Beyond the Testing Paradigm: Towards New Assessment Measures in Knowledge Building Environments

Donald N. Philip

don.philip@utoronto.ca

and

Danielle Truswell

danielle.truswell@utoronto.ca

Abstract

In this paper we introduce a number of possible assessment measures for knowledge building, and illustrate these using data collected from a third-year university science education pre-professional course.

The first measure is a knowledge building profile that uses data mined from the Knowledge Forum online knowledge building environment to create a bar graph that compares an individual's performance on a set of knowledge building measures with the class average. A second kind of knowledge building profile at the group level uses response (build-on) tree size to indicate the level and complexity of responding for individual views or the duration of the class.

The second set of measures uses a social network concentric display of centrality measures to gauge a students influence, outreach to others, and overall importance to the group. There are three levels to the display, and students who are central, medial and peripheral can be easily identified.

Keywords: Online learning, knowledge building, assessment, performance based assessment

Introduction

It has become obvious that the future prosperity of any society is inextricably tied to education (UNESCO, 2005). With an increasing awareness of the importance of education comes a concern for accountability–usually in the form of testing. However, testing is not the only manner of assessment, and once student work is done in an online environment, new possibilities emerge that can take us beyond the testing paradigm. This paper will explore some considerations of assessing online work, and present some candidates for new forms of assessment in online environments.

Christensen's Theory of Disruptive Innovations

One danger of innovation is that as Machiavelli noted (1998, electronic edition, Chapter 6), "[T]he innovator has for enemies all those who have done well under the old conditions, and lukewarm defenders in those who may do well under the new." One of the ways this can manifest itself is a tension between sustaining and disruptive innovations.

Christensen, a business guru, has recently applied his theory of disruptive innovations to education (Christensen, Horn, & Johnson, 2008). *Sustaining innovations* are the most common, and serve to sustain, support, and extend existing practices. *Disruptive innovations*, in contrast, are those that create new paradigms and new ways of doing things. Disruptive innovations often fail or become sustaining innovations due to the phenomenon of *cramming*, a process in which disruptive technologies are inserted into an existing system to co-exist with current practice. This often forces the disruptive innovation to become a sustaining innovation. Christensen et al. (2008) note that this is the case with education: schools have been forced to accommodate disruptive change within their existing traditional practices.

Therefore a tension arises from the desire to reform education for the knowledge age, but at the same time do so by cramming new technologies into the traditional model–an attempt that is doomed to failure. New practices must be assessed in new ways or they will not succeed. Assessments in online environments falls into this category: they must be treated as a disruptive innovations and not allowed to become a sustaining innovations or they will not reveal their true potential.

Assessing Online Learning

Concerns

Speaking broadly, concerns about the assessment of online learning fall into three camps: those concerned about cheating in online tests, those concerned with the effectiveness of online assessment compared to traditional methods, and those concerned about how online learning should be assessed. Figures 1a and 1b demonstrate the first concern:

a.

b.

Figure 1. Scott Adams' Dilbert cartoons highlighting the concerns about cheating in online courses.

Benton (2006) reports a concern with cheating among the Texas school board. Palloff and Pratt (1999) similarly note a concern with cheating.

The effectiveness of online learning is a cause of concern among accreditation bodies (Carnevale, 2001). This often creates a tension between sustaining and disruptive innovation, as online assessment is compared to traditional assessments in such situations. However, Scalise and Gifford (2006) consider online assessment to be the equivalent of testing.

The Horizon Project (Horizon Project, 2007) notes that online work can exist in new forms and that this presents assessment challenges for faculty.

Online Assessment

Many researchers consider that the advantages of online assessment outweigh the disadvantages. A number of researchers consider that an important aspect of online assessment is the measurement of collaborative community formation (Palloff & Pratt, 2001; Wellman, Koku, & Junsinger, 2006). Wellman et al. (2006) note the potential for social network analysis to be useful in measuring community formation.

Davies and Graff (2005) note the importance of participation in the online portion of hybrid classes. They found that students who participated online discourse did better overall than those who did not.

Bennett (2002, p. 4) notes that, "As technology becomes more central to schooling, assessing students in a medium different from the one in which they typically learn will become increasingly untenable." Students working online should be assessed with tools built into that online environment.

Jonassen suggests (1996) that we are obligated to assess constructive outcomes (performance) rather than reproductive learning. Noting that critical thinking is dependent on context, he suggests that different assessment measures need to be developed for different cognitive tools.

Lax et al. (2006) have made use of built-in measures in the Knowledge Forum online knowledge building environment (described below) and found that their use was transformative of student behaviour. Likewise, Philip (2009) similarly found that the use of the build-in analysis tools in Knowledge Forum was transformative of student and teacher behaviour (pp. 180-181).

Crane et al. (2005) are concerned with 21st century skills. They note (p. 10) that we need to use 21st century tools to assess 21st century skills. Silva (2008) notes the difficulty of measuring 21st century skills, and raises an interesting point: the need for open-ended performance based assessment metrics. This concept is discussed further in the next section.

Performance Based Assessment

In discussing online assessment, we are discussing a paradigm shift away from the testing paradigm and towards a performance based metric approach, hence the title of this paper. This represents a considerable change in approach to conventional assessment, and it will be difficult to get support for this. Therefore it is important to note that many people have already had considerable experience with performance based assessment through sports.

Baseball is a good example of this. Currently, there are twenty-one officially recognized baseball statistics, including Hits, At Bats, Batting Average, Batter Strikeouts, Earned Run Average (ERA), and so forth (Mason, 1998). None of these measures generates a numeric mark. Instead they create a *profile* of a player's performance on critical tasks, and one that many are familiar with. Familiarity with such assessments would be useful in introducing the concept of performance based assessment to educational stakeholders.

The University of California has developed a system for generating performance based metrics for use in laboratory projects involving the U.S. Department of Energy (DOE) (Oak Ridge Associated Universities, 2005). This system could be usefully applied to the development of any performance metrics created to assess online learning.

Knowledge Building and the Knowledge Forum Online Knowledge Building Environment

Knowledge building, also called knowledge creation, is a process by which the cultural capital (knowledge) of a society is *increased*, producing new knowledge; learning is a process by which the cultural capital is *distributed* (Scardamalia & Bereiter, 2003).

Typically, the model used in school classes engaging in knowledge building is that of the research community among which ideas are spread, elaborated, and used to create new knowledge.

Fleck (1981, original German publication 1935), describes the process of idea spread in terms of *thought collectives* which are, *"... a community of persons mutually exchanging ideas or maintaining intellectual interactions ... [which] also provided the special 'carrier' for the historical development of any field of thought as well as for the given stock of knowledge and level of culture"* (p. 39, original emphasis). Ideas never spread without some change as Fleck further notes: "Thoughts pass from one individual to another, each time a little transformed, for each individual can attach to them somewhat different associations. Strictly speaking, the receiver never understands the thought exactly in the way that the transmitter intended it to be understood" (p. 42). In fact, unlike traditional views of the knowledge creation process, in which the individual is viewed as creating knowledge in isolation (Dunbar; Sawyer, 2007), Fleck notes that an innovative idea is not the property of a single individual, but rather belongs to the collective (1981).

Knowledge building therefore is a collaborative process for the creation of new knowledge through the elaboration and transformation of ideas among a community engaged in problems solving.

Knowledge Forum: An environment for knowledge building

Starting as *CSILE* (Computer Supported Intentional Learning Environment,) Knowledge Forum was designed to,

> (a) make advanced knowledge processes accessible to all participants, including children, (b) foster the creation and continual improvement of public artifacts or community knowledge (Scardamalia, 2002), and (c) provide a community space for carrying out this knowledge building work collaboratively. (Scardamalia, 2004, p. 183)

Structured as a database, Knowledge Forum creates public work spaces called views in which individuals can post their ideas, theories, information, etc. in the form of notes that are available for other members of the community (often a school class) to read, comment upon, and elaborate.

One key aspect of Knowledge Forum in its current form is the provision of assessment tools so that teachers and students can evaluate their online performance and make changes in it if necessary, a feature termed *embedded, transformative assessment* (Scardamalia, 2004). At present this takes two forms: a set of built-in assessment tools, and the Analytic Toolkit (ATK) that provides more detailed information about student interactions with Knowledge Forum through the use of server log data. The ATK is not readily available to students, and for the most part provides raw, unanalyzed data.

Methodology

SCI396 is the second part of a two-part pre-professional course for undergraduate science students considering a science teaching career. This group, SCI395 (the first part of the course) was their first introduction to online learning. SCI396, the winter term, continued the course, but most of the students have only had one term of exposure to online learning, and some students who took the courses out of order had had no previous exposure.

We used the ATK measures to analyze student interactions with KF and with each other using the metrics proposed earlier. Data were gathered automatically, and anonymized. Of the 24 students in the class, 20 agreed to participate (83%.)

As well as the ATK data collected, the students were surveyed to obtain some basic demographic data.

Results

Demographics

Most students found Knowledge Forum easy to use, with only one students being neutral on this and no one dissenting. The majority of students also agreed that the use of Knowledge Forum enhanced their learning experience, with one student disagreeing.

Results for the Proposed Metrics

This section is not intended to be an exhaustive examination of the SCI396 class, but instead to provide examples, using SCI396 data, of what the proposed assessment measures would look like, and how they might be applied.

Knowledge Building Profile

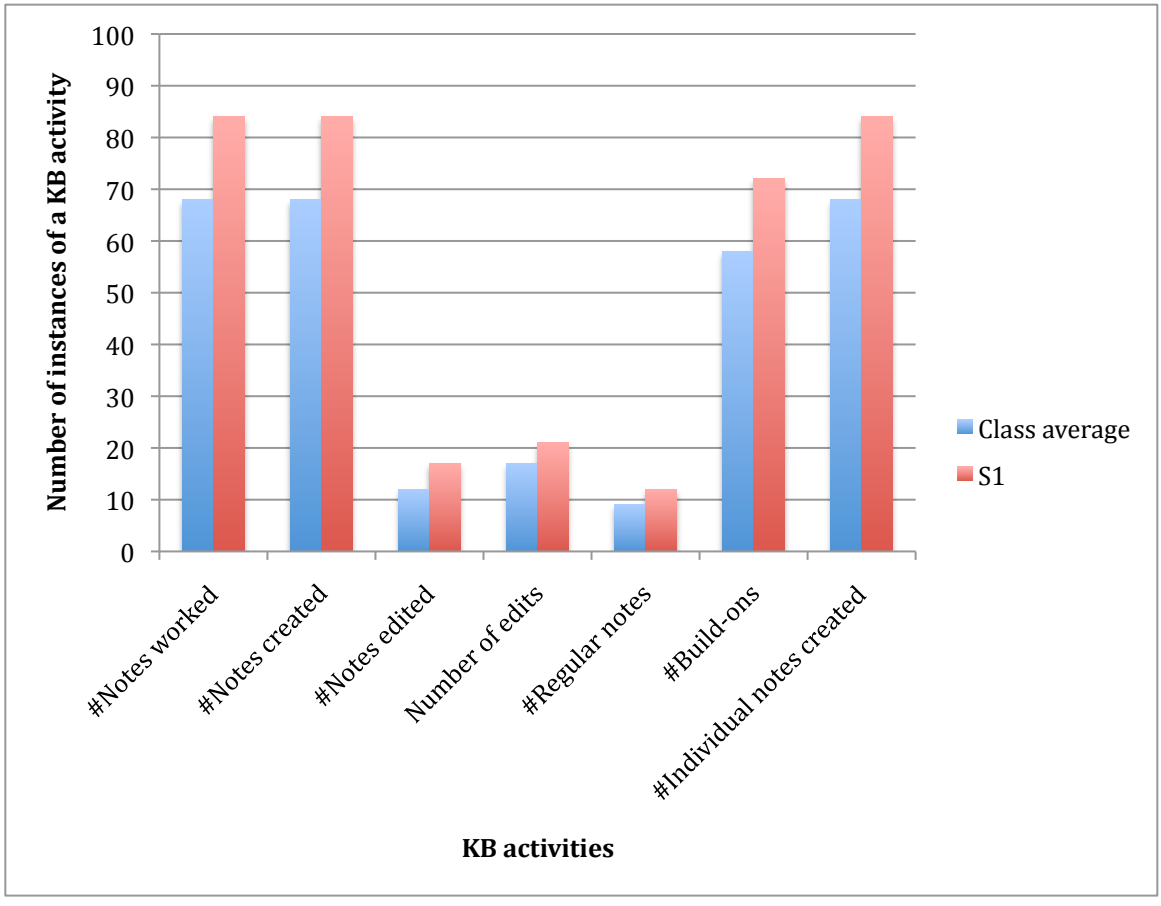

Figure 2. The knowledge building profile for student 1, compared to the class average.

Figure 2 shows the number of times student 1 engaged in knowledge building activities (Y-axis) against the listed KB activities (X-axis). As can be seen, student 1 (red bars) was consistently above class average (blue bars) in all activities.

Such a KB profile appears to us to be an easy measure to read and understand, both for teachers and students. It would provide a teacher with a useful tool to evaluate individual student behaviours against the class average or benchmarks as these become available.

Group Metric: Build-on Tree Profile

A comparison of the size of the build-on trees created in weekly views can provide insight into the complexity of the discourse that is occurring within the community. As can be seen in Figure 3, a comparison of the build-on trees created at the beginning, middle and end of the term demonstrates a progressive increase in size. This information becomes more meaningful when combined with the knowledge of the increase in the ratio of student to instructor postings. The students appear to be carrying out more complex discussions with less direction from the instructor. This is likely a consequence of experience. Tracking the size of build-on trees could be useful for instructors in order to gauge the development of the group's discourse skills and independence.

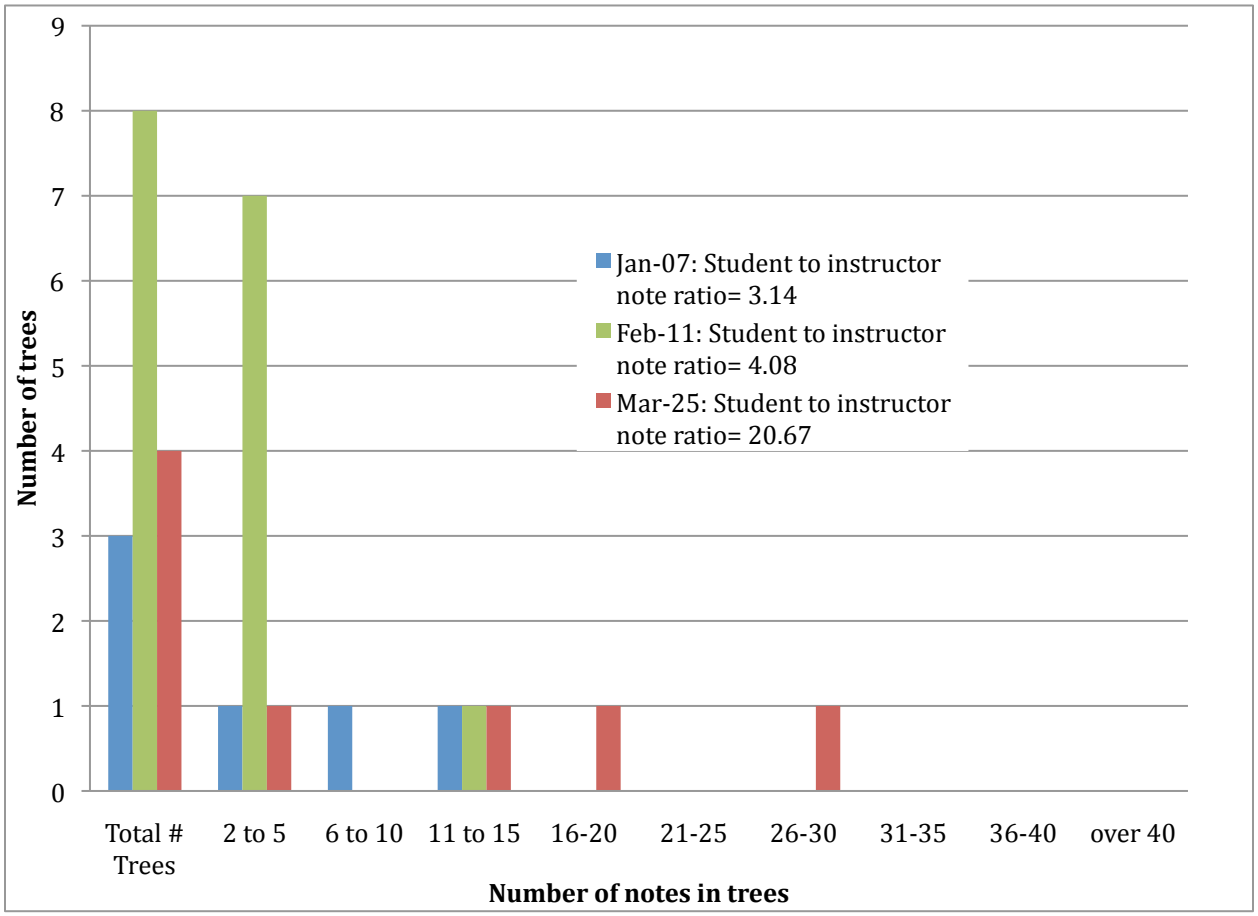

Figure 3. Graph comparing number of build-on trees of various sizes in three weekly views.

Social Network Measures: Influence, Outreach, and Importance

A social network tool is already built into Knowledge Forum. The displays here use much the same data as the social network tool, but display it differently. We feel that such displays will be simple for teachers and students to read and understand, and provide valuable information about individual and class dynamics. There are three we will consider: influence, outreach, and importance.

The influence display is based on indegree centrality; the outreach display is based on outdegree centrality; and the importance display is based on Eigenvector centrality[1] (Lohmann, et al., 2010). Each measure is a well-researched social

[1] Eigenvector centrality is based on a similarity matrix, one that expresses the similarities between two points, rather than on the adjacency matrix used to compute indegree and outdegree centrality. Connections to high-scoring nodes are

network measure in common use. We feel that calling the measures influence, outreach, and importance will be more meaningful to students and teachers than indegree, outdegree, and Eigenvector centralities, as they will be less concerned with the mathematics behind the measures than the interpretation of them. Since all three of these measures look similar, we will consider only one here, and give samples of the others in the Appendix.

Importance

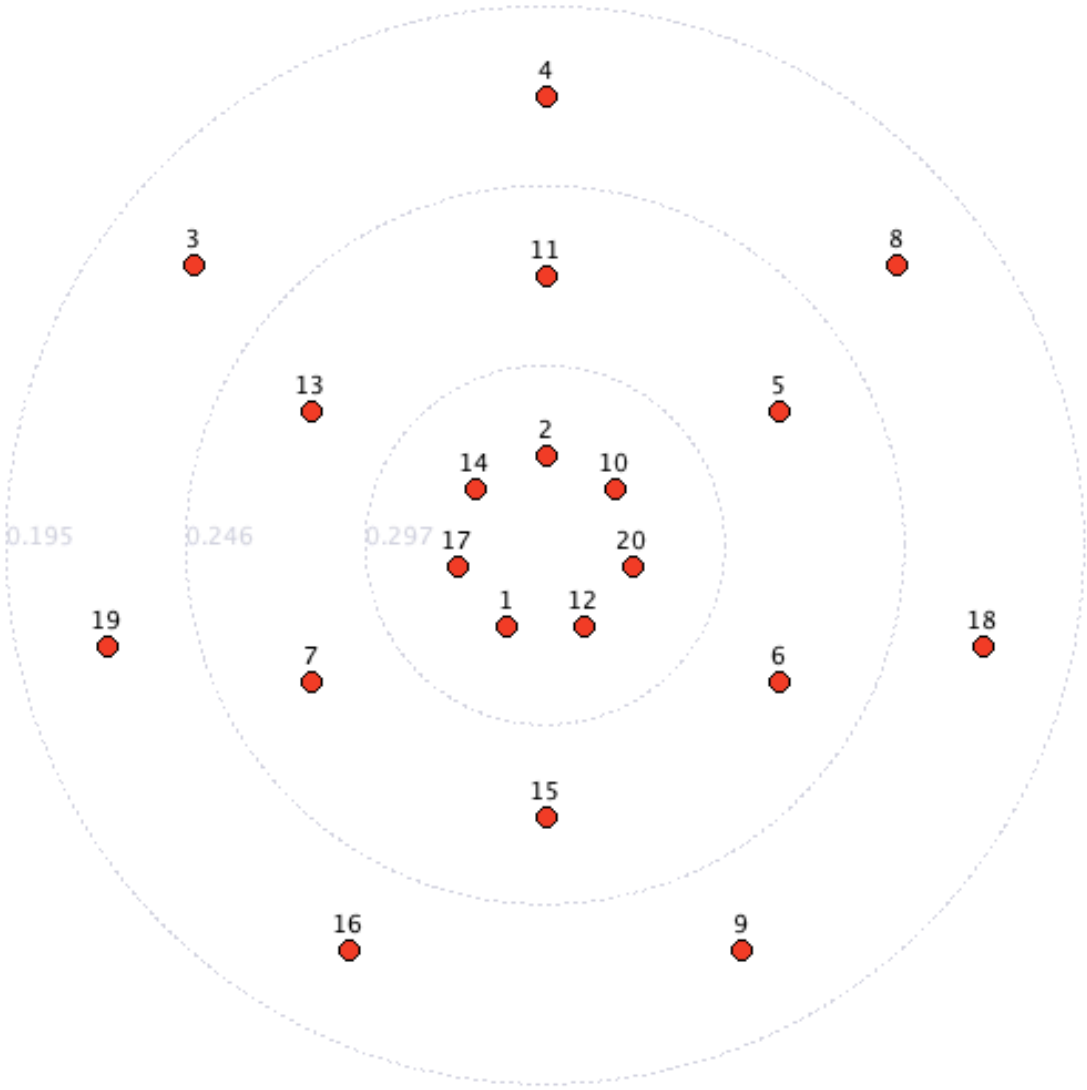

Figure 4. The *importance* concentric social network display.

---

given a higher weight than connections to lower-scoring nodes. Google uses this in its "page rank" algorithm.

The importance display shows the students displayed in three concentric rings: a core group, the most central and therefore important; a medial group; and a peripheral group. The use of three rings was chosen because this aligns with the concept of legitimate peripheral participation (Lave & Wenger, 1999) with which many teachers are familiar. We find such displays easy to read and feel it will potentially useful to teachers who can easily identify students who may not be participating well.

Conclusions

We have presented an introductory exploration of some possible new assessment measures for knowledge building. Exhaustive analysis of these proposed measures was not our intent, as a much larger data set than available to us at this time would have been required. We have aimed instead to introduce a number of possible measures and demonstrate what they can show and how they might be used with other measures. We would need to conduct a much larger study before promulgating these ideas to a larger public. Nonetheless, we feel that the proposed new measures have value and deserve further consideration.

A key consideration in developing metrics such as those proposed is to avoid allowing them to be converted into sustaining innovations. We need to contextualize any new metrics in terms of knowledge building and knowledge building classes, and not compare them to traditional classes and measures. Online assessments and performance based metrics are disruptive innovations and should be treated as such or they will fail.

References

Bennett, R. E. (2002). Inexorable and Inevitable: The Continuing Story of Technology and Assessment. *The Journal of Technology, Learning , and Assessment (JTLA), 1*(1), 1-24.

Benton, J. (2006, June 23). TAKS online testing option raises concerns. *DallasNews.com (The Dallas Morning News)*. Retrieved from http://www.dallasnews.com/sharedcontent/dws/dn/education/stories/062206dnmettaksonline.16cecc8.html

Carnevale, D. (2001). Assessment takes center stage in online learning. *The Chronicle of Higher Education, 47*(31), A43-A45.

Christensen, C. M., Horn, M. B., & Johnson, C. W. (2008). *Disrupting Class. How disruptive Innovation will change the way the world learns.* New York: McGraw Hill.

Crane, T., Wilson, J., Maurizio, A., Bealkowski, S., Bruett, K., Couch, J., et al. (2005). *Learning for the 21st Century. A Report and MILE guide for 21st Century Skills.*: Partnership for 21st Century Skills.

Davies, J., & Graff, M. (2005). Performance in e-learning: online participation and student grades. *British Journal of Educational Technology, 36*(4), 657-663.

Dunbar, K. How scientists think: On-line creativity and conceptual change in science. In T. B. Ward, S. M. Smith & J. Vaid (Eds.), *Conceptual structures and processes: Emergence, discovery, and change*. Washington, D.C.: American Psychological Association Press.

Fleck, L. (1981). *Genesis and development of a scientific fact* (F. Bradley & T. J. Trenn, Trans.). Chicago: University of Chicago Press.

Horizon Project. (2007). *The Horizon Report, 2007 Edition*: The NEW MEDIA CONSORTIUM and EDUCAUSE Learning Initiative.

Lave, J., & Wenger, E. (1999). *Situated learning: Legitimate peripheral participation.* Cambridge: Cambridge University Press.

Lax, L., Singh, A., Scardamalia, M., & Librach, L. (2006). Self-assessment for knowledge building in health care. [Research report]. *QWERTY*, 19-37.

Lohmann, G., Margulies, D. S., Horstmann, A., Pleger, B., Lepsien, J., Goldhahn, D., et al. (2010). Eigenvector Centrality Mapping for Analyzing Connectivity Patterns in fMRI Data of the Human Brain. *PLoS One, 5*(4), e10232.

Machiavelli, N. (1998). The Prince. *Project Gutenberg Etext* Retrieved Nov. 19, 2004, from http://www.gutenberg.org/etext/1232

Mason, P. (1998). Baseball Statistics. The Numbers Used to Legitimize a Boy's Pastime [online]. *Journal of the Cosmic Baseball Association, 17*.

Oak Ridge Associated Universities. (2005). Developing Performance Metrics: University of California Approach [online]. Retrieved May 7, 2010, from http://www.orau.gov/pbm/documents/overview/uc.html

Palloff, R. M., & Pratt, K. (1999). *Building Learning Communities in Cyberspace. Effective Strategies for the Online Classroom*. San Francisco: Jossey-Bass.

Palloff, R. M., & Pratt, K. (2001). *Lessons from the Cyberspace Classroom. The realities of online teaching.* San Francisco: Jossey-Bass.

Philip, D. N. (2009). *Networks and the Spread of Ideas in Knowledge Building Environments.* Ontario Institute for Studies in Education of the University of Toronto, Toronto.

Sawyer, R. K. (2007). *Group genius : the creative power of collaboration*. New York: Basic Books.

Scalise, K., & Gifford, B. (2006). Computer-Based Assessment in E-Learning: A Framework fo Construction "internediate Constraint" Questions and Tasks for Technology Platforms. *The Journal of Technology, Learning , and Assessment (JTLA), 4*(6), 4-44.

Scardamalia, M. (2004). CSILE/Knowledge Forum®. In Kovalchick & K. Dawson (Eds.), *Education and Technology: An encyclopedia* (pp. 183-192). Santa Barbara: ABC-CLIO.

Scardamalia, M., & Bereiter, C. (2003). Knowledge Building. In J. W. Guthrie (Ed.), *Encyclopedia of Education, Second Edition* (pp. 1370-1373). New York: Macmillan Reference, USA.

Silva, E. (2008). *Measuring Skills for the 21st Century*: Education Sector.

UNESCO. (2005). *Towards Knowledge Societies*. Paris: UNESCO Publishing.

Wellman, B., Koku, E., & Junsinger, J. (2006). Networked Scholarship. In J. Weis (Ed.), *Internation Handbook of Virtual Learning Environments* (Vol. 2, pp. 1429-1447). Dordrecht, NL: Springer.

Appendix

Figure A1 shows the *influence* concentric display, based on indegree centrality, for the building-on (response network.) Figure A2 shows the *outreach* concentric display for the same network.

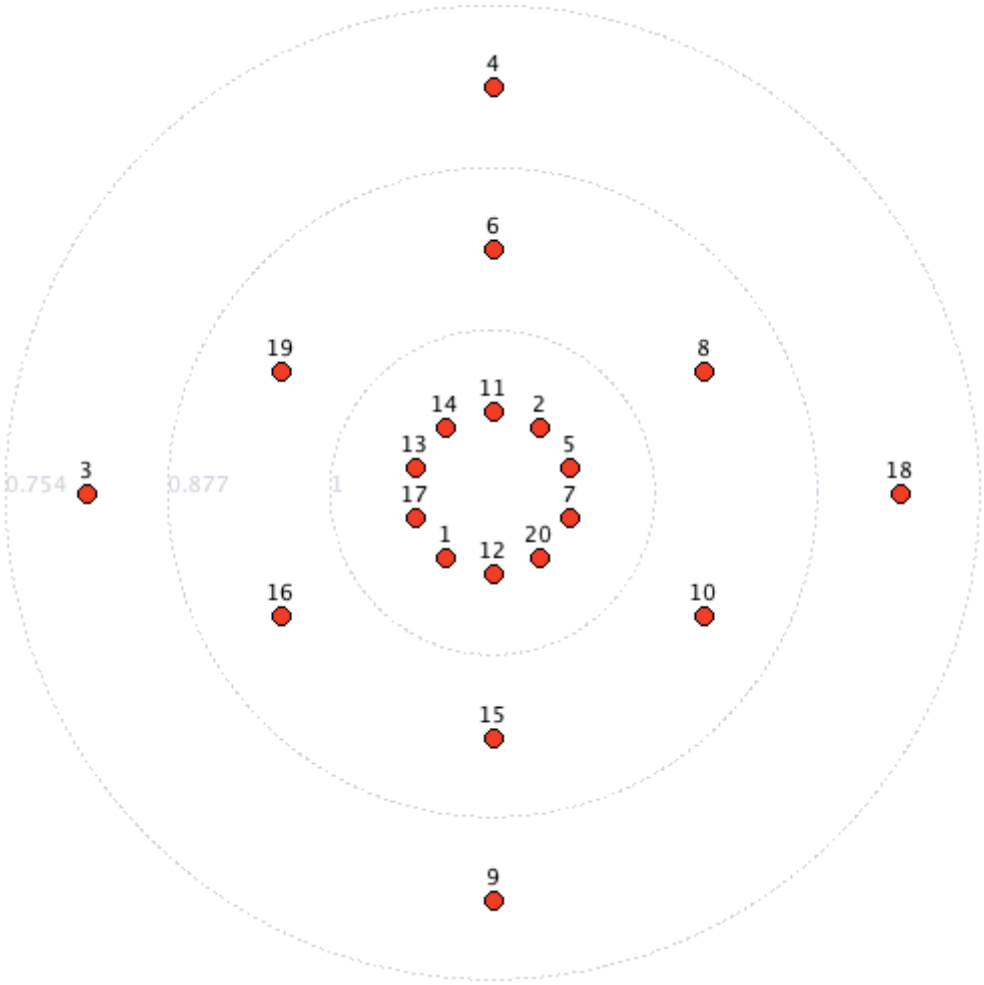

Figure A1. The influence social network concentric display for the building-on network.

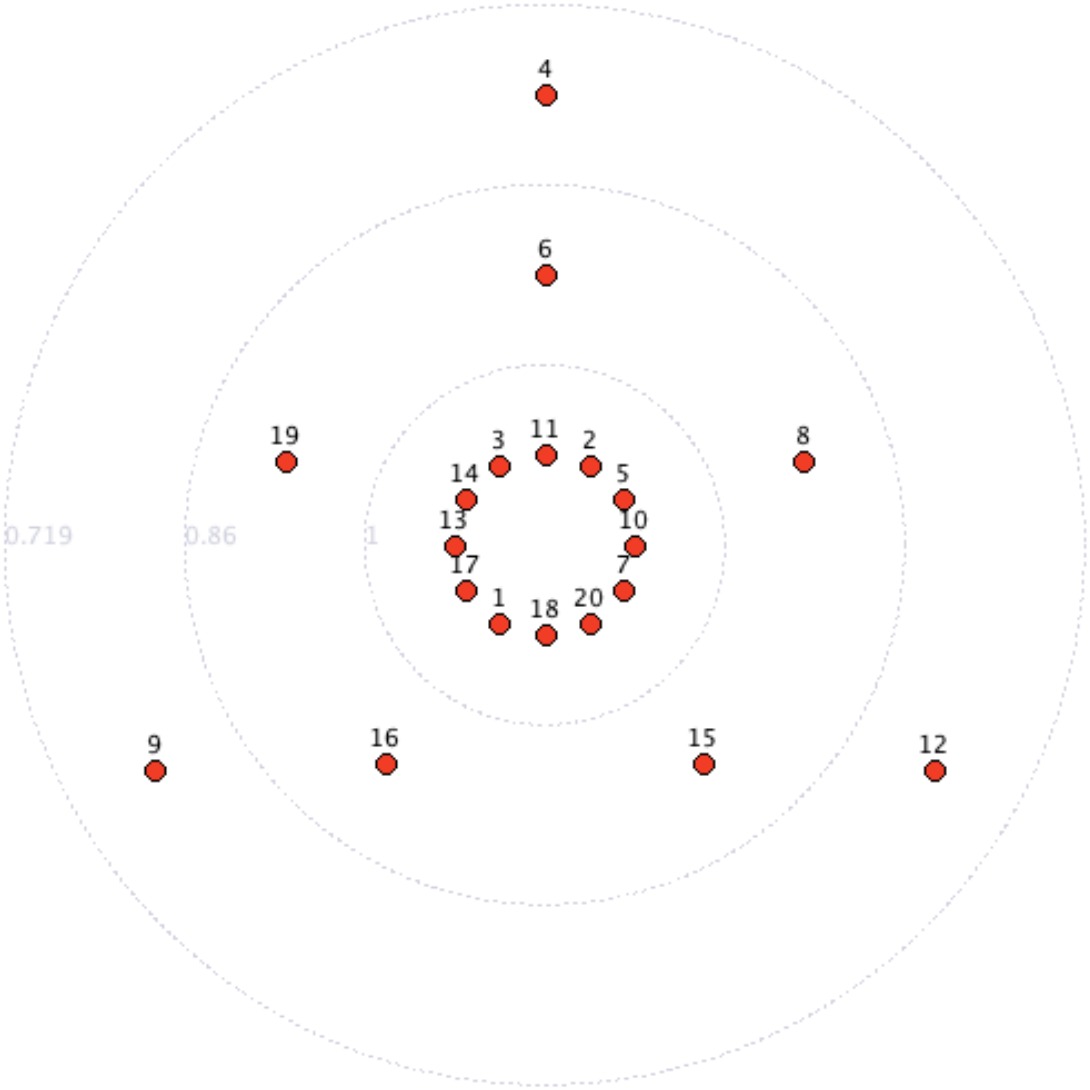

Figure A2. The *outreach* social network concentric display for the building-on network.