Reflection on My Online Class Assessments
Based on a Special Assessment Issue of the
British Journal of Educational Technology

By Janine Lim, July 2009

This file is copyrighted (2010) and a part of my Leadership Portfolio found at http://www.janinelim.com/leadportfolio. It is shared for your learning use only.

Whitelock, D. (2009). Editorial: e-assessment: developing new dialogues for the digital age. *British Journal of Educational Technology, 40*(2), 199-202.

Summary
This issue came out of a European Association for Research on Learning & Instruction Conference with a focus on challenging assessment. The topic is important to the European Union's Programme for International Student Assessment which is considering large scale delivery of exams. The editorial gives an overview of the studies and the various types of assessment covered in the journal, noting that assessment often tells us about the pedagogical principles in practice.

Angus, S. D., & Watson, J. (2009). Does regular online testing enhance student learning in the numerical sciences? Robust evidence from a large data set. *British Journal of Educational Technology, 40*(2), 255-272.

Summary
This study was able to use a large and rich data set to test the hypothesis that exposure to online assessment instruments enhances student learning. The result is that higher exposure to online testing leads to higher student learning as measured by a final proctored exam. They were able to control for prior student aptitude, gender, effort, and other potentially contributing variables.

The class was a first year course applied mathematics course in a School of Economics in Australia. The online quizzes were given every three weeks and were 2% of the grade, with the "best attempt from two 1-hour attempts over a 1-week testing period" used for grades. The quizzes were made from a small set of questions but with the 80 different numerical variations. Students put a number in a blank form, and tolerancing was set in WebCT so that a five-unit error would be accepted as correct, and any complaints were handled by inspecting the student's answer. Students received feedback as soon as their whole attempt was completed. Students could also use a formative multiple choice practice quiz that included fully worked solutions. The quiz questions were written in *Respondus* and then uploaded into WebCT.

Improvement in student learning was measured by "the attainment of a higher final examination mark for students who completed a higher number of the online assessments than the average." They were examining how often the students engaged with the online quizzes.

The study had three major findings. Exposure to a regular online quiz (with a low percent of the grade) has a significant positive effect on student learning. Caution is required in the interpretation of simple regression or correlation studies. And third, that low attainment students are less likely to voluntarily attempt the quizzes, and therefore may need a learning intervention.

Reflection
This study appears particularly useful for the study of mathematics. None of my online courses have knowledge similar to math where the students are learning concepts, knowledge, skills that would be tested like math is in a final exam. My classes are more focused on helping teachers see the possibilities of integrating technology in the curriculum, and to be able to apply that knowledge by creating lessons and/or resources that support their teaching. So this type of testing doesn't really apply to the type of learning that happens in my online classes. Nevertheless, I'm going to share this with colleagues who teach this type of knowledge online.

Barbera, E. (2009). Mutual feedback in e-portfolio assessment: an approach to the netfolio system. *British Journal of Educational Technology, 40*(2), 342-357.

Summary

This exploratory study describes the benefits and functioning of a netfolio, which is an e-portfolio plus personal forums for reflection on the evidences, as well as peer and teacher forums for discussion and reflection on the evidences. In addition, students can easily see their peers' evidences in a particular area because of linking between portfolios.

The study compared two groups of students, one using the e-portfolio and one using the netfolio. The netfolio students revised their work more, and had higher grades than the e-portfolio at the end of the course. The netfolio had more reflection, self-evaluation, and higher cognitive level.

Other notes: "formative evaluation is at the heart of high quality teaching and learning" p. 354.

The study also suggests that in the netfolio, the instructor was silent but approving or affirming the interactions between the students (message read register). Therefore more IT tools are needed to make this "positive presence more explicit and visible" and suggest more "presence" technologies to support this.

Reflection

This study seems to have direct application to the Leadership program in several areas. The LLG process of giving feedback on evidences and peer assessment are supported by this study. However, could some structures or suggested activities be included in instruction to the LLGs so that this reflection and peer assessment would be more likely to happen? This study was also on a research class, and it made me think of the proposal development class. Could some of these principles and ideas be used in that class?

None of my own classes are long and involved enough to include portfolio development. However, the Kid2Kid Videoconference Connections class takes participants through a process of creating and implementing a collaborative project with another class. During this time, lots of peer assessment and feedback occurs, to the benefit of the student participating. This study affirms the design of that class.

Chang, C.-C., & Tseng, K.-H. (2009). Use and performances of Web-based portfolio assessment. *British Journal of Educational Technology, 40*(2), 358-370.

Summary

This study compared junior high students in computer classes (animation and web site development) who were assessed with a portfolio vs. traditional assessment. An experimental design was used, and the results shows that the students assessed with the web portfolio had more reflection, self assessment, continuous improvement, goal setting, problem, solving, data gathering, and peer interaction. However, the peer assessment didn't work very well as the students struggled to find time and to know how to peer assess.

Reflection

If junior high students struggle to peer-assess, do Leadership students also find it difficult? The rubrics help, but are additional instruction or modeling needed for Leadership students to learn how to appropriately assess? It's possible that as our LLG continues through competency check off, we will get better at it. But how will that knowledge get passed down? Only as we continue to take on new members.

Dermo, J. (2009). e-Assessment and the student learning experience: A survey of student perceptions of e-assessment. *British Journal of Educational Technology*, *40*(2), 203-214.

Summary

This study looked at students perceptions, particularly if they feel comfortable doing assessments on the computer, if they think e-assessment contributes to their learning, and if it's considered a good alternative to paper based assessments. In addition, they looked at age and gender to see if there was a difference.

Instructors used e-assessments for a variety of reasons – to increase student engagement, to deliver feedback to large numbers of students, and to save time marking tests.

They found that a major student concern was that the randomized items from question banks were unfair. Students did not think that the e-assessment was unfair, and they expected to have e-assessment at the university. They thought that e-assessment was suitable for their studies. They thought the e-assessments were accessible and that they improved the reliability of grading. They thought that their grades, login, and assessments were secure. They believed that e-assessment added value to learning.

There was no significant difference between age and student ratings of each area. In addition there was no significant difference between gender and the results.

The general impression of the students' open ended comments is that they are positive about online assessments. However, some students do have concerns which should not be ignored.

Reflection

Most of the online assessment in this study referred to short answer and multiple choice assessments. Again, this type of knowledge testing doesn't apply to my online classes in educational technology. I'm not interested in multiple choice questions on terms, etc. The concepts and the application of the concepts to the teaching environment are most useful to me, and I find the discussion space most useful for this type of learning and interaction as well as assessment. Still, this study was interesting and can be used when others ask me about online testing.

Draper, S. W. (2009). Catalytic assessment: understanding how MCQs and EVS can foster deep learning. *British Journal of Educational Technology, 40*(2), 285-293.

Summary/Notes

The article is about using Multiple Choice Questionnaires (MCQ) and Electronic Voting Systems (EVS) and what instructional methods work best with these tools. I liked this one a lot as it has started me on the path of defining my opinions about assessment.

This article is based on a theoretical framework that suggests that formative assessment is "the single most powerful factor in promoting learning (Black & Wiliam, 1998; Nicol & Macfarlane-Dick, 2006; Sadler, 1989, 1998)."

Another good quote: "Summative assessment leads only to questions of validity and reliability of testing as a way of selecting better assessment methods. However, if the aim is to improve learning, we may and should judge assessment techniques in terms of the learning gains associated with them" (p. 286)

Another really good quote: "More than 25 years ago, Clark (1983) argued that it was not technology that caused learning, but the 'instructional method', ie, pedagogical approach or learning design. The papers he reviewed failed to show any benefit of using technology, except where they simultaneously changed the teaching method (thus invalidating the comparison). Has in fact information and communication technology (ICT) in general and e-assessment in particular demonstrated any concrete increase in learning attainments, or only at best favourable attitudes in some participants? (p. 286).

And another: "Shallow learning is characterised by retention of true–false items that are either disconnected or linked in just one way for one use, while deep learning is characterised by learning multiple relationships between items that can support multiple uses (ie, transfer tasks) better. (p. 287). *Maybe this is why I prefer my current way of teaching/assessment as opposed to quizzes etc. I want the participants to make multiple connections with multiple uses and choose what matches their curriculum and teaching style best.*

One example of using multiple choice is to have students vote on an answer, show the chart, and don't tell the answer yet. Then the students discuss with their neighbor why they voted for the option they did. Then they vote again. "peer interaction enhances learning both by stimulating the production of reasons and by metacognition, ie, prompting self-examination of certainty, and follow-up work to resolve uncertainties" (p. 289)

The six ways to use MCQs are all based around peer interaction/instruction – because when you try to explain your reasons, or get past uncertainty, the knowledge sticks in your brain better.

1. Assertion–reason questions, which can be and have been used with EVS.

2. Taking an MCQ and having the learner generate reasons, for and against each response option, rather than simply ticking one. (This is usually done on paper as a private revision technique.)
3. Confidence-based marking, which is normally delivered by ICT and could be done with EVS with some (but not all) software.
4. Mazur's method of using brain-teasers to prompt peer discussion, which is routinely done using EVS.
5. Having students create MCQs as part of presentations using EVS.
6. Having students create MCQs for use in tests that may be administered either by using EVS or on paper.

The powerful learning with EVS/MCQs is from the students interactions after their initial response. So it isn't worth keeping their responses for grades. If the EVS is used for formative evaluation, you don't need to keep their scores. "We need an altogether broader and subtler (and less teacher-centred) conception of assessment and feedback, in which questions (including EVS-administered MCQs) work most powerfully as indirect provocations to learning that are often internal to the learner and do not centre on direct transmission from the teacher. In this respect, they operate like ideal Socratic dialogue, but with even less intervention from a skilled teacher." P. 291

Reflection
How is it that in my classes I find preconceived notions or erroneous ideas and try to correct them? (via formative assessment?) Each of my online classes has readings and application. All of my readings discussion questions require participants to reflect on their current practice as compared with the articles. This usually brings out their current ideas and gives an opportunity to make corrections or plant new ideas as needed.

Need to think more about this paragraph below. I definitely don't subscribe to view a or b in my teaching online; it's more like view c. But could I use MCQs to elicit further discussion as in view d? Maybe in a face to face class, but the online MCQs tend to

> "The notion of (e-)assessment itself thus clearly needs further stretching through a progression of views of assessment. In the simplest commonsense view (view a), often implicitly held by students, assessment is done by teachers to learners to test what they know. In the view represented by Black and Wiliam or Sadler (view b), such summative assessment is extended to include formative feedback, in which the process both judges the learner's answer and gives diagnostic and constructive help. In actual HE practice, this concept needs further extension (view c) to include other inputs by staff to students during the task process (eg, feedforward, discussions and advice during project super-vision), where students' efforts to do the task lead them to seek and receive interactive advice and help before and during the task, as opposed to judgements only of completed tasks. However, in this paper the evidence we have explored implies that the notion of 'assessment' requires still further extension (view d), to events such as brain-teaser MCQs or peer discussion, whose role is only catalytic: the immediate response of the student is not diagnostic, and no further help comes from the staff or software, yet

it leads reliably to important learning gains. This is truly ‘assessment for learning’ but is typically not directly ‘aligned’ to test the course objectives or deep learning because the question format may well be a superficially factual question that in itself gives no indication of the much deeper cognitive learning processes of reasoning and confidence judgements that it will give rise to. This reminds us again of Clark’s assertion that it is not the technology but the teaching method that most strongly determines learning, in this case, the way in which MCQs are used within a learning design more than how they are delivered or even their surface format” (p. 291).

New definition: “‘catalytic assessment’. We might define this as assessment designed to lead to learning later (often after the learner has given an answer to the question) and where that learning typically occurs without formative feedback (ie, external and diagnostic or constructive) but through processes (including metacognition) internal to the learner (p. 291).

Draper, S. W. (2009). What are learners actually regulating when given feedback? *British Journal of Educational Technology, 40*(2), 306-315.

Summary

This article looks at feedback and the six interpretations that students have of that feedback, with the corresponding six actions to take.

The six interpretations are:

1. **Technical knowledge or method:** I did not use the best information or method for the task, but can improve it and do better next time.
2. **Effort:** I did not leave myself enough time to do it well. (Almost everything we do in life is time limited. If it is important enough, then putting more effort in will get a better result. On the other hand, everyone including students has limited time and must save time from some activities to invest in other ones).
3. **Method of learning about the task:** I did not seek the right information to make a good job application; I did not test my paper on the right audience; I should change my revision method for this course; I should have discussed what the criteria really meant before writing the essay.
4. **Ability, trait, aptitude.** This result tells me about relatively unchangeable traits. I should apply for a different kind of job, change the course I am studying on.
5. **Random:** I did the right thing but the process is not deterministic. Another time I will succeed without changing what I do. If it rains when I go for a picnic at a beauty spot, it does not mean either that picnics are bad or that that spot is ugly; not every lottery ticket is a winner; not everyone I ask to fill in a questionnaire for me will agree to.
6. **The judgement process was wrong; I was right.** Appeal the mark the tutor gave me; find the bug in the compiler not my program re-educate my readers; find a different audience. J. K. Rowling had to try many publishers before one would take her first Harry Potter book; revising the book was not what was needed.

The six actions in response are:

1. Work out the appropriate improvements to my knowledge and skill, and adopt them as permanent parts of my future practice.
2. Increase or decrease the time and effort allocated to this task. For example, spend less time on my French lessons (which I amgood at) and more on statistics, which I need but am not doing well at.
3. Seek out new ideas on study methods and try them out; find better information on, or understanding of, the true task criteria.
4. Change the course I am taking.
5. Persist: try, try, try again.
6. Get a second (and third and fourth) opinion.

In light of this, large changes to learning can occur by helping students see which interpretation is appropriate, and by designing activities that help the students know

which interpretation to take. Principles on feedback and assessment then are dependent on context.

> "In universities, institutional requirements focus in effect on summative assessment (4), while many staff instinctively focus on formative assessment (1), but do not address (2) which is the most immediate need for many new students, nor (3) which is actually in the long run the most important (learning how to learn, how to become indeed an independent and lifelong learner)" p. 313.

Notes

I want to keep this quote to consider how best to improve people staying in my online classes and the AVLN online classes.

> *Reducing dropout by emphasising strengths*
> The Open University ran a big, controlled study (Simpson, 2008) over 3 years, involving a total of over 5000 students, and showed that a single phone call made before the course (for the experimental group) reduced dropout by 4–5% over the first year of study compared with the control group (no phone call). They had a highly predictive statistical model of who was at risk, but found the same small but statistically significant ($p < 0.01$) effect for high and low-risk student groups alike. The content of the phone call was derived from the positive psychology literature, and followed a script about emphasising each student's personal strengths at learning (p. 311).

This quote also helps explain my current perspective on assessment in my classes: "shallow learning is essentially characterised by training for a specified task, while deep learning is characterised by seeking to understand" (p. 314).

We need to design feedback and assessment to serve all six types, and tell learners which were applicable in that feedback, and give them suggestions for action.

Reflection

What type of feedback do I give my participants in my online classes? Generally it's answering their questions, letting them know they are on the right track, encouraging them in their effort. Which of these six do I give feedback on and in what ways? How could I change? Need to think more.

Heinze, A., & Heinze, B. (2009). Blended e-learning skeleton of conversation: Improving formative assessment in undergraduate dissertation supervision. *British Journal of Educational Technology, 40*(2), 294-305.

Summary

This study looked at the supervision of undergraduate dissertations. Student perspectives were collected with a focus group and questionnaire. The suggestion is that a combination of face-to-face and electronic communication be used for feedback.

Theories to Investigate Further

> "Gordon Pask (1928–96) and others advanced the constructivist paradigm in the 1970s through conversation theory (Scott, 2001b). The central idea of a dialogue, which is advanced by the conversation theory and represented by the skeleton of conversation, was further developed by the conversational framework (Laurillard, 1993). Although widely cited, the conversational framework is not widely used in practice (Dyke, Conole et al, 2007), for a number of reasons, including a lack of practical considerations and most importantly, its implicit reference to assessment (Heinze et al, 2007). As a result of deficiencies in accommodating assessment and the increasing need to integrate face-to- face sessions with e-learning tools, the blended e-learning skeleton of conversation was developed (Heinze et al, 2007). (p. 296-297)

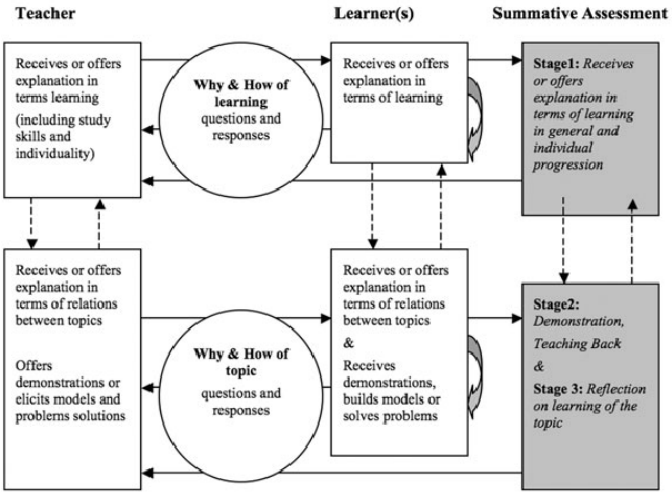

Figure 1: Blended E-Learning Skeleton of Conversation

Reflection

I think that what Newman did with me on my dissertation was this conversation model or something similar.

Interestingly, students wanted to hear back within a day!! The study suggests guidelines for students and supervisors that from there the two can negotiate expectations.

The conversation model is interesting and may be closer to what occurs in the discussion board in my online classes.

Jordan, S., & Mitchell, T. (2009). e-Assessment for learning? The potential of short-answer free-text questions with tailored feedback. *British Journal of Educational Technology*, *40*(2), 371-385.

Summary

This article is a very careful review and evaluation of the Open Mark short answer assessment tool. This tool is based on "natural language" and analyzes sentence structure and placement of verbs and nouns to grade the work. Open University also emphasizes feedback, so there are three levels of feedback for each question. The study compared the grading of the computer vs. the tutors, surveyed student opinions, and observed several students taking the test and thinking outloud to determine what was happening.

It appears to be a very neat system, though there would be a learning curve for faculty. As one invested time and effort in it, over time it could be very valuable. Actual student responses can be used to refine the testing instrument. This is way more sophisticated than the online quizzes inside Desire2Learn, for example.

I found more information online here: http://www.open.ac.uk/openmarkexamples/

Reflection

This type of assessment is admitted by the authors to be used more often with science, math, computing, and health care; and less often with social science and the arts. There is very little that I teach that requires/suggests short answer type assessment. Still, it is good to know that work is being done in this area so when I assist others with their online teaching, I can reference this work.

Loddington, S., Pond, K., Wilkinson, N., & Willmot, P. (2009). A case study of the development of WebPA: An online peer-moderated marking tool. *British Journal of Educational Technology*, *2*, 329-341. Retrieved from 10.1111/j.1467-8535.2008.00922.x

Summary

This article is about the open source online peer assessment tool created at Loughborough University: http://webpaproject.lboro.ac.uk/

Academic tutors find the peer-moderating marking system as improving administrative efficiency and some learning gains. The students found more benefits to the learning process. The institutional benefits were on administrative benefits and quality assurance advantages.

Reflection

Wouldn't it be interesting to have Andrews use this to assess the effective functioning of the Learning and Leadership Groups? It could be used annually to review the LLGs and see how they are doing – right before roundtable to

If I really cared about the short term group work (4 1 hour periods) in the Jazz workshop, this could be useful. However, the group work is so short term, and so few of the participants are actually getting grades for the workshop. I think it would be more of an intrusion and frustration than helpful. Teachers usually deliberately choose how much to contribute to a group, and it's more a function of their own learning and personal needs and situation.

It's interesting to see how this tool developed across the university starting in Excel, then to ASP, then to PHP with user feedback from students and faculty through the 15 year development process. It also moved from being a home grown university tool to having an advisory board, conducting research on the tool, and developing a pedagogical base for the work.

Things to Learn More About

distributed cognitive learning (Firdyiwek, 1999)

Marriott, P. (2009). Students' evaluation of the use of online summative assessment on an undergraduate financial accounting module. *British Journal of Educational Technology*, *40*(2), 237-254.

Summary

In this article, the authors study student responses to 3 years of change in assessments, starting with a paper based culminating summative assessment, to smaller assessments throughout the semester and given online.

Students preferred the smaller bite sized chunks of knowledge, and felt that they learned better than cramming a week before a huge final exam. The immediate results allowed for better learning and better support by the tutors.

This study is similar to the one by Angus that shows students learn better with more frequent assessment.

Notes

> In the UK, the Quality Assurance Agency (QAA, 2006) describes assessment as ‘any processes that appraise an individual’s knowledge, understanding, abilities or skills’. Assessment is multifaceted and includes evaluation, feedback and motivation and has a fundamental affect on students’ learning (QAA, 2006) (p. 237).

Reflection

Nothing new to comment on this one. “For modules that are less technical and more discursive in nature, the method of assessment deployed in this case study may not be appropriate or yield such positive results” (p. 253).

Ng'ambi, D., & Brown, I. (2009). Intended and unintended consequences of student use of an online questioning environment. *British Journal of Educational Technology, 40*(2), 316-328.

Summary

This study used an online anonymous dynamic frequently asked questions (DFAQ) tool to allow students to post questions and receive responses from the lecturer and their peers. The tool is used via the web and also via SMS on their phones. A narrative analysis was used to find the intended and unintended consequences of using the tool.

Students asked questions that ended up modifying the curriculum as the instructor found the feedback useful. In addition, the students complained about some course procedures that could then be addressed while the class was occurring instead of after the fact. Because the posts were anonymous, the students were very confident and honest, and this feedback was very useful to the instructor.

Reflection

This is a really interesting tool. It's a discussion, but not really a traditional discussion forum. It's nice to learn about other tools besides the ones that we see inside Desire2Learn, Blackboard, and Moodle. It seems that this type of Q&A could happen inside my courses – but with an anonymous discussion tool. The same benefits could be realized.

Things to Investigate Further

theory of communicative rationality (Habermas, 1984) "rationality is not about the possession of knowledge but about how 'speaking and acting subjects acquire and use knowledge"

The idea is that students' level of understanding can be deduced from the questions they ask.

Peltenburg, M., van den Heuvel-Panhuizen, M., & Doig, B. (2009). Mathematical power of special-needs pupils: An ICT-based dynamic assessment format to reveal weak pupils' learning potential. *British Journal of Educational Technology*, *40*(2), 273-284.

Summary

This appears to be the one K-12 example in this special issue. The article examines how special needs students interact with a mathematical assessment that gives them visual support as they answer the subtraction questions that require borrowing.

They found that the tool had a positive effect on the student scores, even if they hadn't used the tool before. Not all of the students used the tool to completely answer. For some students, it started their thinking and then they gave the correct answer. The tool helped the instructors see better how the students were thinking and where the errors in their understanding were.

Reflection

This study doesn't apply to my online teaching, but as an instructional technology consultant for K12, it's good to be aware of the different types of tools available and how they affect students of all abilities.

Shephard, K. (2009). e is for exploration: Assessing hard-to-measure learning outcomes. *British Journal of Educational Technology, 40*(2), 386-398.

Summary

This article was a little hard to follow. It was an exploration of assessment techniques and also the hard to measure outcomes such as values and affective concepts. It ended up by suggesting that evaluation is the way to assess values, by evaluating whole groups with surveys. In the middle the author reviewed several types of assessment and evaluation and how digital tools (very briefly) are used in that area.

Interesting Quotes

> Salmon (2005) suggests 'teaching online has almost nothing to do with computers and everything to do with time, motivation, knowledge and the new agency of cyber-experience, as well as good appropriate teaching' (pp. 214–215). (p. 389)

Other kinds of assessment: construct-referenced and ipsative assessment.

http://www.tki.org.nz/r/assessment/two/tech_e.php definition of construct referenced assessment: "the type of assessment that takes place entirely through the medium of coursework, marked by the teachers at the school. This kind of assessment is based on an idea or construct, and can be problematic as it involves assessing complex skills that are often based on creative concepts."

Ipsative assessment definition from wikipedia: the practice of assessing present performance against the prior performance of the person being assessed

The author's definition of the difference between assessment and evaluation: "The terms assessment and evaluation are used interchangeably by some in higher education but not by others. For many, assessment relates to the performance of individuals, while evaluation relates to systems or processes" (p. 395).

> Another interesting quote to consider further:
> Guba and Lincoln (1989, 2001) provide an informative historical description of four generations of evaluation. The first generation was applied to education and was developed primarily to measure various attributes of school children. The school teacher might be expected to be most interested in the performance of individual children in school tests, but the evaluator was interested in the performance of the test, of the teacher and of the school. The second generation of evaluations focused on providing rich descriptions, and the third focused on the need for evaluators to make judgements about the quality of systems and processes. Guba and Lincoln argue the need for a fourth generation of evaluation, one that emphasises negotiation between all stakeholders in determining the intrinsic value of what is being evaluated and that is built on the broad paradigm of constructivism.(p. 395).

Reflection: I'm still processing this article. I think that I can use some of it to set up my reflection paper discussion, but I'm not quite sure how it applies directly to my work.

Sieber, V. (2009). Diagnostic online assessment of basic IT skills in 1st-year undergraduates in the Medical Sciences Division, University of Oxford. *British Journal of Educational Technology, 40*(2), 215-226.

Summary

This study looked at a diagnostic tool for students IT skills. 400 students in 5 courses too, the survey. They found that there was not a correlation between self-evaluation and objective evaluation of IT skills. The weak students seem to not know the capabilities of the software and so don't realize that they are missing knowledge. 25% of the students had received formal training; the majority was self-taught (42%). Interestingly, the students were hostile to the test in the pilot year; and even the faculty are unsure about it.

Reflection

The only way that I see this article fitting into my online teaching is to remember that even "digital natives" may not be aware of all the possibilities in software and may still need instruction and support in their courses. Otherwise it was more about the diagnostic tool & IT skills.

Williams, J. B., & Wong, A. (2009). The efficacy of final examinations: A comparative study of closed-book, invigilated exams and open-book, open-web exams. *British Journal of Educational Technology*, *40*(2), 227-236.

Summary

This study compared closed book pen and paper proctored exams with open book and open web (OBOW) exams. Interestingly, the authors suggest that no literature exists that support the final exam in it's traditional form or defends it.

The OBOW approach is committed to authentic assessment where students apply learning in real-life performance. The OBOW includes the learner as decision maker, and a contemporary real-world problem. The task is highly unique, requires reference to the course materials, and the time frame is tight. This makes it very difficult to engage in any dishonest practices. Students relate to the task because they can see it is meaningful (vs. cramming & dumping knowledge).

In a survey of students, it was found that they found the OBOW format superior to the closed book test in all areas. They found it easier and less stressful to take the exam at a time and place convenient for them. The OBOW is closer to what they would experience in a workplace.

The article ended by suggesting that the OBOW format is more compatible with constructivist learning theory.

Reflection

This last article may assist me in explaining my discomfort with the short answer and multiple choice type assessments. I just don't use them and can't even imagine how I could. It's because I really am constructivist to my core in teaching and learning. These other types of assessment just don't make sense to me.

A quick Google search for "constructivist assessment" found 4 million hits. Guess that's the place to start learning in my next learning cycle!