APPROACHES TO ASSESSMENT OF ONLINE LEARNING: CONCEPTUAL CHALLENGES

Bekir S. Gur and Andrew J. Van Schaack

Approaches to Assessment of Online Learning: Conceptual Challenges

Bekir S. Gur (bekir@cc.usu.edu)
Andrew J. Van Schaack (avanschaack@cc.usu.edu)

Department of Instructional Technology
Utah State University
Logan, UT 84322 U.S.A.

Abstract

Methods of assessment used to evaluate learning within traditional classrooms do not transfer well to online environments, especially those with a social constructivist approach to instruction. Furthermore, very few guidelines have been provided by researchers to instructors regarding valid assessment procedures in the online milieu. This situation calls for the development of new forms of assessment that reflect the instructional strategies and objectives found in these new learning environments. This paper describes the features unique to online instruction and identifies several promising approaches that may provide meaningful assessment. It is concluded that instructional designers and instructors should not underestimate the role of assessment in online learning environments.

Introduction

While there has been a great deal of research and discussion about the opportunities and advantages of online learning, very little information has been provided to instructors on appropriate and effective methods of assessing student learning in collaborative online environments. The standard methods of assessment used to evaluate learning in traditional classrooms do not transfer well to the online environment (Gunawerdena, Lowe, & Carabajal, 2000). This situation calls for the development of new methods of assessment that are aligned with the instructional strategies and objectives of social constructivist learning environments.

It is not uncommon to hear designers and instructors of online learning acknowledge that higher-order skills are constructed through interactions among students, and that through these interactions, meaningful learning occurs. On the other hand, it is equally common to find that these same designers and instructions allocate only a small percentage of each student's grade to the quantity and quality of those interactions.

It is argued that where the instructional strategies and objectives of online courses are different from traditional classroom-based courses, the methods

of assessment should also be different. And since many of the instructional strategies have not been previously practiced, new methods of assessment must be developed. Accordingly, this paper describes the features unique to online instruction and identifies several promising approaches that may provide the type of meaningful assessment desired by instructors and students alike.

Transition from traditional to online settings as a challenge to assessment

New automated digital assessment tools significantly improve the efficiency and convenience of providing feedback and scores to students, and as a result, much of the research in online assessment has focused on automating the scoring process (Hill, Wiley, Nelson, & Han 2004). Although these tools serve a useful purpose, they do not address some of the conceptual challenges in assessing online learning. Furthermore, there are very few guidelines for practitioners as they design assessment processes that match the strategies and objectives of online instruction (McLoughlin & Luca, 2001). This calls for a reconceptualization of learning, pedagogy, and assessment in an online context.

In any instructional context, the three key components of instructional design are “goals,” “instruction,” and “assessment.” “Goals” includes both specific behavioral objectives and general aims; “instruction” covers learning, teaching, resources, and interactions; and “assessment” includes methods and techniques used to evaluate student learning. Instruction should be goal-driven and assessment should be relevant to the instructional strategies and goals (see Figure I).

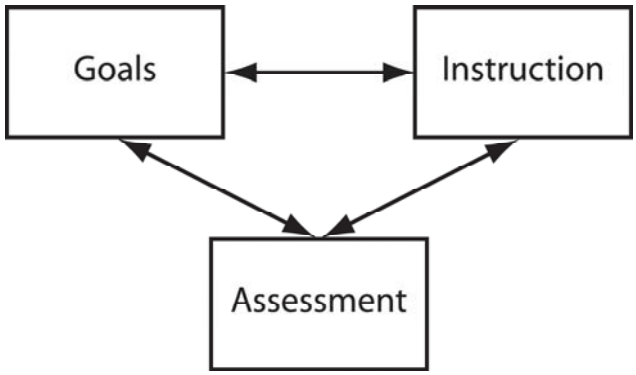

Figure I: Balance of three key components of instructional design

In traditional instructional design, it is the accepted practice to explicitly define the behavioral objectives (i.e., Dick and Carey’s (1996) model)—many instructional modules begin with “goals and objectives.” Conversely, Gunawerdena, et al. (2001) claim that achieving a stated outcome for all

learners within a fixed period of study is not the goal of many online courses—objectives are often phrased in a more general manner and “they could be negotiated, rather than imposed” (Mappin, 2001).

Regarding the second component of Figure I, Gunawerdena et al. (2001) claim that many online learning designs are based on constructivist, learner-centred principles which facilitate the sharing of multiple perspectives and place emphasis on the collaborative construction of knowledge occurring within a group of learners. Many others, however—particularly those not designed by educational institutions—do not use a collaborative approach (Clark & Mayer, 2003). Takle, Sorensen, and Herzmann (2003) argued that the methods of establishing true collaborative dialog in traditional classroom environments may not be fully applicable in online environments. They suggest that researchers should start from scratch to create a design “that emerges ontologically from a true integration of technology and pedagogy.” Generally speaking, online learning is less teacher-dominated and more flexible, with more autonomy and responsibility given to the student, although the definitions of learning, learning outcomes, key activities, curriculum process, and teacher role vary in various online learning models (McLoughlin & Luca, 2001). A critical difference between traditional classroom and online instruction is the role of the instructor. In the online context, the instructor changes roles “from that of information provider and teacher to that of a coach, an orchestrator of resources and a leader and guide in the learning process" (AFLA, no date). Furthermore, if the instruction is based on the constructivist view of learning where 1) the individual learner is expected to take away a unique perspective from the learning experience, 2) learning does not occur in unanimous form, and 3) content is not limited to the subject under study, then traditional methods of measuring learning gains using pre- and post-test scores will not work well (Gunawerdena et al., 2000).

If we acknowledge that there are differences in the instructional strategies and objectives in traditional classroom and online contexts, then the key question becomes, “What are the appropriate methods of assessment for online learning environments based on the constructivist approach?”

Methods and Techniques

Research methodologies developed in the social sciences may be adapted to analyze, evaluate, and report on student interactions, which in the constructivist approach to instruction, form the foundation for meaningful learning.

Content Analysis

Because content analysis can be used to examine any piece of writing, it is widely used to study literature, rhetoric, sociology, and general cognition (The Writing Center, 2003). The process of content analysis consist of six basic steps (Gass, no date):

1. Identify the data set to be used.
2. Develop or borrow a taxonomy (the categories into which the data will be coded).
3. Determine the unit of analysis, called "unitizing" (example: what constitutes an argument).
4. Train coders in the use of the coding scheme.
5. Code the messages into the categories.
6. Analyze the data and interpret the results.

A detailed discussion of nineteen various content analysis studies used in computer mediated communication can be found in Rourke et al. (2001).

Knowledge Construction

Constructivists believe that one way that individuals develop higher-order skills is through active participation in discussions with their peers. Gunawerdena, Lowe, and Anderson (1997) proposed a five-phase model to characterize the construction of knowledge through social negotiation in online learning networks:

Phase I: Sharing/Comparing

Phase II: Dissonance

Phase III: Negotiation/Co-construction

Phase IV: Testing Tentative Constructions

Phase V: Statement/Application of Newly-Constructed Knowledge.

The phases are ordered—moving from lower to higher order mental functioning. In their later studies, Gunawerdena et al., (2001) found that the majority of online discussions do not proceed beyond Phase I. Moreover, they suggest that when participants share their experiences, they prefer to simply agree or disagree, rather than negotiate new meaning as social constructivist theory would suggest (and presumably, their instructors would desire).

Takle, Sorensen, and Herzmann (2003) identify three factors that correlate with the quality of online discussions: the number of comments per student, the length of each comment, and the depth of the discussion threads. One hypothesis that can be generated by combining the outcomes of the previously identified studies is that online discussions with a large number of lengthy comments which are responsive to previous comments quantitatively represents the development of higher order cognitive skills among active participants.

Sequential Analysis

The method of sequential analysis (Bakeman & Gottman, 1986) is based on identifying messages by category names selected from a controlled vocabulary list, and then analyzing each message (along with its following messages) in terms of these categories to reveal processes and outcomes. The category names (i.e. events), used in Jeong's (2001) study were "Position" (+,-), "Agreement," "Disagreement," "Argument," "Experience," "Literature," "Data," "Hypothetical Action," "Evaluation," "Summary," "Negotiation," and "Comments." A Discussion Analysis Tool (DAT), was developed for measuring the processes in group discussions by conducting a quantitative analysis of event sequences. The DAT was found to be effective in identifying patterns in student interactions and measuring the events that follow specific interactions.

One of the main questions in Jeong's study was "What events are likely to follow each type of interaction?" One of the interesting findings was that interaction occurs most often when there is conflict or argument. In addition, "Position–>Agreement" was ten times more likely to occur than "Position-–>Disagreement", which may reflect students' concern for each other's feelings. This hypothesis is supported in Jeong's research by the finding that evaluation of an argument rarely occurred after an argument—rather it occurred after the group had reached a consensus.

Although more studies are required to determine whether sequential analysis would provide a suitable foundation for the assessment of online learning, we can be optimistic about using this method in view of the fact that it provides a means of examining group interaction, processes, and associated outcomes—each of which are essential to online instruction.

Peer Assessment

McLoughlin and Luca (2001) proposed interactive-participatory assessment as a motivating and authentic method of assessment. This approach includes intra-team peer review, inter-team peer review, critique and peer feedback on portfolios, and evaluations of the effectiveness of negotiating roles and task responsibilities. Several studies have demonstrated the value of peer assessment to student achievement. Web-based peer assessment of student portfolios has been used as an innovative assessment method to refine learning (Liu, Liu, Hui, & Horng, 2002). Yu, Liu, and Chan (2002) reported that student assessments of fellow students developing multiples choice test questions facilitated learning. Buckley and Toto (2000) argued that although peer assessment can be helpful, most students believe that their instructor's assessment is much more valuable than that of their peers. Similarly, Eugenia (2002) showed that although students perceived the benefit of receiving multiple opinions through peer assessment, they would like their instructor to provide more comments. Accordingly, instructors should take into consideration their students' attitudes when utilizing peer assessment and view it as a supplement to other forms of assessment and feedback.

Beyond these simple and relatively familiar methods of assessment, there are several other, more advanced methods of analysis that should be considered

for use in an online context. Social Network Analysis (SNA), for example, is a useful method for studying relations among member of social networks. Garton, Haythornwaite, and Wellman (1997) suggested using SNA for analyzing online learner networks. Other researchers have recently investigated the use of network analysis to investigate knowledge construction in asynchronous learning networks (Aviv, Erlich, Ravid, and Aviva, 2003). Although SNA may be useful to researchers of online learning, it is of limited value to instructors who require simpler and faster methods of assessment. Recently, Wiley (2002) has attempted to address this limitation by developing a real-time measure of interactivity—adjusted mean reply depth (*d*)—to obtain standard measures of discussion activity within online environments. His approach is based on the hypothesis that as the level or depth of replies increases, so too does the "intellectual depth" of the discussion. The measure answers: "How much dialog are participants engaging in, comparatively speaking?" (Wiley, 2002). This approach has some shortcomings, however. For instance, the calculation is the same for a forum with 40 participants as it is for a forum with 400 participants. Also, it does not differentiate the variability of interaction in a discussion, (i.e. are participants engaging in different ways).

These limitations are currently being addressed by Wiley and one of the authors of this paper—Gur. If they are successful, that is, if they are able to provide meaningful and easy-to-understand measures of student interaction without sacrificing speed or convenience, this method may be of significant value to instructors of online courses.

Challenges

As we mentioned earlier, in the constructivist environment, the individual learner is expected to have a unique learning experience. Although this individualized learning creates challenges for assessment, there are some strategies that the instructors can use to assess student achievement in online learning environments including: analyzing the transcripts of messages, directly asking participants what they learned, and asking students to keep weekly journals to document their learning.

Thus far, we have reviewed several proposed models and methods to assess interaction in online learning environments. Some of the methods analyze the subject matter of the messages (*Content Analysis*), others analyze the number and length of comments per student (*Knowledge Construction*), or the type of interaction (*Sequential Analysis*). It appears that the main challenge of assessment in online collaborative environments is to quantify the relation between interaction (or process) and learning (or achievement).

Conclusion

Although there are a number of methods of assessment that are available to instructors in traditional classroom environments, relatively few of these are valid in the online milieu. This may explain why instructors spend much more time assessing online learning than face-to-face environments (Oosterhof,

Conrad & Ely, to appear). What is needed, then, are effective, efficient, and practical measures of online interaction—such as Wiley's proposed measure—to reduce the workload of the online instructors and to provide more meaningful feedback to students.

We must acknowledge that one of the factors that drives most students' learning is assessment—"the power of assessment as a means of directing student learning cannot be under-valued" (McKellar, 2002). Accordingly, we should balance the weighting of the factors that contribute to a student's final grade according to our confidence in the theoretical constructs that underlie each factor. In designing the assessment for online courses, we may award more points to postings and discussions (i.e. interactions). The result may be an increase in the quality and quantity of postings, and subsequently, the depth of intellectual attainment.

References

AFLA (Australian Flexible Learning Framework) (no date) "Is teaching online different to teaching face to face?" Retrieved December 2, 2003 from http://www.tafe.net/learnscope/dlo/delivering1.html

Aviv, R., Erlich, Z., Ravid, G., and Aviva G. (2003) "Network Analysis of Knowledge Construction in Asynchronous Learning Networks" Journal of Asynchronous Learning Networks. Volume 7, Issue 3 - September 2003, Retrieved December 2, 2003 from http://www.aln.org/publications/jaln/v7n3/v7n3_aviv.asp

Bakeman, Roger and Gottman, John M. (1986) Observing Interaction: An Introduction to Sequential Analysis. Cambridge University Press.

Buckley, Jeanne and Toto, Roxanne (2000) "Assessment Techniques for Web Based Instruction: Lessons Learned From Teaching a Graduate Course in Instructional Technology at Philadelphia University" International Conference on Learning with Technology, March 9, 2000; Philadelphia, PA. Retrieved October 29, 2003 from http://www.l2l.org/iclt/2000/papers/166a.pdf

Clark, Ruth Colvin and Mayer, Richard E. (2003) *e-Learning and the Science of Instruction*. Pfeiffer.

Dick, W. & Carey, L. (1996). The systematic design of instruction. (4th ed.). New York: Harper Collins College Publishers.

Eugenia, M. W. Ng (2002) "Enhancing Collaborative Learning through Online Discussion and Peer assessment" Proceedings of the International Conference on Computers in Education (ICCE'02). Retrieved October 29, 2003 from http://csdl.computer.org/comp/proceedings/icce/2002/1509/00/15090262.pdf

Garton L., Haythornwaite C., & Wellman B., (1997) "Studying Online Social Networks", Journal of Computer-Mediated Communication, Vol. 3, No. 1, June 1997. Retrieved October 29, 2003 from http://www.ascusc.org/jcmc/vol3/issue1/garton.html

Gass, Robert (no date), "Content Analysis" Retrieved December 2, 2003 http://commfaculty.fullerton.edu/rgass/content_analysis1.htm

Gunawerdena, C.; Carabajal, K.; and Lowe, C. (2001). "Critical Analysis of Models and Methods Used to Evaluate Online Learning Networks" Paper presented at the Annual Meeting of American Educational Research Association (Seattle, WA, April 10-14, 2001)

Gunawerdena, C.; Lowe, C. and Anderson, T. (1997) "Analysis of a global online debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing" Journal of Educational Computing Research, 17(4), 395-429

Gunawerdena, C.; Lowe, C.; and Carabajal, K. (2000). "Evaluating Online Learning: Models and Methods" in Society for Information Technology & Teacher Education International Conference: Proceedings of SITE 2000 (11th, San Diego, California, February 8-12, 2000).

Gunawerdena, Charlotte (1999). "The Challenge of Designing and Evaluating 'Interaction' in Web-based Distance Education" in Webnet 99 World Conference on the WWW and Internet Proceedings (Honolulu, Hawaii, October 22-30, 1999)

Hill, J. R.; Wiley, D.; Nelson, L.M.; and Han, S. (2004) Exploring Research on Internet-Based Learning: From Infrastructure to Interactions. In *The Handbook of Research on Educational Communications and Technology*, Edited By David H. Jonassen, 2nd edition, 433-460.

Jeong, A. (2001). "Sequential Analysis of Group Interaction and Critical Thinking in Online Threaded Discussions". The American Journal of Distance Education, 17(1), 25-43.

Liu, C.; Liu, B.; Hui, T.; and Horng, J.; (2002) "Web Based Peer Assessment Using Knowledge Acquisition Techniques: Tools for Supporting Contexture Awareness" Proceedings of the International Conference on Computers in Education (ICCE'02). Retrieved October 29, 2003 from http://csdl.computer.org/comp/proceedings/icce/2002/1509/00/15091338.pdf

Mappin, David (2001) "Goals and Objectives" Retrieved December 2, 2003 from http://www.quasar.ualberta.ca/edit489/modules/edpy489_8_Goals.htm

McKellar, E.J.K. (2002). Change our assessment practices? Why should we? The theory behind assessment practices. Paper presented at the Learning Communities and Assessment Cultures Conference organised by the EARLI Special Interest Group on Assessment and Evaluation, University of Northumbria, 28-30 August 2002.
http://www.leeds.ac.uk/educol/documents/00002239.htm

McLoughin, Catherine and Luca, Joe (2001). "Quality in Online Delivery: What Does It Mean for Assessment in E-Learning Environments?" in Meeting at the Crossroads. Proceedings of the Annual Conference of the Australasian for Computers in Learning in Tertiary Education (ASCILITE 2001) (18th, Melbourne, Australia)

Oosterhof, A., Conrad, R., & Ely, D. (to appear). *Assessing Online Learners*. Prentice Hall.

Resta, Paul (2003) "Self and Peer Assessment in an Online Collaborative Learning Environment" 18th Annual Teaching Seminar for Experienced Faculty, January 6 & 7, 2003, Retrieved December 2, 2003 from http://www.utexas.edu/academic/cte/efc/efs2003/Resta/index.htm

Rourke, L., Anderson, T., Garrison, D. R., & Archer, W. (2001) Methodological Issues In The Content Analysis Of Computer Conference Transcripts. *International Journal of Artificial Intelligence & Education*, 12. Retrieved November 15, 2003 from
http://computing.unn.ac.uk/staff/cgpb4/ijaied/members01/archive/vol_12/rourke/full.html

Takle, E. S., Sorensen, E. K., and Herzmann, D. (2003) "Online Dialog: What if you build it and they *don't* come" 12th Symposium on Education (12EDU). Retrieved November 12, 2003 from
http://www.meteor.iastate.edu/faculty/takle/AMSEd2003.pdf

The Writing Center (2003), "Content Analysis" Colorado State University, Retrieved December 2, 2003 from
http://writing.colostate.edu/references/research/content/

Wiley, David A. (2002). "A proposed measure of discussion activity in threaded discussion spaces." Working Draft Retrieved October 15, 2003 from http://wiley.ed.usu.edu/docs/discussion09.pdf

Yu, F.; Liu, Y.; and Chan, T. (2002) "The Efficacy of a Web-Based Domain Independent Question-Posing and Peer Assessment Learning System" Proceedings of the International Conference on Computers in Education (ICCE'02). Retrieved October 29, 2003 from
http://csdl.computer.org/comp/proceedings/icce/2002/1509/00/15090641.pdf