment into this model, then you are out of luck. How can we work with this?

What are alternative assessment methods/tools and how do faculty defend them?

Sometimes we are a low production department and we have to defend quality vs. quantity.

First I need to think about what students are going to be taking my course . . . so I am coming at this trying to get a paradigm shift. My goal is that they become competent consumers of research. They are very different, I get all of them except the administration students. So in order to address all their needs, all of the tasks and lessons are online, I set them up to make whole class and small group contributions. I walk them through the process…by the end of the course they have a product, it may not be completely polished… the students review each others work along the way. They have to provide self-assessment, using artifacts of their own work showing they understand each task.

It would be possible for persons from behavioral outlook to retro fit to this type of style.

The trick skill is not taking notes or attending, but instead interpretation and taking the information to another level.

There are certain elements that make a class successful. Self-assessment is a good tool. By talking with another student I start to see how they understand, while learning about what I myself understand. Collaboration. What is meaningful to them is brought out in class, it helps students learn from each other and other perspectives. I am here to learn just as much as you are (professors/instructors). Assessment provides feedback for how students learn certain components of a course. With this information, look at how much we can improve our instruction on the next go around. But we want control, and we don't want that information to get out if we find that our instruction is not working well. There are levels of fear… grades may not tell how much your student learned during your course. You have to demonstrate something as a student. You don't want to be walking around your department saying you're an A professor and you're a B professor. I try to think about that when I am grading my students. So requiring them to demonstrate is good. We set up a model where students

can choose the grade that they want. If we set up assessment so that we can figure out what students learn and how they learn then it helps in the next go around. It makes a difference to me.

Can implement mid-semester evaluations about what works and what doesn't. I didn't realize an innovation that I was using failed miserably until the end of the semester.

Involving students in the decision-making process shifts the authority in the classroom and the instructor must be wiling to accept this sharing of authority.

Examples: students fill in part of the syllabus (decide what to study during that time), have students write some essay questions.

Feedback- formative feedback loop: assessment works at two levels, students learn about what they know and don't, while the instructor learns about the students.

The other level is I don't know how to document. How do we report out this information? How do we get this into a pipeline? How does it get out into the community of practice?

Self-assessment linked to the community of practice and outside evaluation.

It is part of our disciplines…

Outside evaluation
Curriculum-embedded assessment
These are somewhat on opposite ends of the spectrum

Online portfolios: if students post work online then it can be evaluated by other students, instructor, and outside evaluator. It would help demonstrate what my students have done. Presentations (one-on-one with instructor). How do I quantify that? Why must you quantify that? I feel like I must justify the grade.

Hercules class example. Students presented their work… they had a publication. What do faculty members do when they use this type of venue to their students? What occurs in their classroom?

As we begin to examine different ways to teach, we have the opportunity to create the assessment models, etc. It is an opportunity for us to create this work.

Faculty would be laughing and realizing that they need to look into their pedagogy and doing research on their teaching in addition to research in microbiology, etc.

If they are not taking a scholarly approach to their teaching then what type of teaching are they doing? What if it were integrated into the whole system so that it is not an add-on, but instead its part of the process?

Think about how a professor can learn with a technical model and then learn from the process as well…I don't think it has to be separate, I think there is opportunity for them to mesh.

I think that's the idea of a master teacher, that their teaching informs their research and their research informs their teaching. It is the meta-process that I am talking about, it is an extra that someone else may need to look at and assess.

That brings us back to the community of practice… what is working…that brings us back to the assessment process.

How do I justify these grades?

We must demystify the process.

Reduce the anxiety by allowing students to discuss with each other, work in small groups.

Treat students as junior models of the discipline. How better to determine if you want to be part of the community than to act as part of the community.

We put more energy into our students as they progress through the process. If we do it in the survey courses, then we may produce better citizens.

Research vs. Research into teaching

Can we have researchers only and teachers only? NO!?

How do I take the information I am collecting and turn it into something that I can fit into my disciplines restrictions and expectations?

Simulation of the professional environment. (online portfolio).

Using technology to mirror what we have done in the past. It is valuable and essential because these are the baby steps that we need right now.

Example: Web logs

The reason you take a course is to see the world through the instructor's filter. What happens if we make the students become the filter?

How do we present technologically unique information in a traditional way? Is it necessary to present it in a traditional manner, or can their be an innovative means to presenting the information?

I want to assess certain points of progress, but I also want to assess the process of getting from point to point.

Don't force technology into the process- it may not always fit.

What should come out of this?

Link best practices to the portal.

TLT should be a voice different from the GA.

Don't want to see technology move us backward, just because it makes grading easier.

Recommendations:

What can the TLT do to support campuses to develop systems to promote scholarship of teaching activities?

There is not just one way to assess.

How can we help faculty get their mind around that?
Link to examples.

Find ways to communicate between campuses.

The collaborative has to go back to technology at some point. Where does technology fit into this?

Peer-to-peer teaching strategy, time off to visit other campuses.

It is important from the collaborative perspective to continue to focus on the technology issues.

What is the purpose of using technology? Need to continue to ask this question to identify what the goals are for the course and how technology can assist in this goal.

I suspect that our group is talking more about teaching than the other strands. Not really, similar themes appear in each strand.

Technology is not helpful to us if we are not using it toward our goals.

Risk-taking needs to be accepted, not dismissed.

Time needs to be allotted to those implementing new ideas and taking risks.

Departments need to be more accepting of risk-taking.

Tension between entry level courses, class size, and assessment. Redesign and support from TLT…

Example: chemistry class and lab, labs fed the discussion in the lecture system.

Can technology facilitate better understanding?

Performance anxiety: Is there anything that can be done with technology to reduce performance anxiety?

Need to balance the needs of those who need the face-to-face interaction with those who have performance anxiety.

Gender issues

May do students a disservice, because they may need to perform in public.

Can refer back to what they wrote, and ask related questions in class.

May be a scaffolding issue.

Equity issues: gender, ethnicity, economic.

Factors do get in the way, is there an overcoming factor that we can capitalize on?

Students who enroll in and finish an online course are different from those in the average classroom.

Somebody should explore an ongoing online course.

Summary / Goals:
Create a statement about what is important.
Create suggestions that the TLT should do.
What is our position statement?
What are some recommendations that can be made to the TLT?
Observe the demonstrations today and consider if they should be included in the recommendations.
Should the TLT support any of the innovations in the demonstrations?